Spring 5-6-2012

Evolution of the bacteriorhodopsin gene bop in haloarchaea

Maulik Jitesh Jani
University of Connecticut - Storrs, Maulik.J.Jani@gmail.com

Evolution of the bacteriorhodopsin gene *bop* in haloarchaea

Maulik J. Jani

Evolution of the bacteriorhodopsin gene *bop* in haloarchaea

Maulik Jani, R. Thane Papke

Summary:

Rhodopsins are among the most studied protein families. They all incorporate a light based chromophore and an opsin. They are found in all three domains of life and can be found in diverse environments, such as the membrane of the haloarchaea, *Haloarcula (Haa) marismortui*, growing on salt flats in Death Valley, to within human tissues and deep sea water (Briggs and Spudich 2005). Rhodopsins can serve multifarious purposes, from phototaxis away from harmful light to efficient energy generation. Bacteriorhodopsin is the best studied rhodopsin in haloarchaea. The evolutionary pattern of bacteriorhodopsin through its gene, *bop*, is the main focus of this study.

Planned as a comparative study of *bop* against expected phylogeny constructed on 16S rRNA gene sequences, as well as phylogeny based on a tree constructed from core genes; this study also hoped to find possible examples of Horizontal Gene Transfer (HGT). The *bop* gene was amplified and sequenced from multiple strains, and additional sequences were gathered from the National Center for Biotechnology Information (NCBI), bringing a total of 51 sequences from 41 strains covering 15 recognized genera. Additionally, three sensory rhodopsins were included to serve as an out-group. A phylogeny was reconstructed using Maximum Likelihood (ML) with the GTR plus gamma model, creating an ML tree of *bop* based off of bootstrapped multiple sequence alignment.

Acknowledgements:

This study would not have been possible without the guidance and support of Dr. R. Thane Papke. I would like to thank Dr. Lee Aggison Jr., Dr. Peter Gogarten, and Emma White and the Papke Lab for their support. Also thanks to help from Scott Chimileski, Nikhil Mohan, Charles Mackin and Griffin Weigel for their advice and assistance with gathering the data and writing this thesis.

Thank you to the University of Connecticut and the Honors Program for providing me with these wonderful opportunities and for their assistance throughout my time here.

The UConn Bioinformatics Facility provided computing resources and guidance for the Maximum Likelihood phylogenetic analyses performed for this study.

Thanks to the following Funding Sources:

Office of Undergraduate Research (OUR) Grant

Presented to: Maulik Jani

Summer Undergraduate Research Fund (SURF) Grant

Presented to: Maulik Jani

Honors Life Science Thesis Award

Presented to: Maulik Jani

National Science Foundation

Presented to: R. Thane Papke

Table of Contents

List of Figures

List of Tables

1. Introduction:

1.1 Rhodopsins in General

The topic of this study is the evolutionary pattern of bacteriorhodopsin, a member of the rhodopsin family of proteins. Rhodopsins are found in all three domains of life and, though they have many different forms and functions, all have a similar structure. Rhodopsins are now used in everything from genetic studies (Sharma, Walsh et al. 2007) to serving as a form of data storage, though this has only been possible in storing small amounts of data (Stuart, Marcy et al. 2002). All rhodopsins consist of seven trans-membrane helixes with a chromophore covalently linked to a lysine residue always on the seventh helix of the opsin (Birge 1990; Jekely 2009). An opsin moiety is a seven-pass transmembrane helix (Birge 1990). The chromophore is generated from a carotenoid with vitamin A activity. Carotenoids are converted into retinal, which is the chromophore found in all rhodopsins (Birge 1990; Sharma, Spudich et al. 2006; Sung and Chuang 2010). Opsins are 35-55 kDa membrane-bound proteins. The various different opsins trigger different downstream reactions when excited (Sung and Chuang 2010). Some pump ions across the membrane, as in

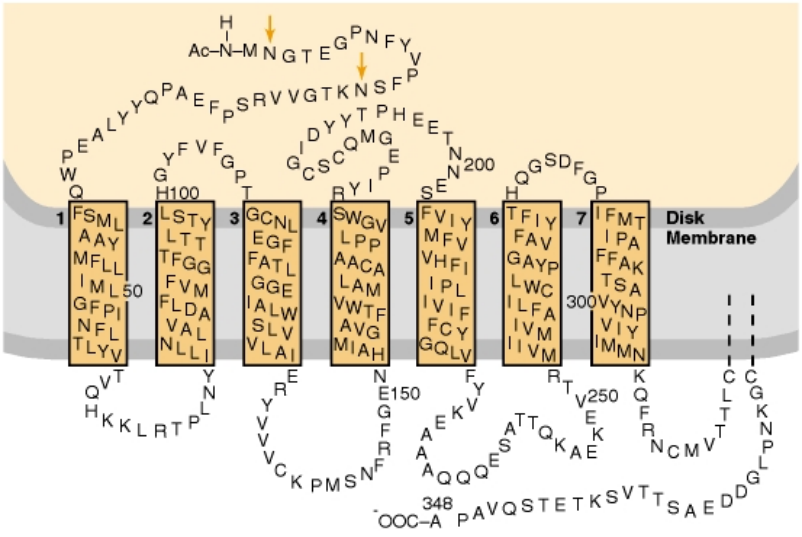

Figure 1 – Rhodopsin Protein Structure
The basic structure of rhodopsin proteins shows the seven helixes and interhelical loops. Additionally the lysine residue can be seen at 296. The image is of bovine type 2 rhodopsin in particular. Image provided by Siegel *et al.* (Siegel 1999)

the case of bacteriorhodopsin, while others directly interact with a variety of attached G-proteins, having effects as diverse as activating flagellar motors and shutting down ion gated channels (Spudich 1998; Sharma, Spudich et al. 2006; Sung and Chuang 2010). An example of a rhodopsin can be seen in Figure 1.

There are a number of different rhodopsin proteins that are part of a large protein family. The Structural Classification of Proteins (SCOP) (Murzin, Brenner et al. 1995) considers rhodopsins the Superfamily, "G protein-coupled receptor-like" which includes two families (Andreeva, Howorth et al. 2004). The first is rhodopsin-like which consists of rhodopsins found in the mammalian eye (Andreeva, Howorth et al. 2004). The second family is bacteriorhodopsin-like which consists of rhodopsins found in haloarchaea such as bacteriorhodopsin, both sensory rhodopsins and halorhodopsin (Andreeva, Howorth et al. 2004). The placement of these two families into one Superfamily suggests that SCOP considers them to have an evolutionary relationship based on their structural and functional similarities (Andreeva, Howorth et al. 2004). However, different families signify a lack of sequence similarity. The SCOP placement means that while these two groups are homologous, they lack sequence similarity support, though this is not required for homology. Even the important residues conserved within the groups do not overlap, however (Spudich, Yang et al. 2000).

Rhodopsins can also be placed into two separate groups: type 1 and 2 rhodopsins (Spudich, Yang et al. 2000). Type 1 rhodopsins are generally archaeal rhodopsins, originally found in haloarchaea (Schlesner, Miller et al. 2009). They are very widespread and can be found in nearly all taxa other than land plants and animals. Type 2 rhodopsins, on the other hand, are generally found in eukaryotic organisms (Spudich, Yang et al. 2000). The key difference between the two types of rhodopsin is the conformational change the chromophore undergoes

when it is activated by light. Type 1 rhodopsins' resting state is all-trans and photoisomerizes into a 13-cis state, which triggers the opsin into action (Schlesner, Miller et al. 2009). On the other hand, the chromophore of type 2 rhodopsins isomerizes from an 11-cis to an all-trans state (Schlesner, Miller et al. 2009). The differences between the isomers can be seen in Figure 2. All type 2 and the sensory type 1 rhodopsins have their retinal found in a single isomer, while the transport type 1 rhodopsins are found in a mix (Hoff, Jung et al. 1997). Size is a major difference between the two types, with type 1 being approximately half the size of type 2 rhodopsins (Spudich, Yang et al. 2000). All rhodopsins reviewed will be presented as members of one of these types.

1.2 Homology Between Type 1 & 2 Rhodopsins

Rhodopsins are similar in structure as well as in general function, though they do have differential sequences. In both types, as light hits the retinal it undergoes a structural change causing the linked opsin to trigger, though the changes are different (Sharma, Spudich et al. 2006; Sung and Chuang 2010). The differences between the opsins leads some researchers to believe the two types are not homologous. This claim is commonly supported by the lack of sequence similarity between the two, which is supported by different photochemistry and transduction pathways of the two types (Schlesner, Miller et al. 2009). However, a lack of sequence similarity does not prove a lack of homology. Other features, such as similar 3D structure, and a number of other similarities suggest homology (Schlesner, Miller et al. 2009). Homology is also likely since the other option is convergent evolution, in which the two evolved from separate ancestors and simply have these features, in common, which is rare on such a large

scale. The significant similarities between the two in terms of mode of action and structure are the primary reasons why it is believed the original rhodopsin existed in the ancestral organism and spread over time through conventional means with assistance from horizontal gene transfer events.

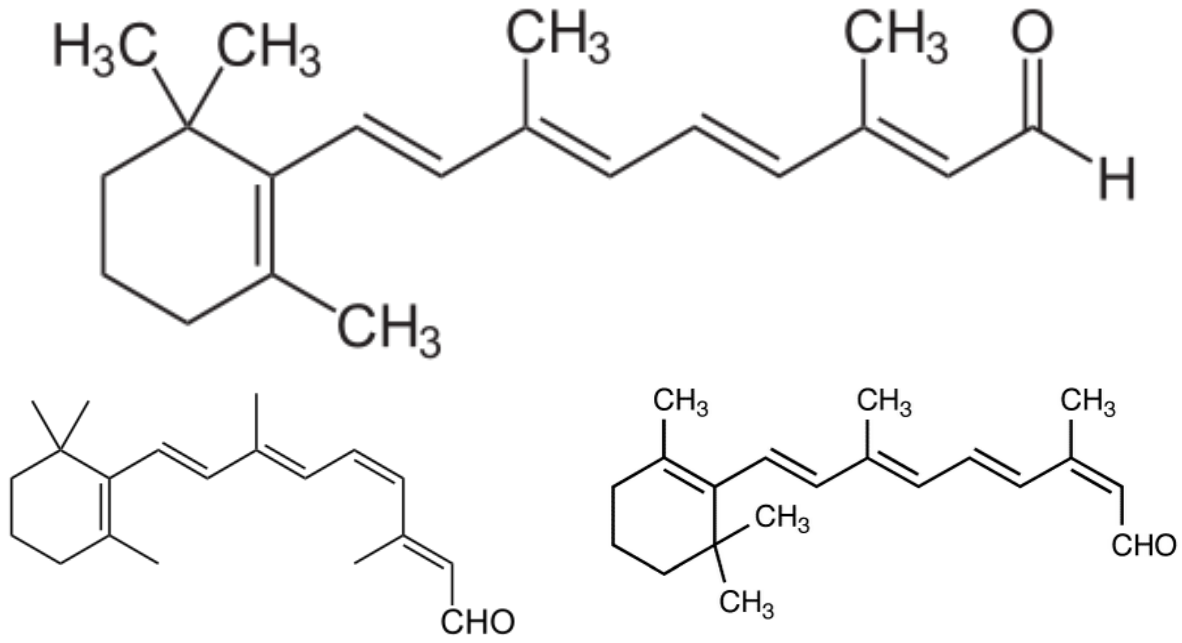

Figure 2 – Retinal Isomers

These are the chemical structures of the retinal isomers that exist in various rhodopsins. On top is all-trans retinal found in both Type 1 & 2 rhodopsins. Below and to the left is 11-cis retinal found in Type 2 rhodopsins and to the right is 13-cis which is found in the Type 1 rhodopsins. Images constructed though eMolecules.com

While most of the support for non-homology between type 1 and 2 rhodopsins is based on sequence similarity some support, however, has come to light that the similarities are from convergent evolution (Jekely 2009). Type 2 rhodopsins can be traced back to a cnidarian–bilaterian ancestor (Jekely 2009). Furthermore, non-opsin G-protein coupled receptors (GPCR) trace back to the eukaryotic last common ancestor (Jekely 2009). These receptors are much more similar to type 2 rhodopsins than they are to type 1 rhodopsins (Jekely 2009). The first true type 2 rhodopsin does not appear in eukaryotes until reaching cnidarians and this rhodopsin is more similar to the non-opsin GPCR than to type 1 rhodopsins (Jekely 2009). Each of these groups is

strongly supported by sequence similarity, but sequence conservation is lacking between the types. This strongly suggests that type 2 rhodopsins evolved from a non-opsin GPCR as opposed to evolving from type 1 rhodopsins (Jekely 2009). And that the similarities between the two types arise from convergent evolution.

1.3 Type 1 – "Microbial Rhodopsins"

Type 1 rhodopsins are rhodopsins commonly found in microbial life. They have also been called archaeal rhodopsins, due to their discovery in halophilic archaea (Ruiz-Gonzalez and Marin 2004). Since then, other rhodopsins with the same retinal have been found. These rhodopsins fall into two large groups: ion transporting rhodopsins and phototaxis oriented rhodopsins (Sharma, Spudich et al. 2006). Ion transporters include either proton based rhodopsins, like proteorhodopsin, bacteriorhodopsin and fungal rhodopsins, or they are chloride pumps, which consist mainly of halorhodopsin. Other type 1 rhodopsins function as sensors for phototaxis, such as the sensory rhodopsins (Ruiz-Gonzalez and Marin 2004).

Proteorhodopsins are light-driven proton pumps that are involved in extensive phototrophy (Slamovits, Okamoto et al. 2011). Proteorhodopsins are suspected to have a significant function in the ecology of the oceans by making energy available for microbial metabolism. The genes for Proteorhodopsin were first revealed through the sequencing and cloning of large 130-kb genomic fragment of uncultured γ-*Proteobacteria* isolated from seawater (Beja, Aravind et al. 2000; Giovannoni, Bibbs et al. 2005). The gene was later found to be phylogenetically varied and extensively distributed, as it was found in the Eastern Pacific Ocean, Southern Ocean, Central North Pacific Ocean and Antarctica, (Beja, Spudich et al. 2001;

Venter, Remington et al. 2004; Giovannoni, Bibbs et al. 2005; Slamovits, Okamoto et al. 2011). These variations present themselves in slightly different forms, each specifically adapted to the light in their environment (Beja, Spudich et al. 2001; Man, Wang et al. 2003). This feature is known as spectral tuning and has been noticed in many rhodopsins but is not well understood (Man, Wang et al. 2003; Briggs and Spudich 2005). In proteorhodopsin analysis, a few possible reasons for spectral tuning were found (Man, Wang et al. 2003). Researchers found one residue was different between the green and blue light proteorhodopsins (Man, Wang et al. 2003). In the green version, residue 105 was a non-polar leucine; in blue proteorhodopsin, the residue is a polar glutamine (Man, Wang et al. 2003). When these residues were swapped through mutation, the absorption spectra of the proteorhodopsins were also swapped (Man, Wang et al. 2003). Variations in residue 105 have been found to underlie differences among a variety of proteorhodopsin sources (Man, Wang et al. 2003).

Although the metabolic role of proteorhodopsins is not certain, researchers infer that they help generate a proton motive force in cells for which the most important mode of metabolism is the assimilation of dissolved organic carbon (Giovannoni, Bibbs et al. 2005; Gomez-Consarnau, Gonzalez et al. 2007). The SAR11 strain HTCC1062 *Pelagibacter ubique*, the first member of the SAR11 clade, carries a proteorhodopsin gene (Giovannoni, Bibbs et al. 2005). Until SAR11, proteorhodopsin had been found either only from DNA isolated from environmental samples and when expressed in non-native carrier organisms such as *E. coli* (Giovannoni, Bibbs et al. 2005; Slamovits, Okamoto et al. 2011). More recent studies have found proteorhodopsins in other organisms, including Flavobacteria isolates (Gomez-Consarnau, Gonzalez et al. 2007). The gene is expressed by cells grown in either constant light, light that follows a natural light cycle or in complete darkness, meaning that light is not needed for gene transcription (Giovannoni, Bibbs et

al. 2005). However, the exact conditions needed for cells to express proteorhodopsin are not exactly known, though its expression undoubtedly involves some manner of biological trade-off (Giovannoni, Bibbs et al. 2005). In *Pelagibacter ubique*, proteorhodopsin does not noticeably change cell growth, though in Flavobacteria, green light was seen to enhance growth while other wavelengths did not (Gomez-Consarnau, Gonzalez et al. 2007). Cells grown in light also appeared to be larger than those grown in darkness (Gomez-Consarnau, Gonzalez et al. 2007). The varying functionality of proteorhodopsins gives some insight into the vast diversity of proteorhodopsins.

Early research suggested proteorhodopsins had spread by horizontal gene transfer to a wide variety of prokaryotes, but found no evidence that proteorhodopsins are present in any eukaryotes. (Spudich 1998; Slamovits, Okamoto et al. 2011). Recent studies have since found proteorhodopsins have been gained through horizontal gene transfer from bacteria on at least two separate events in eukaryotic, dinoflagellate protists (Slamovits, Okamoto et al. 2011). Scientists found a number of rhodopsins in marine, dinoflagellate predator *Oxyrrhis marina* (Slamovits, Okamoto et al. 2011). These rhodopsins separated into two groups upon phylogenetic analysis, one that deeply clustered with proton pumping proteorhodopsins and a second, small group that separated with algal sensory rhodopsins and halophilic archaea (Slamovits, Okamoto et al. 2011). The close grouping of this rhodopsin with other type 1 rhodopsins strongly suggests that some HGT event occurred around the time of dinoflagellate speciation (Slamovits, Okamoto et al. 2011). In *Oxyrrhis marina*, proteorhodopsin is without a doubt the most plentifully expressed nuclear gene, and its product localizes to disconnected cytoplasmic structures (Slamovits, Okamoto et al. 2011). This allows for a number of different possible applications for the proton gradients that proteorhodopsin can create. Currently, photosystems I and II are the only

recognized mechanism for transducing solar energy in eukaryotes (Slamovits, Okamoto et al. 2011). However, it now appears that HGT has allowed the transfer of energy generating rhodopsins to transfer into eukaryotes (Slamovits, Okamoto et al. 2011). Proteorhodopsin may also be used to conserve ATP by supplementing vacuolar ATPases in acidifying digestive vacuoles, rather than directly generating ATP through a proton gradient (Slamovits, Okamoto et al. 2011).

Another bacterial rhodopsin is xanthorhodopsin. Xanthorhodopsin is found in *Salinibacter ruber*, a eubacterium that lives at high salt concentration (Balashov and Lanyi 2007). Similar to bacteriorhodopsin, xanthorhodopsin is a proton transporter (Balashov and Lanyi 2007). It is homologous to both bacteriorhodopsin and proteorhodopsin, which will be discussed later. Unlike other known microbial rhodopsins, Xanthorhodopsin has a second chromophore, a carotenoid which functions as an antenna that assists in collecting light (Balashov and Lanyi 2007). This carotenoid is salinixanthin, the primary carotenoid in *S. ruber*. It is a long, complex molecule (Lanyi and Balashov 2008). The antenna increases the surface area per rhodopsin excitable by light. Once excited the antenna causes a conformational change in the retinal (Lanyi and Balashov 2008). The carotenoid does not have a direct role in the transport of the proton across the membrane though (Lanyi and Balashov 2008). This system gives a more efficient energy generation by increasing the total surface area that can be excited by light (Balashov and Lanyi 2007). Xanthorhodopsin does not have a direct counterpart in the archaeal domain, which suggests that it evolved from a rhodopsin already in the bacterial domain (Balashov and Lanyi 2007).

Actinorhodopsin is a type 2 rhodopsin that is very similar to proteorhodopsin except it is found in non-marine organisms (Sharma, Zhaxybayeva et al. 2008). In 2008, a number of rhodopsins were discovered in Lake Gatun, a freshwater lake, that were absent in all marine samples taken from similar areas during the same study (Sharma, Zhaxybayeva et al. 2008). Later they were found in the Great Lakes of the US, specifically Superior and Erie (Sharma, Sommerfeld et al. 2009). After further analysis, the samples were found to carry a number of other genetic markers common to Actinobacteria, an abundant phylum of bacteria in the sampled environments (Sharma, Zhaxybayeva et al. 2008). Actinorhodopsin was characterized as a rhodopsin from Actinobacteria, since these markers were found only in the samples that had the novel rhodopsin and not in samples where the rhodopsin was missing (Sharma, Zhaxybayeva et al. 2008). Furthermore, the novel rhodopsin has the residues linked with proton pumping, and actinorhodopsin is not associated with the genes responsible for phototransduction. Therefore it is assumed that actinorhodopsin is a proton pump of some sort (Sharma, Zhaxybayeva et al. 2008). It is also highly similar to proteorhodopsin in sequence, and may be non-marine variant. However, actinorhodopsin cluster together and away from other microbial rhodopsins upon phylogenetic analysis (Sharma, Sommerfeld et al. 2009). Further study into the residues shows that there is greater similarity with proteorhodopsins than with archaeal rhodopsins such as bacteriorhodopsin (Sharma, Zhaxybayeva et al. 2008).

1.3.1 Archaeal Rhodopsins

There are four distinct rhodopsins found in archaea, specifically within haloarchaea (Sharma, Spudich et al. 2006). Of the four, two are different types of sensory rhodopsins (SRI &

II), while the other two are ion transporters: one proton transporter, bacteriorhodopsin, and one chloride transporter, halorhodopsin. Halorhodopsin was first described in 1977 by Matsuno-Yagi and Mukohata (Matsuno-Yagi and Mukohata 1977). Halorhodopsin was thought to be unique to haloarchaea until the recent finding in *Salinibacter ruber* (Sharma, Spudich et al. 2006). Halorhodopsin is a chloride ion pump that transfers chloride ions into the cell (Sharma, Spudich et al. 2006). This hyperpolarizes the cell and increases the electrochemical gradient that the proton pumps use to generate ATP. In addition to chloride ions, halorhodopsin is able to transport nitrates and other similarly sized halides across the membrane (Sharma, Spudich et al. 2006; Sharma, Walsh et al. 2007). Halorhodopsin is also used to help keep archaeal osmolarity constant by increasing the cell's salt concentration through chloride uptake so the sodium is kept out of the cell (Sharma, Spudich et al. 2006). In other hypersaline organisms, osmolarity is regulated by ATP driven Na/Cl transporters, but the presence of halorhodopsin allows for the transport to take place without metabolic energy use since halorhodopsin is still a light driven pump (Bamberg, Tittor et al. 1993; Sharma, Spudich et al. 2006). Due to the low primary production zones that haloarchaea live in, conserving energy is very important and can make the difference between a cell's survival and death.

In an example of spectral tuning, halorhodopsin is also capable of functioning as a proton pump when excited by the correct type of light (Bamberg, Tittor et al. 1993). Green light allows halorhodopsin to pump chloride ions, while higher intensity blue light induces proton pumping (Bamberg, Tittor et al. 1993). This is possibly harmful since, if the direction of the pump does not change, the proton would be transported into the cell which is then acidified (Bamberg, Tittor et al. 1993). It was found that while the proton pumping occurred at a slower rate than chloride transport, halorhodopsin is capable of alternating between the two if it is given a light source

containing high levels of blue and green light (Bamberg, Tittor et al. 1993).

The other half of haloarchaeal rhodopsins are sensory rhodopsins I and II (SRI & SRII) which are phototaxis oriented. They were originally discovered in 1982 in a mutant of *Hbt. salinarium* that lacked ion changes induced by light (Bogomolni and Spudich 1982; Spudich 1998), meaning that light caused no change in the ion concentrations either outside or inside the cell and that no light-induced ion transporters were present (Spudich 1998). These two are connected to transducer proteins (Htr I&II) that begin a phosphorylation cascade which ends with modulation of the flagella associated proteins (Spudich 1998; Pebay-Peyroula, Royant et al. 2002). While the nature of the connection between the transduction proteins and rhodopsins is not well understood, it is known that when not connected to transduction proteins such as HtrI & II (Hoff, Jung et al. 1997), these Htr's then interact with histidine kinase (Hoff, Jung et al. 1997). The general ribbon structure of SRII and its connection to its Htr proteins can be seen in Figure 3.

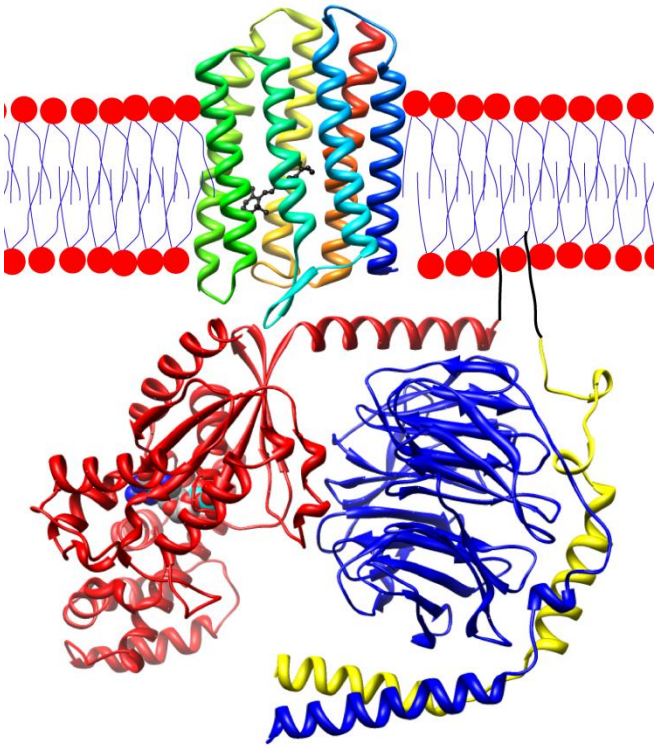

Figure 3 – Ribbon Structure of Sensory Rhodopsin 2 General structure of SRII, embedded in a membrane. The rhodopsin is rainbow colored, with a black retinal embedded. The blue, red and yellow ribbons below the rhodopsin are different subunits of the transducin bound to it. Image is public domain provided by WikiCommons, created by Devon Ryan of UCSF.

SRI is actively transcribed under similar conditions that bacteriorhodopsin and

halorhodopsin are active, when oxygen levels are low, and SRI directs the cell towards light so bacteriorhodopsin and halorhodopsin are able to function optimally (Spudich 1998). SRI is capable of functioning as a proton pump though at much lower rates than bacteriorhodopsin (Hoff, Jung et al. 1997; Pebay-Peyroula, Royant et al. 2002). While SRI is not generally considered a proton pump, it undergoes light activated proton transfers at the same site as bacteriorhodopsin, though the final transfer is blocked by the HtrI interaction (Spudich 1998). When activated, SRI directs the cell toward orange light, which is the optimum range for bacteriorhodopsin and halorhodopsin absorption (Spudich 1998). The sensory function of SRI is also noteworthy because its intermediate state is long lasting (Spudich 1998), unlike the intermediate of bacteriorhodopsin or SRII. The intermediate is very sensitive to UV radiation. When struck by UV light or near-UV light, the cell undergoes a repellent action (Spudich 1998). In order to protect the cell, the repellant behavior of SRI only requires 1-3 activated rhodopsins for the cell to reverse itself (Spudich 1998; Pebay-Peyroula, Royant et al. 2002). This allows for the cell to be guided to regions where bacteriorhodopsin is able to optimally function while still keeping the cell away from possible DNA damage from UV radiation (Spudich 1998).

Once rich aerobic environments are reached, SRI, halorhodopsin and bacteriorhodopsin switch off, and SRII becomes active when there are relatively high levels of oxygen and metabolic substrates (Spudich 1998). When oxygen levels drop off, SRII is down regulated and bacteriorhodopsin is induced. The main purpose of the SRII is to allow haloarchaea to avoid the intense light that is common at haloarchaeal habitats (Spudich 1998; Pebay-Peyroula, Royant et al. 2002). SRII is specifically tuned to absorb blue-green light at 487nm (Spudich 1998). When excited, it drives the cell away from sunlight and towards darkness as efficiently as possible, since bacteriorhodopsin is not being transcribed and strong light is not required (Spudich 1998).

Once again, like SRI, when not bound to its transducer protein HtrII, SRII also demonstrates a proton pumping function (Spudich 1998).

1.3.1.1 Bacteriorhodopsin and bop

The rhodopsin protein of choice for this study is bacteriorhodopsin, first discovered in 1971 by D. Oesterhelt (Oesterhelt and Stoeckenius 1971; Mukohata 1994). It is also the iconic type 1 rhodopsin and is one of the first membrane proteins that had had its crystal structure resolved (Michel and Oesterhelt 1980; Kirchmair, Markt et al. 2008). The primary organism for the study of bacteriorhodopsin is *Hbt. salinarium*, due in part to its ability to have high expression levels of the protein. Expression of bacteriorhodopsin in *Hbt. salinarium* is based on oxygen availability as covered earlier (Spudich 1998). Bacteriorhodopsin is a small 24kDa type 1 rhodopsin (Spudich 1998; Lanyi 2004; Sharma, Walsh et al. 2007). It is a chromophore-containing protein that causes patches of the membrane to turn purple and, in some cases, its presence can turn large portions of the cell membrane purple, as seen in *Hbt. NRC-1* (Spudich 1998; Sharma, Walsh et al. 2007). In some cases bacteriorhodopsin is produced in large amounts to the point where it is the only protein present in 50% of the membrane, and can cause crystallization of the membrane (Spudich 1998). Commonly bacteriorhodopsin is found in trimers consisting of identical monomers, though it can be found in much larger clusters (Slonczewski and Foster 2011).

Bacteriorhodopsin is a protein of interest due to its specialized phototrophic function. Bacteriorhodopsin is used by haloarchaea to generate ATP directly from sunlight by functioning as a proton pump using light at a wavelength between 500 and 650nm, absorbing green light

(Haupts, Tittor et al. 1999). Absorbing green light is the cause of the purple or blue color the cells have when bacteriorhodopsin is highly expressed. As light hits bacteriorhodopsin's attached retinal, it is absorbed and undergoes a conformational change from all-trans to 13-cis retinal conformations as part of a multi-step process to transfer a proton across the membrane (Subramaniam, Gerstein et al. 1993; Haupts, Tittor et al. 1997; Haupts, Tittor et al. 1999). Step one is the conformational change from all-trans to 13-cis, which takes under 500 fs (Subramaniam, Gerstein et al. 1993) and creates intermediate J600 that is then turned into K590(Subramaniam, Gerstein et al. 1993; Haupts, Tittor et al. 1997; Haupts, Tittor et al. 1999). The second step, taking 2 µs, is the conversion to L550 (Subramaniam, Gerstein et al. 1993). From here the proton is transferred to the aspartate-85 for the M state (Subramaniam, Gerstein et al. 1993; Haupts, Tittor et al. 1999). The M state is then converted to N state when the aspartate-96 is re-protonated (Subramaniam, Gerstein et al. 1993; Haupts, Tittor et al. 1999). Conversion of the N to O state requires the isomerization of the retinal from 13-cis to all-trans (Subramaniam, Gerstein et al. 1993; Haupts, Tittor et al. 1999). The final step involves release of the proton by the aspartate-85 and a complete return to its initial state (Subramaniam, Gerstein et al. 1993; Haupts, Tittor et al. 1999). A simplified version of this cycle can be seen in Figure 4.

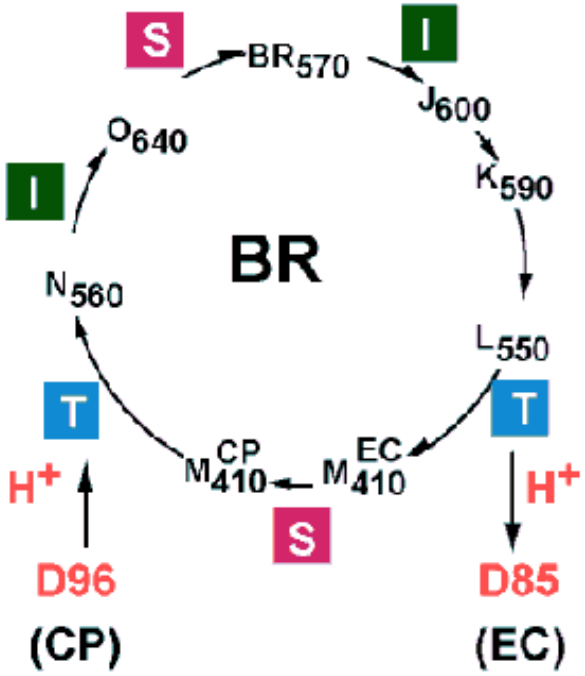

Figure 4 – Catalytic Cycle of Bacteriorhodopsin
The entire photocycle of bacteriorhodopsin is depicted, starting from BR570 and continuing clockwise. It shows the pattern outlined in the text to the left. Image provided by Max Planck Institute of Biochemistry, and Oesterhelt D. <http://www.biochem.mpg.de/en/eg/oesterhelt/web_page_list/Protein_BR/index.html>

Simply put, a proton is then picked up from the cytoplasmic side by an aspartate-96 residue, passed to other residues along the ion channel until ending up at another aspartate-85 and then is released after having crossed the membrane (Subramaniam, Gerstein et al. 1993; Haupts, Tittor et al. 1997; Haupts, Tittor et al. 1999)

The ion transfer rapidly creates a proton gradient, making the cell up to 10,000 times more alkaline than the outside. When the protons reenter the cell through an ATP synthase, the same synthase used during respiration, ATP is generated from ADP. As protons reenter the cell through the synthase, the synthase rotates and undergoes conformational changes that lead to ATP synthesis (Slonczewski and Foster 2011). The generation of a proton motive force creates an efficient system and allows cells to generate energy despite the poor nutrient quality of their environments, which have low primary production. Some genera of haloarchaea are missing *bop* (Boone, Castenholz et al. 2001). Commonly, these strains are able to generate energy in other ways, such as the ability to reduce nitrogen, and are found in light-poor regions, such as low in the water column (Boone, Castenholz et al. 2001). *Haloferax* and other groups lacking *bop* usually have a different energy source when oxygen levels are too low for respiration, such as fermentation or nitrate reduction (Boone, Castenholz et al. 2001; Papke, Douady et al. 2003). The lack of *bop* does not hinder members of this genus for instance, *Hfx. volcanii,* since it lives in regions where light does not strongly penetrate (Mullakhanbhai and Larsen 1975). *Haloferax volcanii* in particular reside at the bottom of the Dead Sea (Mullakhanbhai and Larsen 1975). On the other hand, a number of genera are known to have *bop,* such as *Haloarcula vallismortis* which inhabits shallow salt pools where the sunlight needed for bacteriorhodopsin easily reaches the cell (Gonzalez, Gutierrez et al. 1978; Boone, Castenholz et al. 2001). While some organisms have the ability to follow multiple pathways to generate energy when oxygen levels are too low

for respiration, these paths are mutually exclusive (Papke, Douady et al. 2003).

Bacteriorhodopsin experiences varying degrees of selective pressure. In shallower crystallizer ponds, the pressure is high as archaea with rhodopsin-based energy sources can dominate the environment (Sharma, Walsh et al. 2007). However since bacteriorhodopsin is not required for cellular life, it is under a lower differential selective pressure compared to housekeeping genes such as *EF-2* and *atpB*. Unlike housekeeping genes, the cell will not automatically be killed by any mutations in the bacteriorhodopsin. This allows for bacteriorhodopsin to potentially be used to differentiate within species, as it is less conserved. As a result, bacteriorhodopsin may have a greater specificity when it comes to differentiation both at and below the species level.

The gene that encodes bacteriorhodopsin is *bop*, and the size of the coding regions can vary depending on the strain, some having 666 bases and others having 789 bases (Bilofsky and Burks 1988). The region being amplified and analyzed in this study is from bases 312 to 681 inclusive for *Hbt. salinarium* (Bilofsky and Burks 1988). The amplified regions from other species vary, but in general the regions are of similar size (Bilofsky and Burks 1988). Variance in sequence length is likely due to small insertions and deletions in the sequence that may not have a strong impact on the protein. The amplified region codes for three full alpha helixes and two partial ones which can be seen in Figure 5. While *bop* is found scattered throughout the haloarchaeal genera, there are a number of taxa that typically do not have *bop* (Boone, Castenholz et al. 2001). The well-studied *Haloferax* genus does not, for the most part, have *bop*, though some, such as *Hfx. zhejiangensis*, have been reported to carry the gene (Bilofsky and Burks 1988; Boone, Castenholz et al. 2001; Xu 2004).

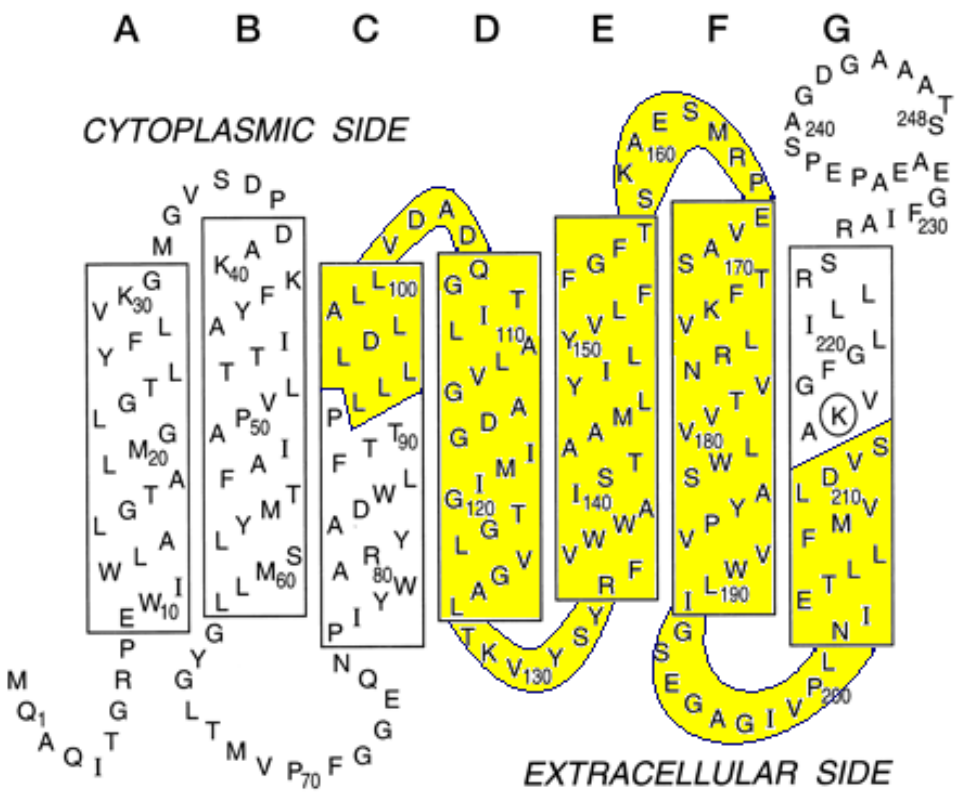

Figure 5 – Bacteriorhodopsin PCR Amplified Region

The above protein is bacteriorhodopsin with the portion studied highlighted in yellow. The lysine residue where retinal binds is circled. Also of note, the indel mentioned later is between residues 130 and 131. The image was provided by Marti and modified by Maulik Jani (Gogarten 1992; Marti 1998).

1.3.2 Homology within Type 1

Type 1 rhodopsins, including bacteriorhodopsin, are generally considered homologous to one another, in that they preserve key amino acids and the retinal isomers (Spudich, Yang et al. 2000). In contrast to its well-known archaeal homolog bacteriorhodopsin, most of the active sites significant to the known bacteriorhodopsin mechanism are preserved in proteorhodopsin (Partha, Krebs et al. 2005). Homologues of the active site residues Arg82, Asp85, Asp212 and Lys216, the retinal Schiff base fastening site, in bacteriorhodopsin are preserved as Arg94, Asp97,

Asp227 and Lys231 respectively in proteorhodopsin (Spudich, Yang et al. 2000; Partha, Krebs et al. 2005; Martinez, Bradley et al. 2007). Preservation of Asp85is important since it is the primary proton acceptor (Partha, Krebs et al. 2005). Also as in bacteriorhodopsin, the retinal chromophore in un-activated proteorhodopsin is primarily but not totally in an all-trans conformation and photoisomerizes to 13-cis upon interaction with light (Partha, Krebs et al. 2005). On the other hand, proteorhodopsin contains no carboxylic acids identifiable as homologous to Glu194 or Glu204, of bacteriorhodopsin, amino acids considered to take part in the proton release to the extracellular surface in bacteriorhodopsin (Partha, Krebs et al. 2005). It is most likely that proteorhodopsin functions all over the earth's oceans as a light-driven H+ pump by a system similar to that of bacteriorhodopsin (Partha, Krebs et al. 2005).

Xanthorhodopsin is proteorhodopsin-like and is considered homologous, based on sequence data, to bacteriorhodopsin, as well as to proteorhodopsin with 23% and 19% sequence similarity respectively (Balashov and Lanyi 2007; Lanyi and Balashov 2008), suggesting that xanthorhodopsin originated from the same ancestral rhodopsin that the archaeal ones did. However, a closer examination of the functional residues of xanthorhodopsin reveals that they are more similar to proteorhodopsin than to bacteriorhodopsin, despite the higher sequence similarity with bacteriorhodopsin (Lanyi and Balashov 2008). Of note is the lack of a homolog for xanthorhodopsin’s glutamic acid to protonate the Schiff base in bacteriorhodopsin, which uses an Asp96 (Lanyi and Balashov 2008). Proteorhodopsin also uses glutamic acid to donate protons onto the Schiff base.

The archaeal rhodopsins are highly similar to one another. The sensory rhodopsins are similar to bacteriorhodopsin at key bases (Spudich, Yang et al. 2000), primarily in the retinal-binding pocket. These bases are 95% similar between SRI and bacteriorhodopsin (Spudich, Yang

et al. 2000). Additionally, the sensory rhodopsins lack a carboxylate residue at position 96, the proton donor site in bacteriorhodopsin (Briggs and Spudich 2005). Halorhodopsin has a number of the same bases as sensory rhodopsins and bacteriorhodopsin, though it has a neutral reside rather than an Asp85 (Spudich, Yang et al. 2000; Váró 2000). Of note, the chloride ion occupies this location in halorhodopsin and when Asp85 is mutated in bacteriorhodopsin, it is also able to pump chloride ions to a degree (Spudich, Yang et al. 2000). When azide was added to halorhodopsin, the protein became capable of pumping protons across the membrane (Váró 2000). Also, like bacteriorhodopsin but unlike the sensory rhodopsins, the retinal is found in a mix of all-trans and 13-cis (Váró 2000).

1.4 Type 2 – "Eukaryotic Rhodopsins"

Type 2 rhodopsins are used primarily for phototaxis and are found primarily in eukaryotic cells. They are often called "Eukaryotic Rhodopsins," though they do not need to be in eukaryotes, nor do they refer to all rhodopsins in the Eukaryota domain (Schlesner, Miller et al. 2009). One difference between the two types is in size. Type 2 rhodopsins are 38kDa, with large interhelical loops, and are much larger than type 1, which are around 24kDa (Spudich, Yang et al. 2000). The most famous example of a type 2 rhodopsin is the rhodopsin found in rod cells of the eye. These rods are used to allow night time vision due to their sensitive nature. Type 2 rhodopsins are embedded in small disks that are contained entirely within the cell (Sung and Chuang 2010). When these rhodopsins are activated, they trigger an associated G protein system (Sung and Chuang 2010). A single rod cell is capable of stimulating hundreds of transducin molecules which in turn induce an equal number of cGMP-phosphodiesterases, each able to

hydrolyze over a thousand cGMP molecules per second, thus lowering cGMP concentrations in a signal transduction cascade (Sung and Chuang 2010). A sudden drop in cGMP levels causes gated cGMP channels to close. As the gated channels close, positive ions are prevented from entering the cell, causing hyperpolarization of the cell's membrane as its potential becomes more negative (Sung and Chuang 2010). Hyperpolarization of the rod cell causes a release of lower levels of glutamate to the next cells in the chain. The end result of this chain is vision.

1.5 Un-typed Rhodopsins

Single celled organisms also use rhodopsins. One interesting euglenoid eukaryote is *Peranema*, which insatiably takes any elements into its body system through phagocytosis (Saranak and Foster 2005). *Peranema's* "feeding behavior" is easily observed under the microscope (Saranak and Foster 2005). *Peranema trichophorum* is a eukaryotic phagotroph of the *Euglenophyta* that spends its life in fresh water (Saranak and Foster 2005). Mature *Peranema* cells glide and can turn or curl impulsively in either light or darkness (Saranak and Foster 2005). When gliding, *Peranema trichophorum* employs its rhodopsin to organize its coiling behavior (Saranak and Foster 2005). From its coiled state, the cell rotates to face a new direction. In a related manner, archaea such as *Halobacterium (Hbt.)* employ light activation of sensory rhodopsins to organize the turnaround of the course of flagella (Saranak and Foster 2005). Each turnaround results in the cell changing its course. The frequency of coiling behavior is actually increased by an improvement in light quality, as evidenced by shortening of time between curls in the presence of light (Saranak and Foster 2005). The optimal wavelength of absorption for this rhodopsin is between 500 and 505 nm, the range which is common for type 2 rhodopsins

(Saranak and Foster 2005). Other than this there is little known about the rhodopsin, and so it was not placed into either type category, though it is likely a type 2 rhodopsin. The structure of bacteriorhodopsin and this rhodopsin have highly similar topology, with helices I to III, VI and VII being particularly similar (Saranak and Foster 2005). Rhodopsins in *Peranema* hold up the conception of a common eukaryotic ancestor having a rhodopsin photoreceptor (Saranak and Foster 2005).

1.6 Archaea

Archaea are the third domain of life, the other two domains being Eubacteria and Eukarya, which represent bacteria and eukaryotes respectively. Until 1977, there were only two domains, eukaryotes and prokaryotes (Woese and Fox 1977; Woese, Magrum et al. 1978). At this point it was proposed by Woese *et al.* that a group of prokaryotes who had significantly divergent rRNA were actually a separate group, Archaebacteria (Woese and Fox 1977; Woese, Magrum et al. 1978), later Archaea (Woese, Kandler et al. 1990). The work of Dr. Woese and others ultimately created the current three domain

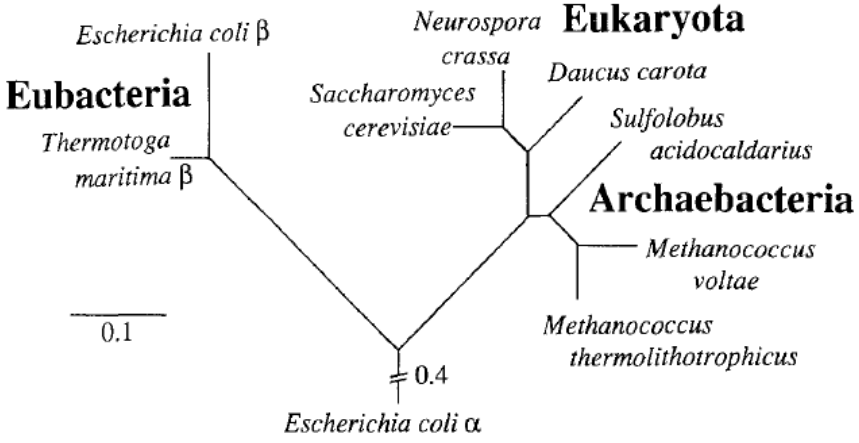

Figure 6 – Rooted Tree of Life
A Tree of Life rooted and based on ATPases. It clearly shows how Archaebacteria (Archaea) are more closely related to Eukaryota (Eukaryotes). Additionally it shows that Archaea are a distinct group separate from Eukaryotes, functioning as a more accurate version of the original tree created by (Woese and Fox 1977). Image provided by Gogarten *et al* (Gogarten 1992).

system consisting of Archaea, Eukaryotes and Bacteria. A conclusion supported by Peter Gogarten's rooting the tree of life, clearly showing that archaea are more closely related to eukaryotes than to bacteria (Gogarten, Kibak et al. 1989; Gogarten 1992). This rooted tree can be seen in Figure 6. Archaebacteria are more closely related to the eukaryotes as opposed to the bacterial branch, with the root appearing between the two (Woese, Kandler et al. 1990; Gogarten 1992). The lack of a number of eukaryotic proteins within the archaebacteria also led to the separation of eukaryotes and archaebacteria into separate domains, despite their similarities (Woese, Kandler et al. 1990). Archaebacteria were later renamed archaea, which is the permutation of the Greek word for "ancient things" (Woese and Fox 1977). The name archaea was chosen due to their ability to survive in an environment that was similar to that of an early earth.

Archaea in general are a widespread and relatively understudied group. While there have been reports of archaeal lipids in sediments as old as 3.8 billion years (Hahn 1986), archaea are still not fully understood by modern scientists. They are single-celled organisms that have no membrane bound organelles and no nucleus; they have also been reported to have chromosomal and cytoplasmic elements (Brown and Doolittle 1997; Slonczewski and Foster 2011). The close relatedness between archaea and eukaryotes is reflected in their similar metabolic pathways and the general enzyme makeup (Brown and Doolittle 1997; Slonczewski and Foster 2011). Specifically, the DNA polymerase and the RNA polymerase of archaea are very similar to those of eukaryotes, as are the associated transcription factors such as TFII, which is present in archaea and eukaryotes but not in bacteria (Brown and Doolittle 1997; Slonczewski and Foster 2011). Additionally, both archaea and eukaryotes use TATA boxes to regulate transcription, which are not used by bacteria (Brown and Doolittle 1997).

Archaeal similarities with bacteria are more easily discernible, and these similarities are the reasons for their original phenotypic grouping with bacteria before the advent of genetic phylogenetics. In appearance they are similar in both general size and their lack of nuclei (Brown and Doolittle 1997). They also usually have a circular chromosome with additional plasmids providing other genes. Additionally, both reproduce primarily by binary fission, and their genes are organized into operons (Brown and Doolittle 1997; Slonczewski and Foster 2011).

Originally archaea were seen primarily as extremophiles, cells adapted to thrive in physically extreme environments. This image was not without reason considering they hold the record for life at highest temperature, lowest pH and highest NaCl concentration (Chaban, Ng et al. 2006). However, many have since been found to live in milder regions as well (Chaban, Ng et al. 2006). Some methanogenic archaea are capable of living in the human mouth as well as the gut (Chaban, Ng et al. 2006). They comprise over 20% of the entire world's microbial ocean biomass (DeLong and Pace 2001). Archaea have been found in many extreme environments, from arctic brine lakes all the way to salted cod (Chaban, Ng et al. 2006). Specializations allow for life at extremes. Some are able to withstand high temperatures such as *Methanopyrus kandleri* 116, which can grow at 122 C, the highest recorded for any organism (Takai, Nakamura et al. 2008). Other species, such as *Picrophilus torridus*, are capable of

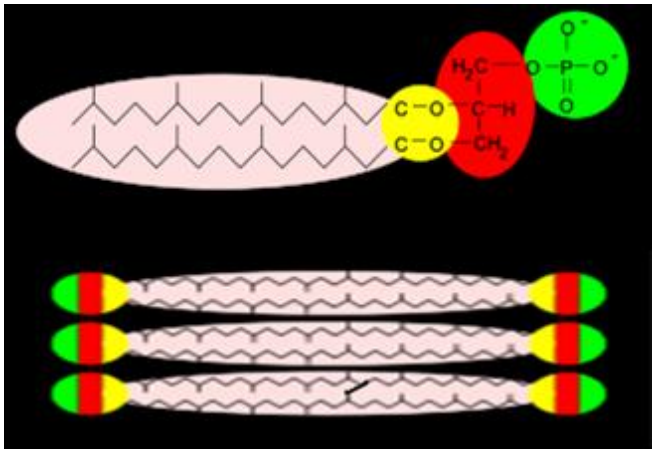

Figure 7 – Examples of Lipids Found in Archaeal Membranes

The top image shows the ether bond between the C-O and the C found in the red glycerol region, forming the C-O-C ether bond. In the lower set of three, the top two are the result of side chains bonding to side chains from the opposite end of the membrane. The lowest molecule is a tetraether due to the cross-linking of the long side chains to each other. Not shown are thermo-stable cyclopentane rings. Image is public domain provided by WikiCommons, created by user Franciscosp2

growing in the crater of sulfur emitting fumaroles where the pH can drop below 0 (Chaban, Ng et al. 2006). Others are able to live in a variety of extreme situations. The primary property of archaea that allows for their survival in harsh conditions is their unique membrane makeup and structure. Bacteria and eukaryotes have lipid bilayers comprised of non-polar fatty acid chains linked to polar head groups (Slonczewski and Foster 2011), however archaea do not have these fatty acids but have branched hydrocarbon chains which can be seen in Figure 7. These are oriented so that the polar heads are facing the aqueous interior and exterior of the cell, while the non-polar chains are facing each other in the center of the membrane. Archaeal membranes can also be specialized for their specific environments though this is not always the case (Slonczewski and Foster 2011). Some contain ether linkages instead of the ester links found in bacteria and eukaryotes (Slonczewski and Foster 2011). These types of links are much more stable and require more energy to break (Slonczewski and Foster 2011). Additionally, in some archaea the side chains are covalently linked to one another, forming a more solid structure (Slonczewski and Foster 2011). In other cases, the side chains are bonded to the side chains of the lipid on the other side of the membrane forming a membrane that is comprised of one molecule (Slonczewski and Foster 2011). Tetraethers are even more stable and occur when the single molecule has a number of interlinks between the side chains (Slonczewski and Foster 2011). Thermophilic archaea in particular use cyclopentane rings incorporated into the side chains, which keep their membranes stable at high temperatures, even though the mechanism is unknown (Slonczewski and Foster 2011). These all create a much more stable membrane and help keep the cell together, despite the environments that archaea live in. Examples of each of these membrane lipids can be seen in Figure 7.

1.6.1 Haloarchaea

Haloarchaea are slow-dividing, prokaryotic, salt-loving members of the archaeal domain, and they are part of the phylum Euryarchaeota and the order Halobacteriales (Boone, Castenholz et al. 2001). Physically, haloarchaea are morphologically diverse. They can vary from a coccoidal shape to very thin squares and rods and flat triangles (Boone, Castenholz et al. 2001; Dyall-Smith 2009). These varied morphologies can be seen in Figure 8. Their cell membrane consists of a single thin layer of glycoproteins commonly called the S-layer (Dyall-Smith 2009; Slonczewski and Foster 2011). This thin membrane is held together by cations, meaning that it can be disrupted by even weak detergents such as EDTA (Boone, Castenholz et al. 2001). Also of note, even in a pure culture cells can have different shapes. The membranes can also be disrupted by simply lowering the salt concentration; however the membranes of coccoid genera are able to remain stable despite lower salt concentrations thanks to their thick cell wall, some of which can be up to 50nm (Boone, Castenholz et al. 2001).

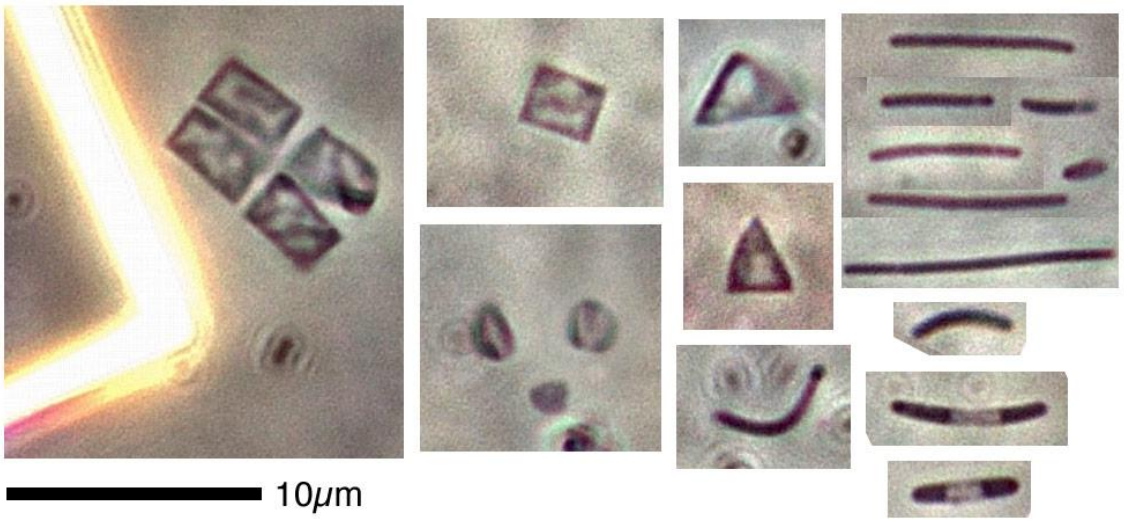

Figure 8 – Variation of Haloarchaeal Cell Morphology
Microphotographs of different haloarchaeal cell shapes as seen through a phase contrast microscope. Squares, coccoidal cells, triangles and both straight and curved rods can be seen. Image is from < http://web.aanet.com.au/aanet/DGBHome/Research/Research.html>.

Haloarchaea are moderately motile organisms. They have both gas vesicles, which allow them to have a great deal of control over their location in the water column, and flagella which they can use to move through it (Boone, Castenholz et al. 2001; Soppa, Baumann et al. 2008; Schlesner, Miller et al. 2009). In general, haloarchaea control their location based on input from the various rhodopsins. Some members have both chemical and light sensors, including SR (Schlesner, Miller et al. 2009). Once again, there are always exceptions. For example, *Hfx. volcanii* does not have gas vesicles, likely because it does not live high in the water column where light penetrates (Mullakhanbhai and Larsen 1975; Boone, Castenholz et al. 2001). Other haloarchaea that do not live in the water column are probably missing gas vesicles as well (Boone, Castenholz et al. 2001). Though some do not, many haloarchaea have flagella (Dyall-Smith 2009), allowing the cells to move away or toward the potentially damaging rays of the sun, which are highly concentrated due to the nature of the high salinity of the water (Boone, Castenholz et al. 2001; Sharma, Spudich et al. 2006).

Haloarchaea are also known for their variety of colors (Boone, Castenholz et al. 2001). These colors are caused by the various carotenoids. The carotenoids that give haloarchaea their color may also be responsible for protecting them from the DNA damaging solar radiation by functioning as oxidation preventers (Slonczewski and Foster 2011). One carotenoid in particular, bacterioruberin, is both the main reason for the red color of haloarchaeal colonies and their main protection from light damage (Slonczewski and Foster 2011). Other carotenoids serve to protect the reaction zone from self oxidation. Colorless haloarchaea have also been found (Boone, Castenholz et al. 2001). They are isolated deeper in the water column where possibly damaging light does not penetrate, while other colorless strains have been cultured from lakes and flats

with low sunlight (Gonzalez, Gutierrez et al. 1978; Dyall-Smith 2009).

Haloarchaea grow in highly saline environments, requiring a minimum of 9% salt, or 1M in the case of some of the more moderate strains such as *Hfx. volcanii* (Boone, Castenholz et al. 2001; Chaban, Ng et al. 2006; Dyall-Smith 2009). If the cells are placed in a lower concentration, the water rushes into the cell, and, in many cases, this influx of water will burst the cell (Boone, Castenholz et al. 2001; Slonczewski and Foster 2011). In the case of *Halococcus,* which does not burst when placed in a hypotonic solution, cell death is caused by an inability to transport cofactors into and around the cell (Boone, Castenholz et al. 2001). Haloarchaea are able to survive high salt concentrations using the salt in strategy, primarily used by haloarchaea and anerobic bacteria, and involves increasing the intercellular concentration of KCl. This is made possible by having a high percentage of acidic residues in the proteins. These residues allow for the intracellular salt concentration to be equal to the external, by attracting KCl, keeping the inside of the cell at equal osmolarity to the outside (Soppa, Baumann et al. 2008; Dyall-Smith 2009). This strategy is different from the more common salt out strategy, which involves constantly pumping water back across the membrane or accumulating small organic solutes to counter-balance the external salinity (Sharma, Walsh et al. 2007; Slonczewski and Foster 2011).

Despite their need for a high salinity environment, haloarchaea have been found throughout the world. Haloarchaea have been found on salted cod, in underground salt deposits and have even been reported inside salt crystals (Boone, Castenholz et al. 2001; Chaban, Ng et al. 2006). Halobacteriales are a very diverse group that are able to survive very harsh environments, from the sun baked salterns to subglacial Antarctic salt lakes and even in high pH lakes in the case of *Natronococcus* (Chaban, Ng et al. 2006; Slonczewski and Foster 2011). They

are commonly found in salt lakes and in such concentrations that they are responsible for the majority of the red tint these lakes commonly have (Dyall-Smith 2009; Slonczewski and Foster 2011). The reddening of the ponds has been tied to salt precipitation in that these blooms increase the water temperature and promote the crystallization of salt (Dyall-Smith 2009). In these crystallizer ponds, haloarchaea are found at concentrations approaching 10^8 cells/ml (Dyall-Smith 2009).

Haloarchaea are also able to utilize a number of sources to generate energy. While all are capable of aerobic respiration, their ability to generate energy without oxygen is varied (Soppa, Baumann et al. 2008). Most are able to ferment amino acids while some are able to use DMSO (Soppa, Baumann et al. 2008). Others are able to use bacteriorhodopsin to generate ATP directly from sunlight, though the distribution of the genes for anaerobic energy generation is patchy (Soppa, Baumann et al. 2008).

The recent studies involving haloarchaea have had varying topics that attempt to clarify the definition of species (Papke, Zhaxybayeva et al. 2007) and horizontal gene transfer (Portillo and Gonzalez 2009), and focus on a specific gene's evolution (Sharma, Spudich et al. 2006; Sharma, Walsh et al. 2007; Papke, White et al. 2011). Haloarchaea were chosen for this study primarily because they are the main organism that contains *bop*, but also because they are known to undergo some form of horizontal gene transfer (HGT), which has been shown to play a large role in evolution (Gogarten, Doolittle et al. 2002).

1.7 Horizontal Gene Transfer

Horizontal gene transfer (HGT), less commonly known as lateral gene transfer (LGT) is

the flow of genetic material from one organism to another without being its direct ancestor. In bacteria, resistance plasmids are the primary focus of horizontal gene transfer studies. Figure 9 shows the three common mechanisms of HGT: transformation, transduction and conjugation (Slonczewski and Foster 2011).

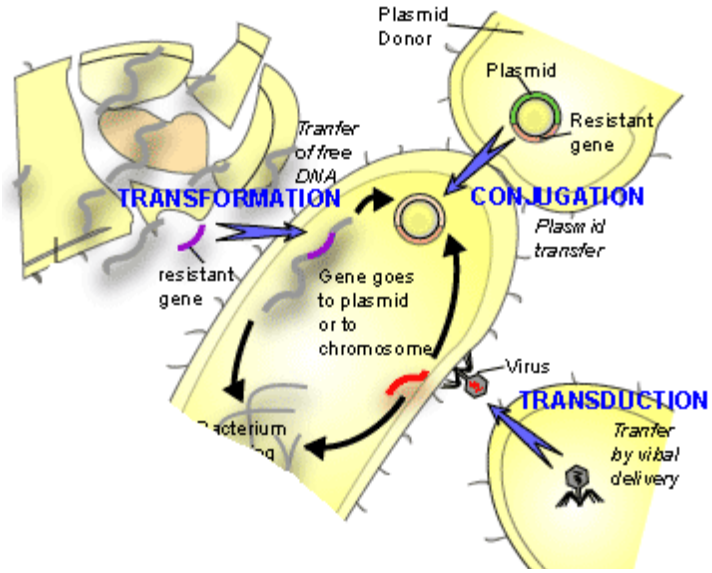

Figure 9 – Horizontal Gene Flow

Above are representations of three mechanisms that HGT may follow. The free DNA required for transformation may come either from lysed cells or may be exported by living cells under certain conditions (Ellen, Zolghadr et al. 2010). Conjugation may occur through a pilus or cytosolic bridge. Image provided by Dr. Kenneth Todar at the University of Wisconsin at Madison

Discovered in 1928 by Frederick Griffith, transformation is the direct uptake of DNA by a competent cell from the environment. A competent cell is a cell that can pick up DNA from the environment (Slonczewski and Foster 2011). If this DNA is then incorporated into the genome, the cell has undergone HGT. It occurs naturally in some cases but can be done artificially through electroporation or chemicals. The natural method varies between Gram-positive and negative cells. Gram-negative cells are generally either always competent or become so once starved (Slonczewski and Foster 2011); Gram-positive cells require a transformasome complex. Transformasomes capture and take in DNA while degrading one of the strands. The complex is induced through the uptake of competence factors that are exported by other cells (Slonczewski and Foster 2011). Chemicals, such as $CaCl_2$, with heat shock can alter the cell membrane, and can also be used to induce competence and allow DNA to cross the membrane (Slonczewski and

Foster 2011). Electroporation is the use of short electric shock to disrupt the membrane and allow for DNA to enter the cell.

Transduction is virus mediated, as the infection leaves behind some of its own DNA that recombines with the host's DNA. This can occur when a virus infects a cell and enters the lysogenic cycle, where the virus is dormant but the viral DNA has recombined with the host's. Additionally, transduction can only occur when the genetic material of the host is the same as that of the infecting virus, in that an RNA virus will not recombine with a DNA based host.

Finally, conjugation involves direct cell-to-cell contact, commonly using a pilus. It is often regarded as bacterial mating since it is a one way exchange of DNA between cells that are in direct contact. In some cases this is mediated by fertility factors that are transferred from one cell to another and allow for conjugation (Slonczewski and Foster 2011). The two cells can be pulled towards each other through contraction of the pilus (Slonczewski and Foster 2011). Through these mechanisms, haloarchaea are able to transfer entire genes to other members of Halobacteriales, though natural transformation has not been demonstrated in haloarchaea (Allers and Mevarech 2005).

Additionally, Gene transfer agents (GTA) can be a mechanism for HGT. GTAs are small phage-like particles that are released by few cells and are taken in by others (Zhao, Wang et al. 2009). The DNA packed into GTAs, however, is for the most part random (Biers, Wang et al. 2008). The gene cluster for GTA seems to be limited to few organisms in Rhodobacteriales and some others in Alphaproteobacteria with infrequent homologs (Biers, Wang et al. 2008; Zhao, Wang et al. 2009).

Identification of HGT events can be difficult. Compositional analysis of the sequences can be used if other homologous sequences are not available. It involves comparison of the

sequence in question with the remaining genome searching for general differences such as GC content and other characteristics (Poptsova and Gogarten 2007). However this is not the primary manner in which HGT is detected. If homologous sequences are available, phylogenetic methods are commonly used (Poptsova and Gogarten 2007). In some cases horizontal gene transfers are identified by the introduction of a non-native gene into the organism (Lawrence and Ochman 1997), such as a bacteriorhodopsin appearing in *Hfx. volcanii*. A current conformation test is the Approximately Unbiased (AU) test. The AU test tries to determine if the given trees are significantly different from the tree constructed from the data (Poptsova and Gogarten 2007). An AU test is normally run on the set of trees from PhyML and generates a P-value for each tree. If this value is below 0.05, the tree is incongruent with the data set (Shimodaira 2002).

1.8 Multi-Locus Sequence Analysis

Multi-Locus Sequence Analysis (MLSA) is the use of multiple gene phylogenies that are combined to create a more clear phylogeny (Papke, White et al. 2011). In MLSA multiple genes are used and then either their trees are combined into a Supertree or all the sequences are concatenated. In a Supertree a number of trees are combined into a final tree that attempts to reconcile any differences between the trees, while a concatenation all the sequences are concatenated into a single long sequence which is used for conventional tree construction. The genes used are conserved genes and should give a clear image of the evolutionary tree (Papke, White et al. 2011). Thanks to the multiple genes used, MLSA is also able to differentiate between similar species even if there is a moderate amount of horizontal gene flow (Papke, White et al. 2011). MLSA is a very efficient system once it has been established and processed

into databanks. A database would have many individual strains each with a set of unique sequences that would identify with similar stains when searched. Additionally, the materials used are easily transported between different labs, since only DNA is needed and not live cells. Labs would also be able to work with more dangerous strains than normal, saving money. The primers and protocols for amplification and the analysis can be exchanged electronically as well.

Hopes are that MLSA will be able to differentiate between different pathogenic strains. However pathogenic strains would still have highly similar housekeeping genes, and so others would need to be used (Chen, Zhang et al. 2005). These additional genes would be the virulence factors which are much more likely to change since they are not required for basic cell life (Chen, Zhang et al. 2005; Poptsova and Gogarten 2007). This technique is Multi-locus Virulence Typing (MSVT) (Chen, Zhang et al. 2005). A similar technique may be required for similar strains that have recently diverged since the housekeeping genes will take some time to differentiate.

Originally the amplification and sequencing of *bop* was a small portion of a larger MLSA project, which also involved a number of other genes. Because it was difficult to get sequence data from *bop* and because it is spottily distributed, it was excluded from the MLSA project. The original project was undertaken to investigate the multiple copies of divergent 16s ribosomal RNA (rRNA) genes found in some haloarchaea (Papke, White et al. 2011). The 16s rRNA gene is the most common genetic marker used to differentiate species, it is also the main method used to assign new strains to a genus. However due to the multiple copies and the frequency of recombination between different species, another approach is needed (Papke, White et al. 2011). A theoretical example of how recombination or HGT event can skew an evolutionary tree is in Figure 10. Additionally, the 16s rRNA gene can be highly conserved and cannot be used to

differentiate between closely related species; it may even recombine between closely related species (Papke, White et al. 2011). The use of MLSA gives a clear image of the evolutionary pattern of the sampled members of *Halorubrum* despite the highly similar and recombinant 16s rRNA genes (Papke, White et al. 2011). The final maximum likelihood tree with bootstrap support is provided further below in Figure 18 (Papke, White et al. 2011). It was constructed using five housekeeping genes: EF-2, radA, secY, atpB and rpoB, though there were some problems with the secY tree (Papke, White et al. 2011). The problems with secY may have been caused by frequent HGT events within clades (Papke, White et al. 2011). Each of these genes was then constructed into a tree as well as a concatenated tree. *Bop* phylogeny was compared to the MLSA tree and any deviations were further investigated at possible HGT events, and then discussed.

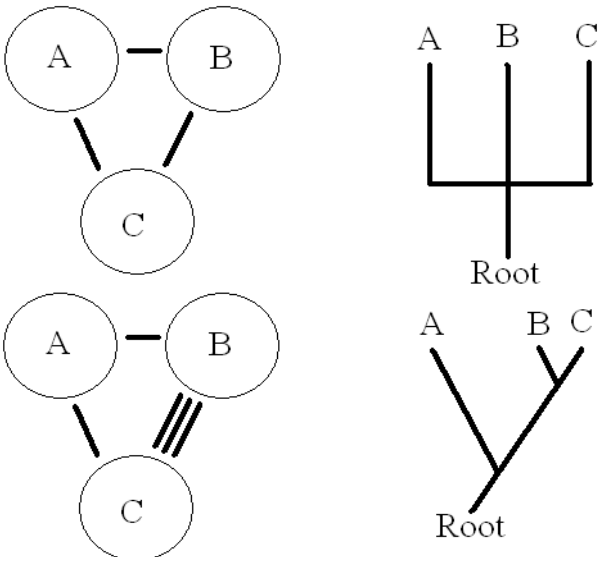

Figure 10 – Phylogenetic HGT (Example)
The top pair is a "normal" species phylogeny, with A, B and C being different strains with a common ancestor and no HGT. The tree to its right shows that they are all equidistant from each other. The bottom pair shows a "gene" tree with HGT between B and C, depicted by the additional lines between them. This is reflected in their much tighter grouping together. This type of discrepancy between the "normal" species tree and the gene is how HGT can be identified. Image provided by Maulik Jani

1.9 Objective of the Thesis

This thesis will attempt to provide additional insight into the evolutionary processes and patterns of *bop* through the use a wide range of genera. It will compare and contrast the *bop*

phylogeny with the established phylogeny based on 16s rRNA gene sequences as well as with the more recent MLSA phylogeny published by (Papke, White et al. 2011). Additionally, this study will search for any possible examples of HGT by looking for strains that do not group closely with others of the same genus. While examples of *bop* gene HGT have been found in haloarchaea before, not all strains in this study were covered (Sharma, Walsh et al. 2007).

2. Methods:

2.1 DNA Isolation

Haloarchaeal cultures were grown to stationary phase in a variety of genera specific media, as described in the Halohandbook (Dyall-Smith 2009) and by Papke *et al.* 2011 (Papke, White et al. 2011). Using diluted lysis buffer, cells were lysed to obtain crude cell contents. Then, the DNA was isolated either by Qiagen or Promega manufactured kits, according to each manufacturer's protocol. DNA was then quantified using a Nanodrop spectrophotometer ND-1000 at 260 nm. DNA was considered pure, and up to sequencing standards, if the 260/280 nm ratio was between 2.3 and 1.5. Pure DNA was stored at -20° C in 10x Tris EDTA buffer.

2.2 PCR

The forward and reverse primers used in this study were *bop* primers, bopF 5'-GAC TGG TTG TTY ACV ACG CC -3', and bopR 5'-AAG CCG AAG CCG AYC TTB GC-3,' described in Papke *et al.* 2007 (Papke, Zhaxybayeva et al. 2007) and originated in Kamekura *et al.*

(Kamekura 1998). Using different polymerase enzymes and reaction mixes, two distinct sets of PCR reactions were carried out. All reactions were carried out in an Eppendorf Mastercycler EPgradient thermocycler, and each set had a different thermal profile. Amplicons were visualized on a 1% agarose gel under UV light for a band of the expected size of 500 base pairs.

The first set of PCR reactions were performed in 25 µl reaction mixtures with *Taq* polymerase using the following reaction mixture: 10× PCR buffer, 2.5 µl; dNTPs [100 mM], 0.5 µl; $MgCl_2$ [50 mM], 0.75 µl; forward and reverse primers [10 µM], 0.75 µl each; *Taq* [5 units/µl], 0.5 µl; genomic DNA [20ng/µl], 0.5 µl; and 18.75 µl of dH_2O. Bacteriorhodopsin gene fragments were amplified with an initial denaturation at 95° C for 2 minutes, followed by 30 cycles of additional denaturation at 95° C for 45 seconds, primer annealing at 54.8° C for 45 seconds and primer extension at 68° C for 30 seconds. After the final cycle, a terminal extension step at 68° C for 5 minutes was performed.

A second set of PCR reactions were performed using Phusion polymerase and modified primers. Primers were modified by adding M13 universal primers to their 5' ends. The M13 regions added to the forward and reverse primers were 5'-TGTAAAACGACGGCCAGTG-3' and 5'-CAGGAAACAGCTATGAC-3', respectively. PCR was performed in 25 µl reaction mixtures containing: Phusion polymerase [2 units/µl], 0.25 µl; Acetamide [25% w/v], 13.25 µl; 5x GC buffer, 5 µl; DMSO [100%], 2.5 µl; dNTPs [100 mM], 1 µl; forward and reverse primers [10 µM], 0.5 µl each; genomic DNA [20ng/µl], 1.0 µl; and 1.0 µl of dH_2O. Fragments were amplified with an initial denaturation at 94° C for 1 minute, followed by 30 cycles of additional denaturation at 94° C for 30 seconds, primer annealing at 56.2° C for 30 seconds, and primer extension at 68° C for 30 seconds. The thermal cycling was completed by terminal extension at 68° C for 5 minutes before storing at 4° C.

2.3 Purification

Both of the reaction sets mentioned were purified using separate protocols and kits from different manufacturers. Following purification, DNA was quantified on a Nanodrop spectrophotometer ND-1000 at 260 nm; any samples under 8 ng/µl were not used. Purity was assessed using the 260/280 nm and 260/230 nm ratios, and samples were re-purified if the ratios were not between 1.8-2.0 and 1.9- 2.3, respectively.

Promega PCR cleanup kit was used for the first set. PCR products were mixed with an equal volume of Membrane Binding Solution, then transferred onto a mini-column with collection tube, and incubated for 1 minute at 50° C. Samples were centrifuged at 16,000 x g for 1 minute, and filtrate was discarded. Membrane Wash Solution (700 µl) was added to the column and then spun at 16,000 x g for 1 minute, flow-through was discarded. The column was washed again with 500 µl of wash solution followed by a 5-minute centrifugation at 16,000 x g. A sterile 1.5 ml centrifuge tube replaced the collection tube, and 50 µl of TE buffer was added to the column. After a 1-minute incubation at 50° C, and 30 seconds of centrifugation, the eluate was collected and stored at 0° C. Samples were also occasionally purified from a gel slice by melting the slice that was cut out and following the above protocol.

The second set of purifications was performed using Qiagen's QIAquick PCR Purification Kit. Entire samples were first run on a 1% agarose gel and the DNA band visualized using ethidium bromide under UV light and excised (approximate size 500bp). This step eliminated any remaining free dNTPs and kept primer dimer formation from interfering with subsequent sequencing reactions. Gel slices were weighed and suspended in QG solubilization

and binding buffer (Guanidinium thiocyanate) at a 1:3 ratio (i.e. 100 mg of gel into 300 µl buffer), then incubated for 10 minutes at 50° C and vortexed until completely dissolved. One gel volume (i.e. 100 µl isopropanol per 100 mg of gel) of isopropanol was added to the mixture and vortexed. Aliquots of 800 µl were then loaded onto the provided filter column with attached collection tube, and the column was spun in a centrifuge at 17,900 x g for 1 minute. The flow-through was discarded. These steps were repeated for samples larger than 800 µl. A 500 µl QG buffer wash was performed with a 1-minute spin, and the flow-through was discarded. This removed any remaining agarose that could interfere with sequencing. To wash samples, 700 µl of PE wash solution (non-chaotropic salts in ethanol) was loaded onto the column, incubated 5 minutes then spun for 1 minute, and the flow-through was discarded. The column was then spun empty for 1 minute to remove traces of PE buffer, which can interfere with sequencing. DNA from the column was eluted using 40 µl of provided Elution Buffer (Tris), incubated between 3 and 8 minutes, and then spun for 1 minute. Eluent containing purified *bop* fragments was collected and stored at 0° C.

2.4 Sequencing and Editing

Most purified PCR products were mailed overnight in non-refrigerated containers to Genewiz Inc. for sequencing. The Genewiz sequencing was performed using BigDye chemistry and an ABI 3730xl DNA Analyzer. Additionally, four samples were successfully sequenced at the University of Connecticut using an ABI 3130xl DNA Analyzer. Sequences were analyzed using the Geneious software package (Drummond AJ 2006) and Finch TV (Geospiza). Finch TV was used for chromatogram based sequence editing and to resolve ambiguous bases.

Furthermore, Geneious allowed for identification of other technical errors including missing bases, leading to a frame shift like errors. After resolving ambiguous bases, Geneious was used to merge the forward and reverse DNA strands into a single consensus sequence, and then primers were trimmed off. Whenever one or both primers could not be found, the sequences were aligned against other similar sequences to identify the primer location(s). In cases where primers could still not be found, samples were reprocessed.

2.5 BLAST Search

In an attempt to identify sequenced amplicons, the Basic Locus Alignment Search Tool (BLAST) (Altschul, Madden et al. 1997) was used to compare all generated sequences to the NCBI GenBank to find similar genes. BLAST was also used to find additional sequences for phylogenetic reconstruction from GenBank (Bilofsky and Burks 1988). David Williams of the Gogarten Lab at UConn provided additional sequences, published and unpublished homologs to *bop*. See Table 1 for a list of sequences. Data mined sequences include accession numbers. All data mined sequences were also blasted to ensure that annotations were correct. Searches were performed at http://blast.ncbi.nlm.nih.gov/Blast.cgi (Altschul, Madden et al. 1997; Altschul, Wootton et al. 2005).

2.6 Alignment Editing

Mesquite software package (Maddison 2010) was used to calculate the codon positions of the trimmed sequences. These positions were used to find the correct open reading frame. Then

nucleotide sequences were translated into protein sequences and aligned using Clustalw2 (Thompson, Higgins et al. 1994) with pseudo-repeat alignments after every step. By overlaying nucleotide data on top of a protein alignment, a Hybrid alignment was created. This alignment had the biological stability of a protein alignment without sacrificing information provided by nucleotides. Regions in the alignment present in less than 33% of all taxa were removed from further phylogenetic analysis because they were not conserved enough to provide phylogenetic resolution, and could skew results.

2.7 Testing for Mutational Saturation

Mutational saturation of the third codon position was determined by the Dambe (Xia and Xie 2001) software package. Calculations were performed based on Xia's algorithm and using first and second codon positions. Dambe computed an Index of Substitution Saturation (Iss) and two Iss.c values, one that assumes a symmetric tree and another that assumes a highly asymmetric tree (Salemi, Vandamme et al. 2009). Differences between Iss and Iss.c values determined if the third codon position was to be included, excluded, or if the sequences should not be used for phylogenetic reconstruction at all (Table 5) (Salemi, Vandamme et al. 2009).

2.8 Phylogenetic Reconstruction

Optimal models of nucleotide evolution were estimated using jModelTest v0.1.1 considering 11 substitution types (Posada 2008). TM1 with estimated gamma (Γ) distribution was the optimum model, based on log likelihood. However, it was not used as it was not

available in the phylogenetic reconstruction programs used. GTR with Γ was the most appropriate model available suggested by jModelTest. ProtTest 3 was used to find the optimum protein model of evolution, VT + Γ in this case (Darriba, Taboada et al. 2011).

TreePuzzle (Schmidt, Strimmer et al. 2002) was used to identify any sequence heterogeneity based on the hybrid alignment. Sequence heterogeneity can skew final trees because evolutionary models assume similar sequence composition (Walsh, Bapteste et al. 2004). Testing was performed using the GTR + Γ model of evolution, estimated nucleotide frequency and TS/TV rates, and used Neighbor-Joining as its estimation algorithm. Only out-groups are intended to fail the test, and so were not included in testing. Sequences that failed the chi square test of sequence homogeneity were excluded from further analysis by PhyML.

Phylogenetic reconstruction of *bop* was performed using PhyML (Guindon and Gascuel 2003) software with maximum likelihood mapping. To avoid reproducing the BioNJ input tree, PhyML was seeded with 5 random trees and used sub-tree pruning and re-grafting, a slower but more comprehensive algorithm than nearest neighbor interactions. Support for the maximum likelihood phylogeny was generated using bootstrap analysis through PhyML. One hundred replicates were taken, and all bootstrap values over 50% are displayed on the tree (Figure 20).

2.9 Approximately Unbiased Testing

Approximately Unbiased (AU) testing was used to determine congruence between the alignment and tree topologies (Shimodaira 2002; Papke, White et al. 2011). Ten different tree topologies were created using TreeView (Page 1996) to compare against the alignment. Site-wise Log Likelihood for the set of maximum likelihood trees were calculated by RAxML

(Stamatakis 2006) using GTR with gamma, then AU testing was performed by Consel (Shimodaira 2002). AU testing generated P-values of significance, a P-value of 1.0 signifies congruence between the alignment and tree, while a P-value below 0.05 indicates incongruence; any value in between 0.05 and 1.0 indicates some degree of congruence. An incongruent result indicates a disagreement between the suggested evolutionary histories and can be interpreted as support for a HGT event (Papke, White et al. 2011). Test trees and log likelihoods are available in the appendix.

3. Results:

3.1 Sequenced Amplicon Identification

Sequences were generated through PCR amplification and sequencing reactions performed by Genewiz Inc. To ensure that they were proton pumps and not another form of rhodopsin, such as sensory rhodopsin, the sequenced amplicons were subjected to BLAST searches against a non-redundant database. BLAST results are presented as statistics: Percent Identity, and E value. Percent Identity refers directly to the similarity between a match and query. Though Percent Identity may be used to characterize match quality, it is not sufficient to locate a significant result, because it does not incorporate match coverage size so short matches can have high identity but low significance. To identify a significant match E value was used instead of Percent Identity. E values represent the likelihood of a match with the same degree of similarity being found by chance when searched in a given database, in this case GenBank. E value is inversely correlated to a match's significance. For all sequences, BLAST results were

highly similar. Table 1 shows BLAST search results of a representative member of the strains included in the final tree, including non-bacteriorhodopsin proton pumps and exact matches from genomes. The highest quality result was always either a complete or partial bacteriorhodopsin with a very low E value, and high identity nearing 0.0 and 90%, respectively. In some cases, identical sequences were found because the strain's genome were already in the database or because of prior use of primers modified from Kamekura *et al.* 1998 (Kamekura 1998), from which the primers used in this study originated (Papke, Zhaxybayeva et al. 2007). Other quality hits include other proton pumps such as cruxrhodopsin and archaerhodopsin. BLAST results showed that the amplified genes are most likely to be proton pumps. Searches were also performed on data mined sequences; the sequences were found to be correctly annotated.

Table 1. Typical BLAST Search Results (*Har. marismortui)*

Accession	Description	Max score	Total score	Query coverage	E value	Max ident
AY596297.1	Haloarcula marismortui ATCC 43049 chromosome I, complete seque	650	650	100%	0.0	100%
D11058.1	Halobacterium salinarum gene for bacteriorhodopsin, partial cds, str	609	609	100%	5e-171	97%
D31882.1	Haloarcula vallismortis cop3 gene for cruxrhodopsin-3, complete cds	506	506	100%	5e-140	91%
CP002921.1	Haloarcula hispanica ATCC 33960 chromosome I, complete sequenc	491	491	100%	1e-135	90%
DQ352132.1	Haloarcula sp. A819 bacteriorhodopsin gene, partial cds	489	489	99%	3e-135	90%
AY279550.1	Haloarcula sp. AJ4 bacteriorhodopsin (bop) gene, partial cds	486	486	99%	4e-134	89%
DQ146999.1	Haloarcula sp. YW016 bacteriorhodopsin (bop) gene, partial cds	475	475	94%	8e-131	90%
AB029320.1	Haloarcula japonica gene for cruxrhodopsin, complete cds	455	455	100%	7e-125	88%
D11057.1	Halobacterium salinarum gene for bacteriorhodopsin, partial cds, str	425	425	100%	1e-115	86%
D31880.1	Haloarcula argentinos cop1 gene for cruxrhodopsin-1, complete cds	419	419	100%	5e-114	85%
CP001688.1	Halomicrobium mukohataei DSM 12286, complete genome	230	230	100%	5e-57	74%
S76743.1	cop-2=cruxrhodopsin-2 [Haloarcula, arg-2, Andes heights isolate, C	226	226	100%	6e-56	73%
EF558555.1	Haloarchaeon TP072 bacteriorhodopsin (bop) gene, partial cds	147	147	99%	5e-32	69%
EF558556.1	Haloarchaeon TP100 bacteriorhodopsin (bop) gene, partial cds	143	143	99%	6e-31	68%
CP001687.1	Halorhabdus utahensis DSM 12940, complete genome	138	138	96%	3e-29	69%
EF558554.1	Halosimplex carlsbadense bacteriorhodopsin (bop) gene, partial cds	127	127	92%	5e-26	68%
EF558559.1	Haloarchaeon TP262 bacteriorhodopsin (bop) gene, partial cds	114	114	94%	3e-22	67%
EF558562.1	Natronorubrum sp. Tenzan-10 bacteriorhodopsin (bop) gene, partia	60.8	60.8	55%	6e-06	67%
AY279548.1	Natrinema ajinwuensis bacteriorhodopsin (bop) gene, partial cds	57.2	57.2	29%	7e-05	71%

The BLAST results, while highly suggestive of a proton pump require additional confirmation. Additional support for these genes' identity was obtained by analyzing key amino acids required for proton pumping and other residues characteristic of *bop*. A protein alignment of all *bop* genes used in this study with these amino acids denoted can be seen in Figure 11. Residues considered characteristic of *bop* (Brown 2001) have a blue diamond placed at the bottom of their column, green diamonds are placed under residues whose augmentation results in

a drastic change of proton pumping function (Brown 2001) (Figure 11). A high level of functional conservation was observed across all genera; only two discrepancies were found and are highlighted in yellow. Based on the strong agreement between these two independent analyses, it was concluded that all genes used in the final tree are highly likely to translate for proton pumps.

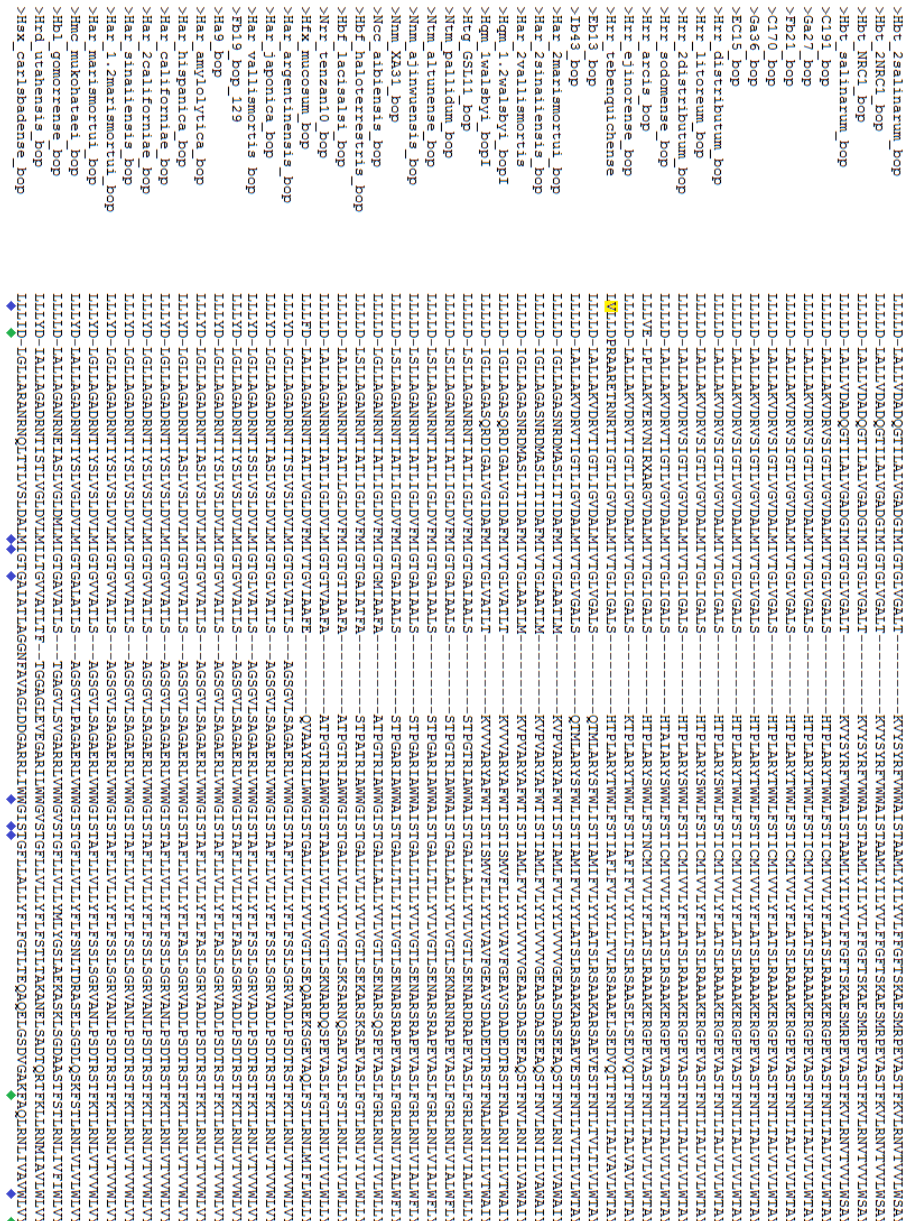

Figure 11 – Protein Alignment of Used bacteriorhodopsin Sequences

A protein alignment of all partial bacteriorhodopsin sequences in the final analysis. Important residues are denoted by diamonds, blue for characteristic *bop* residues, and green for residues whose augmentation results in a drastic change of proton pumping function. Unconserved amino acids are highlighted in yellow. Dashed lines represent gaps introduced into the sequence to compensate for insertions and deletions

One gene not included in the final unrooted tree was amplified from *Hrr. tibetense*, because the gene was concluded to be a sensory rhodopsin II gene (*sopII*) and not a *bop*. BLAST results for *Hrr. tibetense* were different from the others: having high quality bacteriorhodopsin and sensory rhodopsin II hits (Table 2). Though the SRII hits were of lower quality (higher E value, and lower Percent Identity) than the bacteriorhodopsin hit, they were of a significantly low E value. Based on BLAST search data alone the gene may be a *sopII* rather than a *bop* but, it is not sufficient to make a conclusion, and additional confirmation is required.

Table 2. BLAST Search Results of *Hrr. tibetense*

Accession	Description	Max score	Total score	Query coverage	E value	Max ident
CP001365.1	Halorubrum lacusprofundi ATCC 49239 chromosome 1, complete se	354	354	100%	2e-94	81%
AB059748.1	Halorubrum chaoviator aus1pR gene for phoborhodopsin, complete	298	298	98%	1e-77	78%
AM774415.1	Halobacterium salinarum complete genome, strain R1	104	104	96%	5e-19	67%
AE004437.1	Halobacterium sp. NRC-1, complete genome	104	104	96%	5e-19	67%
U62676.1	Halobacterium salinarium phototaxis transducer II and sensory rho	104	104	96%	5e-19	67%
CP002896.1	Amycolatopsis mediterranei S699, complete genome	50.0	50.0	12%	0.010	84%
CP002000.1	Amycolatopsis mediterranei U32, complete genome	50.0	50.0	12%	0.010	84%

Since the BLAST results were not sufficient to confirm the identity of the gene, the functional amino acids of known SRII proteins were compared to the inferred amino acid sequence of the gene amplified form Hrr. tibetense. A protein alignment of known sensory rhodopsins and gene from *Hrr. tibetense* found that a number of key residues (Royant, Nollert et al. 2001) were conserved while others were not (Figure 12). Amino acids essential for SRII function (Royant, Nollert et al. 2001) are highlighted in light blue (Figure 12). Functionally conserved or similar but not identical amino acids at positions characteristic of SRII are marked by a light blue diamond, and unconserved residues are denoted by a black diamond. Residues found in both SRII and bacteriorhodopsin are not shown. The functional amino acid analysis showed that it is highly likely that *sopII* was amplified from *Hrr. tibetense*, and so was not included in the final *bop* phylogenetic analysis.

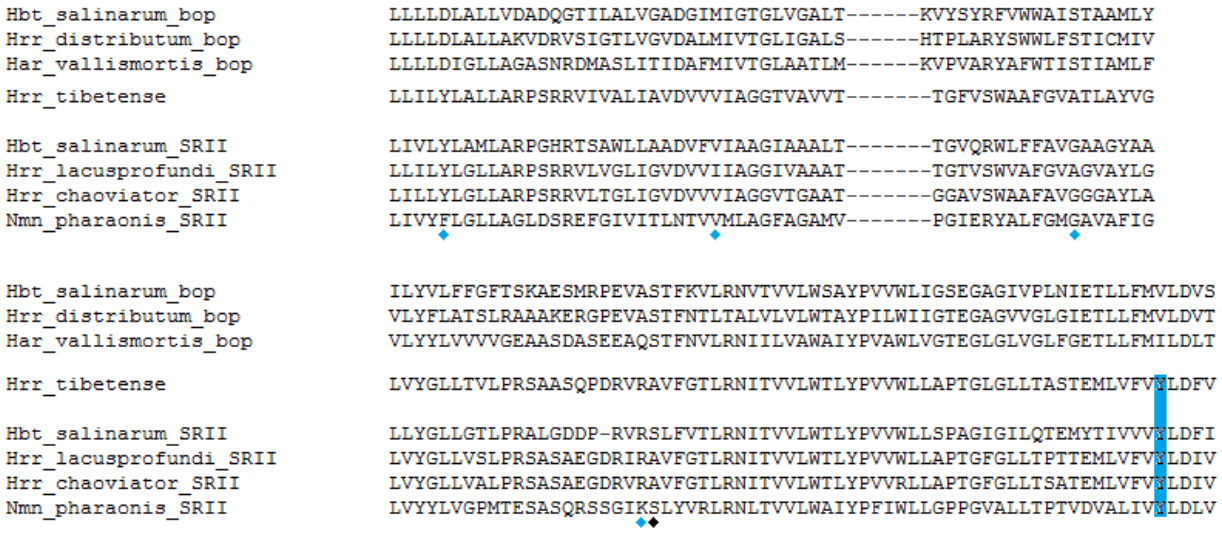

Figure 12 – Comparison of SRII and BR Protein Sequences to *Hrr. tibetense* Gene

A protein alignment of all partial *bop* sequences in the final analysis with key amino acid residues denoted. Amino acids essential for SRII function are highlighted in light blue. Diamonds indicate residues characteristic of SRII, blue for functionally conserved ones and black for unconserved residues. Residues found in both SRII and bacteriorhodopsin are not shown.

3.2 Possible Insertion/Deletion

A seven-amino-acid insertion/deletion (indel) was observed in multiple genera of Halobacteriales specifically *Haloarcula, Halobaculum*, and *Halosimplex (Hsx.),* (Figure 13), both in the generated and data mined sequences. The indel sequence appears to be genus specific, conserved within a genus but varies between genera. All observed indels were localized to an extracellular non-helical region of bacteriorhodopsin after residue 130, between helixes D and E (Figure 5)

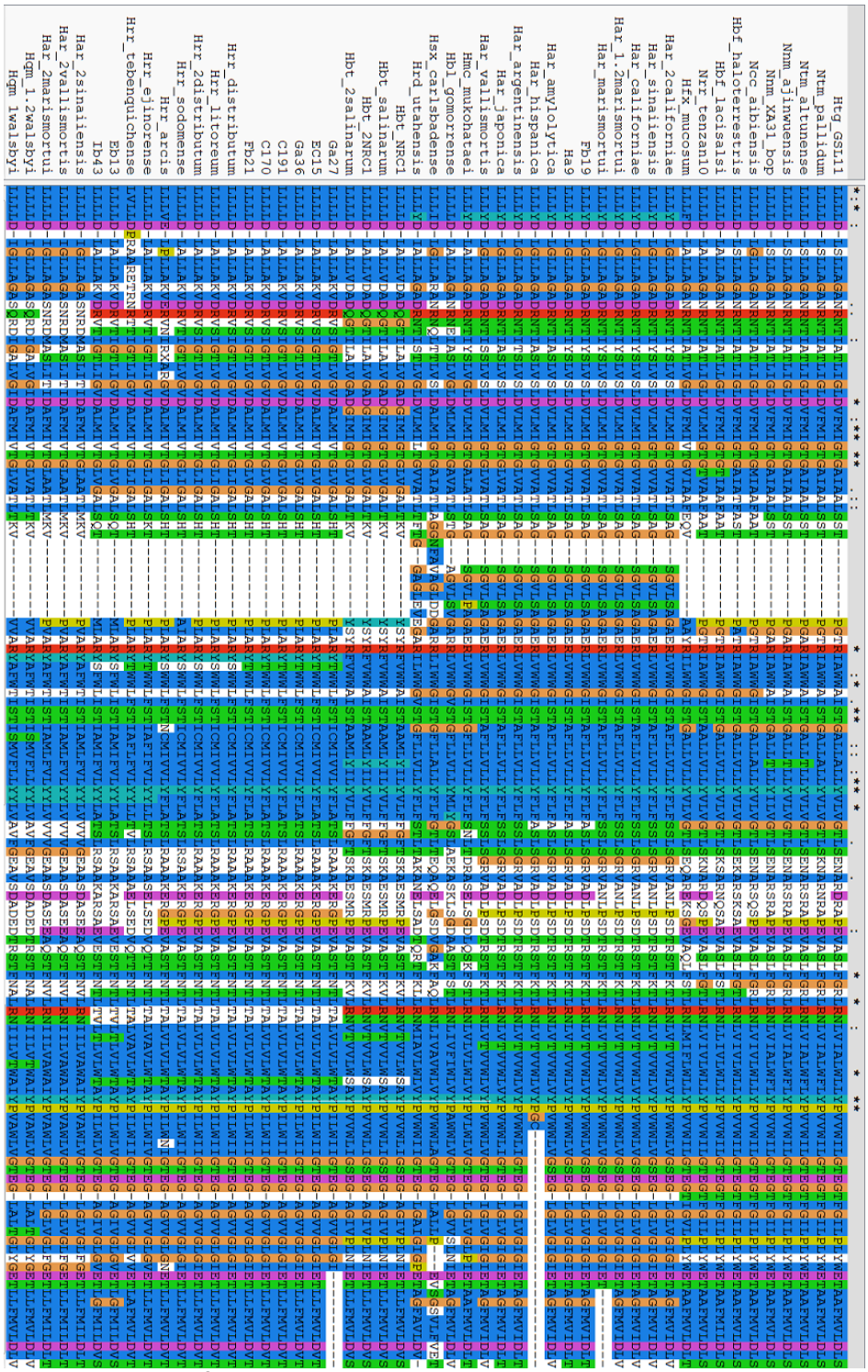

Figure 13 – Protein Alignment of *bop* Genes

The protein alignment was generated using ClustalW2 (Thompson, Higgins et al. 1994). It shows the general sequence similarity across *Haloarchaea* as well as the indel which is primarily localized mainly to *Haloarcula*. The symbols "*", ":", and "." indicate if the amino acids at that position are identical, conserved or semi-conserved respectively. Residues are colored when a set fraction at a position are chemically similar. Blue represents hydrophobic residues, while green corresponds to polar residues with uncharged side chains. Amino acids with negatively charge side chains are colored magenta while positively charged ones are red. Additionally, Gly, Pro, and Cys are shown in orange, yellow and pink respectively. Image provided by Maulik Jani and Clustal.

3.3 New bop in Hbl. gomorrense

A previously unpublished *bop* gene fragment was sequenced from a genomic DNA sample of *Hbl. gomorrense*. When this sequence was subjected to BLAST searches, the top hits corresponded to proton pumps, primarily bacteriorhodopsin (Table 3), suggesting the gene fragment is a *bop*. This sequence was included in the residue analysis in section *"3.1 Sequenced Amplicon Identification"*, where it was found to have all essential residues conserved (Figure 11). Additional studies, including a phenotypic study and a study attempting to amplify the gene from cells not expressing bacteriorhodopsin, are warranted to determine if the gene is silent or generally missing in *Hbl. gomorrense*.

Table 3. BLASTn Results (*Hbl. gomorrense)*

Accession	Description	Max score	Total score	Query coverage	E value	Max ident
EF558556.1	Haloarchaeon TP100 bacteriorhodopsin (bop) gene, partial cds	318	318	99%	1e-85	78%
EF558555.1	Haloarchaeon TP072 bacteriorhodopsin (bop) gene, partial cds	300	300	99%	3e-80	77%
S76743.1	cop-2=cruxrhodopsin-2 [Haloarcula, arg-2, Andes heights isolate, C	259	259	98%	1e-67	75%
CP001688.1	Halomicrobium mukohataei DSM 12286, complete genome	253	253	98%	4e-66	74%
AB029320.1	Haloarcula japonica gene for cruxrhodopsin, complete cds	208	208	98%	2e-52	72%
D21882.1	Haloarcula vallismortis cop3 gene for cruxrhodopsin-3, complete cd	197	197	96%	3e-49	71%
EF558554.1	Halosimplex carlsbadense bacteriorhodopsin (bop) gene, partial cds	188	188	100%	1e-46	71%
EF558553.1	Halobiforma haloterrestris bacteriorhodopsin (bop) gene, partial cds	178	178	97%	3e-43	70%
CP001687.1	Halorhabdus utahensis DSM 12940, complete genome	176	176	87%	9e-43	72%
AY607024.2	Halobiforma lacisalsi bacteriorhodopsin gene, complete cds	172	172	97%	1e-41	70%
AY279551.1	Halobiforma lacisalsi bacteriorhodopsin (bop) gene, partial cds	165	165	98%	2e-39	69%
EF558559.1	Haloarchaeon TP262 bacteriorhodopsin (bop) gene, partial cds	123	123	99%	5e-27	67%
AY279547.1	Natronococcus aibiensis bacteriorhodopsin (bop) gene, partial cds	118	118	98%	2e-25	67%
EF558562.1	Natronorubrum sp. Tenzan-10 bacteriorhodopsin (bop) gene, partia	109	109	98%	1e-22	66%
AY279549.1	Natrinema altunense strain AB3 bacteriorhodopsin (bop) gene, parti	107	107	97%	4e-22	66%

3.4 Composition Bias of bop Gene

Alignments were tested for compositional bias, which can skew later phylogenetic reconstruction. Chi-square test of sequence homogeneity was performed though TreePuzzle on both the nucleotide and protein alignments, any sequences that failed were excluded from tree construction. For the hybrid alignment, *Halorhabdus utahensis, Hfx. zhejiangensis* and the second copy of *bop* (*bop2*) from *Hqm. walsbyi* failed and were excluded from further analysis and tree construction. When the test was performed on the protein alignment *Hsx. carlsbadense*

also failed aside from the strains that failed previously.

3.5 Dambe Test for Third Codon Saturation

To determine whether the third codon position was saturated and should be removed from phylogenetic reconstruction, Dambe (Xia and Xie 2001) was used to test for saturation. Dambe was also used to assess whether the data set was suitable for phylogenetic construction (Xia and Xie 2001). The hybrid alignment's third codon position was found to have some saturation, but it was not substantial enough to warrant removal of the third nucleotide position from further analysis. Testing was carried out as described in section "*2.7 Testing for Mutational Saturation"* and resulted in an Iss and two Iss.c (symmetric and asymmetric) values. The Iss and Iss.c (symmetric) values were calculated to be 0.546 and 0.682, respectively. A symmetric Iss.c value was used over the asymmetric because of the distribution of genera in the *bop* tree (Salemi, Vandamme et al. 2009). Dambe testing showed "Little Saturation" as the Iss was significantly lower than Iss.c. Therefore, all positions were used in the creation of phylogenetic trees. Dambe results can be seen in Table 4.

Table 4. Dambe Results

NumOTU	Iss	Iss.cSym	T	DF	P	Iss.cAsym	T	DF	P
32	0.546	0.682	3.704	220	0.0003	0.355	5.215	220	0.0000

3.6 Phylogenetic Reconstruction of bop

A goal of this study was to use partial *bop* genes, from 15 genera, to generate a relatedness tree in order to better understand the evolutionary history of *bop*. While a number of sequences were generated through amplification and sequencing, others were obtained from NCBI. All sequences and their origins are available in the appendix.

These sequences were used to generate a maximum likelihood phylogeny for the *bop* gene within haloarchaea (Figure 14). Three large clades and two smaller groups were formed by tree construction. The first clade consisted primarily of *Haloarcula*, with a single representative from *Halobaculum, Halosimplex*, and *Halomicrobium*. A second clade was comprised of *Halorubrum* strains and a third (Mixed) clade was a mix of several genera, including *Haloferax, Halobiforma, Natrinema* and others (Figure 15). Of the two remaining groups, one consisted solely of *Halobacterium*. The other group was comprised of *Haloquadratum* and the second copies of *bop* found in *Haloarcula*.

The present study incorporated three second copies of *bop*, from different members of the *Haloarcula* genus. As in previous studies, second copy *bop* genes grouped with the first copy from *Hqm. walsbyi* (Sharma, Walsh et al. 2007). In this study, all three grouped near to *Hqm. walsbyi*, and all three were highly similar in terms of sequence.

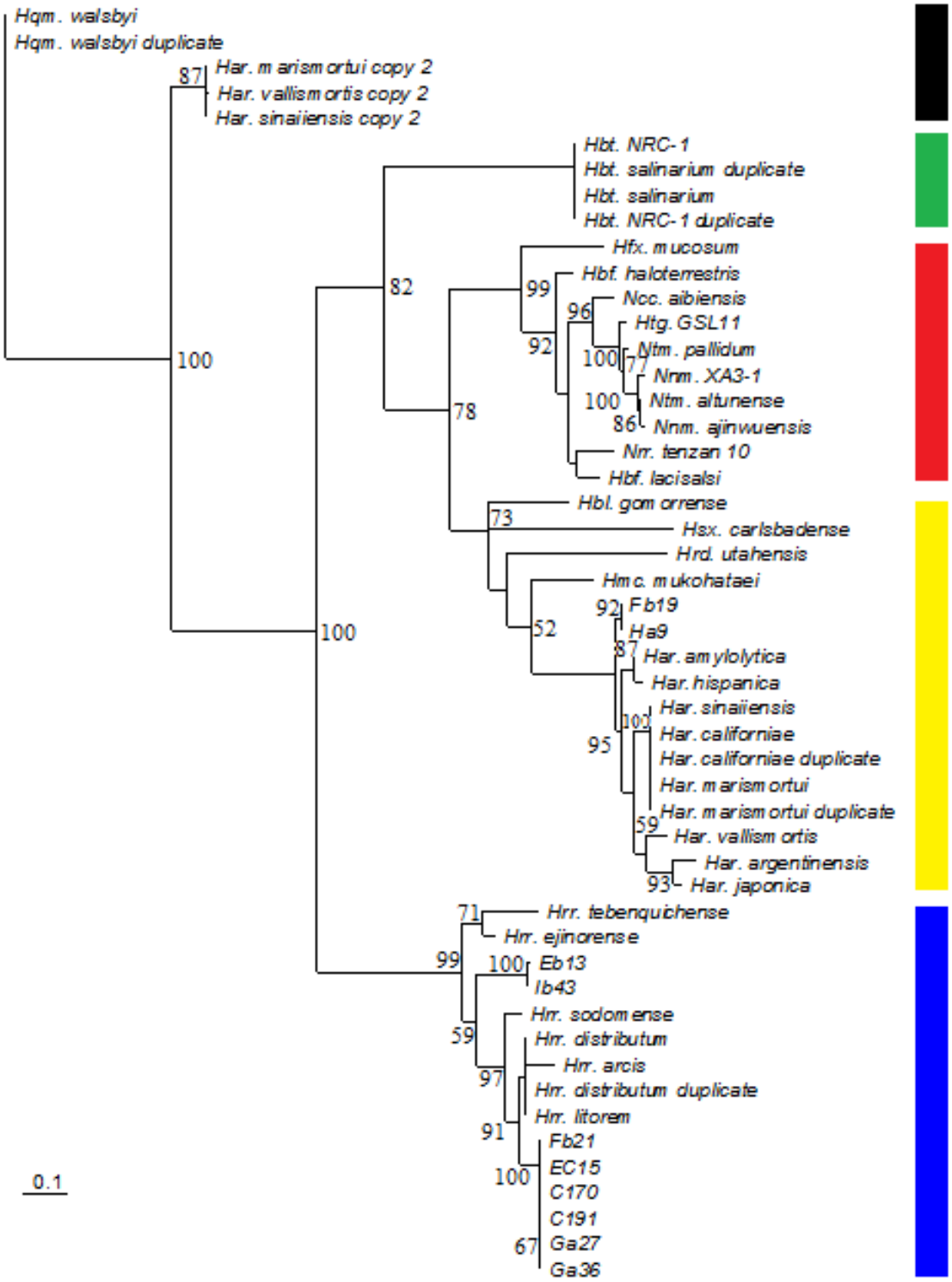

Figure 14 – Single Gene ML Tree from *bop* (Hybrid) – Unrooted

A maximum likelihood tree constructed from a nucleotide data overlaid on top of a protein alignment. Bootstrap values above 50% are shown. Clades are denoted by different colored bars. The black bar represents the *bop2* and *Haloquadratum* clade. The green bar is the *Halobacterium* clade. The blue bar indicates representatives of the *Halorubrum* clade, the red bar indicates member of the Mixed clade, while the yellow bar indicates the clade in which *Haloarcula* is a major member.

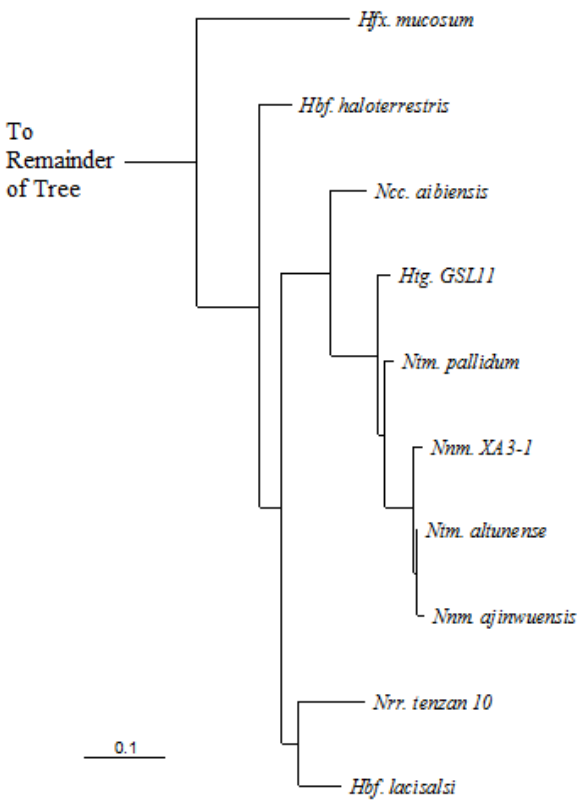

Figure 15 – Mixed Clade
An isolated portion from Figure 14 showing the "Mixed Clade," one of the major clusters formed within the phylogenetic reconstruction of the *bop* gene. Bootstrap values are not shown, but may be seen in Figure 14. Image created by Maulik Jani using TreeView (Page 1996)

3.7 Approximately Unbiased Testing

Approximately Unbiased (AU) testing is used to determine the degree of agreement between an alignment and a phylogenetic tree. It is performed as described in section "*2.11 Approximately Unbiased Testing"*. Here the evolutionary history of ten unique trees was compared to information presented in the hybrid alignment. The set of ten trees included the unmodified constructed tree as a positive control, a randomized tree as a negative control, and

eight test trees designed to test different possible topologies. Of the remaining eight trees, four were based on strict and loose versions of MLSA and 16s rRNA gene topologies. Two other trees were created with particular differences; one involved moving *Hfx. mucosum,* another tree involved the movement of *Halobacterium* so that it grouped closer to *Halorubrum* as seen in other trees. Another tree included moving *Hfx. mucosum*, *Hqm. walsbyi* and the second copies of *bop* to MLSA positions. The final tree involved moving the mixed clade so that it was closer to *Halorubrum* mimicking its position in the phylogeny constructed from 16s rRNA gene sequences. These trees are available in the appendix.

AU testing results are shown in Table 8, the original tree scored highest with a p-value (0.742) while the negative control was strongly rejected ($p<0.005$). The four trees (Tree Item 3, 4, 8, 10) that included the movement of *Haloferax* were also strongly rejected ($p < 0.005$). And, thus, any movement of *Haloferax* from its position is not congruent with the alignment. Additionally, movement of the mixed clade closer to *Halorubrum* (Tree Item 6) rather than with *Haloarcula* was also strongly rejected ($p<0.005$), as was the non-strict 16s rRNA gene tree (Tree Item 9) that included a similar move. A tree (Tree Item 5) that swapped the *Haloquadratum* and *Halobacterium* groups had a p-value (0.105) slightly above rejection. And the tree (Tree Item 7) that grouped *Haloquadratum* with *Halorubrum,* as seen in the MLSA tree, was not rejected (p-value = 0.670), however it was not a significant change to begin with.

Table 5. AU Testing

Rank	Item Number	Obs.	AU	np	bp	pp	kh	sh	wkh	wsh
1	1*	0	0.742	0.402	0.402	0.5	0.51	0.965	0.51	0.965
2	7~	0	0.67	0.433	0.433	0.5	0.49	0.954	0.49	0.954
3	5°°	6.9	0.105	0.102	0.101	5.00E-04	0.111	0.725	0.11	0.37
4	9~~~	7.8	0.067	0.06	0.064	2.00E-04	0.077	0.708	0.077	0.281
5	6°°°	23.7	0.002	0.001	0.001	2.00E-11	0.008	0.329	0.008	0.028
6	4°	52.5	7.00E-06	1.00E-05	0	8.00E-24	1.00E-04	0.124	1.00E-04	0.001
7	8~~	52.5	7.00E-06	1.00E-05	0	8.00E-24	1.00E-04	0.124	1.00E-04	0.001
8	3***	52.5	7.00E-06	1.00E-05	0	8.00E-24	1.00E-04	0.124	1.00E-04	0.001
9	10[h]	84.3	3.00E-29	2.00E-13	0	1.00E-37	0	0.023	0	0
10	2**	909.3	0.002	8.00E-07	0	0	0	0	0	0

*	Original Maximum Likelihood Tree
**	Negative Control (Original with random changes)
***	*Haloferax* Grouped with Halorubrum Clade
°	Second Copies of *bop* from *Halorubrum* and the *Haloferax* Sequence Grouped with Halorubrum Clade
°°	*Halobacterium* Clade Grouped with *Halorubrum* Clade
°°°	Mixed Clade Grouped with *Halorubrum* Clade
~	Modeled After MLSA phylogeny
~~	Modeled After MLSA phylogeny (Strict Version)
~~	Modeled After 16s rRNA phylogeny
[h]	Modeled After 16s rRNA phylogeny (Strict Version)

4. Discussion:

4.1 Difficulties in Obtaining Sequence Data

A single amino acid can be specified by multiple nucleotide combinations. As a result nucleotides can change without altering the amino acid. In some cases, including that of bacteriorhodopsin, even amino acids may change without greatly affecting the protein. To compensate, degenerate primers, a mixture of similar primers in equal concentrations intended to include all likely permutations of the gene's priming regions, are used. This degeneracy allows for the amplification of a gene across a wide variety of genera. However, if the primer is too

degenerate, the likelihood of the correct sequence binding to the template strand decreases; if a primer does not bind during early PCR cycles, there may not be sufficient amplicons to be visualized; and lastly, if the primers are not degenerate enough, the correct primer may not be present in the mixture. The third scenario would result in false negative results, particularly when attempting to amplify from multiple genera.

A number of PCR difficulties were encountered in this study. Initially, PCR resulted in bands of expected size (500 bases) but did not amplify well across all genera, despite acetamide, DMSO, and temperature optimization. To improve PCR efficiency, primers were redesigned from a nucleotide alignment created from *bop* genes available through GenBank. New primers had fewer ambiguous bases as reflected in the improved conservation of the priming region in the new alignment, but the results were ultimately mixed. While PCR efficiency increased and additional strains successfully amplified, sequence chromatograms were ambiguous and amplicons more frequently appeared as a smear. Different PCR techniques were then used, including nested and touchdown PCR, without significant improvement to results. In addition, the polymerase was changed from native *Taq* to a variety of hot-start mixes and other polymerases based on the success another lab member had with different enzymes to solve similar problems. Phusion polymerase (designed by Finnzymes) was eventually adopted due to 1) Phusion's ability to amplify *bop* genes that had previously failed and 2) Phusion's higher fidelity (50x that of *Taq*) (Gilje, Heikkila et al. 2008) eliminated random gaps, making chromatograms easier to read and reduced errors. Additionally, Phusion either gave a single clean band or no band at all, as opposed to the multiple bands or an inconclusive smear that could appear when samples were amplified using *Taq*. Thus, the use of Phusion with redesigned primers greatly increased the likelihood of a successful PCR amplification.

Nevertheless, certain samples did not amplify using the *bop* primers. While these failed consistently, it is still not possible to claim these strains do not contain a *bop* gene, though it is a strong possibility. There are a number of alternative possibilities, including mutations as well as divergent priming regions, as seen in *Haloquadratum*. The strains that failed to amplify are listed in Table 6. A positive amplification result was characterized by the presence of a correct sized band (approximately 500 bases), identified through gel electrophoresis.

Low sequencing efficiency was originally attributed to primer degeneracy, because the low concentrations of primers used during sequencing reactions reduces the likelihood of a correct primer-DNA match. Therefore, M13 regions, derived from M13 bacteriophage, were added to the 5' end of existing primers, (forward and reverse) as described in section "*2.2 PCR*". These non-degenerate "sequencing" primer sections allowed for purified amplicons to be shipped directly to Genewiz for sequencing, which simplified the sequencing process, increased efficiency, and eliminated errors caused by improper mixing of primer and amplicons. While PCR efficiency was very high and reliable, with failed PCRs resulting from a lack of the gene rather than experiment failure, sequencing continued to fail. Other attempts to improve success rates have also largely failed, despite the good quality of template DNA and samples being sent out no more than a day after amplification. This issue is currently being further investigated. Primer and M13 sequences can be found in Methods section "*2.2 PCR"*.

Across repeated successful amplifications, multiple PCR products listed in Table 6 failed to sequence, even though a variety of different protocols were attempted including additional purifications, use of M13 primers and overnight shipping. Finally, numerous unprocessed gel cutouts were sent to Genewiz to be purified and then sequenced. These strains are detailed Table 6.

A possible explanation for these failures is that a gene with similar priming regions was amplified either instead of or in addition to *bop*; the more stringent sequencing conditions then resulted in a failed reaction or the sequencing of two genes results in an un-rectifiable chromatogram. Though supported by the ability of these primers to amplify genes other than *bop*, such as *sopII* (sensory rhodopsin II gene) as seen in section "*3.1 Sequenced Amplicon Identification*", this explanation is unlikely since multiple bands would be present if two genes were amplified. However, if the similar gene was the same size as *bop* (both around 500 bases), individual bands may not be distinguishable during gel electrophoresis. As their cause is unknown, the failure of some of these sequences should be further analyzed.

Table 6. Amplification and Sequencing Results

Species	Strain Identifier	Amplification[x][a]	Sequencing[x][h]	Species	Strain Identifier	Amplification[x][a]	Sequencing[x][h]
Haladaptatus paucihalophilus	JCM 13897	x	x	Halorubrum terrestre	JCM 10247	√	x
Halalkalicoccus jeotgali	JCM 14584	x	x	Halorubrum tibetense	JCM 11889	√	√
Halalkalicoccus tibetensis	JCM 11890	√	x	Halorubrum trapanicum	JCM 10477	x	x
Haloarcula amylolytica	JCM 13557	√	√	Halorubrum vacuolatum	JCM 9060	√	x
Haloarcula argentinensis	JCM 9739	√	√	Halorubrum xinjiangense	JCM 12388	√	x
Haloarcula californiae	JCM 8912	√	√	Halosarcina pallida	JCM 14848	√	x
Haloarcula hispanica	ATCC 33960	√	√	Halosimplex carlsbadense	JCM 11222	√	√
Haloarcula japonica	JCM 7785	√	√	Halostagnicola larsenii	JCM 13463	x	x
Haloarcula marismortui	ATCC 43049	√	√	Haloterrigena sp.	GSL-11	√	√
Haloarcula quadrata	JCM 11048	√	x	Haloterrigena turkmenica	ATCC 51198	x	x
Haloarcula sinaiiensis	ATCC 33800	√	x	Halovivax asiaticus	JCM 14624	√	x
Haloarcula vallismortis	ATCC 29715	√	√	Halovivax rubber	JCM 13892	x	x
Halobacterium jilantaiense	JCM 13558	√	√	Natrialba aegyptia	JCM 11194	x	x
Halobacterium noricense	JCM 15102	x	x	Natrialba asiatica	JCM 9576	x	x
Halobacterium salinarum	NCIMB 2288	√	√	Natrialba magadii	JCM 8861	x	x
Halobacterium sp.	NRC-1	√	√	Natrinema altunense	JCM 12890	√	√
Halobaculum gomorrense	JCM 9908	√	√	Natrinema pallidum	NCIMB 784	√	√
Halobiforma haloterrestris	JCM 11627	√	x	Natronococcus jeotgali	JCM 14583	√	x
Halobiforma lacisalsi	JCM 12983	√	x	Natronomonas pharaonis	JCM 8858	x	x
Halobiforma nitratireducens	JCM 10879	x	x	Natronorubrum aibiense	JCM 13488	√	x
Halococcus dombrowskii	JCM 12289	√	x	Natronorubrum sp.	Tenzan-10	√	x
Halococcus hamelinensis	JCM 12892	x	x	Unnamed	GA27	√	√
Halococcus morrhuae	NRC 16008	x	x	Unnamed	FB21	√	√
Halococcus qingdaonensis	JCM 13587	x	x	Unnamed	Ha9	√	√
Halococcus saccharolyticus	ATCC 49257	x	x	Unnamed	EC15	√	√
Halococcus salifodinae	JCM 9578	x	x	Unnamed	GA2P	√	x
Haloferax alexandrinus	JCM 10717	x	x	Unnamed	Fb19	√	√
Haloferax denitrificans	ATCC 35960	x	x	Unnamed	Eb13	√	√
Haloferax volcanii	NCIMB 2287	x	x	Unnamed	Ib43	√	√
Halogeometricum borinquense	JCM 10706	√	x	Unnamed	Ea1	√	x
Halomicrobium mukohatei	DSM 12286	√	√	Unnamed	Ga66	x	x
Haloplanus natans	JCM 14081	x	x	Unnamed	C191	√	√
Haloquadratum walsbyi	DSM 16790	√	x	Unnamed	Ga36	√	√
Halorhabdus utahensis	DSM 12940	√	x	Unnamed	C170	√	√
Halorubrum aidingense	JCM 13560	x	x				
Halorubrum arcis	JCM 13916	√	√	[x] Positive results are denoted by "√" while negative results are represented by a "x"			
Halorubrum coriense	JCM 9275	√	x				
Halorubrum distributum	JCM 9100	√	√	[a] Positive results were characterized by the presence of a correctly sized band			
Halorubrum ejinorense	JCM 14265	√	√	(~500 bases), identified through gel electrophoresis			
Halorubrum lacusprofundi	JCM 8891	x	x	[a] Negative results were characterized by the lack of the correct band			
Halorubrum lipolyticum	JCM 13559	x	x				
Halorubrum litoreum	JCM 13561	√	√	[h] Positive results were characterized by the presence of both primers and			
Halorubrum saccharovorum	JCM 8865	x	x	a sequence resembling bop			
Halorubrum sodomense	JCM 8880	√	√	[h] A negative result were recorded when either primer could not be located			
Halorubrum tebenquichense	JCM 12290	√	√	or the sequence was incomplete, of poor quality, or contained other notable errors			

4.2 Rhodopsin Identity

Based on BLAST search results and functional residue analysis, all sequenced amplicons other than that of *Hrr. tibetense* are likely to be *bop* genes. Analysis of functional amino acid residues is strongly supportive as all sites were conserved in the majority of strains. There were two cases of disagreement, in both the correct amino acid was substituted for a valine. Leucine-93 was substituted for a valine in *Hrr. Tebenquichense;* however, since both are branched non-polar (hydrophobic) amino acids, the change is not likely to have a large functional effect. A second case involved substitution of aspertine-212 in *Hsx. carlsbadense.* Unlike leu-93, an important residue for isomerization, asp-212 is essential for proton release (Brown 2001). Changing from aspartate, a polar, acidic residue, to the non-polar valine may have a significant effect on bacteriorhodopsin function. Nevertheless, in both *Hsx. carlsbadense* and *Hrr. tebenquichense* sequences, only one site was not conserved in each sequence, though their function may be augmented. Therefore, it is highly likely that all generated sequences except the sequence from *Hrr. tibetense* are *bop* genes.

The BLAST results from *Hrr. tibetense* were inconclusive, and residue analysis was not definitive as most characteristic residues for SRII are also present in bacteriorhodopsin. This is most likely due to the homology of SRII with a bacteriorhodopsin-like proton pump (Royant, Nollert et al. 2001; Sudo, Furutani et al. 2006). Additionally, many essential residues for SRII function are outside the region amplified (Sudo, Furutani et al. 2006). However, when a number of sensory rhodopsin II genes (*sopII*) were taken from NCBI and aligned against the sequence, unique/characteristic residues (Royant, Nollert et al. 2001) in the *Hrr. tibetense* sequence were found to be conserved or semi-conserved. The conservation of these residues strongly suggests the gene is not a *bop* but a *sopII*.

4.3 Indel

Since the indel was found in both data mined and generated sequences, it is probably not an artifact. Because the indel is not part of a helix but rather an extra cellular inter-helical region, its effect is currently unknown but is unlikely to be responsible for a loss of function. However, the indel may cause modification of the rhodopsin's efficiency or its three-dimensional structure. Any functional differences caused by the indel have not been evaluated.

4.4 New bop in Hbl. gomorrense

The discovery of a *bop* gene in *Hbl. gomorrense* is surprising given that past phenotypic analysis of *Hbl. gomorrense* reported no bacteriorhodopsin despite growing the cells in an environment conducive to bacteriorhodopsin production (Oren, Gurevich et al. 1995). Even so, this finding by Oren *et al.* does not discount the existence of a *bop* since a protein's absence does not necessarily correspond to a missing gene. Based on the findings of this study a *bop* is present in the *Hbl. gomorrense* genome but not expressed. One explanation for this is that a mutation in a region not covered by this study is preventing the generation of a functional protein by either preventing transcription or proper folding. Another explanation for the lack of expression is that another gene necessary for bacteriorhodopsin production is not properly functioning. One gene that is necessary for *bop* transcription is the bacteriorhodopsin activator gene (*bat*) (Gropp and Betlach 1994). If *bat* was non-functional the organism would not produce bacteriorhodopsin despite the presence of *bop* and in favorable conditions (Leong, Boyer et al. 1988; Gropp and

Betlach 1994).

4.5 Phylogenetic Reconstruction

The major goal of this study was to analyze the evolutionary relationships between *bop* genes amplified from multiple strains across 15 genera. During phylogenetic reconstruction, decisions were made to determine the optimum tree to use for later comparisons to other trees. First, all three data sets were used to determine if there were any significant differences. Next, rooted and unrooted versions were generated to determine if root removal improved resolution.

4.5.1 Nucleotide vs. Protein Trees

Three rooted trees were constructed from the different data sets: nucleotide, protein and a hybrid, generated by overlaying nucleotides on protein alignment. The topology of these trees was similar with regard to the clusters being formed, though how clusters were related to one another was different. In the nucleotide tree, the *Haloarcula* and *Halorubrum* clusters were more closely related than to the mixed cluster. However, in the protein tree, the mixed cluster was more closely related to the *Haloarcula* cluster. Finally, the hybrid tree has a similar pattern to the nucleotide tree but is more asymmetric. Due to the remarkable similarities between the three and their strong bootstrap support, all are possible options for the gene's true evolutionary tree, though the nucleotide tree has notably lower bootstrap support. Based on these results, the nucleotide tree was not used for later comparisons. These trees are available in the appendix.

4.5.2 Rooted vs. Unrooted Trees

To determine if a root should be included in the final tree, rooted and unrooted versions of both the nucleotide and protein trees were generated and compared. A rooted tree includes an out-group, a small set of homologous but distinctly different sequences. For this study, sensory rhodopsin II genes served as the out-group. In both instances, unrooted and rooted trees were functionally identical and are thus equally likely to represent the true relatedness between *bop* genes. Removing the root helped clarify relatedness between strains and between clusters by reducing total difference between sequences. The unrooted versions of both data sets had higher bootstrap support than their rooted counterparts. Both unrooted trees had identical branching patterns, and so the unrooted tree was used to make comparisons to other trees. Comparison demonstrated that the trees are identical, though the hybrid phylogeny (Figure 14) has higher bootstrap support and so was used for all further comparisons. The unrooted hybrid phylogeny is shown here while others are available in the appendix.

4.6 Phylogenetic Comparisons

The final unrooted tree based on a hybrid alignment was used for comparisons to other studies. Comparisons were performed to understand how the evolutionary pattern of *bop* differed from MLSA and 16s rRNA gene based phylogenies. Additionally, the constructed final tree was compared to a previous *bop* study to determine if there was a need to reexamine either the sequences used or the tree as a whole.

The final unrooted phylogeny based on a hybrid alignment was compared to previously published results. To determine how the evolutionary history of the organisms and *bop* differed, the constructed phylogeny was compared to a MLSA phylogeny (Papke, White et al. 2011) described in section "*1.8 Multi-Locus Sequence Analysis".* This study's results were also compared to a 16s rRNA gene phylogeny (Boone, Castenholz et al. 2001). To ensure the validity of this study's predicted phylogeny, the resulting tree **was** compared to results from another bacteriorhodopsin study (Sharma, Walsh et al. 2007).

4.6.1 Comparison to Other bop Phylogeny

In order to validate the constructed *bop* phylogeny, a comparison was made to previous attempts by other researchers. When compared to a *bop* tree constructed by Sharma *et al.* (Sharma, Walsh et al. 2007) (Figure 16), the trees are similar. Three basic clusters, *Haloarcula, Halorubrum* and Mixed, as described in "*3.6 Phylogenetic Reconstruction of bop*", still appear in both trees, though some genera are missing from the analysis performed by Sharma *et al.*, and these clusters are related to each other in the same pattern. While bootstrap support and sequences included differ, the branching pattern in both trees is nearly

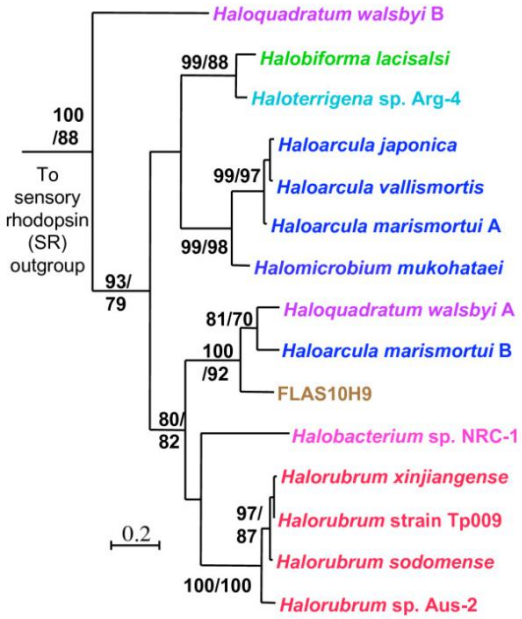

Figure 16 – Sharma *et al. bop* Tree

A *bop* phylogeny, rooted with sensory rhodopsins, constructed by Sharma *et al*, using primers highly similar to the ones used in this study. Bootstrap values over 70% are shown. Clades that represent a single genus are color coded. Image provided by Sharma *et al*. 2007 (Sharma, Walsh et al. 2007)

identical. A difference between the two trees lies in the respective positions of *Halorubrum* and Haloquadratum/*Haloarcula* second copy clusters, which are switched. However, as it is essentially a swap of two small clusters and only one unique sequence was used for *Halobacterium* between the trees, this is not incompatible with the data. This is reflected in the results from AU testing, where the tree was not rejected ($P > 0.005$). The similarities between these two *bop* trees support the validity of both trees as legitimate representatives of *bop* gene history. AU testing results are shown in Table 5 (Item Number 5).

4.6.2 Comparison Between bop and MLSA Phylogenies

The constructed *bop* tree was compared against the MLSA tree generated by Papke *et al.* (Papke, White et al. 2011) to compare *bop* evolutionary history to the evolutionary history given by MLSA, which reflects true evolution of the organisms. A number of similarities between the concatenated MLSA tree (Figure 17), and the unrooted hybrid *bop* tree (Figure 14) exist. In both phylogenies, equal clades of a similar pattern are formed consisting of similar genera and strains. Grouping was similar for the Mixed and *Haloarcula* clades, while differing in location for *Haloquadratum, Halobacterium* and *Haloferax*. In the *bop* phylogeny, *Haloquadratum* groups with the second copy *bop* genes from *Haloarcula,* while in MLSA it groups with *Haloferax*. *Haloferax* in the MLSA phylogeny groups with *Haloquadratum* and *Halorubrum* with significant support though in the *bop* phylogeny *Haloferax* is part of the Mixed clade, while *Halobacterium* separates from other genera. While the position of *Halobacterium* may seem different, this false impression is caused by *Haloquadratum* and the *bop2* group's position, because the genes do not appear in the MLSA phylogeny. The position of the genes is correct,

and is seen in other phylogenetic reconstructions of *bop* including Sharma *et al.* (Sharma, Walsh et al. 2007).

However, the different location of *Haloferax* has a number of possible explanations. Most obviously, this is an example of how *bop* evolutionary history is different from the true evolutionary path. A second possibility involves an HGT event from a member of the mixed clade, either to an ancestor *Haloferax* or to *Hfx. mucosum* in particular. However, only one representative of *Haloferax* is in the *bop* tree, so conclusions cannot be made. A final option is that *Haloferax mucosum* was not correctly classified based on 16s rRNA gene and phenotypic analysis (Allen, Goh et al. 2008), and a possible MLSA analysis of the strain may find that it was incorrectly characterized. The lack of this strain in the MLSA construction prevents conclusions.

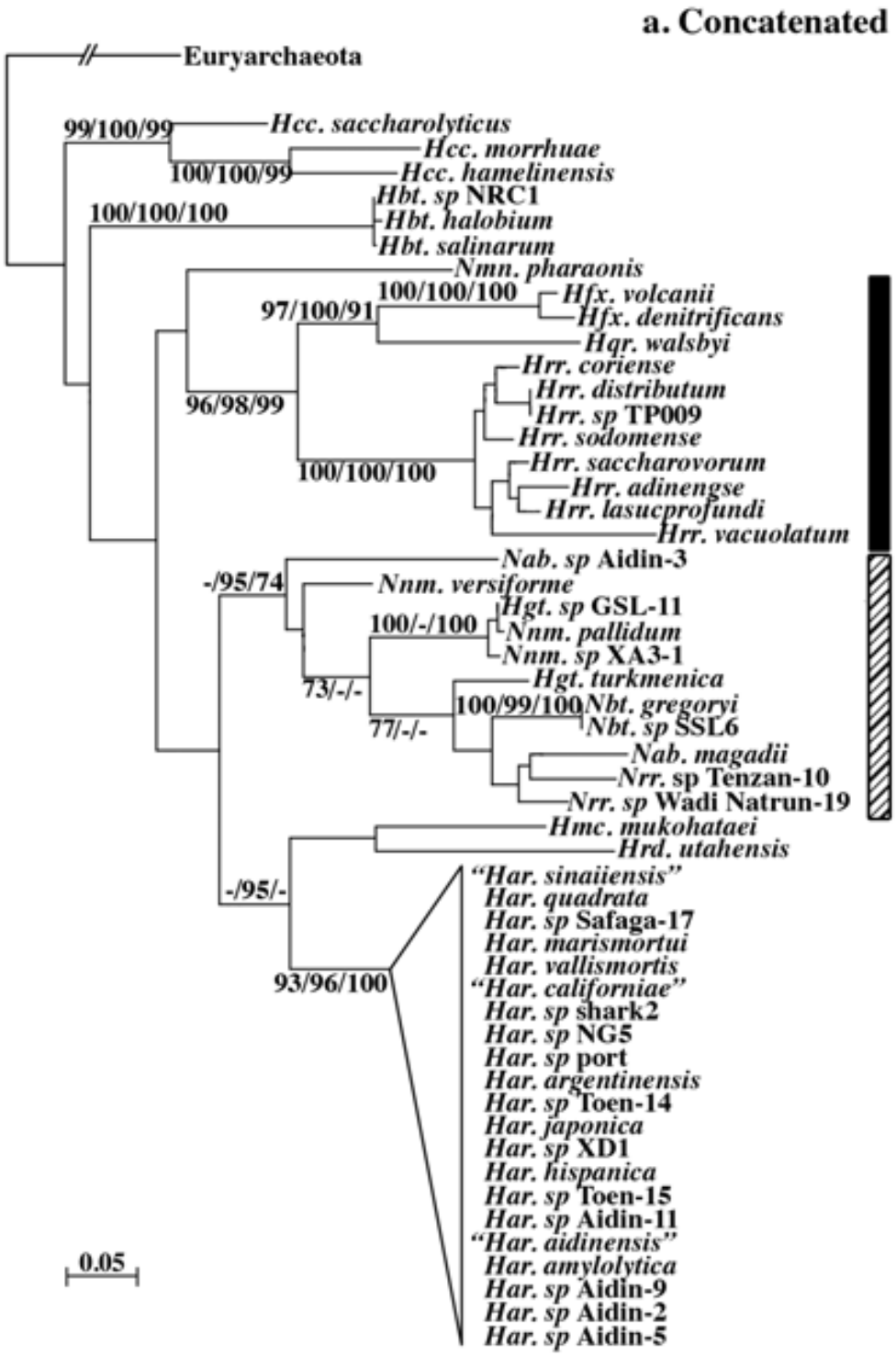

Figure 17 – Concatenated MLSA Tree

This tree was built using five housekeeping genes, specifically, *secY, EF-2, rpoB, radA* and *atpB*. Their respective sequences were concatenated to create a final tree. The out-group consisted of genes from other members of the Euryarchaeota, obtained through NCBI. The striped and solid bars represent clades 1 and 2 respectively, which correspond to Mixed and Halorubrum clusters of this study. Image provided by Papke *et al. (*Papke, White et al. 2011)

4.6.3 Comparison to 16s rRNA Gene Phylogeny

Relationships given by the *bop* tree (Figure 14) differ from those based on analysis of 16s rRNA genes. There are a number of disagreements between the tree published in Bergey's Manual (Figure 18) and constructed *bop* phylogeny. Again, relatedness within clades is high, and differences are in the relatedness between clades: as in the MLSA tree (Figure 17), the 16s rRNA gene phylogeny groups *Haloferax* closer to *Halorubrum* than to *Haloarcula* also disagrees with the *bop* and MLSA trees on the position of the mixed clade, which groups with *Halorubrum* rather than with *Haloarcula*.

Differences between 16s rRNA gene tree and the two trees place doubt on the strength of the 16s rRNA gene tree. While the trees agree on what strains group together into a genus, they disagree on how strains within genera are related. If this difference was only present in the *bop* tree, it could be due to differential gene evolution; however, since it was also present in MLSA tree, the differences are likely due to the evolutionary pattern of 16S rRNA gene failing to reflect true relatedness of the genera. One factor may be HGT, which MLSA takes into account, but 16s rRNA gene's analysis does not. While originally 16s rRNA genes were not considered to undergo HGT, recent studies show that HGT has played a significant role in the 16S rRNA gene phylogeny (Gogarten, Doolittle et al. 2002). Due to HGT, the 16s rRNA gene phylogeny should be reevaluated, particularly given the differences when compared to MLSA phylogeny.

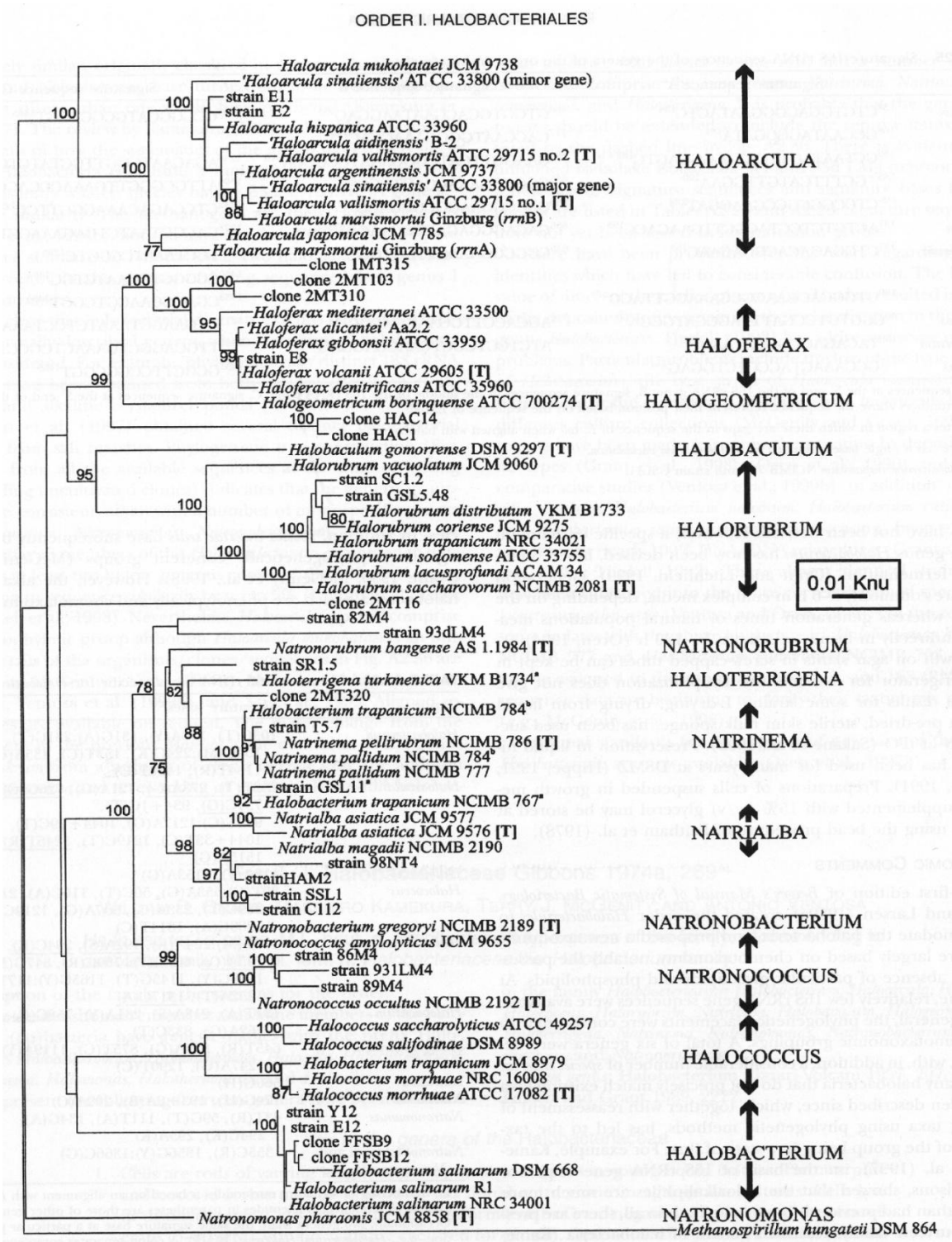

Figure 18 – Bergey's Manual *Halobacteriales* 16s rRNA Gene Tree

Tree constructed using a 16 s rRNA nucleotide alignment of a large portion of known Haloarchaeal species. Bootstrap values above 75% are shown at the nodes. The out-group consisted of 16srRNA gene from a methanogen. Knu, shortened from of Knuc is a measurement of the average extent of sequence change at any position in two homologous sequences (Hori and Osawa 1979) as cited in (Ventura and Zink 2003). Constructed in 1999. Tree obtained from Bergey's Manual (Boone, Castenholz et al. 2001).

4.7 Multiple bop Genes

As the presence of multiple copies of *bop* in a strain has been known for some time (Sharma, Walsh et al. 2007), this study incorporated three second copies of *bop* (*bop2)* from different members of *Haloarchaea* in addition to the primary copies of *bop* from these strains. In previous studies, these *bop2* genes grouped with the *bop* of *Hqm. walsbyi* (Sharma, Walsh et al. 2007). In this study, all three *bop2* genes grouped with *Hqm. walsbyi,* and all three were highly similar in terms of sequence (Figure 14). All three *bop2* genes grouping together suggest the event that created them occurred at some point before the differentiation of these three strains. This duplicated gene could then have mutated over time to their current state where they are similar to the *bop* of *Hqm. walsbyi,* making these copies orthologs since they have the same purpose as the original. Possibly, at some point before the speciation of the *Haloarcula* genus, there was a HGT event with the secondary copy originating from the ancestor of *Hqm. walsbyi.* If true, then the copies are not orthologs but rather xenologs due to HGT. It is unlikely that the second copies have originated due to multiple HGT events due to their tight grouping and general sequence similarities. Therefore, it is likely that second copy *bop* genes appeared in a *Haloarcula* ancestor and persisted in some strains.

4.8 Hrr. tibetense and HGT

A secondary goal for this study was to search for a HGT event. Originally, the gene amplified (assumed to be a *bop*) from *Hrr. tibetense* was observed in an odd phylogenetic location based on its separation from the other *Halorubrum*. Since this strain is not covered in

other *bop* phylogenetic studies, its location was not easily rectified, and it was concluded that *bop* of *Hrr. tibetense* was an example of HGT. This conclusion was supported by trees with a halorhodopsin gene out-group, where *Hrr. tibetense* grouped with the mixed cluster with significant bootstrap support, though also a long branch length, despite attempts to correct for it. Two other possibilities were also considered. Since *Hrr. tibetense* was discovered in Tibet (Boone, Castenholz et al. 2001), its *bop* may be better suited to these conditions, becoming a divergent *bop*. As suggested by the BLAST results, the gene may not be a *bop,* though its clustering with known *bop* genes makes this less likely. A second sequence was obtained from cells stored in glycerol at -80˚C which was identical to the first, excluding a possible technical error in PCR, sequencing or contamination of the initial cell sample.

However, when trees were constructed using sensory rhodopsins as an out-group, the position of *Hrr. tibetense* changed, grouping with sensory rhodopsins with significant bootstrap support. These results, when coupled with BLAST results and amino acid analysis, strongly suggest the gene amplified in this case was a sensory rhodopsin. These new findings revealed that the original conclusion was likely incorrect and the gene amplified from *Hrr. tibetense* was not a *bop* and was more likely a *sopII* gene. As a result, the sequence was not included in the final tree. Trees including *Hrr. tibetense* are included in the appendix.

4.9 Haloferax mucosum

A possible HGT event involves *Hfx. mucosum,* whose position in the *bop* phylogeny (Figure 14) is unusual based on the position of *Haloferax* in both the MLSA (Figure 17) and 16s rRNA gene trees (Figure 18). In the MLSA and 16s rRNA gene based phylogenies *Haloferax* is

closely related to *Halorubrum*, while in the *bop* tree it is more closely related to the Mixed clade. There are a number of possible explanations. One is that both trees are correct, and the evolutionary pattern of *Hfx. mucosum* is different for *bop* and 16s rRNA gene phylogenies, likely due to different selective pressures on *bop*. Another possibility is that the *bop* sequenced from *Hfx. mucosum* is a result of a HGT event from a member of the mixed clade. Additional AU testing was conducted to determine if the location was correct, which found the location of *Hfx. mucosum* in the *bop* tree was correct for the given data. However as the sequence used is currently unpublished it has not been used in other studies so the location cannot be rectified.

4.10 Conclusions

The creation of a *bop* tree has produced enlightening, though limited, findings. Interestingly, as shown by *Hrr. tibetense*, the primers used were found to be capable of amplifying genes other than *bop*. Most importantly, creating the tree provided a clear image of the relatedness not only between the different strains and genera of haloarchaea but of the evolutionary pattern of *bop*. While not covering all genera, a number of the well studied genera are covered. When compared to other trees, the differences found were primarily due to the gene's evolution, though in some cases differences were still possibly due to HGT. One such case is that of the *bop* amplified from *Haloferax*, although the *bop* position is correct as supported by AU testing since all cases of its movement result in rejection. It remains unclear if its location is representative of the genus as a whole or exclusively this particular *bop*. Though its placement is correct, without other members of the genus, the possibility of a HGT event cannot be determined.

Additionally, while MLSA and *bop* phylogenies agree with the location of the mixed clade, they both disagree with the 16s rRNA gene tree's placement. Based on this disagreement, both *bop* evolution and the history presented by MLSA analysis differs from the version presented by the 16s rRNA gene sequences. The amplification of a sensory rhodopsin, rather than a *bop*, from *Hrr. tibetense* was also interesting as it showed the primers used were able to amplify genes other than *bop* and care must be taken to ensure correct gene annotation. While no solid examples of HGT were found, this study does provide some insight into the relatedness of *bop* genes across a number of genera.

5. Appendix:

See Attached

6. References:

Allen, M. A., F. Goh, et al. (2008). "Haloferax elongans sp. nov. and Haloferax mucosum sp. nov., isolated from microbial mats from Hamelin Pool, Shark Bay, Australia." International journal of systematic and evolutionary microbiology **58**(Pt 4): 798-802.

Allers, T. and M. Mevarech (2005). "Archaeal genetics - the third way." Nat Rev Genet **6**(1): 58-73.

Altschul, S. F., T. L. Madden, et al. (1997). "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." Nucleic acids research **25**(17): 3389-3402.

Altschul, S. F., J. C. Wootton, et al. (2005). "Protein database searches using compositionally adjusted substitution matrices." FEBS J **272**(20): 5101-5109.

Andreeva, A., D. Howorth, et al. (2004). "SCOP database in 2004: refinements integrate structure and sequence family data." Nucleic acids research **32**(Database issue): D226-229.

Balashov, S. P. and J. K. Lanyi (2007). "Xanthorhodopsin: Proton pump with a carotenoid antenna." Cellular and molecular life sciences : CMLS **64**(18): 2323-2328.

Bamberg, E., J. Tittor, et al. (1993). "Light-driven proton or chloride pumping by halorhodopsin." Proceedings of the National Academy of Sciences of the United States of America **90**(2): 639-643.

Beja, O., L. Aravind, et al. (2000). "Bacterial rhodopsin: evidence for a new type of phototrophy in the sea." Science **289**(5486): 1902-1906.

Beja, O., E. N. Spudich, et al. (2001). "Proteorhodopsin phototrophy in the ocean." Nature **411**(6839): 786-789.

Biers, E. J., K. Wang, et al. (2008). "Occurrence and expression of gene transfer agent genes in marine bacterioplankton." Applied and environmental microbiology **74**(10): 2933-2939.

Bilofsky, H. S. and C. Burks (1988). "The GenBank genetic sequence data bank." Nucleic acids research **16**(5): 1861-1863.

Birge, R. R. (1990). "Photophysics and molecular electronic applications of the rhodopsins." Annual review of physical chemistry **41**: 683-733.

Bogomolni, R. A. and J. L. Spudich (1982). "Identification of a third rhodopsin-like pigment in phototactic Halobacterium halobium." Proceedings of the National Academy of Sciences of the United States of America **79**(20): 6250-6254.

Boone, D. R., R. W. Castenholz, et al. (2001). Bergey's manual of systematic bacteriology. New York, Springer.

Briggs, W. R. and J. L. Spudich (2005). Handbook of photosensory receptors. Weinheim, Wiley-VCH.

Brown, J. R. and W. F. Doolittle (1997). "Archaea and the prokaryote-to-eukaryote transition." Microbiology and molecular biology reviews : MMBR **61**(4): 456-502.

Brown, L. S. (2001). "Proton transport mechanism of bacteriorhodopsin as revealed by site-specific mutagenesis and protein sequence variability." Biochemistry (Mosc) **66**(11): 1249-1255.

Chaban, B., S. Y. Ng, et al. (2006). "Archaeal habitats--from the extreme to the ordinary." Canadian journal of microbiology **52**(2): 73-116.

Chen, Y., W. Zhang, et al. (2005). "Multi-virulence-locus sequence typing clarifies epidemiology of recent listeriosis outbreaks in the United States." Journal of clinical microbiology **43**(10): 5291-5294.

Darriba, D., G. L. Taboada, et al. (2011). "ProtTest 3: fast selection of best-fit models of protein evolution." Bioinformatics **27**(8): 1164-1165.

DeLong, E. F. and N. R. Pace (2001). "Environmental diversity of bacteria and archaea." Systematic biology **50**(4): 470-478.

Drummond AJ, K. M., Heled J, Moir R, Thierer T, Ashton B, Wilson A (2006). " Geneious v2.0.10."

Dyall-Smith, M. (2009). The Halohandbook: Protocols for Halobacterial Genetics. M. Dyall-Smith.

Ellen, A. F., B. Zolghadr, et al. (2010). "Shaping the archaeal cell envelope." Archaea **2010**: 608243.

Geospiza, I. "FinchTV 1.4.0."

Gilje, B., R. Heikkila, et al. (2008). "High-fidelity DNA polymerase enhances the sensitivity of a peptide nucleic acid clamp PCR assay for K-ras mutations." J Mol Diagn **10**(4): 325-331.

Giovannoni, S. J., L. Bibbs, et al. (2005). "Proteorhodopsin in the ubiquitous marine bacterium SAR11." Nature **438**(7064): 82-85.

Gogarten, J. P., W. F. Doolittle, et al. (2002). "Prokaryotic evolution in light of gene transfer." Molecular biology and evolution **19**(12): 2226-2238.

Gogarten, J. P., H. Kibak, et al. (1989). "Evolution of the vacuolar H+-ATPase: implications for the origin of eukaryotes." Proceedings of the National Academy of Sciences of the United States of America **86**(17): 6661-6665.

Gogarten, J. P. T., L. (1992). "Evolution of proton pumping ATPases: Rooting the tree of life." Photosynthesis research **33**: 137-146.

Gomez-Consarnau, L., J. M. Gonzalez, et al. (2007). "Light stimulates growth of proteorhodopsin-containing marine Flavobacteria." Nature **445**(7124): 210-213.

Gonzalez, C., C. Gutierrez, et al. (1978). "Halobacterium vallismortis sp. nov. An amylolytic and carbohydrate-metabolizing, extremely halophilic bacterium." Canadian journal of microbiology **24**(6): 710-715.

Gropp, F. and M. C. Betlach (1994). "The bat gene of Halobacterium halobium encodes a trans-acting oxygen inducibility factor." Proceedings of the National Academy of Sciences of the United States of America **91**(12): 5475-5479.

Guindon, S. and O. Gascuel (2003). "A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood." Systematic biology **52**(5): 696-704.

Hahn, J. H., P. (1986). "Traces of Archaebacteria in ancient sediments." Systematic and applied microbiology **7**: 178-183.

Haupts, U., J. Tittor, et al. (1997). "General concept for ion translocation by halobacterial retinal proteins: the isomerization/switch/transfer (IST) model." Biochemistry **36**(1): 2-7.

Haupts, U., J. Tittor, et al. (1999). "Closing in on bacteriorhodopsin: progress in understanding the molecule." Annual review of biophysics and biomolecular structure **28**: 367-399.

Hoff, W. D., K. H. Jung, et al. (1997). "Molecular mechanism of photosignaling by archaeal sensory rhodopsins." Annual review of biophysics and biomolecular structure **26**: 223-258.

Hori, H. and S. Osawa (1979). "Evolutionary change in 5S RNA secondary structure and a phylogenic tree of 54 5S RNA species." Proceedings of the National Academy of Sciences of the United States of America **76**(1): 381-385.

Jekely, G. (2009). "Evolution of phototaxis." Philosophical transactions of the Royal Society of London. Series B, Biological sciences **364**(1531): 2795-2808.

Kamekura, M. (1998). "Diversity of extremely halophilic bacteria." Extremophiles : life under extreme conditions **2**(3): 289-295.

Kirchmair, J., P. Markt, et al. (2008). "The Protein Data Bank (PDB), its related services and software tools as key components for in silico guided drug discovery." J Med Chem **51**(22): 7021-7040.

Lanyi, J. K. (2004). "Bacteriorhodopsin." Annual review of physiology **66**: 665-688.

Lanyi, J. K. and S. P. Balashov (2008). "Xanthorhodopsin: a bacteriorhodopsin-like proton pump with a carotenoid antenna." Biochimica et biophysica acta **1777**(7-8): 684-688.

Lawrence, J. G. and H. Ochman (1997). "Amelioration of bacterial genomes: rates of change and exchange." Journal of molecular evolution **44**(4): 383-397.

Leong, D., H. Boyer, et al. (1988). "Transcription of genes involved in bacterio-opsin gene expression in mutants of a halophilic archaebacterium." J Bacteriol **170**(10): 4910-4915.

Maddison, W. P. M., D.R. (2010). Mesquite: a modular system for evolutionary analysis. Sunderland, Massachusetts.

Man, D., W. Wang, et al. (2003). "Diversification and spectral tuning in marine proteorhodopsins." The EMBO journal **22**(8): 1725-1731.

Marti, T. (1998). "Refolding of bacteriorhodopsin from expressed polypeptide fragments." The Journal of biological chemistry **273**(15): 9312-9322.

Martinez, A., A. S. Bradley, et al. (2007). "Proteorhodopsin photosystem gene expression enables photophosphorylation in a heterologous host." Proceedings of the National Academy of Sciences of the United States of America **104**(13): 5590-5595.

Matsuno-Yagi, A. and Y. Mukohata (1977). "Two possible roles of bacteriorhodopsin; a comparative study of strains of Halobacterium halobium differing in pigmentation." Biochem Biophys Res Commun **78**(1): 237-243.

Michel, H. and D. Oesterhelt (1980). "Three-dimensional crystals of membrane proteins: bacteriorhodopsin." Proceedings of the National Academy of Sciences of the United States of America **77**(3): 1283-1285.

Mukohata, Y. (1994). "Comparative studies on ion pumps of the bacterial rhodopsin family." Biophysical chemistry **50**(1-2): 191-201.

Mullakhanbhai, M. F. and H. Larsen (1975). "Halobacterium volcanii spec. nov., a Dead Sea halobacterium with a moderate salt requirement." Archives of microbiology **104**(3): 207-214.

Murzin, A. G., S. E. Brenner, et al. (1995). "SCOP: a structural classification of proteins database for the investigation of sequences and structures." Journal of molecular biology **247**(4): 536-540.

Oesterhelt, D. and W. Stoeckenius (1971). "Rhodopsin-like protein from the purple membrane of Halobacterium halobium." Nature: New biology **233**(39): 149-152.

Oren, A., P. Gurevich, et al. (1995). "Halobaculum gomorrense gen. nov., sp. nov., a novel extremely halophilic archaeon from the Dead Sea." International journal of systematic bacteriology **45**(4): 747-754.

Page, R. D. M. (1996). "Tree View: An application to display phylogenetic trees on personal computers." Comput Appl Biosci **12**(4): 357-358.

Papke, R. T., C. J. Douady, et al. (2003). "Diversity of bacteriorhodopsins in different hypersaline waters from a single Spanish saltern." Environmental microbiology **5**(11): 1039-1045.

Papke, R. T., E. White, et al. (2011). "A Multilocus Sequence Analysis (MLSA) Approach to Halobacteriales Phylogeny and Taxonomy." International journal of systematic and evolutionary microbiology.

Papke, R. T., O. Zhaxybayeva, et al. (2007). "Searching for species in haloarchaea." Proceedings of the National Academy of Sciences of the United States of America **104**(35): 14092-14097.

Partha, R., R. Krebs, et al. (2005). "Weakened coupling of conserved arginine to the proteorhodopsin chromophore and its counterion implies structural differences from bacteriorhodopsin." Biochimica et biophysica acta **1708**(1): 6-12.

Pebay-Peyroula, E., A. Royant, et al. (2002). "Structural basis for sensory rhodopsin function." Biochimica et biophysica acta **1565**(2): 196-205.

Poptsova, M. S. and J. P. Gogarten (2007). "The power of phylogenetic approaches to detect horizontally transferred genes." BMC evolutionary biology **7**: 45.

Portillo, M. C. and J. M. Gonzalez (2009). "CRISPR elements in the Thermococcales: evidence for associated horizontal gene transfer in Pyrococcus furiosus." Journal of applied genetics **50**(4): 421-430.

Posada, D. (2008). "jModelTest: phylogenetic model averaging." Molecular biology and evolution **25**(7): 1253-1256.

Royant, A., P. Nollert, et al. (2001). "X-ray structure of sensory rhodopsin II at 2.1-A resolution." Proceedings of the National Academy of Sciences of the United States of America **98**(18): 10131-10136.

Ruiz-Gonzalez, M. X. and I. Marin (2004). "New insights into the evolutionary history of type 1 rhodopsins." Journal of molecular evolution **58**(3): 348-358.

Salemi, M., A.-M. Vandamme, et al. (2009). The phylogenetic handbook : a practical approach to phylogenetic analysis and hypothesis testing. Cambridge, UK ; New York, Cambridge University Press.

Saranak, J. and K. W. Foster (2005). "Photoreceptor for curling behavior in Peranema trichophorum and evolution of eukaryotic rhodopsins." Eukaryotic cell **4**(10): 1605-1612.

Schlesner, M., A. Miller, et al. (2009). "Identification of Archaea-specific chemotaxis proteins which interact with the flagellar apparatus." BMC microbiology **9**: 56.

Schmidt, H. A., K. Strimmer, et al. (2002). "TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing." Bioinformatics **18**(3): 502-504.

Sharma, A. K., K. Sommerfeld, et al. (2009). "Actinorhodopsin genes discovered in diverse freshwater habitats and among cultivated freshwater Actinobacteria." The ISME journal **3**(6): 726-737.

Sharma, A. K., J. L. Spudich, et al. (2006). "Microbial rhodopsins: functional versatility and genetic mobility." Trends in microbiology **14**(11): 463-469.

Sharma, A. K., D. A. Walsh, et al. (2007). "Evolution of rhodopsin ion pumps in haloarchaea." BMC evolutionary biology **7**: 79.

Sharma, A. K., O. Zhaxybayeva, et al. (2008). "Actinorhodopsins: proteorhodopsin-like gene sequences found predominantly in non-marine environments." Environmental microbiology **10**(4): 1039-1056.

Shimodaira, H. (2002). "An approximately unbiased test of phylogenetic tree selection." Systematic biology **51**(3): 492-508.

Siegel, G. J. (1999). Basic neurochemistry : molecular, cellular, and medical aspects. Philadelphia, Lippincott Williams & Wilkins.

Slamovits, C. H., N. Okamoto, et al. (2011). "A bacterial proteorhodopsin proton pump in marine eukaryotes." Nature communications **2**: 183.

Slonczewski, J. and J. W. Foster (2011). Microbiology : an evolving science. New York, W.W. Norton.

Soppa, J., A. Baumann, et al. (2008). "Genomics and functional genomics with haloarchaea." Archives of microbiology **190**(3): 197-215.

Spudich, J. L. (1998). "Variations on a molecular switch: transport and sensory signalling by archaeal rhodopsins." Molecular microbiology **28**(6): 1051-1058.

Spudich, J. L., C. S. Yang, et al. (2000). "Retinylidene proteins: structures and functions from archaea to humans." Annual review of cell and developmental biology **16**: 365-392.

Stamatakis, A. (2006). "RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models." Bioinformatics **22**(21): 2688-2690.

Stuart, J. A., D. L. Marcy, et al. (2002). "Volumetric optical memory based on bacteriorhodopsin." Synthetic Metals **127**(1-3): 3-15.

Subramaniam, S., M. Gerstein, et al. (1993). "Electron diffraction analysis of structural changes in the photocycle of bacteriorhodopsin." The EMBO journal **12**(1): 1-8.

Sudo, Y., Y. Furutani, et al. (2006). "Functional importance of the interhelical hydrogen bond between Thr204 and Tyr174 of sensory rhodopsin II and its alteration during the signaling process." The Journal of biological chemistry **281**(45): 34239-34245.

Sung, C. H. and J. Z. Chuang (2010). "The cell biology of vision." The Journal of cell biology **190**(6): 953-963.

Takai, K., K. Nakamura, et al. (2008). "Cell proliferation at 122 degrees C and isotopically heavy CH4 production by a hyperthermophilic methanogen under high-pressure cultivation." Proceedings of the National Academy of Sciences of the United States of America **105**(31): 10949-10954.

Thompson, J. D., D. G. Higgins, et al. (1994). "CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice." Nucleic acids research **22**(22): 4673-4680.

Váró, G. (2000). "Analogies between halorhodopsin and bacteriorhodopsin." Biochimica et Biophysica Acta (BBA) - Bioenergetics **1460**(1): 220-229.

Venter, J. C., K. Remington, et al. (2004). "Environmental genome shotgun sequencing of the Sargasso Sea." Science **304**(5667): 66-74.

Ventura, M. and R. Zink (2003). "Comparative sequence analysis of the tuf and recA genes and restriction fragment length polymorphism of the internal transcribed spacer region sequences supply additional tools for discriminating Bifidobacterium lactis from Bifidobacterium animalis." Applied and environmental microbiology **69**(12): 7517-7522.

Walsh, D. A., E. Bapteste, et al. (2004). "Evolution of the RNA polymerase B' subunit gene (rpoB') in Halobacteriales: a complementary molecular marker to the SSU rRNA gene." Molecular biology and evolution **21**(12): 2340-2351.

Woese, C. R. and G. E. Fox (1977). "Phylogenetic structure of the prokaryotic domain: the primary kingdoms." Proceedings of the National Academy of Sciences of the United States of America **74**(11): 5088-5090.

Woese, C. R., O. Kandler, et al. (1990). "Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya." Proceedings of the National Academy of Sciences of the United States of America **87**(12): 4576-4579.

Woese, C. R., L. J. Magrum, et al. (1978). "Archaebacteria." Journal of molecular evolution **11**(3): 245-251.

Xia, X. and Z. Xie (2001). "DAMBE: software package for data analysis in molecular biology and evolution." J Hered **92**(4): 371-373.

Xu, X. H. W., M. (2004). Partial sequence of a novel bR protein. Hangzhou, Zhenjiang, Zhejiang University.

Zhao, Y., K. Wang, et al. (2009). "Gene transfer agent (GTA) genes reveal diverse and dynamic Roseobacter and Rhodobacter populations in the Chesapeake Bay." The ISME journal **3**(3): 364-373.

5. Appendix:

5.1 All Sequences Used in Analysis

Species	Strain Identifier	Accession	Gene	Sequence
Haloarcula amylolytica	JCM 13557	Self Generated	*bop*	TTACTGCTGTATGACCTCGGACTGCTTGCAGGGGCAG ACCGCAACACCATCGCCTCGCTCGTCAGCCTCGACGT GCTGATGATCGGTACCGGTGTGGTCGCGACGCTAAGC GCAGGGAGCGGCGTGCTGTCGGCCGGCGCGGAACGG CTGGTCTGGTGGGGTATCAGCACCGCCTTCCTGCTGG TCCTGCTGTACTTCCTGTTCGCCTCGCTGTCCGGCCGG GTCGCAGACCTGCCCAGTGACACGCGGAGCACGTTC AAGACGCTGCGCAACCTCGTGACGGTCGTGTGGCTGG TCTACCCGGTCTGGTGGCTCGTCGGCAGTGAAGGACT CGGCCTCGTCGGTATTGGGATCGAGACGGCCGGCTTC ATGGTCATCGACCTCGTC
Haloarcula argentinensis	JCM 9739	Self Generated	*bop*	TTACTGTTGTACGACCTCGGGTTGCTTGCCGGGGCAG ATCGTAACACTATCACCTCGCTCGTCAGCCTCGACGT GCTGATGATCGGGACCGGGCTGGTCGCGACGCTGAG CGCAGGAAGCGGCGTGCTGTCGGCCGGTGCAGAACG GCTGGTCTGGTGGGGCATCAGTACCGCCTTCCTGCTG GTCCTGCTGTACTTCCTGTTCAGCTCGCTGTCCGGTCG AGTCGCAGACTTGCCCAGTGACACGCGTAGCACCTTC AAGACGCTACGAAACCTCGTGACTGTCGTGTGGTTGG TGTACCCGGTGTGGTGGCTCATCGGGACTGAGGGGAT CGGCCTCGTCGGGATCGGTATCGAGACGGCCGGCTTC ATGGTCATCGACCTGACC
Haloarcula californiae	JCM 8912	Self Generated	*bop*	TTGCTGCTGTATGACCTCGGACTGCTTGCAGGTGCAG ACCGCAACACCATTTACTCGCTCGTCAGCCTCGACGT GCTGATGATCGGTACTGGTGTGGTCGCAACGCTGAGT GCAGGCAGTGGCGTGCTGTCTGCCGGTGCAGAACGG CTGGTCTGGTGGGGTATCAGCACTGCCTTCCTGCTGG TCCTGCTGTACTTCCTGTTCAGTTCGCTGTCCGGCCGG GTTGCAAACCTGCCCAGTGACACGCGAAGCACGTTCA AGACGTTGCGCAACCTCGTCACTGTCGTCTGGCTGGT CTACCCAGTCTGGTGGCTCGTCGGCAGTGAGGGGCTT GGCCTCGTCGGTATCGGAATCGAGACGGCCGGCTTCA TGGTCATCGACCTCGTC
Haloarcula hispanica	ATCC 33960	Self Generated	*bop*	TTACTGCTGTATGACCTCGGACTGCTTGCAGGGGCAG ACCGCAACACCATCGCCTCGCTCGTCAGCCTTGACGT GCTGATGATCGGTACCGGTGTGGTCGCGACGCTAAGC GCAGGGAGCGGCGTGCTGTCGGCCGGCGCGGAACGG CTGGTCTGGTGGGGTATCAGCACCGCCTTCCTGCTGG TCCTGCTGTACTTCCTGTTCGCCTCGCTGTCCGGTCGG GTCGCAGACCTGCCCAGCGACACGCGGAGCACGTTC AAGACGCTGCGAAACCTCGTGACGGTCGTGTGGCTG GTCTACCCGGGCTGC

Haloarcula japonica	JCM 7785	Self Generated	*bop*	TTGCTGTTGTACGACCTCGGGTTGCTTGCGGGGGCAG ACCGCAACACCATCGCCTCGCTCGTCAGCCTCGACGT GCTGATGATCGGGACCGGACTGGTCGCGACGCTGAG CGCAGGCAGTGGTGTGCTGTCGGCCGGCGCGGAACG GCTGGTCTGGTGGGGTATCAGTACCGCCTTCCTGCTG GTCCTGCTGTACTTCCTGTTCAGTTCGCTGTCCGGTCG GGTCGCAGACCTGCCCAGTGACACGCGTAGCACCTTC AAGACGCTACGGAACCTCGTAACCGTCGTGTGGTTGG TGTACCCGGTGTGGTGGCTCATCGGGACTGAGGGTCT CGGCCTCGTCGGTATCGGGATCGAGACGGCCGGCTTC ATGGTCATCGACCTGACC
Haloarcula marismortui	ATCC 43049	Self Generated	*bop*	TTGCTGCTGTATGACCTCGGACTGCTTGCAGGTGCAG ACCGCAACACCATTTACTCGCTCGTCAGCCTCGACGT GCTGATGATCGGTACTGGTGTGGTCGCAACGCTGAGT GCAGGCAGTGGCGTGCTGTCTGCCGGTGCAGAACGG CTGGTCTGGTGGGGTATCAGCACTGCCTTCCTGCTGG TCCTGCTGTACTTCCTGTTCAGTTCGCTGTCTGGCCGG GTTGCAAACCTGCCCAGTGACACGCGAAGCACGTTCA AGACGTTGCGCAACCTCGTCACTGTCGTCTGGCTGGT CTACCCAGTCTGGTGGCTCGTCGGCAGTGAGGGGCTT GGCCTCGTCGGTATCGGAATCGAGACG
Haloarcula vallismortis	ATCC 29715	Self Generated	*bop*	TTGCTGTTGTATGACCTCGGACTGCTTGCCGGGGCAG ACCGTAACACCATTAGCTCGCTCGTCAGCCTCGACGT GCTGATGATCGGGACCGGTCTGGTCGCGACGCTGAGT GCGGGAAGTGGCGTGCTGTCGGCCGGCGCGGAACGG CTGGTCTGGTGGGGCATCAGTACCGCCTTCCTGCTGG TCCTGCTGTACTTCCTGTTCAGTTCGCTATCCGGCCGG GTCGCGGACCTGCCAAGCGATACGCGTAGCACGTTCA AGACGTTGCGCAACCTCGTCACTGTCGTCTGGCTGGT CTACCCAGTCTGGTGGCTCGTCGGCACTGAGGGCATC GGCCTCGTCGGTATCGGCATCGAGACGGCCGGTTTCA TGGTCATCGACCTCGTT
Halobacterium salinarum	NCIMB 2288	Self Generated	*bop*	CTGTTGTTGTTAGACCTCGCGTTGCTCGTTGACGCGG ATCAGGGAACGATCCTTGCGCTCGTCGGTGCCGACGG CATCATGATCGGGACCGGCCTGGTCGGCGCACTGACG AAGGTCTACTCGTACCGCTTCGTGTGGTGGGCGATCA GCACCGCAGCGATGCTGTACATCCTGTACGTGCTGTT CTTCGGGTTCACCTCGAAGGCCGAAAGCATGCGCCCC GAGGTCGCATCCACGTTCAAAGTACTGCGTAACGTTA CCGTTGTGTTGTGGTCCGCGTATCCCGTCGTGTGGCT GATCGGCAGCGAAGGTGCGGGAATCGTGCCGCTGAA CATCGAGACGCTGCTGTTCATGGTGCTTGACGTGAGC

Halobacterium sp.	NRC-1	Self Generated	*bop*	CTGTTGTTGTTAGACCTCGCGTTGCTCGTTGACGCGG ATCAGGGAACGATCCTTGCGCTCGTCGGTGCCGACGG CATCATGATCGGGACCGGCCTGGTCGGCGCACTGACG AAGGTCTACTCGTACCGCTTCGTGTGGTGGGCGATCA GCACCGCAGCGATGCTGTACATCCTGTACGTGCTGTT CTTCGGGTTCACCTCGAAGGCCGAAAGCATGCGCCCC GAGGTCGCATCCACGTTCAAAGTACTGCGTAACGTTA CCGTTGTGTTGTGGTCCGCGTATCCCGTCGTGTGGCT GATCGGCAGCGAAGGTGCGGGAATCGTGCCGCTGAA CATCGAGACGCTGCTGTTCATGGTGCTTGACGTGAGC
Halobaculum gomorrense	JCM 9908	Self Generated	*bop*	CTGTTGCTGCTCGATCTCGCTCTCCTTGCGGGGGCGA ACCGCAACGAGATCGCGTCGCTGGTCGGCCTCGACAT GCTGATGATCGGCACGGGCGCCGTCGCGACGCTGAG CACGGGCGCGGGGGGTGCTCTCGGTCGGCGCTCGACG CCTCGTCTGGTGGGGCGTCTCGACGGGCTTCCTGCTC GTCCTCCTGTACATGCTGTACGGCTCGCTCGCCGAGA AAGCCAGCAAGCTCTCGGGCGACGCCGCCTCGACGTT CAGCACGCTGCGGAACCTGATCGTGTTCATCTGGCTG GTGTACCCGGCCTGGTGGCTCGTCGGCACCGAAGGGC TCGGCGTCGTCTCGCTGAACATCGAGACGGCCGGCTT CATGGTGCTGGACCTGGTC
Halomicrobium mukohatei	DSM 12286	Self Generated	*bop*	TTGCTGCTGTACGACCTCGCCCTGCTGGCGGGGGCCG ACCGGAACACGATCTACTCGCTGGTCGGTCTGGACGT GCTGATGATCGGTACCGGCGCGCTCGCGACGCTGTCG GCGGGTTCGGGCGTCCTCCCGGCGGGTGCCGAACGCC TCGTCTGGTGGGGTATCAGCACCGGCTTCCTGCTGGT CCTGCTGTACTTCCTGTTCAGCAACCTCACCGACCGA GCGAGCGAGCTCTCGGGTGACTTGCAGTCGAAGTTCT CGACGCTGCGCAACCTGGTGCTGGTCCTGTGGCTCGT GTACCCCGTCCTGTGGTTGGTCGGGACCGAGGGCCTC GGACTCGTCGGCCTCCCGATCGAGACGGCCGCGTTCA TGGTGCTCGACCTGACC
Halorubrum arcis	JCM 13916	Self Generated	*bop*	CTGCTGCTTGTCGAACTCCCGCTGCTGGCGAAAGTCG AACGCGTGAACATTCGGGASGCTCGTGGCGTCGACGC GCTGATGATCGTCACCGGCCTCATCGGCGCGCTCTCG CACACGCCGCTCGCGCGGTACTCCTGGTGGCTGTTCA GCACGAACTGCATGATCGTCGTGCTGTACTTCCTCGC CACGAGCCTGCGCGCCGCGGCGAAAGAGCGCGGCCC CGAAGTCGCGAGCACGTTCAACACGCTGACCGCGTTG GTGCTCGTCCTCTGGACCGCGTACCCGATCCTCTGGA ACATCGGTACTGAGGGAGCCGGCGTCGTCGGCCTCG GCAACGAAACCCTCCTGTTCATGGTTCTCGACGTGAC C

Halorubrum distributum	JCM 9100	Self Generated	*bop*	CTGCTGCTGCTCGACCTCGCGCTGCTGGCGAAGGTCG ACCGCGTGAGCATCGGGACGCTCGTCGGCGTCGACG CGCTGATGATCGTCACCGGCCTCATCGGCGCGCTCTC GCACACGCCGCTCGCGCGGTACTCCTGGTGGCTGTTC AGCACGATCTGCATGATCGTCGTGCTGTACTTCCTCG CCACGAGCCTGCGCGCCGCGGCGAAGGAGCGCGGCC CCGAAGTCGCGAGCACGTTCAACACGCTGACCGCGTT GGTGCTCGTCCTCTGGACCGCGTACCCGATCCTCTGG ATCATCGGTACTGAGGGAGCCGGCGTCGTCGGCCTCG GCATCGAGACCCTCCTGTTCATGGTTCTCGACGTGAC C
Halorubrum ejinorense	JCM 14265	Self Generated	*bop*	CTGCTGCTGCTCGACCTCGCGCTCCTCGCGAAGGTCG ACCGCGTGACCATCGGCACGCTCATCGGCGTCGACGC GCTGATGATCGTCACCGGTCTCATCGGTGCACTGTCG AAGACGCCGCTCGCGCGGTACACCTGGTGGCTGTTCA GCACGATCGCGTTCATCTTCGTGCTCTACTACCTCCTC ACGAGCCTCCGGAGCGCGGCTTCGGAGCTCTCTGAGG ACGTACAGACCACCTTCAACACGCTGACCGCACTGGT CGCCGTCCTCTGGACGGCGTACCCGATCCTGTGGATC GTCGGGACCGAGGGAGCCGGCGTCGTCGGCCTCGGC GTCGAGACGCTCCTGTTCATGGTGCTCGACGTGACG
Halorubrum litoreum	JCM 13561	Self Generated	*bop*	CTGCTGCTGCTCGACCTCGCGCTGCTGGCGAAGGTCG ACCGCGTGAGCATCGGGACGCTCGTCGGCGTCGACG CGCTGATGATCGTCACCGGCCTCATCGGCGCGCTCTC GCACACGCCGCTCGCGCGGTACTCCTGGTGGCTGTTC AGCACGATCTGCATGATCGTCGTGCTGTACTTCCTCG CCACGAGCCTGCGCGCCGCGGCGAAGGAGCGCGGCC CCGAAGTCGCGAGCACGTTCAACACGCTGACCGCGTT GGTGCTCGTCCTCTGGACCGCGTACCCGATCCTCTGG ATCATCGGTACTGAGGGAGCCGGCGTCGTCGGCCTCG GCATCGAGACCCTCCTGTTCATGGTTCTCGACGTGAC C
Halorubrum sodomense	JCM 8880	Self Generated	*bop*	CTGCTGCTGCTCGACCTCGCGCTGCTCGCAAAGGTCG ACCGCGTCACCATCGGGACGCTCGTCGGCGTCGACGC GCTGATGATCGTCACCGGCCTCATCGGCGCGCTCTCG CACACGGCGATCGCGCGGTACTCCTGGTGGCTGTTCA GCACGATTTGCATGATCGTCGTGCTGTACTTCCTCGC CACGAGCCTCCGGAGCGCGGCGAAGGAGCGCGGACC TGAAGTCGCGAGCACCTTCAACACGTTGACCGCGCTG GTCCTGGTGCTCTGGACGGCCTACCCGATCCTGTGGA TCATCGGAACCGAGGGCGCCGGCGTCGTCGGCCTCG GCATCGAGACCCTCCTGTTCATGGTTCTCGACGTGAC G

Halorubrum tebenquichense	JCM 12290	Self Generated	*bop*	CTGGTGCTGCTTGACCCTCGCGCTGCTCGCGAAACGC GGAACCGCACCACCATCGGGACGCTCATCGGCGTCG ACGCGCTGATGATCGTCACCGGTCTCATCGGTGCGCT CTCGCACACGCCGCTCGCGCGGTACACCTGGTGGCTG TTCAGCACGATCGCGTTCCTGTTCGTGCTGTACTACCT CCTGACGGTGCTCCGCAGCGCGGCCGCGGAGCTCTCG GAGGACGTGCAGACCACCTTCAACACGCTAACGGCG CTCGTCGCCGTCCTCTGGACGGCGTACCCGATCCTGT GGATCATCGGGACCGAAGGCGCCGGCGTCGTCGGCC TGGTCGTCGAGACGCTCGCGTTCATGGTGCTCGACGT GACG
Halosimplex carlsbadense	JCM 11222	Self Generated	*bop*	CTGTTGCTGATCGACCTCGGACTGCTGGCGCGGGCCA ACCGCAACCAGCTGACGACGCTGGTCAGCCTCGACG CGCTGATGATCGGCACCGGCGCGATCGCGACGCTGG CCGGCGGGAACGTCGCCGGCCTCGACGACGGCGCCC GCCGGCTGATCTGGTGGGGAATCAGCACCGGCTTCCT GCTCGCGTTGCTGTACTTCCTGTTCGGGACGCTCACC GAGCAGGCACAGGAGCTAGGTAGTGACGTGGGCGCG AAGTTCGCGCAGTTGCGCAACCTGATCGTCGCGGTCT GGCTGGTCTACCCCGTATGGTGGCTGGCCGGTACCGA GGGACTTGCGCTGCTCCCCGAGGTCAGCGGGAGCGTC CTGTTCGTCGAGACG
Haloterrigena sp.	GSL-11	Self Generated	*bop*	CTGCTGTTGCTCGACCTCTCGCTGCTAGCCGGGGCGA ACCGAAACACGATCGCGACGCTGATCGGCCTCGACG TCTTCATGATCGGGACCGGCGCGATCGCAGCGCTCTC GTCCACCCCGGGTACCCGGATCGCCTGGTGGGCGATC AGCACCGGTGCTCTGCTCGCCTTGCTGTACGTCCTCG TCGGGACGCTCTCCGAGAACGCTCGCGACCGGGCCCC CGAGGTCGCATCGCTGTTCGGGAGACTCCGCAACCTG GTGATCGCGCTGTGGCTCCTCTATCCGGTGGTCTGGA TCCTCGGCACGGAAGGGACGTTCGGCATCCTCCCGCT GTACTGGGAGACCGCGGCGTTCATGGTGCTCGACCTC TCG
Natrinema altunense	JCM 12890	Self Generated	*bop*	CTGCTGTTGCTCGACCTCTCGCTGTTAGCCGGGGCGA ACCGAAACACGATCGCGACGCTGATCGGCCTCGATGT CTTCATGATCGGTACCGGCGCGATCGCAGCGCTCTCG TCCACCCCGGGTGCTCGGATCGCCTGGTGGGCGATCA GCACCGGTGCCCTGCTCACCCTACTGTACGTCCTCGT CGGGACGCTCTCCGAGAACGCGCGCAGTCGGGCCCC CGAGGTCGCGTCGCTGTTCGGGAGACTCCGCAATCTG GTTATCGCGCTGTGGTTCCTCTACCCGGTGGTCTGGA TCCTCGGCACGGAAGGGACGTTCGGCATCCTGCCGCT GTACTGGGAAACCGCGGCGTTCATGGTGCTCGACCTC TCG

Natrinema pallidum	NCIMB 784	Self Generated	*bop*	CTGCTGTTGCTCGACCTCTCGCTGCTCGCCGGGGCGA ACCGAAACACGATCGCGACGCTGATCGGCCTCGACG TTTTCATGATCGGGACCGGCGCGATCGCAGCGCTCTC GTCCACCCCGGGTACCCGGATCGCCTGGTGGGCGATC AGCACCGGTGCTCTGCTCGCCCTGCTGTACGTCCTCG TCGGGACGCTCTCCAAGAACGCGCGCAACCGGGCCC CCGAGGTCGCATCGCTGTTCGGGAGACTCCGCAACCT GGTTATCGCGCTGTGGTTCCTCTACCCGGTGGTCTGG ATCCTCGGCACGGAAGGGACGTTCGGCATCCTTCCGC TGTACTGGGAAACCGCGGCGTTCATGGTGCTCGACCT CTCG
Unnamed	GA27	Self Generated	*bop*	CTGCTGCTGCTCGACCTCGCGCTGCTGGCGAAGGTCG ATCGCGTGAGCATCGGCACGCTCGTCGGCGTCGACGC GCTGATGATCGTCACCGGACTCGTCGGCGCGCTCTCG CACACGCCGCTCGCGCGGTACACGTGGTGGCTGTTCA GCACGATCTGCATGATCGTCGTGCTGTACTTCCTCGC CACGAGCCTGCGCGCCGCTGCGAAGGAGCGCGGTCC CGAGGTCGCGAGCACGTTCAACACCCTTACGGCGTTG GTGTTGGTGCTCTGGACGGCCTACCCGATCCTGTGGA TCATCGGTACCGAGGGTGCCGGCGTCGTCGGCCTCGG CATC
Unnamed	FB21	Self Generated	*bop*	CTGCTGCTGCTCGACCTCGCGCTGCTGGCGAAGGTCG ACCGCGTGAGCATCGGCACGCTCGTCGGCGTCGACGC GCTGATGATCGTCACCGGACTCGTCGGCGCGCTCTCG CACACGCCGCTCGCGCGGTACACGTGGTGGCTGTTCA GCACGATCTGCATGATCGTCGTGCTGTACTTCCTCGC CACGAGCCTGCGCGCCGCTGCGAAGGAGCGCGGTCC CGAGGTCGCGAGCACGTTCAACACCCTTACGGCGTTG GTGTTGGTGCTCTGGACGGCCTACCCGATCCTGTGGA TCATCGGTACCGAGGGTGCCGGCGTCGTCGGCCTCGG CATCGAGACCCTCCTGTTCATGGTGCTCGACGTGACC
Unnamed	Ha9	Self Generated	*bop*	TTGCTGCTGTATGACCTCGGACTGCTCGCAGGTGCAG ACCGCAATACCATTTACTCGCTCGTCAGCCTCGACGT GCTGATGATCGGTACCGGTGTGGTCGCGACGCTGAGC GCAGGCAGCGGCGTGCTCTCGGCCGGTGCGGAACGG CTGGTCTGGTGGGGTATCAGTACCGCCTTCCTGCTGG TCCTGCTGTACTTCCTGTTCGCATCGCTGTCCGGTCGG GTCGCAGACCTGCCCAGTGACACGCGGAGCACGTTC AAGACGTTGCGCAACCTCGTGACGGTCGTGTGGCTGG TCTACCCGGTCTGGTGGCTCGTCGGCAGTGAAGGACT CGGCCTCGTCGGTATCGGGATCGAGACGGCCGGCTTC ATGGTCATCGACCTGACC

Unnamed	EC15	Self Generated	*bop*	CTGCTGCTGCTCGACCTCGCGCTGCTGGCGAAGGTCG ACCGCGTGAGCATCGGCACGCTCGTCGGCGTCGACGC GCTGATGATCGTCACCGGACTCGTCGGCGCGCTCTCG CACACGCCGCTCGCGCGGTACACGTGGTGGCTGTTCA GCACGATCTGCATGATCGTCGTGCTGTACTTCCTCGC CACGAGCCTGCGCGCCGCTGCGAAGGAGCGCGGTCC CGAGGTCGCGAGCACGTTCAACACCCTTACGGCGTTG GTGTTGGTGCTCTGGACGGCCTACCCGATCCTGTGGA TCATCGGTACCGAGGGTGCCGGCGTCGTCGGCCTCGG CATCGAGACCCTCCTGTTCATGGTGCTCGACGTGACC
Unnamed	Fb19	Self Generated	*bop*	TTGCTGCTGTATGACCTCGGACTGCTCGCAGGTGCAG ACCGCAATACCATTTACTCGCTCGTCAGCCTCGACGT GCTGATGATCGGTACCGGTGTGGTCGCGACGCTGAGC GCAGGCAGCGGCGTGCTCTCGGCCGGTGCGGAACGG CTGGTCTGGTGGGGTATCAGTACCGCCTTCCTGCTGG TCCTGCTGTACTTCCTGTTCGCATCGCTGTCCGGTCGG GTCGCAGACCTGCCCAGTGACACGCGGAGCACGTTC AAGACGTTGCGCAACCTCGTGACGGTCGTGTGGCTGG TCTACCCGGTCTGGTGGCTCGTCGGCAGTGAAGGACT CGGCCTCGTCGGTATCGGGATCGAGACGGCCGGCTTC ATGGTCATCGACCTGACC
Unnamed	Eb13	Self Generated	*bop*	CTGCTGTTGCTCGACCTCGCGCTCCTCGCGAAGGTCG ACCGCGTGACGATCGGGACGCTCATCGGCGTCGACG CGCTGATGATCGTCACCGGCCTCGTCGGTGCTCTCTC ACAGACGATGCTTGCGCGGTACTCGTTCTGGCTCATC AGCACGATCGCGATGATCTTCGTGCTGTACTACCTCG CCACGAGCCTCCGCAGCGCGGCGAAGGCGCGCTCGG CTGAGGTGGAGAGCACGTTCAACACGTTGACCGTGTT GACGTTGGTCCTTTGGACCGCGTACCCGATCCTGTGG CTCGTCGGAACCGAAGGGGCCGGCATCGTGGGTCTC GGCGTCGAGACGCTCGGCTTCATGATCCTCGACGTGA GC
Unnamed	Ib43	Self Generated	*bop*	CTGCTGTTGCTCGACCTCGCGCTCCTCGCGAAGGTCG ACCGCGTGACGATCGGGACGCTCATCGGCGTCGACG CGCTGATGATCGTCACCGGCCTCGTCGGTGCGCTCTC GCAGACGATGCTTGCGCGGTACTCGTTCTGGCTCATC AGCACGATCGCGATGATCTTCGTGCTGTACTACCTCG CCACGAGCCTCCGCAGCGCGGCGAAGGCGCGCTCGG CTGAGGTGGAGAGCACGTTCAACACGTTGACCGTGTT GACGTTGGTCCTTTGGACCGCGTACCCGATCCTGTGG CTCGTCGGAACCGAAGGGGCCGGCATCGTGGGTCTC GGCGTCGAGACGCTCGGCTTCATGATCCTCGACGTGA GC

Unnamed	C191	Self Generated	*bop*	CTGCTGCTGCTCGACCTCGCGCTGCTGGCGAAGGTCG ACCGCGTGAGCATCGGCACGCTCGTCGGCGTCGACGC GCTGATGATCGTCACCGGACTCGTCGGCGCGCTCTCG CACACGCCGCTCGCGCGGTACACGTGGTGGCTGTTCA GCACGATCTGCATGATCGTCGTGCTGTACTTCCTCGC CACGAGCCTGCGCGCCGCTGCGAAGGAGCGCGGTCC CGAGGTCGCGAGCACGTTCAACACCCTTACGGCGTTG GTGTTGGTGCTCTGGACGGCCTACCCGATCCTGTGGA TCATCGGTACCGAGGGTGCCGGCGTCGTCGGCCTCGG CATCGAGACCCTCCTGTTCATGGTGCTCGACGTGACC
Unnamed	Ga36	Self Generated	*bop*	CTGCTGCTGCTCGACCTCGCGCTGCTGGCGAAGGTCG ATCGCGTGAGCATCGGCACGCTCGTCGGCGTCGACGC GCTGATGATCGTCACCGGACTCGTCGGCGCGCTCTCG CACACGCCGCTCGCGCGGTACACGTGGTGGCTGTTCA GCACGATCTGCATGATCGTCGTGCTGTACTTCCTCGC CACGAGCCTGCGCGCCGCTGCGAAGGAGCGCGGTCC CGAGGTCGCGAGCACGTTCAACACCCTTACGGCGTTG GTGTTGGTGCTCTGGACGGCCTACCCGATCCTGTGGA TCATCGGTACCGAGGGTGCCGGCGTCGTCGGCCTCGG CATCGAGACCCTCCTGTTCATGGTGCTCGACGTGACC
Unnamed	C170	Self Generated	*bop*	CTGCTGCTGCTCGACCTCGCGCTGCTGGCGAAGGTCG ACCGCGTGAGCATCGGCACGCTCGTCGGCGTCGACGC GCTGATGATCGTCACCGGACTCGTCGGCGCGCTCTCG CACACGCCGCTCGCGCGGTACACGTGGTGGCTGTTCA GCACGATCTGCATGATCGTCGTGCTGTACTTCCTCGC CACGAGCCTGCGCGCCGCTGCGAAGGAGCGCGGTCC CGAGGTCGCGAGCACGTTCAACACCCTTACGGCGTTG GTGTTGGTGCTCTGGACGGCCTACCCGATCCTGTGGA TCATCGGTACCGAGGGTGCCGGCGTCGTCGGCCTCGG CATCGAGACCCTCCTGTTCATGGTGCTCGACGTGACC
Halorhabdus utahensis	DSM 12940	CP001687.1	*bop*	CTGTTGTTGTACGACATCGCGTTGCTCGCCGGTGCTG ACCGCAACACGATCAGTACACTTGTTGGACTCGACGT GTTGATGATCCTCACCGGCGTCGTCGCGACGCTGACG TTCACCGGCGGGGCCGGGCTGGAAGTCGAGGGCGCT CGCATCCTCTGGTGGGGCGTCTCCACCGGCTTCCTCC TGGTGTTGCTGTACTTCCTGTTCTCGACGCTGACGGC AAAAGCCAACGAGCTGTCGGCTGACACACAGCGTAC GTTCAAACTTCTGCGCAACATGATCGCAGTCCTCTGG CTGGTCTACCCGGTCTGGTGGATCATCGGCACCGAAG GACTCGGCGCTATCGGGATCGGCCCGGAGACCGCCG GTTTCGCCGTGCTTGACGTC

Natrinema altunense sp.	ajinwuensis AJ13	AY570921.1	*bop*	CTGTTGTTGTTAGACCTCTCGCTGTTAGCCGGGGCGA ACCGAAACACGATCGCGACGCTGATCGGCCTCGATGT CTTCATGATCGGTACCGGCGCGATCGCAGCGCTCTCG TCCACCCCGGGTGCTCGGATCGCCTGGTGGGCGATCA GCACCGGTGCCCTGCTCACCCTACTGTACGTCCTCGT CGGGACGCTCTCCGAGAACGCGCGCAGTCGGGCCCC CGAGGTCGCGTCGCTGTTCGGGAGACTCCGCAATCTG GTTATCGCGCTGTGGTTCCTCTACCCGGTGGTCTGGA TCCTCGGCACGGAAGGGACGTTCGGCATCCTGCCGCT GTACTGGGAAACCGCGGCGTTCATGGTGCTCGACCTC TCG
Natrinema sp.	XA3-1	EF558563.1	*bop*	CTGCTGTTGCTCGACCTCTCGCTGTTAGCCGGGGCGA ACCGAAACACGATCGCGACGCTGATCGGCCTCGACG TCTTCATGATCGGAACCGGCGCGATCGCAGCGCTCTC GTCTACCCCGGGCGCTCGGATCGCCTGGTGGGCGATC AGCACCGGTGCCCTGCTCACCCTACTGTACATCCTCG TCGGGACGCTCTCCGAGAACGCGCGCAGTCGGGCCC CCGAGGTCGCGTCGCTGTTCGGGAGACTCCGCAATCT GGTTATCGCGCTGTGGTTCCTCTACCCGGTGGTCTGG ATCCTCGGCACAGAAGGGACGTTCGGCATCCTGCCGC TGTACTGGGAAACCGCGGCGTTCATGGTGCTCGACCT CTCG
Natronococcus aibiensis	AB1	AY279547	*bop*	CTGTTGTTGCTAGACCTCGGACTGCTGGCAGGGGCGA ACCGAAACACGATCGCGACGCTGATCGGACTCGACG TGTTCATGATCGGGACCGGCATGATCGCGGCGTTCGC CGCCACCCCGGGCACCCGGATCGCCTGGTGGGGCATC AGCACCGGCGCCCTGCTCGCCCTGCTGTACGTCCTCG TCGGAACGCTCTCCGAGAACGCGCGCAGTCAGTCCCC CGAGGTCGCATCGCTGTTCGGGCGACTCCGTAACCTG GTCATCGTGCTGTGGCTCCTCTACCCGGTGGTCTGGA TCCTCGGTACGGAGGGGACGTTCGGCATCCTCCCGCT GTACTGGGAGACCGCCGCGTTCATGGTACTCGACCTC TCG
Halobiforma haloterrestris	JCM 11627	EF558553	*bop*	CTCCTGTTGCTCGACCTGTCGCTGCTCGCCGGAGCGA ACCGCAACACGATCGCGACGCTGATCGGACTCGACG TGTTCATGATCGGCACCGGCGCGATCGCGACGTTCGC GTCGACTCCCGCGACCCGCATCGCGTGGTGGGGCATC AGCACCGGCGCCCTGCTGGTCCTCCTGTACGTCCTCG TGGGGACCCTCTCCGAGAAGGCCCGAAGCAAGTCCG CCGAGGTCGCGTCGCTGTTCGGCACCCTTCGTAACCT GGTCATCGTGCTGTGGCTCCTGTACCCGGTCGTCTGG ATCCTCGGCACCGAAGGAACGTTCGGCATCCTCCCGC TGTACTGGGAGACGGCCGCGTTCATGGTGCTCGACCT CTCG

Halobiforma lacisalsi	Unknown	AY607024	*bop*	CTCCTGCTGCTCGACCTCGCACTGCTCGCCGGAGCGA ACCGCAACACGATCGCGACGCTGCTCGGTCTCGACGT CTTCATGATCGGGACCGGCACGATCGCGGCGTTCGCG GCGACCCCTGGCACCCGCATCGCGTGGTGGGGCATCA GCACCGGCGCCCTGCTCGTGCTCCTGTACGTCCTCGT CGGGACGCTCTCCAAGAGTGCCCGAAACCAGTCCGC GGAGGTAGCCTCGCTGTTCAGCACGCTGCGAAACCTG CTCATCGTGCTGTGGCTGCTCTACCCCGTCGTCTGGAT CCTCGGCACCGAAGGAACGTTCGGCATCCTCCCGCTG TACTGGGAGACGGCCGCGTTCATGATCCTCGACCTCT CG
Haloquadratum walsbyi Copy 1	DSM 16790	NC008212	*bop*	CTGCTTCTGCTTGACATTGGCCTCCTTGCCGGTGCAA GTCAGCGCGATATTGGTGCGCTTGTCGGTATTGACGC ATTCATGATCGTGACTGGTCTTGTGGCAACGTTGACA AAGGTCGTGGTCGCACGATATGCTTTCTGGACGATAA GCACTATTTCAATGGTTTTCCTCCTATACTACCTGGTT GCCGTCTTCGGTGAGGCAGTTAGTGATGCAGATGAGG ACACGCGGTCGACATTCAACGCGCTTCGAAATATTAT CCTTGTAACGTGGGCAATATACCCAGTTGCATGGCTT GTTGGAACTGAAGGGCTTGCACTCACCGGTCTCTACG GTGAAACGCTTCTCTTCATGGTTCTTGATTTGGTC
Haloquadratum walsbyi Copy 2	DSM 16790	NC008212	*bop*	CTTTTGATAATTGATTTAGCATTAGTTGCAGGCGCGA GAAAGCAGACACTGTATAAATTGATTATCATTGATGC AATCATGATTCTCGGTGGACTCGCCGGGTCGATGATG CAGCAAGGTGCCGTCATTAGAATCGTATGGTGGGCA GTGAGTACTGCGGCATTTATTATTTTATTATATTATCT ACTCGGTGAACTCTCAGAACGCGCTAGAAGCCGATC AGCAGAGACGGGTATAGTATTCAACCGACTTCGTAAT ATTACACTCGGACTCTGGGCACTATACCCAATCGTAT GGATTCTTGGAACAGGCGGTGGATTTGGTATAATCGC TGTCACCACAGAAATAATGCTATACGTTATGCTTGAT ATTGGA
Natronorubrum sp.	Tenzan-10	EF558562.1	*bop*	CTGCTGTTGCTCGACCTCGCGCTGCTCGCCGGCGCGA ACCGCAACACGATCGCGACGCTAATCGGACTCGACG TCCTCATGATCGGAACCGGTACGGTCGCGGCGTTCGC GGCGACCCCCGGCACCCGGATCGCGTGGTGGGGCAT CAGCACCGCCGCCCTGCTCGTGTTGCTGTACGTCCTT GTCGGGACCCTCTCGAAGAACGCGCGTGACCAGTCTC CGGAAGTCGCGTCGCTATTCGGCACCCTCCGCAACCT GGTCATCGTGCTCTGGTTGCTCTACCCCGTCGTCTGG ATCCTCGGGACCGAGGGCACGTTCGGCATCCTCCCGC TGTACTGGGAGACGGCCGCGTTCATGGTCCTCGATCT CTCG

Haloferax mucosum	Unknown	Hmu5122174	*bop*	TTGTTACTGTTTGACCTCGCGCTGTTAGCCGGCGCGA ATCGAAACACCATCGCGACGCTCGTCGGTCTTGACGT GTTCATGATCGTGACCGGCGTCATCGCCGCGTTCGAA CAGGTCGCCGCGTACCGCATCCTCTGGTGGGGTATCA GCACGGGTGCGCTGTTGGTCCTGCTGTACGTCCTCGT CGGTACGCTGTCCGAACAGGCCCGTGAAAAGTCGGG CGAGGTCGCACAACTGTTCAGCACGCTTCGAAACCTG ATGATATTCCTGTGGCTCCTGTACCCCGTGGTCTGGA TTCTCGGGACCGAGGGAACCATCGGCGTCCTCCCGCT GTACTGGGAGACGGCGGCGTTCATGGTGCTCGACCTC ACC
Haloarcula californiae duplicate	Unknown	Hca7992155	*bop*	TTGCTGCTGTATGACCTCGGACTGCTTGCAGGTGCAG ACCGCAACACCATTTACTCGCTCGTCAGCCTCGACGT GCTGATGATCGGTACTGGTGTGGTCGCAACGCTGAGT GCAGGCAGTGGCGTGCTGTCTGCCGGTGCAGAACGG CTGGTCTGGTGGGGTATCAGCACTGCCTTCCTGCTGG TCCTGCTGTACTTCCTGTTCAGTTCGCTGTCCGGCCGG GTTGCAAACCTGCCCAGTGACACGCGAAGCACGTTCA AGACGTTGCGCAACCTCGTCACTGTCGTCTGGCTGGT CTACCCAGTCTGGTGGCTCGTCGGCAGTGAGGGGCTT GGCCTCGTCGGTATCGGAATCGAGACGGCCGGCTTCA TGGTCATCGACCTCGTC
Haloarcula sinaiiensis	Unknown	Hsi8002227	*bop*	TTGCTGCTGTATGACCTCGGACTGCTTGCAGGTGCAG ACCGCAACACCATTTACTCGCTCGTCAGCCTCGACGT GCTGATGATCGGTACTGGTGTGGTCGCAACGCTGAGT GCAGGCAGTGGCGTGCTGTCTGCCGGTGCAGAACGG CTGGTCTGGTGGGGTATCAGCACTGCCTTCCTGCTGG TCCTGCTGTACTTCCTGTTCAGTTCGCTGTCCGGCCGG GTTGCAAACCTGCCCAGTGACACGCGAAGCACGTTCA AGACGTTGCGCAACCTCGTCACTGTCGTCTGGCTGGT CTACCCAGTCTGGTGGCTCGTCGGCAGTGAGGGGCTT GGCCTCGTCGGTATCGGAATCGAGACGGCCGGCTTCA TGGTCATCGACCTCGTC
Haloarcula marismortui duplicate	ATCC 43049	AY596297.1	*bop*	TTGCTGCTGTATGACCTCGGACTGCTTGCAGGTGCAG ACCGCAACACCATTTACTCGCTCGTCAGCCTCGACGT GCTGATGATCGGTACTGGTGTGGTCGCAACGCTGAGT GCAGGCAGTGGCGTGCTGTCTGCCGGTGCAGAACGG CTGGTCTGGTGGGGTATCAGCACTGCCTTCCTGCTGG TCCTGCTGTACTTCCTGTTCAGTTCGCTGTCTGGCCGG GTTGCAAACCTGCCCAGTGACACGCGAAGCACGTTCA AGACGTTGCGCAACCTCGTCACTGTCGTCTGGCTGGT CTACCCAGTCTGGTGGCTCGTCGGCAGTGAGGGGCTT GGCCTCGTCGGTATCGGAATCGAGACGGCCGGCTTCA TGGTCATCGACCTCGTC

Halobacterium salinarium duplicate	DSM 671	AM774415.1	*bop*	CTGTTGTTGTTAGACCTCGCGTTGCTCGTTGACGCGG ATCAGGGAACGATCCTTGCGCTCGTCGGTGCCGACGG CATCATGATCGGGACCGGCCTGGTCGGCGCACTGACG AAGGTCTACTCGTACCGCTTCGTGTGGTGGGCGATCA GCACCGCAGCGATGCTGTACATCCTGTACGTGCTGTT CTTCGGGTTCACCTCGAAGGCCGAAAGCATGCGCCCC GAGGTCGCATCCACGTTCAAAGTACTGCGTAACGTTA CCGTTGTGTTGTGGTCCGCGTATCCCGTCGTGTGGCT GATCGGCAGCGAAGGTGCGGGAATCGTGCCGCTGAA CATCGAGACGCTGCTGTTCATGGTGCTTGACGTGAGC
Halobacterium sp. duplicate	NRC-1	AE004437.1	*bop*	CTGTTGTTGTTAGACCTCGCGTTGCTCGTTGACGCGG ATCAGGGAACGATCCTTGCGCTCGTCGGTGCCGACGG CATCATGATCGGGACCGGCCTGGTCGGCGCACTGACG AAGGTCTACTCGTACCGCTTCGTGTGGTGGGCGATCA GCACCGCAGCGATGCTGTACATCCTGTACGTGCTGTT CTTCGGGTTCACCTCGAAGGCCGAAAGCATGCGCCCC GAGGTCGCATCCACGTTCAAAGTACTGCGTAACGTTA CCGTTGTGTTGTGGTCCGCGTATCCCGTCGTGTGGCT GATCGGCAGCGAAGGTGCGGGAATCGTGCCGCTGAA CATCGAGACGCTGCTGTTCATGGTGCTTGACGTGAGC
Haloarcula marismortui Copy 2	ATCC 43049	AY596297.1	*bop*	CTGTTGTTGCTTGATATCGGCCTCCTGGCCGGCGCAA GCAACCGCGACATGGCGTCGCTCATCACCATCGACGC CTTCATGATCGTTACCGGCCTCGCCGCGACGCTGATG AAGGTGCCTGTGGCCCGGTACGCCTTCTGGACTATCA GCACCATCGCGATGCTGTTCGTCCTGTACTACCTCGT CGTCGTTGTCGGCGAGGCCGCGAGCGATGCCAGCGA AGAAGCACAGTCGACGTTCAACGTCCTTCGAAACATC ATCCTCGTCGCGTGGGCCATCTACCCCGTCGCGTGGC TTGTTGGGACCGAGGGACTCGGACTCGTCGGCCTGTT CGGCGAAACCCTGCTGTTCATGATTCTGGACCTGACC
Haloarcula sinaiiensis Copy 2	Unknown	Hsi8002753	*bop*	CTGTTGTTGCTTGATATCGGCCTCCTGGCCGGCGCAA GCAACCGCGACATGGCGTCGCTCATCACCATCGACGC CTTCATGATCGTTACCGGCCTCGCCGCGACGCTGATG AAGGTGCCTGTGGCCCGGTACGCCTTCTGGACTATCA GCACCATCGCGATGCTGTTCGTCCTGTACTACCTCGT CGTCGTTGTCGGCGAGGCCGCGAGCGATGCCAGCGA AGAAGCACAGTCGACGTTCAACGTCCTTCGAAACATC ATCCTCGTCGCGTGGGCCATCTACCCCGTCGCGTGGC TTGTTGGGACCGAGGGACTCGGACTCGTCGGCCTGTT CGGCGAAACCCTGCTGTTCATGATTCTGGACCTGACC

Haloarcula vallismortis Copy 2	Unknown	Hva7150169	*bop*	CTGCTGTTGCTTGATATCGGCCTCCTGGCCGGCGCAA GCAACCGCGACATGGCGTCGCTCATCACCATCGACGC CTTCATGATCGTTACCGGCCTCGCCGCGACGCTGATG AAGGTGCCTGTGGCCCGGTACGCCTTCTGGACTATCA GCACCATCGCGATGCTGTTCGTCCTGTACTACCTCGT CGTCGTTGTCGGCGAGGCCGCGAGCGATGCCAGCGA AGAAGCACAGTCGACGTTCAACGTCCTTCGAAACATC ATCCTCGTCGCGTGGGCTATCTACCCCGTCGCGTGGC TTGTTGGGACCGAGGGACTCGGACTCGTCGGCCTGTT CGGCGAAACCCTGCTGTTCATGATTCTGGACCTGACC
Haloquadratum walsbyi Copy 1 duplicate	DSM 16790	AM180088.1	*bop*	CTGCTTCTGCTTGACATTGGCCTCCTTGCCGGTGCAA GTCAGCGCGATATTGGTGCGCTTGTCGGTATTGACGC ATTCATGATCGTGACTGGTCTTGTGGCAACGTTGACA AAGGTCGTGGTCGCACGATATGCTTTCTGGACGATAA GCACTATTTCAATGGTTTTCCTCCTATACTACCTGGTT GCCGTCTTCGGTGAGGCAGTTAGTGATGCAGATGAGG ACACGCGGTCGACATTCAACGCGCTTCGAAATATTAT CCTTGTAACGTGGGCAATATACCCAGTTGCATGGCTT GTT
Halorubrum tibetense	JCM 11889	Self Generated	*sopII*	CTTTTGATCCTCTACCTCGCGTTGCTCGCCCGGCCGTC CCGGCGGGTCATCGTCGCGCTGATCGCGGTCGACGTG GTCGTCATCGCGGGCGGCACCGTCGCCGTCGTCACGA CCGGGTTCGTCTCGTGGGCCGCGTTCGGCGTCGCGAC CCTCGCGTACGTCGGACTCGTCTACGGGTTGCTCACC GTACTTCCGCGTTCGGCGGCGAGCCAGCCGGACCGG GTACGGGCGGTTTTCGGAACGCTCCGGAACATCACGG TGGTCCTGTGGACGCTGTACCCGGTCGTCTGGCTGCT CGCGCCGACCGGGCTGGGCCTGTTGACCGCCTCGACG GAGATGCTCGTGTTCGTCTACCTGGATTTCGTC
Halorubrum tibetense	JCM 11889	Self Generated	*sopII*	CTTTTGATCCTCTACCTCGCGTTGCTCGCCCGGCCGTC CCGGCGGGTCATCGTCGCGCTGATCGCGGTCGACGTG GTCGTCATCGCGGGCGGCACCGTCGCCGTCGTCACGA CCGGGTTCGTCTCGTGGGCCGCGTTCGGCGTCGCGAC CCTCGCGTACGTCGGACTCGTCTACGGGTTGCTCACC GTACTTCCGCGTTCGGCGGCGAGCCAGCCGGACCGG GTACGGGCGGTTTTCGGAACGCTCCGGAACATCACGG TGGTCCTGTGGACGCTGTACCCGGTCGTCTGGCTGCT CGCGCCGACCGGGCTGGGCCTGTTGACCGCCTCGACG GAGATGCTCGTGTTCGTCTACCTGGATTTCGTC

Haloferax zhejiangensis	ZJ204	AY838280	*bop*	CCGCTGCTGCTGTTCGACCTCGGGCAGGCACCGACGG TCACGACCGACCCCGAGGAAGCTATCGACGGCGCCG ACGTGGTCTACACCGACGTGTGGATTTCGATGGGCCA AGAGGACCAACGCGACGAAAAGCTCGAAGCGTTCGA CGGGTTCCAACTGAATGGCGAACTCCTCGACGGGAG CGACGCGCAAGTCATGCACTGCCTGCCCGCCCACCGC GGCGAGGAGATTACGGGTGACGTGCTCGAAGGCGAG CAGTCGCTCGTCTGGGAGCAAGCGGAGAACCGCCTG CACGCCCAGAAGGGCCTCATCGTCGAGTTGCTGGAA GAAAAGGTCGGGTTCGGCTTCATCCT
Halorubrum lacusprofundi	ATCC 49239	CP001365.1	*sopII*	CTTTTGATCCTTTACCTCGGGCTGCTCGCCCGGCCGTC CCGACGCGTGCTCGTCGGGCTGATCGGGGTCGACGTG GTGATCATCGCCGGCGGTATCGTCGCCGCTGCCACCA CCGGAACGGTTTCGTGGGTCGCCTTCGGCGTCGCCGG AGTCGCGTACCTCGGACTTGTCTACGGGCTCCTCGTC TCGCTCCCGCGGTCGGCCTCGGCGGAGGGGGACCGG ATCCGCGCCGTCTTCGGTACGCTCCGGAACATCACGG TGGTGTTGTGGACGCTGTACCCGGTCGTGTGGCTGCT CGCGCCCACCGGGTTCGGCCTGCTGACGCCGACGACG GAGATGCTCGTGTTCGTCTACCTCGATATCGTC
Halobacterium salinarium	DSM 671	AM774415.1	*sopII*	CTGATCGTGTTGTACCTCGCGATGCTCGCCCGCCCCG GCCACCGCACGTCCGCCTGGCTGCTCGCCGCCGACGT GTTCGTCATCGCCGCCGGGATCGCCGCCGCGCTCACC ACGGGCGTGCAGCGCTGGCTGTTTTTCGCCGTCGGCG CTGCCGGCTACGCCGCGCTCCTGTACGGCCTCCTGGG GACGCTGCCGCGCGCGCTCGGCGACGACCCCCGCGT GCGCTCGCTGTTCGTCACGCTTCGCAACATCACGGTC GTCCTCTGGACGCTGTACCCCGTGGTCTGGCTGCTGT CCCCGGCCGGCATCGGCATCCTCCAAACCGAGATGTA CACCATCGTCGTGGTCTACCTGGATTTCATC
Halobacterium sp.	NRC-1	AE004437.1	*sopII*	CTGATCGTGTTGTACCTCGCGATGCTCGCCCGCCCCG GCCACCGCACGTCCGCCTGGCTGCTCGCCGCCGACGT GTTCGTCATCGCCGCCGGGATCGCCGCCGCGCTCACC ACGGGCGTGCAGCGCTGGCTGTTTTTCGCCGTCGGCG CTGCCGGCTACGCCGCGCTCCTGTACGGCCTCCTGGG GACGCTGCCGCGCGCGCTCGGCGACGACCCCCGCGT GCGCTCGCTGTTCGTCACGCTTCGCAACATCACGGTC GTCCTCTGGACGCTGTACCCCGTGGTCTGGCTGCTGT CCCCGGCCGGCATCGGCATCCTCCAAACCGAGATGTA CACCATCGTCGTGGTCTACCTGGATTTCATC

Halobacterium salinarium	DSM 671	AM774415.1	*hop*	GTAAAACGCCCACCGGAAGAGCAGCGCCGACGTGGT CATCGCCGCCGCCAACCCCGTCACGCACATCCCGATG TCGGCCGCGATCACGGTAAACAGACTGCCGAGATCG ACGTCCGCCAGCAGCCCCAGCGCCAGCAGTATCATCG GCGTCGACAGCGCCCACGTGAGATACCGCCCCCACTG ACTGCGCACCATCTCGCCGGCCAGCGCGTGCCCGGCG GGCATCTCGATCATCCCCACGGTGAGCCCCGACAGCA GCCCGAGGTAGCTGGAGATCGACACCAGCGGGATCA TCAGCGTCGCCCCCCAGATGAGCCGCGGTCGTCCCGG TCTGATGGTGCGTCCCATATACACGAACACGAGGATC GCGATCCCCGCGAGCGCGACGTTCACCCACAGCGAC GAACTCAACAGCGCGTTCTCGCGGACCGCGGCCGCC GATTGCGCGCCCAGCACCCCCGCATCGACCACACCGG G

5.2 Trees Used for Hybrid Approximately Unbiased (AU) Testing

50 10

(Hqm_1walsb,Hqm_1.2wal,((Har_2maris,(Har_2valli,Har_2sinai)),(((Hbt_2NRC1_,(Hbt_2salin,(Hbt_salina,Hbt_NRC1_b))),((Hfx_mucosu,(Hbf_halote,((Ncc_aibien,(Htg_GSL11_,(Ntm_pallid,(Nnm_XA31_b,(Ntm_altune,Nnm_ajinwu))))),(Nrr_tenzan,Hbf_lacisa)))),(Hbl_gomorr,(Hsx_carlsb,(Hrd_utahen,(Hmc_mukoha,((Fb19_bop,Ha9_bop),((Har_amylol,Har_hispan),((Har_sinaii,(Har_califo,(Har_2calif,(Har_marism,Har_1.2mar)))),(Har_vallis,(Har_argent,Har_japoni))))))))))),((Hrr_tebenq,Hrr_ejinor),((Eb13_bop,Ib43_bop),(Hrr_sodome,(((Hrr_2distr,Hrr_arcis_),(Hrr_distri,Hrr_litore)),(Fb21_bop,(EC15_bop,(C170_bop,(C191_bop,(Ga27_bop,Ga36_bop))))))))));

(Hqm_1walsb,((Hqm_1.2wal,(((Hrr_sodome,(Fb21_bop,(EC15_bop,(Hbt_2NRC1_,(Hbt_2salin,(Hbt_salina,Hbt_NRC1_b)))))),(Eb13_bop,Ib43_bop)),(Hrr_tebenq,Hrr_ejinor))),(((Hsx_carlsb,(Hbl_gomorr,(Har_2maris,(Har_2valli,Har_2sinai)))),((Hmc_mukoha,(((((Ncc_aibien,(Htg_GSL11_,(Ntm_pallid,(Nnm_XA31_b,(Ntm_altune,Nnm_ajinwu))))),(Har_sinaii,(Har_califo,(Har_2calif,(Har_marism,Har_1.2mar))))),(Har_vallis,(Har_argent,Har_japoni))),(Har_amylol,Har_hispan)),(Fb19_bop,Ha9_bop))),(Hrd_utahen,((Hrr_2distr,Hrr_arcis_),(Hrr_distri,Hrr_litore))))),((Hfx_mucosu,(Hbf_halote,(Nrr_tenzan,Hbf_lacisa))),(C170_bop,(C191_bop,(Ga27_bop,Ga36_bop)))))));

(Hqm_1walsb,Hqm_1.2wal,((Har_2maris,(Har_2valli,Har_2sinai)),(((Hbt_2NRC1_,(Hbt_2salin,(Hbt_salina,Hbt_NRC1_b))),((Hbf_halote,((Ncc_aibien,(Htg_GSL11_,(Ntm_pallid,(Nnm_XA31_b,(Ntm_altune,Nnm_ajinwu))))),(Nrr_tenzan,Hbf_lacisa))),(Hbl_gomorr,(Hsx_carlsb,(Hrd_utahen,(Hmc_mukoha,((Fb19_bop,Ha9_bop),((Har_amylol,Har_hispan),((Har_sinaii,(Har_califo,(Har_2calif,(Har_marism,Har_1.2mar)))),(Har_vallis,(Har_argent,Har_japoni))))))))))),(Hfx_mucosu,((Hrr_tebenq,Hrr_ejinor),((Eb13_bop,Ib43_bop),(Hrr_sodome,(((Hrr_2distr,Hrr_arcis_),(Hrr_distri,Hrr_litore)),(Fb21_bop,(EC15_bop,(C170_bop,(C191_bop,(Ga27_bop,Ga36_bop)))))))))));

(((Hfx_mucosu,((Hqm_1walsb,Hqm_1.2wal),(Har_2maris,(Har_2valli,Har_2sinai)))),((Hrr_tebenq,Hrr_ejinor),((Eb13_bop,Ib43_bop),(Hrr_sodome,(((Hrr_2distr,Hrr_arcis_),(Hrr_distri,Hrr_litore)),(Fb21_bop,(EC15_bop,(C170_bop,(C191_bop,(Ga27_bop,Ga36_bop))))))))),((Hbt_2NRC1_,(Hbt_2salin,(Hbt_salina,Hbt_NRC1_b))),((Hbf_halote,((Ncc_aibien,(Htg_GSL11_,(Ntm_pallid,(Nnm_XA31_b,(Ntm_altune,Nnm_ajinwu))))),(Nrr_tenzan,Hbf_lacisa))),(Hbl_gomorr,(Hsx_carlsb,(Hrd_utahen,(Hmc_mukoha,((Fb19_bop,Ha9_bop),((Har_amylol,Har_hispan),((Har_sinaii,(Har_califo,(Har_2calif,(Har_marism,Har_1.2mar)))),(Har_vallis,(Har_argent,Har_japoni))))))))))));

(Hqm_1walsb,Hqm_1.2wal,((Har_2maris,(Har_2valli,Har_2sinai)),(((Hbt_2NRC1_,(Hbt_2salin,(Hbt_salina,Hbt_NRC1_b))),((Hrr_tebenq,Hrr_ejinor),((Eb13_bop,Ib43_bop),(Hrr_sodome,(((Hrr_2distr,Hrr_arcis_),(Hrr_distri,Hrr_litore)),(Fb21_bop,(EC15_bop,(C170_bop,(C191_bop,(Ga27_bop,Ga36_bop))))))))),((Hfx_mucosu,(Hbf_halote,((Ncc_aibien,(Htg_GSL11_,(Ntm_pallid,(Nnm_XA31_b,(Ntm_altune,Nnm_ajinwu))))),(Nrr_tenzan,Hbf_lacisa)))),(Hbl_gomorr,(Hsx_carlsb,(Hrd_utahen,(Hmc_mukoha,((Fb19_bop,Ha9_bop),((Har_amylol,Har_hispan),((Har_sinaii,(Har_califo,(Har_2calif,(Har_marism,Har_1.2mar)))),(Har_vallis,(Har_argent,Har_japoni)))))))))));
(Hqm_1walsb,Hqm_1.2wal,((Har_2maris,(Har_2valli,Har_2sinai)),((((Hbt_2NRC1_,(Hbt_2salin,(Hbt_salina,Hbt_NRC1_b))),(Hfx_mucosu,(Hbf_halote,((Ncc_aibien,(Htg_GSL11_,(Ntm_pallid,(Nnm_XA31_b,(Ntm_altune,Nnm_ajinwu))))),(Nrr_tenzan,Hbf_lacisa))))),((Hrr_tebenq,Hrr_ejinor),((Eb13_bop,Ib43_bop),(Hrr_sodome,(((Hrr_2distr,Hrr_arcis_),(Hrr_distri,Hrr_litore)),(Fb21_bop,(EC15_bop,(C170_bop,(C191_bop,(Ga27_bop,Ga36_bop))))))))),(Hbl_gomorr,(Hsx_carlsb,(Hrd_utahen,(Hmc_mukoha,((Fb19_bop,Ha9_bop),((Har_amylol,Har_hispan),((Har_sinaii,(Har_califo,(Har_2calif,(Har_marism,Har_1.2mar)))),(Har_vallis,(Har_argent,Har_japoni)))))))))));
(((((Hqm_1walsb,Hqm_1.2wal),(Har_2maris,(Har_2valli,Har_2sinai))),((Hrr_tebenq,Hrr_ejinor),((Eb13_bop,Ib43_bop),(Hrr_sodome,(((Hrr_2distr,Hrr_arcis_),(Hrr_distri,Hrr_litore)),(Fb21_bop,(EC15_bop,(C170_bop,(C191_bop,(Ga27_bop,Ga36_bop))))))))),((Hfx_mucosu,(Hbf_halote,((Ncc_aibien,(Htg_GSL11_,(Ntm_pallid,(Nnm_XA31_b,(Ntm_altune,Nnm_ajinwu))))),(Nrr_tenzan,Hbf_lacisa)))),(Hbl_gomorr,(Hsx_carlsb,(Hrd_utahen,(Hmc_mukoha,((Fb19_bop,Ha9_bop),((Har_amylol,Har_hispan),((Har_sinaii,(Har_califo,(Har_2calif,(Har_marism,Har_1.2mar)))),(Har_vallis,(Har_argent,Har_japoni)))))))))),(Hbt_2NRC1_,(Hbt_2salin,(Hbt_salina,Hbt_NRC1_b))));
((((Hfx_mucosu,((Hqm_1walsb,Hqm_1.2wal),(Har_2maris,(Har_2valli,Har_2sinai)))),((Hrr_tebenq,Hrr_ejinor),((Eb13_bop,Ib43_bop),(Hrr_sodome,(((Hrr_2distr,Hrr_arcis_),(Hrr_distri,Hrr_litore)),(Fb21_bop,(EC15_bop,(C170_bop,(C191_bop,(Ga27_bop,Ga36_bop))))))))),((Hbf_halote,((Ncc_aibien,(Htg_GSL11_,(Ntm_pallid,(Nnm_XA31_b,(Ntm_altune,Nnm_ajinwu))))),(Nrr_tenzan,Hbf_lacisa))),(Hbl_gomorr,(Hsx_carlsb,(Hrd_utahen,(Hmc_mukoha,((Fb19_bop,Ha9_bop),((Har_amylol,Har_hispan),((Har_sinaii,(Har_califo,(Har_2calif,(Har_marism,Har_1.2mar)))),(Har_vallis,(Har_argent,Har_japoni)))))))))),(Hbt_2NRC1_,(Hbt_2salin,(Hbt_salina,Hbt_NRC1_b))));
(((((Hqm_1walsb,Hqm_1.2wal),(Har_2maris,(Har_2valli,Har_2sinai))),((Hrr_tebenq,Hrr_ejinor),((Eb13_bop,Ib43_bop),(Hrr_sodome,(((Hrr_2distr,Hrr_arcis_),(Hrr_distri,Hrr_litore)),(Fb21_bop,(EC15_bop,(C170_bop,(C191_bop,(Ga27_bop,Ga36_bop))))))))),((Hbt_2NRC1_,(Hbt_2salin,(Hbt_salina,Hbt_NRC1_b))),(Hfx_mucosu,(Hbf_halote,((Ncc_aibien,(Htg_GSL11_,(Ntm_pallid,(Nnm_XA31_b,(Ntm_altune,Nnm_ajinwu))))),(Nrr_tenzan,Hbf_lacisa))))),(Hbl_gomorr,(Hsx_carlsb,(Hrd_utahen,(Hmc_mukoha,((Fb19_bop,Ha9_bop),((Har_amylol,Har_hispan),((Har_sinaii,(Har_califo,(Har_2calif,(Har_marism,Har_1.2mar)))),(Har_vallis,(Har_argent,Har_japoni)))))))))));
((((Hfx_mucosu,((Hqm_1walsb,Hqm_1.2wal),(Har_2maris,(Har_2valli,Har_2sinai)))),(Hbl_gomorr,((Hrr_tebenq,Hrr_ejinor),((Eb13_bop,Ib43_bop),(Hrr_sodome,(((Hrr_2distr,Hrr_arcis_),(Hrr_distri,Hrr_litore)),(Fb21_bop,(EC15_bop,(C170_bop,(C191_bop,(Ga27_bop,Ga36_bop))))))))),((Hbt_2NRC1_,(Hbt_2salin,(Hbt_salina,Hbt_NRC1_b))),(Hbf_halote,((Ncc_aibien,(Htg_GSL11_,(Ntm_pallid,(Nnm_XA31_b,(Ntm_altune,Nnm_ajinwu))))),(Nrr_tenzan,Hbf_lacisa))))),(Hsx_carlsb,(Hrd_utahen,(Hmc_mukoha,((Fb19_bop,Ha9_bop),((Har_amylol,Har_hispan),((Har_sinaii,(Har_califo,(Har_2calif,(Har_marism,Har_1.2mar)))),(Har_vallis,(Har_argent,Har_japoni)))))))));

5.*3 Tree Key for AU Testing Tree Set*

1: Original Maximum Likelihood Tree
2: Negative Control (Original with random changes)
3: Grouped with Halorubrum Clade
4: Secondary Copies of *bop* from Halorubrum and Haloferax Sequences Grouped with Halorubrum Clade
5: Halobacterium Clade Grouped with Halorubrum Clade
6: Mixed Clade Grouped with Halorubrum Clade
7: Modeled After MLSA phylogeny
8: Modeled After MLSA phylogeny (Strict Version)
9: Modeled After 16s rRNA phylogeny
10: Modeled After 16s rRNA phylogeny (Strict Version)

5.4 Site Wise Log Likelihoods for AU Testing

10 402
tr1 -11.991771 -3.178918 -23.020959 -32.357208 -3.178918 -11.223634 -25.186238 -3.178918 -9.297133 -25.176642 -10.184068 -32.148904 -2.822538 -3.968040 -16.560927 -1.136585 -1.136585 -1.513334 -11.576637 -8.613616 -11.773389 -19.888632 -16.496119 -30.291766 -16.913720 -7.488678 -17.694760 -12.674621 -8.613616 -37.545284 -2.822538 -11.464636 -34.185472 -9.540665 -13.065447 -35.661158 -7.412069 -11.441755 -21.508549 -16.733030 -7.178916 -27.994735 -8.869556 -9.296224 -27.733653 -14.470397 -14.893428 -23.027294 -13.185268 -23.763951 -21.775512 -7.767736 -3.178918 -25.044360 -32.857797 -26.816766 -24.132527 -24.126378 -9.238104 -6.589858 -8.191928 -8.149368 -23.970046 -24.174418 -9.290933 -11.349236 -13.971165 -7.673247 -28.111402 -10.355888 -6.264313 -18.821947 -2.822538 -3.968040 -10.898053 -7.287929 -11.529651 -24.588931 -18.134377 -3.178918 -10.600111 -3.968040 -3.178918 -2.822538 -3.968040 -3.178918 -2.876176 -6.336921 -16.182530 -36.516782 -3.968040 -2.876176 -15.083232 -2.822538 -2.822538 -30.032019 -20.835988 -19.216287 -16.875940 -24.144429 -7.776123 -6.428470 -2.822538 -8.994599 -25.296523 -15.655461 -2.876176 -19.909715 -14.275649 -3.178918 -19.545350 -22.601158 -17.864488 -32.172011 -31.563582 -17.760137 -29.613199 -14.133226 -16.453885 -30.814124 -3.432976 -3.432976 -1.597615 -0.000000 -0.000000 -0.000000 -0.000000 -0.000000 -0.000000 -9.273786 -8.433714 -13.878979 -2.053512 -6.446631 -10.401195 -2.053512 -7.889031 -7.314983 -2.087769 -2.421174 -9.861884 -9.250089 -8.406689 -11.776781 -2.053512 -8.824979 -5.825350 -22.306812 -18.893334 -31.697760 -15.812795 -14.974637 -36.376548 -21.643922 -18.929174 -23.239281 -2.876176 -2.822538 -21.087153 -15.118045 -8.934182 -12.410534 -24.116068 -11.156889 -23.337901 -3.178918 -12.779927 -10.481977 -3.178918 -2.822538 -2.822538 -9.338263 -14.188026 -30.019805 -18.542875 -3.178918 -7.454231 -13.147171 -9.929577 -18.350199 -3.968040 -2.876176 -19.521801 -7.522778 -21.063017 -21.668193 -18.118217 -12.542309 -24.031455 -14.500267 -3.178918 -11.243206 -17.066820 -3.178918 -21.497528 -18.839896 -14.132613 -14.584061 -23.610624 -3.178918 -23.648455 -2.876176 -3.178918 -12.049760 -3.178918 -3.968040 -2.876176 -21.374813 -14.383439 -9.957440 -2.876176 -3.178918 -15.783577 -16.081148 -13.531726 -15.303765 -21.656141 -20.460737 -31.531836 -19.740568 -19.027719 -23.128386 -12.185656 -3.178918 -21.821702 -19.287974 -

7.133846 -28.724196 -25.393658 -19.836842 -28.495525 -17.951082 -16.608455 -31.500100 -
2.822538 -10.217015 -36.055391 -23.158616 -20.349684 -33.341128 -25.565645 -24.028981 -
24.921392 -22.012220 -20.367614 -18.147249 -24.375415 -16.925890 -37.403757 -22.844733 -
19.714373 -33.170998 -2.822538 -3.968040 -28.578031 -15.843086 -14.180308 -26.999576 -
15.350133 -22.446369 -32.610319 -19.906976 -18.320760 -17.454602 -8.696702 -15.484690 -
23.136112 -3.178918 -3.178918 -2.876176 -20.073054 -17.093994 -19.306834 -18.897523 -
19.052385 -26.834339 -25.102294 -3.178918 -30.660274 -7.627910 -6.434325 -43.133030 -
7.179681 -12.477055 -19.119604 -20.783506 -3.178918 -17.609712 -22.305431 -10.084647 -
37.068378 -25.702388 -13.805247 -34.132304 -12.944411 -10.731778 -22.874071 -19.081349 -
6.500747 -25.643210 -3.178918 -2.822538 -2.822538 -24.780392 -7.293610 -32.301160 -
10.438051 -8.802994 -28.838926 -3.178918 -3.968040 -12.719409 -2.876176 -2.876176 -
27.523336 -6.746200 -13.871654 -20.867628 -14.872579 -16.048805 -20.927104 -3.175784 -
2.819342 -2.819342 -22.070973 -9.637497 -17.040203 -23.188308 -6.272535 -6.587982 -
2.819342 -2.819342 -33.406898 -3.964397 -14.830754 -25.320995 -2.819342 -3.964397 -
30.673037 -2.819342 -2.819342 -46.209958 -2.628034 -1.673965 -4.264061 -24.440094 -
8.793625 -16.223228 -2.819342 -9.047788 -17.257244 -19.378965 -6.290523 -6.405257 -
14.843186 -6.397485 -22.392881 -13.229036 -11.330482 -22.272023 -12.015144 -3.175783 -
21.449106 -18.255711 -21.965424 -23.026598 -20.320380 -20.073525 -12.626306 -2.819342 -
3.964397 -18.127295 -6.990144 -6.388567 -19.549018 -10.239783 -10.074008 -18.202948 -
14.654911 -20.586670 -21.989806 -7.729703 -7.315995 -2.872903 -9.675734 -6.415647 -
9.073355 -18.186215 -3.175783 -29.717993 -11.110792 -3.175783 -16.792470 -2.819342 -
8.292124 -11.065538 -15.032087 -7.729703 -19.299384 -23.934164 -23.051731 -22.451622
tr2 -14.706015 -3.631661 -23.476321 -34.762836 -3.631661 -13.544328 -28.084869 -
3.631661 -10.082437 -29.188370 -17.330204 -34.968741 -3.272086 -4.525268 -15.746989 -
1.136585 -1.136585 -1.513334 -16.199214 -9.433748 -13.654524 -21.286459 -19.077496 -
33.142210 -18.351521 -8.294200 -19.351689 -14.042658 -9.433748 -37.473550 -3.272086 -
11.996200 -36.232494 -15.622899 -16.431350 -36.432384 -8.205647 -17.308524 -25.818948 -
20.865969 -11.771151 -29.944630 -7.581443 -12.672870 -27.422899 -18.952003 -21.875846 -
26.140288 -16.716257 -27.898991 -26.057724 -8.545553 -3.631661 -27.353575 -36.202574 -
29.882469 -26.375574 -26.824012 -10.026269 -7.365706 -8.978199 -8.941005 -27.535910 -
27.334716 -10.090163 -12.640623 -15.279632 -6.443664 -31.227784 -15.893491 -6.265753 -
21.020988 -3.272086 -4.525268 -11.684473 -7.950406 -15.496532 -25.159929 -21.900845 -
3.631661 -15.749944 -4.525268 -3.631661 -3.272086 -4.525268 -3.631661 -3.343422 -7.120794
-23.034721 -38.576683 -4.525268 -3.343422 -16.842911 -3.272086 -3.272086 -34.024238 -
25.712299 -21.145483 -19.014415 -28.202134 -6.543758 -8.478408 -3.272086 -11.440526 -
27.241495 -18.936516 -3.343422 -20.876213 -16.515033 -3.631661 -24.983633 -29.138568 -
19.855762 -34.736166 -35.292545 -24.137796 -31.669921 -21.140827 -22.017306 -34.395588 -
3.453787 -3.453787 -1.580827 -0.000000 -0.000000 -0.000000 -0.000000 -0.000000 -0.000000 -
9.193844 -8.488413 -13.967728 -2.047035 -6.420527 -10.340763 -2.047035 -7.881548 -
7.377265 -2.081128 -2.414815 -9.914818 -9.256067 -8.343735 -11.813052 -2.047035 -8.546410
-5.937978 -25.293479 -23.364183 -32.433177 -18.570697 -21.824140 -38.410930 -25.030704 -
23.556388 -26.165155 -3.343422 -3.272086 -22.119726 -22.584912 -20.385552 -15.671216 -
28.413316 -17.925922 -21.350291 -3.631661 -17.394925 -14.030164 -3.631661 -3.272086 -
3.272086 -16.182960 -20.030070 -31.838455 -26.307523 -3.631661 -9.631289 -14.434817 -

11.103639 -20.172353 -4.525268 -3.343422 -23.410846 -15.945794 -27.475058 -22.784184 -
22.269998 -21.368161 -26.255625 -22.695145 -3.631661 -12.327365 -22.734869 -3.631661 -
24.354650 -22.992330 -12.467286 -15.539094 -26.432728 -3.631661 -24.647965 -3.343422 -
3.631661 -13.578888 -3.631661 -4.525268 -3.343422 -24.209008 -17.087359 -10.343070 -
3.343422 -3.631661 -20.534661 -23.560354 -20.050593 -17.290478 -26.291903 -23.287906 -
34.672104 -25.102927 -23.003172 -25.805466 -12.282795 -3.631661 -26.654219 -24.142915 -
14.569407 -30.874745 -24.538742 -23.571580 -32.572318 -24.047185 -24.784038 -34.188139 -
3.272086 -13.480304 -37.118407 -26.453509 -24.944677 -33.819788 -28.971130 -26.514555 -
27.666571 -24.873331 -27.819394 -20.901631 -28.830554 -20.469916 -38.283182 -24.096163 -
24.637360 -35.337634 -3.272086 -4.525268 -31.508712 -19.431228 -18.741605 -30.433457 -
20.388829 -26.090129 -33.861822 -27.980825 -24.202745 -21.180099 -12.260646 -18.754308 -
24.169934 -3.631661 -3.631661 -3.343422 -22.943831 -20.610123 -17.928104 -22.765233 -
17.051654 -26.814191 -28.125062 -3.631661 -34.654771 -16.136195 -12.751702 -46.024858 -
13.730669 -19.956399 -23.471866 -26.071046 -3.631661 -20.431032 -26.600700 -10.549956 -
40.223552 -29.964680 -17.258251 -33.593219 -14.238127 -11.559300 -21.893131 -21.759037 -
10.527759 -29.128393 -3.631661 -3.272086 -3.272086 -30.729039 -12.669624 -30.383455 -
15.161751 -10.952925 -28.250701 -3.631661 -4.525268 -13.195205 -3.343422 -3.343422 -
28.667336 -13.507185 -15.075040 -20.573042 -21.393783 -23.380050 -20.061543 -3.629444 -
3.269800 -3.269800 -28.771856 -10.434104 -19.998558 -28.121061 -7.003269 -10.696688 -
3.269800 -3.269800 -36.179539 -4.522476 -17.458633 -21.782018 -3.269800 -4.522476 -
34.111010 -3.269800 -3.269800 -44.723020 -3.118899 -2.053259 -5.122858 -28.253343 -
10.944623 -17.430587 -3.269800 -10.931620 -21.065792 -23.438190 -7.131083 -7.241008 -
17.582511 -8.569028 -24.047956 -15.853684 -14.712525 -26.042389 -15.720197 -3.629443 -
23.907867 -23.026329 -25.671116 -26.154391 -24.411484 -24.651871 -15.599450 -3.269800 -
4.522475 -21.681342 -7.752980 -7.114846 -20.846487 -15.503034 -15.921867 -18.536712 -
19.461142 -26.880473 -25.195558 -8.476079 -8.087275 -3.341057 -10.975502 -7.211123 -
10.297044 -19.023406 -3.629443 -31.599646 -14.601132 -3.629443 -16.943229 -3.269800 -
9.071764 -12.746370 -17.263407 -8.476079 -21.053460 -28.865660 -25.438504 -24.983143
tr3 -12.187902 -3.229575 -23.263834 -32.222646 -3.229575 -11.363611 -24.924322 -
3.229575 -9.407462 -25.091559 -10.201073 -31.834539 -2.870950 -4.029554 -16.653030 -
1.136585 -1.136585 -1.513334 -11.090445 -8.762934 -11.911348 -20.113153 -16.237376 -
30.314226 -16.930568 -7.634101 -18.139319 -12.850034 -8.762934 -37.516359 -2.870950 -
11.359602 -34.188806 -9.132562 -12.911938 -35.442880 -7.550624 -11.022551 -21.204675 -
16.278811 -6.650894 -27.879392 -8.961663 -9.056109 -27.681702 -14.652477 -14.634743 -
22.671092 -12.926506 -23.453638 -21.517765 -7.915483 -3.229575 -25.073489 -32.913069 -
26.636938 -23.987641 -23.814801 -9.350110 -6.692353 -8.303609 -8.265652 -23.809238 -
24.189312 -9.408609 -11.525381 -14.014346 -7.744636 -28.087577 -11.410728 -6.050470 -
19.031574 -2.870950 -4.029554 -11.004385 -7.451626 -12.426761 -24.440097 -17.746265 -
3.229575 -9.542450 -4.029554 -3.229575 -2.870950 -4.029554 -3.229575 -2.926786 -6.440764 -
14.720227 -36.627011 -4.029554 -2.926786 -15.334908 -2.870950 -2.870950 -30.201740 -
22.079068 -18.933991 -16.549248 -24.365141 -7.876320 -6.508271 -2.870950 -8.469643 -
24.746259 -14.989620 -2.926786 -19.763931 -16.800102 -3.229575 -20.547057 -24.097945 -
17.925793 -32.279287 -31.711800 -16.310562 -29.866433 -13.504138 -15.129400 -31.012072 -
3.443022 -3.443022 -1.589392 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 -

9.288925 -8.466265 -13.948596 -2.035577 -6.462583 -10.402529 -2.035577 -7.913921 -
7.339271 -2.069216 -2.402972 -9.905565 -9.296685 -8.424363 -11.803038 -2.035577 -8.855859
-5.820928 -22.077112 -18.576744 -31.630854 -15.218052 -15.814667 -36.367154 -21.689854 -
18.857686 -22.876995 -2.926786 -2.870950 -21.375714 -18.674189 -11.132531 -12.638213 -
24.037799 -11.352981 -23.108842 -3.229575 -12.236695 -10.036083 -3.229575 -2.870950 -
2.870950 -8.676420 -14.972991 -29.789677 -18.195745 -3.229575 -7.535264 -13.421299 -
10.149388 -18.609351 -4.029554 -2.926786 -19.145556 -9.071177 -22.205619 -21.614609 -
17.534674 -12.015950 -23.681861 -17.691902 -3.229575 -11.182835 -16.695205 -3.229575 -
21.406318 -21.584772 -14.049523 -14.845601 -24.007336 -3.229575 -23.399018 -2.926786 -
3.229575 -12.078455 -3.229575 -4.029554 -2.926786 -22.904752 -15.982698 -9.867350 -
2.926786 -3.229575 -15.721164 -15.800125 -14.074716 -15.523873 -21.263323 -20.923569 -
31.389790 -19.493624 -18.719854 -23.151336 -11.766780 -3.229575 -22.043009 -20.917751 -
8.368168 -28.868255 -24.965735 -19.258917 -28.239116 -18.111544 -17.347074 -31.515511 -
2.870950 -10.374001 -36.254225 -25.328330 -20.337418 -33.173945 -25.375842 -24.096940 -
24.395077 -21.342284 -22.527421 -18.460760 -24.166327 -16.732734 -37.298721 -22.912197 -
19.578416 -33.099502 -2.870950 -4.029554 -28.498793 -15.936253 -13.834368 -27.164721 -
14.951916 -23.053924 -32.707193 -19.537436 -18.636106 -17.082146 -13.425847 -18.103795 -
23.125077 -3.229575 -3.229575 -2.926786 -20.172362 -19.675856 -18.904514 -18.565301 -
19.063868 -26.709140 -25.396083 -3.229575 -30.966565 -7.034301 -5.869104 -42.936927 -
6.613273 -11.926303 -18.687984 -20.430072 -3.229575 -17.266065 -21.875444 -9.942851 -
37.225653 -26.213561 -13.833970 -33.971107 -13.179202 -10.943675 -22.680301 -18.822909 -
5.984937 -25.542566 -3.229575 -2.870950 -2.870950 -24.712812 -9.828308 -32.313257 -
13.118136 -8.208878 -28.920974 -3.229575 -4.029554 -12.582824 -2.926786 -2.926786 -
27.360257 -6.213471 -14.110458 -21.129231 -13.740362 -14.960318 -23.340189 -3.226606 -
2.867908 -2.867908 -22.422471 -9.754631 -16.425476 -24.510238 -6.407318 -6.072792 -
2.867908 -2.867908 -33.308264 -4.026087 -14.762918 -25.580006 -2.867908 -4.026087 -
30.741817 -2.867908 -2.867908 -46.578554 -2.658829 -1.697692 -4.066873 -24.265950 -
8.200136 -16.166774 -2.867908 -9.208300 -17.018318 -18.748997 -6.412200 -6.521425 -
17.274190 -6.482473 -22.411412 -14.084862 -12.050762 -22.452496 -12.186591 -3.226606 -
21.061718 -18.323719 -21.731551 -23.278228 -22.860880 -21.216867 -15.493995 -2.867908 -
4.026087 -17.830769 -7.133524 -6.518837 -19.627753 -12.461996 -12.611258 -17.576706 -
15.989657 -21.565887 -22.142530 -7.868483 -7.460821 -2.923671 -9.901113 -6.526456 -
9.259123 -18.264540 -3.226605 -29.779355 -11.320846 -3.226605 -16.570837 -2.867908 -
8.440168 -11.213354 -14.765151 -7.868483 -21.658293 -24.097791 -23.135302 -22.256915
tr4 -12.187916 -3.229582 -23.263846 -32.222635 -3.229582 -11.363600 -24.924324 -
3.229582 -9.407469 -25.091565 -10.201079 -31.834539 -2.870957 -4.029562 -16.653027 -
1.136585 -1.136585 -1.513334 -11.090437 -8.762941 -11.911340 -20.113164 -16.237369 -
30.314220 -16.930576 -7.634109 -18.139314 -12.850046 -8.762941 -37.516323 -2.870957 -
11.359608 -34.188804 -9.132568 -12.911943 -35.442856 -7.550631 -11.022559 -21.204681 -
16.278821 -6.650883 -27.879390 -8.961672 -9.056101 -27.681698 -14.652465 -14.634734 -
22.671105 -12.926499 -23.453633 -21.517755 -7.915490 -3.229582 -25.073476 -32.913080 -
26.636932 -23.987627 -23.814785 -9.350117 -6.692359 -8.303616 -8.265660 -23.809229 -
24.189310 -9.408616 -11.525390 -14.014341 -7.744644 -28.087569 -11.410717 -6.050474 -
19.031569 -2.870957 -4.029562 -11.004392 -7.451632 -12.426749 -24.440084 -17.746270 -

3.229582 -9.542456 -4.029562 -3.229582 -2.870957 -4.029562 -3.229582 -2.926793 -6.440770 -14.720223 -36.626985 -4.029562 -2.926793 -15.334900 -2.870957 -2.870957 -30.201741 -22.079065 -18.934005 -16.549236 -24.365145 -7.876329 -6.508259 -2.870957 -8.469646 -24.746262 -14.989634 -2.926793 -19.763938 -16.800095 -3.229582 -20.547056 -24.097957 -17.925786 -32.279282 -31.711809 -16.310581 -29.866441 -13.504142 -15.129398 -31.012062 -3.443021 -3.443021 -1.589394 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 -9.288924 -8.466259 -13.948589 -2.035579 -6.462580 -10.402528 -2.035579 -7.913916 -7.339269 -2.069218 -2.402974 -9.905563 -9.296678 -8.424360 -11.803034 -2.035579 -8.855856 -5.820929 -22.077105 -18.576733 -31.630831 -15.218058 -15.814660 -36.367129 -21.689855 -18.857684 -22.876984 -2.926793 -2.870957 -21.375702 -18.674183 -11.132522 -12.638204 -24.037816 -11.352976 -23.108835 -3.229582 -12.236685 -10.036075 -3.229582 -2.870957 -2.870957 -8.676412 -14.972980 -29.789647 -18.195758 -3.229582 -7.535251 -13.421309 -10.149395 -18.609366 -4.029562 -2.926793 -19.145572 -9.071169 -22.205612 -21.614603 -17.534660 -12.015958 -23.681856 -17.691896 -3.229582 -11.182847 -16.695202 -3.229582 -21.406305 -21.584763 -14.049529 -14.845611 -24.007325 -3.229582 -23.399019 -2.926793 -3.229582 -12.078446 -3.229582 -4.029562 -2.926793 -22.904748 -15.982690 -9.867355 -2.926793 -3.229582 -15.721151 -15.800132 -14.074721 -15.523867 -21.263331 -20.923555 -31.389786 -19.493616 -18.719848 -23.151323 -11.766759 -3.229582 -22.043022 -20.917743 -8.368158 -28.868254 -24.965741 -19.258927 -28.239132 -18.111544 -17.347067 -31.515498 -2.870957 -10.373991 -36.254197 -25.328338 -20.337426 -33.173922 -25.375852 -24.096955 -24.395079 -21.342278 -22.527414 -18.460747 -24.166311 -16.732713 -37.298692 -22.912204 -19.578412 -33.099489 -2.870957 -4.029562 -28.498782 -15.936246 -13.834362 -27.164708 -14.951912 -23.053914 -32.707198 -19.537455 -18.636121 -17.082155 -13.425860 -18.103797 -23.125077 -3.229582 -3.229582 -2.926793 -20.172378 -19.675864 -18.904523 -18.565292 -19.063858 -26.709131 -25.396096 -3.229582 -30.966565 -7.034306 -5.869109 -42.936911 -6.613279 -11.926312 -18.687975 -20.430066 -3.229582 -17.266044 -21.875461 -9.942858 -37.225656 -26.213560 -13.833984 -33.971091 -13.179213 -10.943683 -22.680293 -18.822892 -5.984928 -25.542576 -3.229582 -2.870957 -2.870957 -24.712805 -9.828298 -32.313235 -13.118129 -8.208882 -28.920960 -3.229582 -4.029562 -12.582830 -2.926793 -2.926793 -27.360243 -6.213476 -14.110468 -21.129227 -13.740370 -14.960313 -23.340183 -3.226613 -2.867915 -2.867915 -22.422465 -9.754639 -16.425449 -24.510240 -6.407325 -6.072782 -2.867915 -2.867915 -33.308247 -4.026096 -14.762926 -25.580002 -2.867915 -4.026096 -30.741822 -2.867915 -2.867915 -46.578545 -2.658831 -1.697694 -4.066873 -24.265961 -8.200140 -16.166764 -2.867915 -9.208289 -17.018309 -18.748996 -6.412206 -6.521432 -17.274181 -6.482461 -22.411419 -14.084855 -12.050755 -22.452484 -12.186600 -3.226613 -21.061713 -18.323727 -21.731553 -23.278244 -22.860875 -21.216865 -15.493983 -2.867915 -4.026095 -17.830766 -7.133531 -6.518843 -19.627746 -12.461985 -12.611251 -17.576678 -15.989649 -21.565890 -22.142525 -7.868489 -7.460827 -2.923679 -9.901122 -6.526462 -9.259130 -18.264534 -3.226612 -29.779349 -11.320854 -3.226612 -16.570816 -2.867915 -8.440175 -11.213345 -14.765139 -7.868489 -21.658298 -24.097789 -23.135297 -22.256926

tr5 -11.997624 -3.190503 -23.061758 -32.413057 -3.190503 -11.229991 -25.153875 -3.190503 -9.316783 -25.428410 -10.335820 -32.062612 -2.833828 -3.981910 -16.535036 -1.136585 -1.136585 -1.513334 -11.418249 -8.642433 -11.796392 -19.924916 -16.389686 -30.245096 -16.894984 -7.516729 -18.331058 -12.685155 -8.642433 -37.506300 -2.833828 -

11.427627 -34.105640 -9.501079 -12.346236 -35.576870 -7.439203 -11.401462 -21.556018 -
16.356105 -7.003295 -27.965708 -8.874241 -9.070143 -27.714650 -14.148125 -14.718470 -
22.762851 -13.043635 -23.634186 -22.096276 -7.794751 -3.190503 -25.037413 -32.886270 -
26.848899 -24.126251 -23.955050 -9.258036 -6.609577 -8.211808 -8.169797 -23.980474 -
24.462354 -9.311583 -11.380714 -13.954525 -7.673606 -28.157243 -9.997186 -6.208027 -
18.851773 -2.833828 -3.981910 -10.918123 -7.292384 -11.645065 -24.483002 -17.931555 -
3.190503 -10.494016 -3.981910 -3.190503 -2.833828 -3.981910 -3.190503 -2.887903 -6.356792
-17.472394 -36.493375 -3.981910 -2.887903 -15.102108 -2.833828 -2.833828 -30.067239 -
20.794014 -19.239642 -17.212572 -24.105606 -7.781418 -6.445093 -2.833828 -7.964334 -
25.109699 -15.209747 -2.887903 -19.840993 -14.253744 -3.190503 -19.485158 -22.530328 -
17.816437 -32.200573 -31.546172 -17.711159 -29.688103 -14.106824 -16.338452 -31.054587 -
3.435771 -3.435771 -1.595305 -0.000000 -0.000000 -0.000000 -0.000000 -0.000000 -0.000000 -
9.275928 -8.437646 -13.874870 -2.052130 -6.446813 -10.391465 -2.052130 -7.890197 -
7.320862 -2.086341 -2.419781 -9.860958 -9.252837 -8.405710 -11.777857 -2.052130 -8.823742
-5.830971 -22.279436 -18.987839 -31.739703 -15.407149 -15.944734 -36.323017 -21.728874 -
19.532868 -23.207266 -2.887903 -2.833828 -21.483509 -15.012556 -11.546787 -12.442772 -
24.144239 -12.173304 -23.332874 -3.190503 -12.597287 -10.330838 -3.190503 -2.833828 -
2.833828 -9.454262 -14.050531 -29.851483 -18.476888 -3.190503 -7.471177 -13.168416 -
9.944957 -18.373143 -3.981910 -2.887903 -19.480584 -9.200568 -21.439191 -21.620864 -
18.151838 -12.579003 -24.023123 -14.281692 -3.190503 -11.205685 -17.058227 -3.190503 -
21.530928 -18.733128 -14.098327 -14.626329 -23.653550 -3.190503 -23.528772 -2.887903 -
3.190503 -12.049259 -3.190503 -3.981910 -2.887903 -21.393555 -16.387327 -9.895740 -
2.887903 -3.190503 -16.012890 -16.423956 -13.622924 -15.342500 -21.479986 -20.382206 -
31.493169 -19.539651 -18.616781 -23.404228 -12.033441 -3.190503 -21.850274 -19.414593 -
8.816435 -28.758371 -25.194696 -19.567283 -28.513087 -18.603548 -17.950970 -31.457503 -
2.833828 -10.236597 -36.186540 -23.081288 -20.045797 -33.280229 -25.703932 -24.004392 -
24.791536 -21.980071 -20.346517 -18.477617 -24.269727 -16.782110 -37.387122 -22.841705 -
19.841577 -33.189868 -2.833828 -3.981910 -28.590308 -15.964210 -14.043065 -27.105514 -
15.481209 -22.699089 -32.645439 -19.857971 -18.397950 -17.047331 -8.662696 -15.484396 -
23.137744 -3.190503 -3.190503 -2.887903 -20.161937 -17.105228 -19.387353 -18.719277 -
18.970107 -26.796728 -25.158072 -3.190503 -30.513261 -7.547897 -6.359436 -43.071639 -
7.101170 -12.396476 -19.083820 -20.632970 -3.190503 -17.439505 -22.301772 -10.031502 -
37.110511 -25.770026 -13.781797 -34.083961 -12.937004 -10.769046 -22.896792 -18.945329 -
6.332908 -25.631475 -3.190503 -2.833828 -2.833828 -24.970382 -7.162941 -32.260695 -
10.253725 -7.901182 -28.841265 -3.190503 -3.981910 -12.674778 -2.887903 -2.887903 -
27.473539 -6.679488 -13.890600 -20.899433 -14.815415 -15.909540 -20.910788 -3.187401 -
2.830663 -2.830663 -21.970167 -9.658081 -16.899593 -23.081385 -6.297324 -6.424595 -
2.830663 -2.830663 -33.596522 -3.978301 -14.786834 -25.360880 -2.830663 -3.978301 -
30.821031 -2.830663 -2.830663 -46.209436 -2.628803 -1.674572 -4.259005 -24.108238 -
7.893928 -16.190163 -2.830663 -9.078062 -17.048334 -19.309909 -6.313287 -6.427454 -
14.832201 -6.414797 -22.363663 -13.449128 -12.005208 -22.556208 -12.059444 -3.187401 -
21.486303 -18.112762 -22.088170 -23.078409 -20.695837 -19.856951 -12.605952 -2.830663 -
3.978301 -18.029973 -7.016406 -6.412857 -19.578672 -10.139517 -9.967810 -18.060692 -
14.575949 -20.552982 -22.015875 -7.755174 -7.342521 -2.884662 -9.716623 -6.436443 -

9.108468 -18.180595 -3.187400 -29.825329 -11.138877 -3.187400 -16.664356 -2.830662 -
8.319205 -11.095341 -14.598496 -7.755174 -19.294486 -24.012214 -23.011148 -22.456078
tr6 -12.126595 -3.204719 -23.128384 -32.342539 -3.204719 -11.324640 -25.087036 -
3.204719 -9.346811 -25.339453 -10.549699 -32.068649 -2.847058 -3.999432 -16.461525 -
1.136585 -1.136585 -1.513334 -11.155967 -8.698093 -11.831436 -19.990633 -16.187385 -
30.142065 -16.923640 -7.570899 -18.178589 -12.802420 -8.698093 -37.483592 -2.847058 -
11.401306 -33.967361 -9.282269 -13.137773 -35.554250 -7.490482 -11.187596 -21.331406 -
16.739443 -6.721861 -27.990287 -8.897960 -9.212001 -27.757131 -14.993732 -15.030512 -
22.907524 -12.781955 -23.347857 -22.069934 -7.834795 -3.204719 -24.968198 -32.931255 -
26.826256 -24.143934 -23.719088 -9.288794 -6.654021 -8.242481 -8.202544 -24.361785 -
24.394541 -9.344714 -11.444898 -13.922463 -7.687049 -28.079502 -11.731841 -6.133957 -
18.959889 -2.847058 -3.999432 -10.950442 -7.288740 -12.737060 -24.452873 -17.641345 -
3.204719 -9.783984 -3.999432 -3.204719 -2.847058 -3.999432 -3.204719 -2.901846 -6.401977 -
17.443779 -36.356831 -3.999432 -2.901846 -15.107337 -2.847058 -2.847058 -30.128859 -
22.228321 -19.617467 -17.094715 -23.997190 -7.809352 -6.460051 -2.847058 -8.786677 -
25.196587 -15.078676 -2.901846 -19.775176 -14.367602 -3.204719 -19.303907 -22.361331 -
17.823200 -32.263676 -31.572790 -18.661315 -29.728868 -13.857984 -16.136270 -31.062123 -
3.443570 -3.443570 -1.588951 -0.000000 -0.000000 -0.000000 -0.000000 -0.000000 -0.000000 -
9.279741 -8.463584 -13.892709 -2.044645 -6.457592 -10.368816 -2.044645 -7.906978 -
7.317283 -2.078598 -2.412186 -9.876557 -9.269432 -8.415360 -11.787672 -2.044645 -8.827688
-5.820636 -22.445411 -18.812823 -31.671848 -15.522863 -16.004006 -36.237484 -21.571590 -
19.277849 -22.906990 -2.901846 -2.847058 -21.441054 -16.134605 -11.455693 -12.500144 -
24.086865 -11.863050 -23.195043 -3.204719 -12.302228 -10.083614 -3.204719 -2.847058 -
2.847058 -8.980813 -15.307560 -30.057881 -18.240136 -3.204719 -7.486348 -13.211793 -
9.971961 -18.401534 -3.999432 -2.901846 -19.321443 -9.333010 -22.458003 -21.686065 -
17.913476 -12.392766 -24.591099 -17.651645 -3.204719 -11.323284 -16.912179 -3.204719 -
21.411029 -21.692090 -14.058706 -14.705130 -24.158870 -3.204719 -23.444705 -2.901846 -
3.204719 -11.998091 -3.204719 -3.999432 -2.901846 -20.282892 -16.256204 -9.806912 -
2.901846 -3.204719 -15.785081 -16.115041 -13.540567 -15.410857 -21.478862 -21.082879 -
31.451246 -19.405273 -19.017195 -23.296485 -11.874051 -3.204719 -21.981931 -19.077809 -
8.666767 -28.772057 -25.145319 -19.491561 -28.474167 -18.478746 -17.594723 -31.372051 -
2.847058 -10.278716 -36.149352 -23.024251 -20.324846 -33.205205 -25.556256 -24.028828 -
24.518774 -21.177067 -20.135770 -18.247232 -24.126187 -16.777578 -37.375401 -22.839361 -
19.946215 -33.025470 -2.847058 -3.999432 -28.422856 -16.107555 -13.793966 -26.943222 -
15.438258 -22.349866 -32.644358 -19.732420 -18.466425 -17.401039 -8.796126 -15.620979 -
23.159046 -3.204719 -3.204719 -2.901846 -20.191924 -17.463733 -19.334904 -18.608712 -
19.064420 -26.753309 -25.212111 -3.204719 -30.368842 -7.259704 -6.083652 -42.931754 -
6.823001 -12.125612 -18.807501 -20.517724 -3.204719 -17.368713 -22.123570 -9.992190 -
37.221480 -25.699578 -13.933910 -34.114020 -13.024627 -10.827597 -22.759689 -18.751297 -
6.057719 -25.469712 -3.204719 -2.847058 -2.847058 -25.002048 -10.035266 -32.069742 -
10.107505 -8.515112 -28.830804 -3.204719 -3.999432 -12.626291 -2.901846 -2.901846 -
27.455497 -6.420006 -13.908704 -21.015819 -14.292239 -15.427513 -21.011710 -3.201668 -
2.843939 -2.843939 -21.778565 -9.691067 -16.744238 -23.149907 -6.333213 -6.147837 -
2.843939 -2.843939 -33.593146 -3.995878 -14.769775 -25.409099 -2.843939 -3.995878 -

30.813598 -2.843939 -2.843939 -46.302075 -2.621587 -1.668563 -4.273947 -24.281877 -
8.506094 -16.153157 -2.843939 -9.110733 -17.116895 -18.968565 -6.364111 -6.475999 -
14.976975 -6.432318 -22.373360 -12.639059 -11.352629 -22.477059 -12.153690 -3.201668 -
21.246394 -17.882945 -21.920958 -23.147971 -20.634001 -19.709562 -12.743363 -2.843938 -
3.995878 -17.826150 -7.055321 -6.446708 -19.525035 -12.602631 -12.571483 -18.042486 -
16.134344 -21.406652 -22.034663 -7.792065 -7.380956 -2.898653 -9.796298 -6.484427 -
9.173883 -18.156247 -3.201667 -29.678408 -11.202153 -3.201667 -16.532218 -2.843938 -
8.359341 -11.125300 -14.961666 -7.792065 -19.450216 -24.003162 -23.001784 -22.423596
tr7 -11.991678 -3.178866 -23.020912 -32.357236 -3.178866 -11.223570 -25.186255 -
3.178866 -9.297107 -25.176646 -10.184086 -32.148959 -2.822486 -3.967979 -16.560951 -
1.136585 -1.136585 -1.513334 -11.576606 -8.613594 -11.773329 -19.888593 -16.496119 -
30.291881 -16.913766 -7.488656 -17.694794 -12.674547 -8.613594 -37.545527 -2.822486 -
11.464651 -34.185681 -9.540659 -13.065478 -35.661335 -7.412048 -11.441763 -21.508537 -
16.733104 -7.178887 -27.994736 -8.869567 -9.296233 -27.733630 -14.470426 -14.893477 -
23.027268 -13.185243 -23.763993 -21.775515 -7.767700 -3.178866 -25.044393 -32.857838 -
26.816827 -24.132606 -24.126409 -9.238077 -6.589819 -8.191902 -8.149342 -23.970070 -
24.174442 -9.290907 -11.349205 -13.971194 -7.673257 -28.111437 -10.355896 -6.264314 -
18.821861 -2.822486 -3.967979 -10.898010 -7.287900 -11.529685 -24.588931 -18.134404 -
3.178866 -10.600097 -3.967979 -3.178866 -2.822486 -3.967979 -3.178866 -2.876122 -6.336881
-16.182486 -36.516940 -3.967979 -2.876122 -15.083200 -2.822486 -2.822486 -30.032041 -
20.835995 -19.216273 -16.875920 -24.144419 -7.776133 -6.428431 -2.822486 -8.994621 -
25.296544 -15.655444 -2.876122 -19.909750 -14.275641 -3.178866 -19.545385 -22.601170 -
17.864494 -32.172078 -31.563604 -17.760106 -29.613233 -14.133255 -16.453887 -30.814154 -
3.432995 -3.432995 -1.597600 -0.000000 -0.000000 -0.000000 -0.000000 -0.000000 -0.000000 -
9.273789 -8.433769 -13.879099 -2.053480 -6.446641 -10.401224 -2.053480 -7.889052 -
7.314991 -2.087736 -2.421142 -9.861938 -9.250147 -8.406699 -11.776812 -2.053480 -8.824993
-5.825326 -22.306842 -18.893355 -31.697912 -15.812824 -14.974625 -36.376805 -21.643932 -
18.929198 -23.239277 -2.876122 -2.822486 -21.087159 -15.118043 -8.934149 -12.410528 -
24.116047 -11.156905 -23.337949 -3.178866 -12.779920 -10.481966 -3.178866 -2.822486 -
2.822486 -9.338262 -14.188059 -30.020036 -18.542880 -3.178866 -7.454193 -13.147129 -
9.929545 -18.350118 -3.967979 -2.876122 -19.521783 -7.522748 -21.063062 -21.668171 -
18.118193 -12.542274 -24.031447 -14.500207 -3.178866 -11.243131 -17.066819 -3.178866 -
21.497559 -18.839911 -14.132625 -14.584004 -23.610612 -3.178866 -23.648447 -2.876122 -
3.178866 -12.049743 -3.178866 -3.967979 -2.876122 -21.374829 -14.383425 -9.957448 -
2.876122 -3.178866 -15.783586 -16.081135 -13.531721 -15.303727 -21.656167 -20.460779 -
31.531922 -19.740640 -19.027755 -23.128385 -12.185696 -3.178866 -21.821694 -19.288036 -
7.133818 -28.724185 -25.393690 -19.836936 -28.495482 -17.951008 -16.608405 -31.500185 -
2.822486 -10.216998 -36.055557 -23.158685 -20.349686 -33.341297 -25.565592 -24.028957 -
24.921413 -22.012228 -20.367650 -18.147249 -24.375463 -16.925945 -37.404117 -22.844765 -
19.714410 -33.171080 -2.822486 -3.967979 -28.578079 -15.843071 -14.180317 -26.999653 -
15.350133 -22.446387 -32.610360 -19.906978 -18.320747 -17.454623 -8.696685 -15.484666 -
23.136089 -3.178866 -3.178866 -2.876122 -20.073031 -17.093968 -19.306769 -18.897536 -
19.052410 -26.834380 -25.102304 -3.178866 -30.660365 -7.627903 -6.434318 -43.133310 -
7.179672 -12.477056 -19.119624 -20.783544 -3.178866 -17.609784 -22.305452 -10.084661 -

37.068496 -25.702398 -13.805201 -34.132506 -12.944350 -10.731746 -22.874106 -19.081386 -
6.500719 -25.643257 -3.178866 -2.822486 -2.822486 -24.780427 -7.293599 -32.301253 -
10.438048 -8.803018 -28.839040 -3.178866 -3.967979 -12.719398 -2.876122 -2.876122 -
27.523409 -6.746192 -13.871605 -20.867600 -14.872549 -16.048752 -20.927084 -3.175731 -
2.819290 -2.819290 -22.070995 -9.637471 -17.040189 -23.188365 -6.272502 -6.587954 -
2.819289 -2.819289 -33.407041 -3.964335 -14.830787 -25.320979 -2.819289 -3.964335 -
30.672986 -2.819289 -2.819289 -46.210194 -2.628011 -1.673945 -4.264071 -24.440116 -
8.793648 -16.223242 -2.819289 -9.047756 -17.257259 -19.378968 -6.290483 -6.405217 -
14.843191 -6.397446 -22.392902 -13.229055 -11.330476 -22.272046 -12.015127 -3.175731 -
21.449076 -18.255724 -21.965448 -23.026545 -20.320408 -20.073544 -12.626295 -2.819289 -
3.964335 -18.127278 -6.990109 -6.388534 -19.549034 -10.239777 -10.074002 -18.202959 -
14.654942 -20.586679 -21.989812 -7.729670 -7.315962 -2.872849 -9.675689 -6.415607 -
9.073319 -18.186237 -3.175730 -29.718045 -11.110779 -3.175730 -16.792515 -2.819289 -
8.292088 -11.065493 -15.032123 -7.729670 -19.299447 -23.934247 -23.051808 -22.451629
tr8 -12.187909 -3.229577 -23.263825 -32.222636 -3.229577 -11.363602 -24.924322 -
3.229577 -9.407470 -25.091560 -10.201071 -31.834535 -2.870953 -4.029557 -16.653037 -
1.136585 -1.136585 -1.513334 -11.090438 -8.762942 -11.911335 -20.113155 -16.237362 -
30.314210 -16.930573 -7.634110 -18.139334 -12.850038 -8.762942 -37.516329 -2.870953 -
11.359613 -34.188805 -9.132567 -12.911956 -35.442858 -7.550632 -11.022569 -21.204679 -
16.278828 -6.650890 -27.879403 -8.961685 -9.056110 -27.681693 -14.652492 -14.634762 -
22.671119 -12.926494 -23.453640 -21.517764 -7.915480 -3.229577 -25.073477 -32.913065 -
26.636927 -23.987640 -23.814795 -9.350118 -6.692351 -8.303617 -8.265660 -23.809229 -
24.189319 -9.408617 -11.525385 -14.014342 -7.744656 -28.087559 -11.410738 -6.050475 -
19.031563 -2.870953 -4.029557 -11.004383 -7.451624 -12.426765 -24.440086 -17.746266 -
3.229577 -9.542456 -4.029557 -3.229577 -2.870953 -4.029557 -3.229577 -2.926788 -6.440761 -
14.720218 -36.626984 -4.029557 -2.926788 -15.334891 -2.870953 -2.870953 -30.201740 -
22.079069 -18.933981 -16.549235 -24.365140 -7.876341 -6.508258 -2.870953 -8.469657 -
24.746274 -14.989634 -2.926788 -19.763949 -16.800101 -3.229577 -20.547064 -24.097952 -
17.925789 -32.279273 -31.711806 -16.310555 -29.866431 -13.504137 -15.129391 -31.012059 -
3.443020 -3.443020 -1.589394 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 -
9.288924 -8.466259 -13.948589 -2.035580 -6.462580 -10.402528 -2.035580 -7.913916 -
7.339269 -2.069219 -2.402975 -9.905563 -9.296677 -8.424360 -11.803033 -2.035580 -8.855856
-5.820930 -22.077125 -18.576737 -31.630827 -15.218071 -15.814665 -36.367127 -21.689853 -
18.857681 -22.876979 -2.926788 -2.870953 -21.375701 -18.674189 -11.132537 -12.638195 -
24.037824 -11.352984 -23.108844 -3.229577 -12.236697 -10.036086 -3.229577 -2.870953 -
2.870953 -8.676426 -14.972998 -29.789659 -18.195759 -3.229577 -7.535251 -13.421293 -
10.149382 -18.609348 -4.029557 -2.926788 -19.145576 -9.071182 -22.205629 -21.614610 -
17.534662 -12.015954 -23.681857 -17.691916 -3.229577 -11.182844 -16.695219 -3.229577 -
21.406309 -21.584787 -14.049521 -14.845593 -24.007330 -3.229577 -23.399010 -2.926788 -
3.229577 -12.078451 -3.229577 -4.029557 -2.926788 -22.904750 -15.982711 -9.867351 -
2.926788 -3.229577 -15.721148 -15.800136 -14.074715 -15.523867 -21.263325 -20.923566 -
31.389775 -19.493622 -18.719869 -23.151335 -11.766777 -3.229577 -22.043014 -20.917750 -
8.368173 -28.868255 -24.965746 -19.258933 -28.239118 -18.111549 -17.347076 -31.515497 -
2.870953 -10.373986 -36.254211 -25.328344 -20.337425 -33.173926 -25.375835 -24.096944 -

24.395074 -21.342280 -22.527428 -18.460752 -24.166317 -16.732731 -37.298723 -22.912196 -
19.578418 -33.099502 -2.870953 -4.029557 -28.498798 -15.936237 -13.834354 -27.164724 -
14.951918 -23.053922 -32.707175 -19.537448 -18.636118 -17.082157 -13.425851 -18.103791 -
23.125082 -3.229577 -3.229577 -2.926788 -20.172351 -19.675851 -18.904507 -18.565287 -
19.063858 -26.709132 -25.396087 -3.229577 -30.966573 -7.034312 -5.869115 -42.936913 -
6.613285 -11.926322 -18.687987 -20.430083 -3.229577 -17.266061 -21.875454 -9.942862 -
37.225653 -26.213579 -13.833975 -33.971106 -13.179206 -10.943669 -22.680302 -18.822892 -
5.984934 -25.542564 -3.229577 -2.870953 -2.870953 -24.712810 -9.828316 -32.313251 -
13.118135 -8.208891 -28.920965 -3.229577 -4.029557 -12.582826 -2.926788 -2.926788 -
27.360238 -6.213481 -14.110455 -21.129218 -13.740366 -14.960308 -23.340189 -3.226609 -
2.867911 -2.867911 -22.422465 -9.754639 -16.425450 -24.510250 -6.407316 -6.072788 -
2.867910 -2.867910 -33.308259 -4.026090 -14.762918 -25.579992 -2.867910 -4.026090 -
30.741823 -2.867910 -2.867910 -46.578552 -2.658831 -1.697694 -4.066876 -24.265960 -
8.200150 -16.166761 -2.867910 -9.208284 -17.018311 -18.749001 -6.412197 -6.521423 -
17.274180 -6.482461 -22.411413 -14.084859 -12.050763 -22.452491 -12.186586 -3.226608 -
21.061716 -18.323721 -21.731565 -23.278221 -22.860886 -21.216870 -15.493984 -2.867910 -
4.026090 -17.830769 -7.133522 -6.518835 -19.627749 -12.462000 -12.611262 -17.576687 -
15.989676 -21.565902 -22.142515 -7.868480 -7.460818 -2.923674 -9.901107 -6.526454 -
9.259117 -18.264539 -3.226608 -29.779364 -11.320840 -3.226608 -16.570825 -2.867910 -
8.440165 -11.213339 -14.765145 -7.868480 -21.658289 -24.097764 -23.135294 -22.256923
tr9 -12.060487 -3.185863 -23.043252 -32.366099 -3.185863 -11.290480 -25.177781 -
3.185863 -9.309413 -25.057406 -10.336663 -32.098346 -2.829081 -3.976514 -16.532425 -
1.136585 -1.136585 -1.513334 -11.574420 -8.633487 -11.799850 -19.915001 -16.489621 -
30.260012 -16.776360 -7.508082 -17.464875 -12.739081 -8.633487 -37.547712 -2.829081 -
11.287936 -34.081608 -9.567557 -13.336925 -35.658457 -7.430259 -11.476887 -21.462124 -
16.822999 -7.170749 -27.908799 -8.863631 -9.536818 -27.671122 -15.028804 -15.306544 -
23.023577 -13.161503 -23.692760 -21.718335 -7.785430 -3.185863 -25.044208 -32.801078 -
26.802919 -24.062658 -23.966967 -9.250723 -6.611531 -8.204515 -8.162799 -24.189183 -
24.048279 -9.304509 -11.378368 -13.974135 -7.666943 -28.018317 -11.425737 -6.065154 -
18.894620 -2.829081 -3.976514 -10.912140 -7.271299 -11.853285 -24.554501 -17.948516 -
3.185863 -10.344007 -3.976514 -3.185863 -2.829081 -3.976514 -3.185863 -2.883045 -6.358946
-15.912869 -36.388871 -3.976514 -2.883045 -15.078285 -2.829081 -2.829081 -30.067080 -
21.657609 -19.462294 -16.711891 -24.146353 -7.771815 -6.445577 -2.829081 -9.099374 -
25.346577 -15.481020 -2.883045 -19.750733 -14.368811 -3.185863 -19.581713 -22.736369 -
17.890350 -32.206988 -31.657145 -18.548343 -29.630855 -14.113641 -16.535449 -30.789553 -
3.437271 -3.437271 -1.594072 -0.000000 -0.000000 -0.000000 -0.000000 -0.000000 -0.000000 -
9.277459 -8.446770 -13.880309 -2.050955 -6.452361 -10.381686 -2.050955 -7.898019 -
7.315120 -2.085123 -2.418573 -9.866796 -9.256690 -8.411194 -11.780362 -2.050955 -8.823915
-5.824650 -22.357352 -18.755165 -31.709535 -15.742639 -14.747511 -36.324798 -21.392269 -
18.634002 -23.124434 -2.883045 -2.829081 -20.979082 -15.981397 -8.513528 -12.442718 -
24.164701 -10.806038 -23.128568 -3.185863 -12.772731 -10.475151 -3.185863 -2.829081 -
2.829081 -9.223705 -15.015183 -30.115048 -18.594615 -3.185863 -7.471662 -13.151923 -
9.928646 -18.347004 -3.976514 -2.883045 -19.550637 -7.176550 -21.404655 -21.557237 -
18.098585 -12.536401 -24.389327 -16.789999 -3.185863 -11.324553 -17.128887 -3.185863 -

21.483338 -20.651363 -13.931355 -14.620376 -23.842505 -3.185863 -23.627542 -2.883045 -
3.185863 -12.044736 -3.185863 -3.976514 -2.883045 -20.287145 -13.982601 -9.735307 -
2.883045 -3.185863 -15.630634 -15.880154 -13.472050 -15.339961 -21.561951 -20.869455 -
31.542632 -19.662593 -19.256646 -23.049482 -12.062565 -3.185863 -21.897581 -19.225379 -
6.768307 -28.746438 -25.270834 -19.706776 -28.550857 -17.799622 -16.264120 -31.473009 -
2.829081 -10.242635 -36.034910 -22.962707 -20.474020 -33.302076 -25.408119 -23.863580 -
24.828335 -21.547944 -20.370733 -17.932910 -24.312426 -16.974407 -37.410905 -22.708277 -
19.578711 -33.107537 -2.829081 -3.976514 -28.495923 -15.959931 -14.141559 -26.865272 -
15.363142 -22.082605 -32.554897 -20.008014 -18.373908 -17.562348 -8.807657 -15.584172 -
23.190938 -3.185863 -3.185863 -2.883045 -20.064527 -17.350967 -19.183718 -18.856482 -
19.083997 -26.788721 -25.124635 -3.185863 -30.653471 -7.663046 -6.466698 -43.039175 -
7.220370 -12.536112 -19.124406 -20.760016 -3.185863 -17.552285 -22.321078 -9.904753 -
37.129622 -25.692750 -13.901913 -34.133335 -12.981468 -10.755883 -22.676415 -18.953286 -
6.489016 -25.610754 -3.185863 -2.829081 -2.829081 -24.694148 -8.986463 -32.209085 -
10.503238 -8.824343 -28.848895 -3.185863 -3.976514 -12.536324 -2.883045 -2.883045 -
27.511155 -6.772703 -13.862348 -20.937492 -14.805924 -15.983558 -21.011986 -3.182745 -
2.825899 -2.825899 -22.114583 -9.651015 -16.932559 -23.151404 -6.288418 -6.572851 -
2.825899 -2.825899 -33.390309 -3.972888 -14.655633 -25.344163 -2.825899 -3.972888 -
30.696540 -2.825899 -2.825899 -46.064655 -2.625839 -1.672160 -4.270946 -24.488943 -
8.814954 -16.058156 -2.825899 -9.072633 -17.431874 -19.208424 -6.314918 -6.428777 -
14.954396 -6.415385 -22.250484 -12.529021 -10.874755 -22.186921 -12.064997 -3.182744 -
21.361675 -18.224555 -21.974876 -23.054524 -20.268025 -20.136331 -12.737899 -2.825899 -
3.972888 -18.113460 -7.007334 -6.403633 -19.543107 -11.669006 -11.624663 -18.191710 -
15.642876 -21.172899 -22.007865 -7.746016 -7.332977 -2.879787 -9.712639 -6.438795 -
9.104215 -18.196413 -3.182744 -29.673374 -11.135131 -3.182744 -16.680407 -2.825898 -
8.309848 -11.087962 -15.152106 -7.746016 -19.397701 -23.892180 -22.933571 -22.462282
tr10 -11.631125 -3.238885 -23.109789 -31.851195 -3.238885 -11.018475 -24.868115 -
3.238885 -9.421651 -24.581229 -10.588406 -31.568003 -2.880655 -4.040287 -16.395006 -
1.136585 -1.136585 -1.513334 -11.156701 -8.780607 -11.858131 -20.022226 -16.192275 -
30.509105 -17.028227 -7.651691 -18.075767 -12.464330 -8.780607 -37.041695 -2.880655 -
11.421980 -34.964297 -9.208695 -13.077170 -35.359834 -7.569343 -11.153065 -21.110602 -
16.619516 -6.686948 -27.735664 -8.842369 -9.129649 -27.473563 -14.865895 -14.874215 -
22.969024 -12.694068 -23.430714 -21.446121 -7.884658 -3.238885 -25.067340 -32.848287 -
26.814164 -24.167692 -23.940382 -9.364131 -6.726645 -8.317590 -8.278669 -23.967226 -
24.260216 -9.421759 -11.569421 -13.891675 -7.635594 -27.798156 -11.635427 -6.098695 -
18.717862 -2.880655 -4.040287 -11.018180 -7.226863 -12.599958 -24.569538 -19.002220 -
3.238885 -12.385913 -4.040287 -3.238885 -2.880655 -4.040287 -3.238885 -2.936693 -6.475855
-15.949314 -36.839718 -4.040287 -2.936693 -15.157222 -2.880655 -2.880655 -30.369345 -
22.338720 -19.182740 -15.822614 -25.073281 -7.758109 -6.561578 -2.880655 -8.637200 -
24.860972 -15.884560 -2.936693 -19.829567 -16.716675 -3.238885 -21.131086 -23.870914 -
17.195913 -32.265213 -31.311268 -19.423302 -29.879939 -13.110595 -17.432488 -31.049093 -
3.444223 -3.444223 -1.588425 -0.000000 -0.000000 -0.000000 -0.000000 -0.000000 -0.000000 -
9.327372 -8.469315 -13.957200 -2.069437 -6.487812 -10.113633 -2.069437 -7.962759 -
7.392213 -2.104285 -2.437395 -9.954425 -9.225865 -8.423946 -11.872004 -2.069437 -8.717224

-5.954444 -22.220226 -20.555942 -31.740780 -15.418140 -16.116451 -36.414855 -21.397672 -18.550909 -24.197774 -2.936693 -2.880655 -21.449841 -19.999217 -11.296389 -12.595206 -23.632817 -11.067991 -22.569015 -3.238885 -12.284020 -10.075890 -3.238885 -2.880655 -2.880655 -8.818216 -15.216090 -29.442464 -18.163572 -3.238885 -7.590081 -13.210194 -9.966685 -18.464513 -4.040287 -2.936693 -18.645231 -9.220168 -22.413966 -21.333286 -21.874640 -15.954564 -24.345367 -17.912399 -3.238885 -11.567987 -16.485804 -3.238885 -21.434580 -21.910182 -14.049151 -14.866686 -24.198334 -3.238885 -22.992852 -2.936693 -3.238885 -12.123971 -3.238885 -4.040287 -2.936693 -22.676031 -16.170806 -9.680065 -2.936693 -3.238885 -17.768545 -18.500398 -13.349169 -15.479382 -21.510865 -21.124658 -30.955490 -19.270128 -18.998621 -23.229802 -11.816713 -3.238885 -21.650838 -21.049083 -8.524453 -29.021641 -25.152734 -19.422781 -28.095036 -17.386336 -17.410724 -32.049468 -2.880655 -10.373405 -36.318608 -25.025465 -20.418726 -32.956284 -24.624874 -24.050560 -24.226950 -22.227023 -22.748385 -18.545829 -24.172347 -16.780334 -37.062445 -22.899206 -19.272005 -33.168275 -2.880655 -4.040287 -29.809015 -16.189085 -13.957543 -27.113900 -15.469138 -22.859488 -32.596789 -19.203475 -17.966699 -16.940494 -12.413095 -17.927382 -23.387037 -3.238885 -3.238885 -2.936693 -19.639441 -19.015992 -18.384773 -18.642745 -19.077829 -26.514365 -25.330487 -3.238885 -31.334265 -7.157028 -5.987197 -42.519946 -6.732556 -12.068823 -18.848688 -20.474731 -3.238885 -17.336737 -21.307860 -9.997531 -36.832991 -26.565829 -14.166083 -33.972568 -12.502577 -10.918822 -22.144421 -18.836353 -6.018829 -25.824457 -3.238885 -2.880655 -2.880655 -24.932463 -10.057679 -32.989691 -12.489144 -8.367902 -28.911712 -3.238885 -4.040287 -12.638710 -2.936693 -2.936693 -27.206967 -6.326132 -13.893960 -20.713442 -17.927982 -18.471118 -22.972230 -3.235944 -2.877643 -2.877643 -23.588374 -9.767765 -15.778675 -24.627494 -6.378923 -6.107062 -2.877642 -2.877642 -33.185482 -4.036846 -14.765674 -25.571131 -2.877642 -4.036846 -30.545455 -2.877642 -2.877642 -46.699990 -2.772290 -1.785530 -3.723642 -24.191167 -8.359267 -16.233080 -2.877642 -9.225259 -17.139253 -18.173401 -6.441133 -6.551286 -16.886499 -6.534493 -22.457880 -14.138892 -11.873313 -22.487102 -12.229179 -3.235943 -20.760283 -19.055725 -21.725952 -23.162134 -22.831493 -21.259821 -15.021223 -2.877642 -4.036845 -17.871944 -7.104151 -6.491110 -19.591269 -12.672866 -12.874915 -17.309309 -16.174794 -22.936081 -22.382810 -7.838862 -7.429137 -2.933607 -9.891112 -6.560303 -9.243277 -18.239985 -3.235943 -29.762126 -11.241348 -3.235943 -16.434026 -2.877642 -8.409481 -11.268522 -14.937751 -7.838862 -21.285968 -23.199271 -23.111760 -21.930883

5.5 Unrooted Maximum Likelihood Tree (Hybrid)

(Hqm. walsbyi:0.0000000001,Hqm. walsbyi duplicate:0.0000003649,((Har. marismortui copy 2:0.0000000001,(Har. vallismortis copy 2:0.0055638941,Har. sinaiiensis copy 2:0.0000001250)20:0.0000001250)87:0.0795515072,(((Hbt. NRC-1:0.0000000001,(Hbt. salinarium duplicate:0.0000000001,(Hbt. salinarium:0.0000000001,Hbt. NRC-1 duplicate:0.0000000001)19:0.0000000001)25:0.0000000001)100:0.4317896317,((Hfx. mucosum:0.1877868546,(Hbf. haloterrestris:0.0384358825,((Ncc. aibiensis:0.0454414131,(Htg. GSL11:0.0141317077,(Ntm. pallidum:0.0105174656,(Nnm. XA31:0.0123403684,(Ntm. altunense:0.0000001564,Nnm. ajinwuensis:0.0083920538)86:0.0048061719)100:0.0344376179)77:0.0075612823)100:0.06092

09910)96:0.0596699648,(Nrr. tenzan 10:0.0838107930,Hbf.
lacisalsi:0.0518838237)47:0.0195190936)60:0.0276717090)92:0.0776872647)99:0.1648366700
,(Hbl. gomorrense:0.1850772945,(Hsx. carlsbadense:0.4273207332,(Hrd. utahensis
:0.3716899392,(Hmc.
mukohataei:0.1397648596,((Fb19:0.0000000001,Ha9:0.0000000001)92:0.0125251753,((Har.
amylolytica:0.0000000001,Har. hispanica:0.0194993746)87:0.0265230572,((Har.
sinaiiensis:0.0000000001,(Har. californiae:0.0000001450,(Har. californiae
duplicate:0.0000000001,(Har. marismortui:0.0000003074,Har. marismortui
duplicate:0.0000000001)59:0.0026074546)18:0.0000001863)20:0.0000001450)100:0.0360122
288,(Har. vallismortis:0.0442508992,(Har. argentinensis:0.0478429571,Har.
japonica:0.0184763725)93:0.0657149074)45:0.0299139761)34:0.0276216082)24:0.014991522
7)98:0.1916377632)52:0.0555173008)32:0.0436928033)29:0.0004730006)73:0.0877948482)78
:0.1466364170)82:0.1558623965,((Hrr. tebenquichense:0.1281304320,Hrr.
ejinorense:0.0277595040)71:0.0465497399,((Eb13:0.0056454279,Ib43:0.0000000001)100:0.11
39434370,(Hrr. sodomense:0.0350927659,(((Hrr. distributum:0.0000000001,Hrr.
arcis:0.0661737939)26:0.0000000001,(Hrr. distributum duplicate:0.0000001083,Hrr.
litorem:0.0000001340)14:0.0000001083)91:0.0145309150,(Fb21:0.0000000001,(EC15:0.00000
00001,(C170:0.0000000001,(C191:0.0000001775,(Ga27:0.0000001438,Ga36:0.0000000001)67:
0.0028257437)4:0.0000002999)9:0.0000002027)15:0.0000001492)100:0.0471801814)89:0.030
2524668)97:0.0655038236)59:0.0329947285)99:0.3316552010)100:0.3347009271)100:0.3765
427867);

5.6 Unrooted Maximum Likelihood Tree (Protein)

(Har. sinaiiensis copy 2:0.0000000001,Har. marismortui copy 2:0.0000000001,(Har. vallismortis
copy 2:0.0000000001,((Hqm. walsbyi duplicate:0.0000000001,Hqm.
walsbyi:0.0000000001)100:0.1477266572,(((Hrr. ejinorense:0.0000007557,Hrr.
tebenquichense:0.1336414695)77:0.0411543467,((Eb13:0.0000000001,Ib43:0.0000000001)99:
0.1169916685,(Hrr. sodomense:0.0144216001,(Hrr. distributum duplicate:0.0000000001,(Hrr.
distributum:0.0000002187,(Hrr. litoreum:0.0000008253,(Hrr.
arcis:0.0912953154,(EC15:0.0000000001,(Ga27:0.0000004292,((Fb21:0.0000000001,C191:0.00
00003287)5:0.0000008612,(Ga36:0.0000000001,C170:0.0000000002)8:0.0000000001)1:0.0000
004292)10:0.0000008019)78:0.0146253622)6:0.0000003716)5:0.0000002187)13:0.000000840
6)88:0.0149686288)95:0.0745037056)53:0.0387457480)92:0.2347209310,((Hbt. NRC-1
duplicate:0.0000000001,(Hbt. salinarium:0.0000000001,(Hbt. NRC-1:0.0000000001,Hbt.
salinarium
duplicate:0.0000000002)25:0.0000000001)31:0.0000004306)100:0.4091737444,((Hfx.
mucosum:0.1198787739,(Hbf. haloterrestris:0.0078841928,((Ncc. aibiensis:0.0165663234,(Htg.
GSL11:0.0070096520,(Ntm. pallidum:0.0145323191,(Nnm. XA31:0.0074617635,(Nnm.
ajinwuensis:0.0000000001,Ntm.
altunense:0.0000003651)48:0.0000000001)78:0.0160695410)61:0.0077947966)92:0.05592043
99)70:0.0168370729,(Nrr. tenzan 10:0.0285287039,Hbf.
lacisalsi:0.0516020303)55:0.0273885889)60:0.0386446163)62:0.0721293197)91:0.1958758378
,(Hsx. carlsbadense:0.4542840004,(Hbl. gomorrense:0.1407645460,(Hmc.

mukohataei:0.0862536117,(Hrd. utahensis:0.3496962612,((Har. japonica:0.0000004770,Har. argentinensis:0.0140624192)41:0.0040887915,(Har. vallismortis:0.0139423323,(Har. hispanica:0.0171674106,(Har. amylolytica :0.0000001745,((Fb19:0.0000000001,Ha9:0.0000000001)52:0.0070243859,(Har. marismortui duplicate:0.0000000001,(Har. marismortui:0.0000004492,(Har. sinaiiensis:0.0000000001,(Har. californiae:0.0000000001,Har. californiae duplicate:0.0000000001)13:0.0000000001)12:0.0000000001)12:0.0000000001)69:0.0140343137)37:0.0070530062)14:0.0000001745)66:0.0212016674)34:0.0100364679)97:0.1131536963)61:0.0401205786)36:0.0846178914)10:0.0002676907)61:0.1077390571)90:0.2214983634)43:0.1776830170)100:0.3244308621)95:0.0813947557)29:0.0000000001);

5.7 Maximum Likelihood Tree (Nucleotide) with Hrr. tibetense Rooted with sensory rhodopsin genes

(Ntm. pallidum:0.0107037823,(Htg. GSL11:0.0141291678,(Ncc. aibiensis:0.0473879403,((Hbf. lacisalsi:0.0521955586,Nrr. tenzan 10:0.0854266900)37:0.0209594168,(Hbf. haloterrestris:0.0384790385,(Hfx. mucosum:0.1870322602,(((Hbt. NRC-1 sopI:0.0000002692,Hbt. salinarium sopI:0.0000001551)100:0.3289068844,(Hrr. lacusprofundi sopII:0.1498950282,(Hrr. tibetense:0.0485635973,Hrr. tibetense duplicate:0.0013526084)92:0.1139788299)93:0.2991027784)100:0.7482203339,(((Hrd. utahensis:0.3526213618,(Hmc. mukohataei:0.1571489722,((Har. argentinensis:0.0549443964,Har. japonica:0.0125234344)88:0.0556069012,((Har. hispanica:0.0171204799,Har. amylolytica:0.0032459990)88:0.0215204216,((Fb19:0.0000003243,Ha9:0.0000001462)96:0.0158948171,(Har. vallismortis:0.0732558055,(Har. californiae duplicate:0.0000000160,((Har. marism:0.0000002854,Har. marismortui duplicate:0.0000001344)69:0.0026401963,(Har. californiae:0.0000000805,Har. sinaiiensis:0.0000001213)10:0.0000000321)22:0.0000000160)100:0.0288637643)55:0.0414805627)28:0.0164485557)34:0.0162317103)95:0.1714884817)42:0.0484936723)48:0.0527518885,(Hbl. gomorrense:0.1722299458,Hsx. carlsbadense:0.4318209933)23:0.0025833005)75:0.0728508887,(((((Hrr. sodomense:0.0368513380,((C191:0.0000001498,((Ga36:0.0000003198,Ga27:0.0000001442)66:0.0028476273,(EC15:0.0000001361,(Fb21:0.0000000001,C170:0.0000002928)7:0.0000002785)12:0.0000003389)21:0.0000001498)100:0.0479056855,(Hrr. arcis:0.0352517934,(Hrr. litoreum:0.0000003219,(Hrr. distributum duplicate:0.0000003138,Hrr. distributum:0.0000001399)17:0.0000003138)6:0.0000001806)92:0.0144038010)83:0.0296581036)94:0.0656729200,(Ib43:0.0000002823,Eb13:0.0057002643)100:0.1154694020)52:0.0327422530,(Hrr. ejinorense:0.0357819567,Hrr. tebenquichense:0.0858139081)70:0.0442392981)100:0.3169684824,((Har. vallismortis copy 2:0.0055822503,(Har. marismortui copy 2:0.0000000356,Har. sinaiiensis copy 2:0.0000000809)63:0.0000579535)85:0.0900718378,(Hqm. walsbyi:0.0198952347,Hqm. walsbyi duplicate:0.0064791971)100:0.3642178032)100:0.3640392617)64:0.1205023225,(Hbt. NRC-1:0.0000001172,(Hbt. NRC-1 duplicate:0.0000001150,(Hbt. salinarium duplicate:0.0000001968,Hbt.

salinarium:0.0000002513)17:0.0000001217)22:0.0000001150)100:0.4351118329)64:0.1376825695)70:0.1367145432)55:0.0270018254)62:0.0816182211)40:0.0265706890)92:0.0603201533)100:0.0605838905)80:0.0077119026,(Nnm. XA31:0.0125726466,(Ntm. altunense:0.0000001568,Nnm. ajinwuensis:0.0084847273)88:0.0047756407)99:0.0346776042);

5.8 Maximum Likelihood Tree (Hybrid) with Hrr. tibetense Rooted with sensory rhodopsin genes

(Ntm. pallidum:0.0106646975,(Htg. GSL11:0.0140799547,(Ncc. aibiensis:0.0463056993,((Nrr. tenzan 10:0.0845232991,Hbf. lacisalsi:0.0518551650)38:0.0201148594,(Hbf. haloterrestris:0.0380587038,(Hfx. mucosum:0.1682128745,((((((Har. vallismortis copy 2:0.0055835398,(Har. marismortui copy 2:0.0000000720,Har. sinaiiensis copy 2:0.0000000544)62:0.0000003194)88:0.0815727941,(Hqm. walsbyi duplicate:0.0000003336,Hqm. walsbyi:0.0000003276)100:0.3728509428)100:0.3204317062,(((Eb13:0.0056586271,Ib43:0.0000003767)99:0.1124552478,(Hrr. sodomense:0.0364003131,((Hrr. arcis:0.0661876788,(Hrr. litoreum:0.0000002623,(Hrr. distributum duplicate:0.0000000300,Hrr. distributum:0.0000000767)16:0.0000000767)13:0.0000000600)95:0.0142886345,((Ga27:0.0000002009,Ga36:0.0000002563)70:0.0028270582,(C170:0.0000000626,(Fb21:0.0000001206,(C191:0.0000000626,EC15:0.0000003051)9:0.0000001531)12:0.0000003002)9:0.0000000626)100:0.0474205911)91:0.0291929117)98:0.0667674154)63:0.0350395274,(Hrr. ejinorense:0.0285424767,Hrr. tebenquichense:0.1275597164)68:0.0438455325)99:0.3062578821)82:0.1404098571,(Hbt. salinarium:0.0000001153,(Hbt. salinarium duplicate:0.0000001718,(Hbt. NRC-1 duplicate:0.0000000001,Hbt. NRC-1:0.0000000297)19:0.0000000297)23:0.0000000593)100:0.4368073527)53:0.1195776477,((Hbl. gomorrense:0.1896547043,(Hmc. mukohataei:0.1318537262,(((Har. amylolytica:0.0000002946,Har. hispanica:0.0196170794)93:0.0273304964,((Har. vallismortis:0.0442268392,(Har. japonica:0.0192613179,Har. argentinensis:0.0476737212)92:0.0662393180)42:0.0311048332,(Har. sinaiiensis:0.0000001348,(Har. californiae:0.0000000001,(Har. californiae duplicate:0.0000001542,(Har. marism:0.0000002944,Har. marismortui duplicate:0.0000003227)71:0.0026361572)13:0.0000001542)23:0.0000001531)100:0.0360011166)28:0.0277745694)20:0.0158203829,(Ha9:0.0000002096,Fb19:0.0000002988)96:0.0113474251)100:0.2002087810)57:0.0865460773)26:0.0185091060,(Hrd. utahensis:0.2966555018,Hsx. carlsbadense:0.3871624395)25:0.0462906833)82:0.0722188238)62:0.1317433560,((Hbt. salinarium sopI:0.0224045965,Hbt. NRC-1 sopI:0.0113331039)100:0.3091468427,(Hrr. lacusprofundi sopII:0.1292943803,(Hrr. tibetense duplicate:0.0000001501,Hrr. tibetense:0.0000002892)99:0.1331017741)98:0.3297921302)100:0.6984053284)58:0.0362948484)70:0.0940669208)48:0.0272428619)91:0.0600140162)99:0.0608843035)75:0.0076187642,(Nnm. XA31:0.0124526287,(Ntm. altunense:0.0000001260,Nnm. ajinwuensis:0.0084336016)81:0.0047610443)98:0.0344777565);

5.9 Maximum Likelihood Tree (Protein) with Hrr. tibetense Rooted with sensory rhodopsin genes

((Hmc. mukohataei:0.0966066064,(Hrd. utahensis:0.3383971822,(Har. japonica:0.0000008979,(Har. argentinensis:0.0085147731,(Har. vallismortis:0.0080499253,((Har. hispanica:0.0180397766,Har. amylolytica:0.0000003963)36:0.0000010374,((Har. californiae duplicate:0.0000001393,(Har. californiae:0.0000004755,(Har. marismortui duplicate:0.0000006054,(Har. sinaiiensis:0.0000006413,Har. marismortui:0.0000004777)9:0.0000003684)5:0.0000004755)13:0.0000001393)70:0.0147180596,(Ha9:0.0000002109,Fb19:0.0000004276)51:0.0073662404)36:0.0073970987)66:0.0288316536)40:0.0151100236)17:0.0063839395)95:0.1248053171)40:0.0458187232)46:0.0890419894,Hsx. carlsbadense:0.5274298690,(((((((Hrr. ejinorense:0.0003563042,Hrr. tebenquichense:0.1415125400)78:0.0841399687,(Hrr. sodomense:0.0152416145,((Hrr. litoreum:0.0000008968,Hrr. distributum:0.0000005390)12:0.0000003229,((Fb21:0.0000001467,(Ga36:0.0000006984,(C191:0.0000001474,(C170:0.0000004215,(EC15:0.0000007835,Ga27:0.0000004753)15:0.0000004215)3:0.0000004069)8:0.0000001474)14:0.0000001467)72:0.0156091886,(Hrr. distributum duplicate:0.0000002414,Hrr. arcis:0.0979089250)14:0.0000002414)11:0.0000008347)80:0.0161152349)94:0.0749486259)39:0.0242301457,(Ib43:0.0000008323,Eb13:0.0000007848)100:0.1058463685)94:0.2622445250,((Hqm. walsbyi:0.0000006088,Hqm. walsbyi duplicate:0.0000000001)99:0.1494633762,(Har. vallismortis copy 2:0.0000005697,(Har. marismortui copy 2:0.0000003067,Har. sinaiiensis copy 2:0.0000006730)24:0.0000007829)95:0.0932441117)99:0.3373383670)38:0.1771535574,(Hbt. NRC-1:0.0000005515,(Hbt. salinarium duplicate:0.0000000001,(Hbt. salinarium:0.0000001903,Hbt. NRC-1 duplicate:0.0000002773)17:0.0000002773)20:0.0000004135)100:0.4352220998)45:0.2451751662,(Hfx. mucosum:0.1378336015,(Hbf. haloterrestris:0.0000031792,((Ncc. aibiensis:0.0186147883,(Htg. GSL11:0.0074937062,(Ntm. pallidum:0.0155273568,(Ntm. altunense:0.0000001938,(Nnm. ajinwuensis:0.0000008148,Nnm. XA31:0.0080181237)16:0.0000008572)90:0.0174449374)65:0.0084321867)96:0.0583859180)77:0.0170536047,(Hbf. lacisalsi:0.0558217823,(Nrr. tenzan 10:0.0103935769,((Hrr. lacusprofundi SRII:0.1777982983,(Hrr. tibetense:0.0000003602,Hrr. tibetense duplicate:0.0000003786)80:0.0472661181)95:0.2559166738,(Hbt. NRC-1 SRI:0.0297530532,Hbt. salinarium SRI:0.0154572336)100:0.4236501163)100:0.9071349571)27:0.0197826915)18:0.0322635360)27:0.0482554694)34:0.0750860471)44:0.1883511982)69:0.1080319900,Hbl. gomorrense:0.1361000488)18:0.0000033398);

5.10 Unrooted Maximum Likelihood Tree (Hybrid) with Hrr. tibetense

(Hrr. litoreum:0.0000000598,(Hrr. distributum duplicate:0.0000000559,Hrr. arcis:0.0692087330)32:0.0000000559,(Hrr. distributum:0.0000002170,(((C191:0.0000001065,C170:0.0000002084)8:0.0000001891,(EC15:0

.0000000945,(Fb21:0.0000001804,(Ga36:0.0000001448,Ga27:0.0000001723)67:0.0029146456)8:0.0000003465)6:0.0000001557)97:0.0487408142,(Hrr. sodomense:0.0343931286,((Eb13:0.0057720427,Ib43:0.0000002406)98:0.1204319697,((Hqm. walsbyi:0.7874199311,((Hbt. salinarium:0.0000003084,Hbt. NRC-1:0.0000003292)100:0.4577476122,((Hfx. mucosum:0.1896495775,(Hbf. haloterrestris:0.0404352220,((Hbf. lacisalsi:0.0529802388,(Nrr. tenzan 10:0.0000319798,(Hrr. tibetense duplicate:0.0000002449,Hrr. tibetense:0.0000000001)100:1.7385843148)28:0.0909256329)13:0.0209947213,(Ncc. aibiensis:0.0456187799,(Htg. GSL11:0.0148946534,(Ntm. pallidum:0.0107353932,(Nnm. XA31:0.0127855999,(Ntm. altunense:0.0000001555,Nnm. ajinwuensis:0.0085312975)80:0.0047834810)100:0.0354632932)81:0.0076624082)100:0.0656325751)88:0.0643112595)24:0.0293108141)53:0.0913982068)61:0.1980692451,(Hsx. carlsbadense:0.4946545583,(Hbl. gomorrense:0.2103108660,(Hmc. mukohataei:0.1606861530,(Hrd. utahensis:0.4248487046,(((Har. hispanica:0.0199260499,Har. amylolytica:0.0000004859)95:0.0305904559,((Har. vallismortis:0.0441142221,(Har. argentinensis:0.0488423608,Har. japonica:0.0191359278)91:0.0697243120)36:0.0324603806,(Har. californiae:0.0000001356,Har. marismortui:0.0027302177)93:0.0341634670)25:0.0288580748)24:0.0172323155,(Ha9:0.0000001358,Fb19:0.0000001364)98:0.0081698229)92:0.1811390229)36:0.0581437497)37:0.0856793297)9:0.0000029061)46:0.0937564597)37:0.2014006354)39:0.1739930958)99:0.4227104189,(Hrr. ejinorense:0.0290517719,Hrr. tebenquichense:0.1378549131)80:0.0489484678)44:0.0385692448)93:0.0714941432)88:0.0337734457)94:0.0162953817)24:0.0000000598);

5.11 Maximum Likelihood Tree (Hybrid) with Hrr. tibetense, Rooted with halorhodopsin genes

(Hbt. salinarium hop:0.00063222,((Ncc. aibiensis:0.127387,Nnm. ajinwuensis:6.4751e-05)'67':12.7091,(Hmc. mukohataei:5.72136e-05,((((Haa. hispanica:0.0159841,Haa. amylolytica:3.567e-07)'80':0.0278481,((Haa. californiae:1.222e-07,Haa. marismortui:0.00267416)'99':0.0328029,(Haa. vallismortis:0.0384923,(Haa. japonica:0.0207514,Haa. argentinensis:0.0426865)'87':0.0617639)'39':0.0286917)'33':0.0277809)'21':0.0120097,(Fb19:4e-10,Ha9:7.07e-08)'95':0.0131046)'82':0.201054,(Hsx. carlsbadense:0.282676,(Hbl. gomorrense:0.155462,(((Hbf. haloterrestris:0.0492725,(Htg. GSL11:0.0127878,(Ntm. pallidum:0.00975663,Ntm. altunense:0.0373798)'60':0.00899447)'97':0.116998)'94':0.127745,(Hrr. tibetense:2.086e-07,Hrr. tibetense duplicate:2.828e-07)'94':1.26559)'28':0.0990768,((Hbt. salinarium:2.785e-07,Hbt. NRC-1:2.628e-07)'87':0.384912,((Hrr. ejinorense:0.0280741,Hrr. tebenquichense:0.128013)'80':0.0449152,((Eb13:0.00551496,Ib43:3.46e-07)'98':0.108756,(Hrr. sodomense:0.0317945,((Hrr. litoreum:1.067e-07,(Hrr. distributum:1.542e-07,(Hrr. distributum duplicate:6.8e-08,Hrr. arcis:0.0654504)'12':2.979e-07)'12':2.977e-07)'94':0.0156384,((Ga36:1e-10,Ga27:2.92e-07)'64':0.00276809,((C191:1e-10,(C170:1e-10,Fb21:1.254e-07)'6':1.497e-

07)'2':1.526e-07,EC15:1.399e-07)'14':1.526e-
07)'100':0.0446652)'81':0.0288433)'93':0.0641361)'46':0.0332074)'96':0.41243)'41':0.17332)'1
5':0.11548)'10':0.0393516)'14':0.0902489)'2':0.123919)'79':15.5569):0);