Microbial Ecology

Effects of Culturing on the Population Structure of a Hyperthermophilic Virus

J.C. Snyder[1], J. Spuhler[1], B. Wiedenheft[1], F.F. Roberto[2], T. Douglas[1] and M.J. Young[1]

(1) Thermal Biology Institute, Montana State University, Bozeman, MT 59717, USA
(2) Department of Biotechnologies, INEEL, Idaho Falls, ID 83415, USA

Received: 7 November 2003 / Accepted: 4 June 2004 / Online publication: 9 November 2004

Abstract

The existence of a culturing bias has long been known when sampling organisms from the environment. This bias underestimates microbial diversity and does not accurately reflect the most ecologically relevant species. Until now no study has examined the effects of culture bias on viral populations. We have employed culture-independent methods to assess the diversity of *Sulfolobus* spindle–shaped viruses (SSVs) from extremely hyperthermal environments. This diversity is then compared to the viral diversity of cultured samples. We detected a clear culturing bias between environmental samples and cultured isolates. This is the first study identifying a culture bias in a viral population.

Introduction

A limiting factor in environmental microbiology is the inability to culture, either in pure culture or as a consortium, the organisms that inhabit a particular environment [1]. This is especially apparent in culturing organisms that inhabit extreme environments because of our rudimentary understanding of the factors required to support microbial growth from these environments. The advent of non-culture-dependent methods has allowed us to examine the microbial diversity present in high-temperature environments [2, 3, 6, 20]. These studies demonstrate that only a limited number of the total organisms in that environment have been successfully cultured. It has been estimated that as little as 0.1% of the microorganisms in a particular environment have been grown in pure culture. The inability to culture a majority of thermophiles greatly limits our ability to understand the biology and the ecological roles these organisms play in the environment.

It has long been recognized that a culturing bias exists when isolating microorganisms from the environment. The selective pressures induced by culturing are thought to be the principal mechanism for this bias. Often pure cultures established from extreme environments do not represent the dominant organism found in that particular environment [1, 21]. For example, non-culture-dependent estimations of the microbial diversity in Octopus Springs suggest that *Synechococcus lividus* B' is a dominant microorganism, whereas *Synechococcus lividus* P3 is the organism that is typically cultured [21]. In addition, the organism adapted to growth in culture often does not faithfully maintain the biochemical and ecological function of its parents operating in the environment.

To date, no studies have determined if there is a strong selection bias when viruses of extremophiles are replicated in culture as compared to their natural environments. Most investigators use viruses replicating in culture to examine fundamental aspects of viral replication and virus–host interactions. However, viruses replicating in culture may not accurately represent the virus behavior and ecology found in the natural environment. Interest in the role of viruses in microbial ecology has been heightened by the completion of multiple microbial genomes, which has implicated the importance of viruses in horizontal gene transfer [7, 9, 11, 18].

Sulfolobus spindle-shaped viruses (SSVs) provide an ideal model system for investigating the effect of culturing on viral diversity. SSVs are nonlytic temperate viruses that infect *Sulfolobus* species inhabiting high temperature (>80°C) acidic (pH < 4) environments [15, 24, 25]. SSVs are 60 × 90 nm spindle-shaped particles which encapsidate a circular dsDNA viral genome of approximately 15 kb [10, 17, 22]. SSV genomes are stably maintained in three different forms. The viral genome integrates into

Correspondence to: M.J. Young; E-mail: myoung@montana.edu

the host chromosome and the episomal forms can be both positively or negatively supercoiled. SSVs have been isolated from multiple locations around the world and are the most common virus type isolated from high-temperature acidic environments [14, 22]. We have previously reported the isolation of SSVs from multiple high-temperature acidic locations within Yellowstone National Park (YNP) [14]. The broad collection of SSV genomes available has allowed a detailed comparison of the similarities and differences between geographically distributed SSVs [22]. SSV ORFs show little or no sequence similarity to genes in the public database. However, there is a common set of 18 genes shared by all four sequenced SSV isolates. In addition to this common set of genes, each SSV genome contains ORFs that are unique to that particular isolate. The distinct genetic signature of these viruses and the novel spindle-shaped morphology has resulted in the creation of a new viral family, the *Fusselloviridae* (http://www.ncbi.nlm.nih.gov/ICTV/). To date, all of the sequenced SSVs have been determined from cultured isolates. However, we do not know if these cultured isolates are representative of SSVs found in the environment.

A detailed understanding of the ecological role of SSVs in high-temperature acidic environments requires an understanding of the diversity of viruses present in that environment. As a first step in understanding that diversity, we were interested in determining whether a culture bias exists for these hyperthermophilic viruses in the acidic hot springs of Yellowstone National Park. Here we report the effects of culturing on SSV diversity from the environment. We use two separate genetic markers in the SSV genomes to monitor changes in virus populations as a consequence of culturing environmental samples.

Methods

Environmental Location. Environmental samples were collected from a hot spring in the Midway Geyser Basin (44° 31.287′N, 110° 48.647′W) and a hot spring in the Norris Geyser Basin (44° 43.653′N, 110° 42.862′W) in Yellowstone National Park. The hot spring in the Midway Geyser Basin (Rabbit Creek site) was characterized by a range of temperatures from 71 to 91°C and pH values from 2.8 to 4.2. The hot spring in the Norris Geyser Basin (Ragged Hills site) ranges from 68 to 78°C and pH 2.6–3.7.

Sampling. Water samples (500 mL) were collected from each hot spring using methods previously described [14, 24]. Three hours after collection, ~1 mL of the environmental sample was inoculated into 60 mL of enriching medium (DSMZ media 88; http://www.dsmz.de/dsmzhome.htm). The remaining environmental sample was used for total nucleic acid extraction. The cultures were incubated at 80°C and pH 3 for 7–10 days until cultures were visibly turbid. After the 7 days of incubation, 1 mL of the established culture was passaged into 60 mL of fresh medium. The remainder of the primary culture was used for DNA extraction. The first passage culture was grown for 7–10 days, after which DNA was extracted.

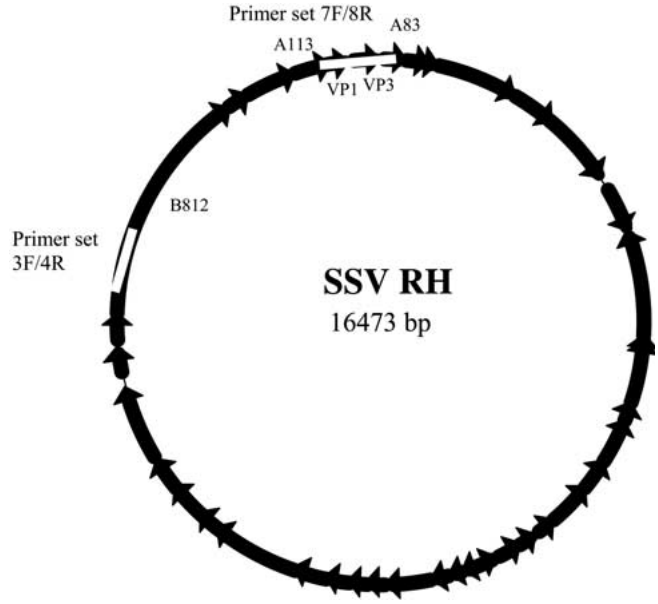

Figure 1. Genomic map of SSV RH from the Ragged Hills area in Norris Geyser Basin (Yellowstone National Park). The locations of the PCR primers and approximate amplicon regions are indicated in white.

DNA Isolation. Total nucleic acids were extracted using the Mo-Bio UltraClean Water DNA Extraction Kit according to manufacturer's instructions (Mo-Bio, Solana Beach, CA). DNA was extracted from the enrichment cultures using a rapid alkaline lysis method described previously [16, 23]. The yield of DNA extracted from the environment and the cultures was estimated by ethidium bromide staining of agarose gels.

PCR Amplification of SSV Genome Segments. Whole genome alignments of four geographically distributed SSVs isolated were used to identify regions of conserved nucleotides [22]. Two primer sets (Fig. 1) were identified to amplify SSV sequences. Primers UnvSSV#7F (5′ ATTCAGATTCTGWATWCA GAAC) and UnvSSV#8R (5′ TCSCCTAACGCACTC ATC) are located in SSVRH ORFs b170 and a83, respectively. These primers flank the viral coat proteins and amplify 1100 to 650 nucleotides, depending on the presence or absence of the VP2 gene in the isolate [22]. Primer sets UnvSSV#3F (5′ CAATGCCATAGG

CTACGG) and UnvSSV#4R (5′ GTACCGTTATAGTAR TAAACG) amplify ~256 nucleotides of the largest ORF conserved in all SSV genomes (SSV RH ORF B812) [22]. PCR was performed with a Perkin-Elmer DNA Thermal Cycler 9700, using 40 cycles of 94°C for 1 min, 45°C for 1 min, and 72°C for 1.5 min. Reaction mixtures contained 10× *Taq* buffer, 10 mM dNTPs, 50 pmol/μL of both the forward and reverse primer, ~10 ng of total extracted DNA, and 1 unit of *Taq* DNA polymerase. PCR products were visualized using agarose gels to estimate the size and concentration of each amplicon.

Cloning and Sequencing. The PCR products were cloned using the TOPO TA Cloning Kit as described by the manufacturer (Invitrogen, Carlsbad, CA). A 2 mL deep 96-well plate was inoculated with *E. coli* transformed with the cloned PCR products. Plasmids were purified from *E. coli* using the Millipore 96-well plasmid prep kit (Millipore, Billerica, MA). The plasmid DNA was sequenced using BigDye Terminator Mix (Applied Biosystems, Foster City, CA) on an ABI9700 automated sequencer (Applied Biosystems).

Phylogenetics. Sequences were edited and analyzed using Sequencher 4.1 (Gene Codes Corp., Ann Arbor, MI). Sequence alignments were performed using ClustalX [4, 8]. The alignments were subsequently analyzed by PAUP 4.0b10 [19] and a maximum parsimony analysis was conducted. Parsimony bootstrap values were obtained by resampling the data 10,000 times. A maximum likelihood analysis was conducted using Modeltest to select the nucleotide substation model [12]. Clade credibility values were derived using MrBayes where 11 million permutations were sampled every 10,000 generations [5].

Results and Discussion

SSV diversity was monitored directly in the environment and in subsequent enrichment cultures by phylogenetic analysis of two conserved regions within SSV viral genomes (Fig. 1). We detected a strong culturing bias in SSVs between the environment and the primary enrichment culture. Two hundred fifty-six nucleotides within the largest conserved SSV ORF (SSV RH ORF B812) were amplified from total environmental DNA using universal SSV primers 3F and 4R. Seventy-four PCR generated SSV clones from the Midway Geyser Basin (Rabbit Creek site) were phylogenetically analyzed. SSV diversity in the Midway Geyser Basin is dominated by two major clusters and one minor cluster (Fig. 2). Phylogenetic analysis of 31 viral clones from the primary enrichment culture reveals an obvious culture bias (blue in Fig. 2). SSV diversity observed in the hot spring was reduced to a single cluster when grown in an enrichment culture. The primary enrichment culture was associated with only one of the dominant environmental clusters. This trend was maintained when 20 additional clones were analyzed from the first passage of the primary enrichment culture (green in Fig. 2). The enrichment culture clones may form a separate cluster distinct from those observed in the environment, but this separation is not well supported by high bootstrap values and more clones will have to be considered (Fig. 2). These results identify a strong culture bias when comparing the diversity of SSVs in the environment to SSVs in culture.

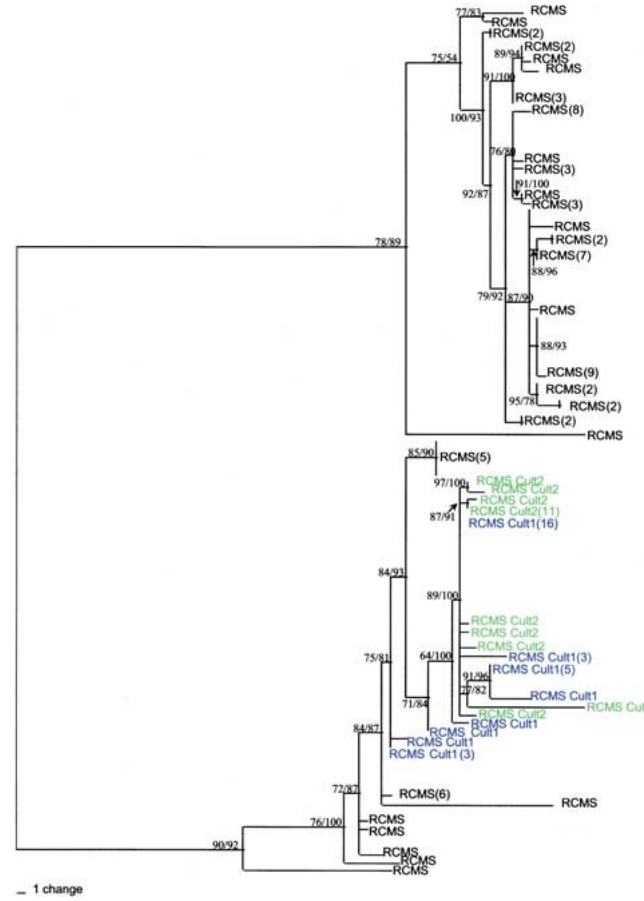

Figure 2. Phylogenetic tree of SSV sequences from the Rabbit Creek site in the Midway Geyser Basin. Seventy-five environmental clones (black) represent 53 unique sequences amplified using universal SSV primers 3F and 4R. Blue represents 31 sequences (seven unique) obtained from the original enrichment culture using the same primer set. Green represents 20 sequences (nine unique) obtained from the first passage of the enrichment culture. The frequency of each environmental clone is indicated directly following the clone name. Both maximum parsimony and maximum likelihood analyses resulted in the same tree topology. The first number at each node is the boostrap value obtained from resampling the data 10,000 times in the maximum parsimony analysis. The second number at the nodes represents the Bayesian clade credibility from the maximum likelihood analysis.

Figure 3. Phylogenetic tree of environmental clones from Ragged Hills in the Norris Geyser Basin comparing environmental sequences and the sequences of four cultured SSV isolates (gray) from geographically distinct locations. Black represents 43 environmental clones (21 unique) amplified with primers 3F and 4R. Both maximum parsimony and maximum likelihood analyses resulted in the same tree topology. The numbers at the node represent support from both analyses (see Fig. 1 legend for description).

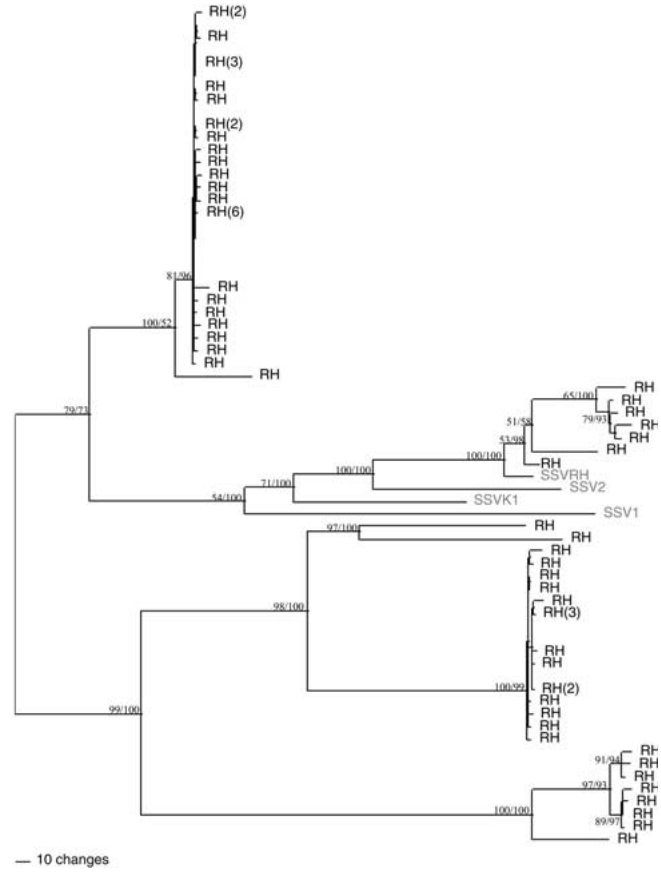

Figure 4. Phylogenetic tree of environmental clones from Ragged Hills and four sequenced isolates from geographically distinct locations. Black represents 63 environmental clones (40 unique) amplified with primers 7F and 8R. Gray represents the four sequenced isolates grown in culture (for a description see Fig. 3 legend and [22]). Both maximum parsimony and maximum likelihood analyses resulted in the same tree topology. The numbers at the node represent support from both analyses (see Fig. 1 legend).

We employed a similar approach to a second geothermal feature in Yellowstone National Park. Forty-three environmental PCR clones were analyzed from the Ragged Hills site within the Norris Geyser Basin. As at the Rabbit Creek site, a diverse population structure was observed (Fig. 3). At least five distinct clusters were detected and supported by high bootstrap values. These 43 environmental clones were compared to the four sequenced SSV isolates that have been cultured from distinct locations around the world [22]. It is worth mentioning that one of these four sequenced isolates was originally cultured from this same thermal pool in Ragged Hills. This Ragged Hills isolate (SSV RH from *Sulfolobus solfataricus*) clusters more closely together with cultured isolates from Japan (SSV1 from *S. shibatae*), Russia (SSV K1 from *S. solfataricus*) and Iceland (SSV2 from *S. islandicus*) than it does with environmental clones from the Ragged Hills thermal pool (Fig. 3; sequenced isolates represented in gray). All characterized SSVs are viruses of *Sulfolobus*, but the *Sulfolobus* species hosting the four sequenced isolates varies. Apparently this viral culturing bias is independent of geographic location and is not solely attributable to selection of a particular host. The only discernible common thread among these four sequenced isolates is that they were established from pure cultures using similar protocols.

A culturing bias is also apparent when analyzing sequence from the SSV viral coat protein region from the Ragged Hills site. Direct analysis of 63 environmental clones identifies five dominant clusters (Fig. 4). In agreement with the B812 marker analysis, phylogenetic analysis using the coat protein marker places all four sequenced viral isolates in the same cluster, despite their geographic separation (Fig. 4; sequenced isolates repre-

sented in gray). These results suggest that culturing imposes a strong selective pressure in these viruses regardless of geographical origin.

Overall, similar results were obtained using two independent markers from two different geothermal features. There was a significant reduction in SSV diversity observed in the primary enrichment culture (Fig. 2). This suggests that culturing SSVs from the environment alters the population structure. It is likely that culturing introduces a selection pressure that results in a bottleneck of SSV sequences that are not representative of the diversity originally present in the natural environment. Taken together, these results indicate that describing the diversity and population structure based on viruses isolated from cultures would severely underestimate the viral diversity actually present in a particular environment.

There are several possibilities that may explain the observed culturing bias in the SSV populations. One possibility is that the host range of SSVs extends beyond the *Sulfolobus* genus. 16S rDNA sequence analysis indicates that *Acidianus* and *Stygiolobus* are the dominant members of the Archaeal community in the two high-temperature acidic environments sampled (data not shown). However, these organisms were not detected in the enrichment cultures. All enrichment cultures were exclusively *Sulfolobus solfataricus* as determined by 16S rDNA sequence analysis (data not shown). In contrast, Rachel et al. [13] established cultures from environmental samples from a hot springs in the Crater Hills region of YNP (44° 39′ 13.3″N and 110° 28′ 39.8″W) using similar protocols that were exclusively *Acidianus* by 16S rDNA analysis [13]. A 60 × 90 nm SSV-like spindle-shaped virus like morphology was observed in these enrichment cultures. However, it has not yet been determined if these viruses are members of the *Fusselloviridae*.

Another possibility for the apparent high diversity in the environmental samples is errors may have occurred during PCR, cloning, and/or sequencing. The PCR reactions were performed using *Taq* polymerase, which does not possess the capability to proofread. However, the expected error rate due to *Taq* polymerase would be low because of the short length of the amplicons. The PCR products were then cloned, which could also induce errors, although the diversity could actually be underestimated because of the limited number of clones that were screened. It is also pertinent to note that two different DNA extraction methods were used to extract viral DNA from the environment and from cultures. However, it is unlikely that different DNA extractions would account for the bias detected.

The PCR assays used in this investigation do not necessarily ensure that we are detecting replicating viral sequences. The sequenced regions could be remnants of sequences found in the genomes of host organisms. However, the sequences are likely to be from replicating virus because the coding capacity, high sequence similarity, and size of the amplicon significantly match those of known replicating SSVs (data not shown). If anything, PCR primer bias may be underestimating the diversity of these viruses.

It has long been known that there is a culturing bias in microorganisms; however, this phenomenon has not been shown in viruses until now. Clearly, microbial diversity in enrichment cultures seldom reflects the diversity of microbial life present in the environment. As shown by this study, the same appears to be true for a hyperthermophilic virus. We have shown that when viruses are cultured from the environment, the viral sequences may not be representative of the viruses in the natural environment. Therefore, conclusions regarding viruses based on culturing experiments must be made cautiously.

Acknowledgments

We thank Matthew Lavin and Mary Bateson for critically reading this manuscript. This research was conducted with permission from Yellowstone National Park research permit YELL-05090. This work is supported by grants from the National Science Foundation (MCB01322156) and the National Aeronautics and Space Administration (NAG5-8807).

References

1. Amann, RI, Ludwig, W, Schleifer, KH (1995) Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. Microbiol Rev 59: 143–169
2. Barns, SM, Delwiche, CF, Palmer, JD, Pace, NR (1996) Perspectives on archaeal diversity, thermophily, and monophyly from environmental rRNA sequences. Proc Natl Acad Sci USA 93: 9188–9193
3. Barns, SM, Fundyga, RE, Jeffries, MW, Pace, NR (1994) Remarkable archaeal diversity detected in a Yellowstone National Park hot spring environment. Proc Natl Acad Sci USA 91: 1609–1613
4. Higgins, DG, Thompson, JD, Gibson, TJ (1996) Using CLUSTAL for multiple sequence alignments. Methods Enzymol 266: 383–402
5. Huelsenbeck, JP, Ronquist, F (2001) MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17: 745–755
6. Hugenholtz, P, Pitulle, C, Hershberger, KL, Pace, NR (1998) Novel division level bacterial diversity in a Yellowstone hot spring. J Bacteriol 180: 366–376
7. Hughes, AL, Friedman, R (2003) Genome-wide survey for genes horizontally transferred from cellular organisms to baculoviruses. Mol Biol Evol 20: 979–987
8. Jeanmougin, F, Thompson, JD, Gouy, M, Higgins, DG, Gibson, TJ (1998) Multiple sequence alignment with Clustal X. Trends Biochem Sci 23: 403–405
9. Nelson, KE, Clayton, RA, Gill, SR, Gwinn, ML, Dodson, RJ, Haft, DH, Hickey, EK, Peterson, JD, Nelson, WC, Ketchum, KA, McDonald, L, Utterback, TR, Malek, JA, Linher, KD, Garrett, MM, Stewart, AM, Cotton, MD, Pratt, MS, Phillips, CA, Richardson, D,

Heidelberg, J, Sutton, GG, Fleischmann, RD, Eisen, JA, Fraser, CM, et al. (1999) Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of *Thermotoga maritima.* Nature 399: 323–329

10. Palm, P, Schleper, C, Grampp, B, Yeats, S, McWilliam, P, Reiter, WD, Zillig, W (1991) Complete nucleotide sequence of the virus SSV1 of the archaebacterium *Sulfolobus shibatae.* Virology 185: 242–250
11. Peng, X, Blum, H, She, Q, Mallok, S, Brugger, K, Garrett, RA, Zilling, W, Pragishvili, D (2001) Sequences and replication of genomes of the archaeal rudiviruses SIRV1 and SIRV2: relationships to the archaeal lipothrixvirus SIFV and some eukaryal viruses. Virology 291: 226–234
12. Posada, D, Crandall, KA (1998) MODELTEST: testing the model of DNA substitution. Bioinformatics 14: 817–818
13. Rachel, R, Bettstetter, M, Hedlund, BP, Haring, M, Kessler, A, Stetter, KO, Prangishvili, D (2002) Remarkable morphological diversity of viruses and virus-like particles in hot terrestrial environments. Arch Virol 147: 2419–2429
14. Rice, G, Stedman, K, Snyder, J, Wiedenheft, B, Willits, D, Brumfield, S, McDermott, T, Young, MJ (2001) Viruses from extreme thermal environments. Proc Natl Acad Sci USA 98: 13341–13345
15. She, Q, Singh, RK, Confalonieri, F, Zivanovic, Y, Allard, G, Awayez, MJ, Chan-Weiher, CC-Y, Clausen, IG, Curtis, BA, Moors, AD, Erauso, G, Fletcher, C, Gordon, PMK, Jong, IH-d, Jeffries, AC, Kozera, CJ, Medina, N, Peng, X, Thi-Ngoc, HP, Redder, P, Schenk, ME, Theriault, C, Tolstrup, N, Charlebois, RL, Doolittle, WF, Duguet, M, Gaasterland, T, Garrett, RA, Ragan, MA, Sensen, CW, Oost, JVd (2001) The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. Proc Natl Acad Sci USA 98: 7835–7840
16. Stedman, KM, Schleper, C, Rumpf, E, Zillig, W (1999) Genetic requirements for the function of the archaeal virus SSV1 in *Sulfolobus solfataricus*: construction and testing of viral shuttle vectors. Genetics 152: 1397–1405
17. Stedman, KM, She, Q, Phan, H, Arnold, HP, Holz, I, Garrett, RA, Zillig, W (2003) Relationships between fuselloviruses infecting the extremely thermophilic archaeon *Sulfolobus*: SSV1 and SSV2. Res Microbiol 154: 295–302
18. Sullivan, MB, Waterbury, JB, Chisholm, SW (2003) Cyanophages infecting the oceanic cyanobacterium *Prochlorococcus.* Nature 424: 1047–1051
19. Swofford, DL (2002) PAUP*4.0 Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4 beta 10 ed., Sinauer Associates, Sunderlund, MA
20. Ward, DM, Ferris, MJ, Nold, SC, Bateson, MM (1998) A natural view of microbial biodiversity within hot spring cyanobacterial mat communities. Microbiol Mol Biol Rev 62: 1353–1370
21. Ward, DM, Weller, R, Bateson, MM (1990) 16S rRNA sequences reveal numerous uncultured microorganisms in a natural community. Nature 345: 63–65
22. Wiedenheft, B, Stedman, K, Roberto, F, Willits, D, Gleske, A-K, Zoeller, L, Snyder, J, Douglas, T, Young, M (2004) Comparative genomic analysis of hyperthermophilic archaeal fuselloviridae viruses. J Virol 78: 1954–1961
23. Yeats, S, McWilliam, P, Zillig, W (1982) A plasmid in the archaebacterium *Sulfolobus acidocaldarius.* EMBO J 1: 1035–1038
24. Zillig, W, Kletzin, A, Schleper, C, Holz, I, Janekovic, D, Hain, J, Lanzedorfer, M, Kristjansson, JK (1994) Screening for *Sulfolobales*, their plasmids, and their viruses in Icelandic solfataras. Syst Appl Microbiol 16: 609–628
25. Zillig, W, Stetter, KO, Wunderl, S, Schulz, W, Preiss, H, Scholz, H (1980) The *Sulfolobus*–''Caldariella'' group: taxonomy on the basis of the structure of DNA-dependent RNA polymerases. Arch Microbiol 125: 259–269