The Red Layer Microbial Observatory Database: A Model for the Integration and Dissemination of Biological and Geochemical Data via the World Wide Web

Sarah M. Boomer[1]* | Will M. Drury[2] | Bryan E. Dutton[1]
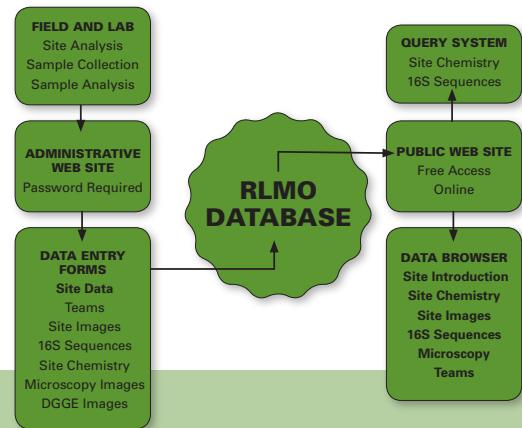Daniel P. Lodge[1] | Melissa S. Boschee[1] | William M. Kernan[2]

[1]Department of Biology, Western Oregon University, Monmouth
[2]University Computing Services, Western Oregon University, Monmouth

*Corresponding Author:
Department of Biology
Western Oregon University
345 Monmouth Ave
Monmouth, OR 97361

Phone: 503.838.8209   Fax: 503.838.8072   E-mail: boomers@wou.edu

ABSTRACT

The Red Layer Microbial Observatory (RLMO) aims to survey and compare alkaline mat communities throughout Yellowstone National Park, all of which contain distinct layers of red, filamentous bacteria we have identified as atypical red Chloroflexi using 16S rRNA, pigment spectroscopy, and microscopic methods. The RLMO Database Project is a Web-accessible Oracle™ database that integrates physical, chemical, and molecular data from RLMO sites. This application, in its second year of development, was written using PL/SQL, HTML, and Javascript. Each research site in this study is assigned a unique identifier that is linked to the following tables: Geochemical Data (pH, 15 common salts and metals); DNA Sequence Data (16S clone name, GenBank-linked accession number, BLAST-inferred identity); Macroscopic and Microscopic Image Data; and Student Collection Team Information. Using the administrative URL, all data can be entered and edited through restricted Web access. Using the public URL, users can view and query data. Site Query results can be formatted to display any combination of geochemical parameters across one or more years and/or sites. Sequence Query results can be formatted to display inferred identity and GenBank-linked accession numbers across one or more sites. The RLMO Database, in its current form, is designed to accommodate physical, chemical, and molecular information as part of a five-year study to address whether observed variation in geothermal ground water chemistry affects the population distribution of genetically distinct strains of Chloroflexi in red layer communities.

Key Words

Chloroflexi
database
green non-sulfur (GNS)
microbial mat
Oracle
16S rRNA™
World Wide Web

1.0 RED LAYER MICROBIAL COMMUNITIES

Moderately thermophilic (45-60°C) microorganisms such as those thriving in hot spring microbial mats are hypothesized to be modern homologs of the most ancient forms of life on Earth (Walsh and Lowe 1985; Schopf and Packer 1987). Well-characterized cyanobacterial mat communities are largely composed of filamentous photosynthetic bacteria arranged in layers according to light-harvesting and metabolic properties; typically, cyanobacteria occupy the surface and green Chloroflexi (formerly Green nonsulfur, GNS) reside in a secondary underlayer (Castenholz 1984; Pierson and Castenholz 1992). Phylogenetically, these photosynthetic bacteria comprise diverse genera: green Chloroflexi are anoxygenic phototrophs that contain both bacteriochlorophyll (Bchl) *a* and *c* (Oyaizu et al. 1987). In contrast, cyanobacteria, oxygenic phototrophs that contain chlorophyll *a*, comprise a distinct phylum predicted to have arisen after Chloroflexi (Woese 1987). While these phototrophs appear to be the predominant organisms in the mat, a remarkable diversity of microorganisms has been observed via the application of molecular methodology to hot spring communities (Ward et al. 1998).

An unusual red layer-containing microbial mat (pH 8.0-8.2, 30-50°C) was first reported near Rabbit Creek, Midway Geyser Basin (Castenholz 1984). We studied this red layer community until its supporting geyser source dried up in 1996. In contrast to typical cyanobacterial mats, the Rabbit Creek community was composed of three layers: cyanobacteria (top), traditional green Chloroflexi (middle), and a distinct red layer (bottom) (Boomer et al. 2000). The red layer was dominated by filaments that contained only Bchl *a* and performed sulfide-independent photoheterotrophy, suggesting they were related to Chloroflexi bacteria (Boomer et al. 2000). Unfortunately, the Rabbit Creek red filaments were not amenable to isolation in pure culture, and 16S rRNA-based identification methods were not applied to the since-exhausted samples.

Examples of, and interest in, atypical Chloroflexi bacteria that contain only Bchl *a* have increased over the last 20 years. In 1985, *Heliothrix oregonensis*, the first Chloroflexi containing only Bchl *a*, was isolated in co-culture from unusual orange mats in Oregon (pH 8.0, 40°C; Pierson et al. 1984, 1985). In 1992, *Heliothrix*-like 16S rRNA sequences were retrieved from Octopus Spring (Yellowstone) using a general bacterial amplification strategy (Weller et al. 1992). In 1998, we began to survey thermal basins throughout Yellowstone for new examples of red layer communities. In early 2002, we published our first molecular characterization of five Red Layer Microbial Observatory (RLMO) sites, generated from both general bacterial and Chloroflexi-targeted 16S rRNA libraries (Boomer et al. 2002). Coincidentally, *Roseiflexus castenholzii* was isolated from red surface mats (pH 8.0, 40°C) in Japan (Hanada et al. 2002). In late 2002, Chloroflexi-targeted 16S libraries were retrieved from Mushroom Spring (Yellowstone), and Red Chloroflexi sequences were linked, via fluorescent *in situ* hybridization, directly to observed filaments (Nubel et al. 2002). It should be noted that when we began our survey of Yellowstone in 1998, there were only two *Heliothrix*-like Chloroflexi sequences and no Red Chloroflexi-like sequences in GenBank; today there are more than 300, including some Red Chloroflexi-like sequences retrieved using general bacterial amplification strategies from thermal groundwater and communities that lacked visibly obvious red layers (Bonheyo et al. 2001).

RLMO-derived Chloroflexi and Japanese *Roseiflexus* comprise a unique red subgroup that is distinct from both traditional green (e.g., *Chloroflexus*) and orange (e.g., *Heliothrix*) Chloroflexi groups (Boomer et al. 2000, 2002). Within this red subgroup, we have observed heterogeneity and site-specific sub-clustering, suggesting selection by as-yet-undetermined site feature(s) (Boomer et al. 2002). Consequently, we performed more thorough water chemistry assessments at four RLMO sites to assess the relationship between site-specific chemical signatures and observed selection patterns. Preliminary characterization of 15 common salts and metals suggested that RLMO sites, despite similar temperatures and pHs, exhibit distinct chemical signatures; unpublished representative site chemistry data from 2002-2003, queried and downloaded directly from the RLMO database, is shown in **Table 1 (next page)**. Assessing site water chemistry and genetic variation at RLMO sites, with an emphasis on Chloroflexi bacteria and their relative distribution in red

layer communities, remains the central hypothesis we are testing over the next several years.

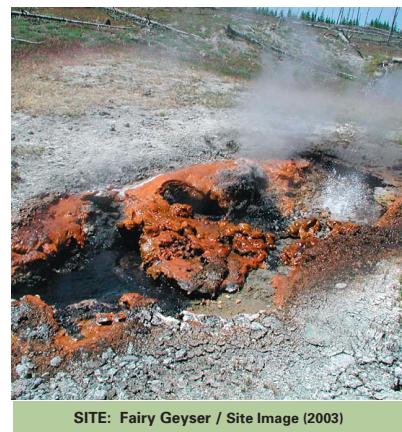## 2.0 WEB-ACCESSIBLE DATABASES AND YELLOWSTONE DATABASE PROJECTS

For a general overview of concepts regarding Web-accessible tools and database terminology, we recommend *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins* (Baxevanis and Ouellette 1998). Although biased towards the NCBI/GenBank database, this text serves as an appropriate starting point for defining basic terms and considering database structure and data entry issues. Online databases come in many formats: from simple Web sites that provide only links for related Web sites, to intermediately complex Web sites that provide stand-alone data tables (e.g., Microsoft Word™, Microsoft Excel™), to complex collections of inter-related tables constructed using database software (e.g., Microsoft Access™, Oracle™) and programmed for online access (e.g., using PL/SQL). Software packages like Microsoft SQL Server™ provided integrated database and online access features.

To examine the extent to which Yellowstone scientists were using the Internet to provide data, we reviewed the National Park Service Investigators' Annual Report (**http://www.nps.gov/yell/publications/pdfs/investigatorsindex.htm**). Searching Yellowstone IARs for titles and objectives that contained the word "database," we located 13 projects, including our RLMO Database (**Table 2**). While three IARs cited online database addresses, the remaining IARs provided

Table 1. **Current RLMO Sites and Water Chemistry.** To compile this table, we downloaded images and data directly from our RLMO Database. Site images (all 2003) have not been published elsewhere. Preliminary water chemistry data gathered in 2002 and 2003 illustrates a representative subset of information being surveyed as part of our project to examine the relationship between genetic variation and site chemistry.

SITE: Hillside Springs / Site Image (2003)

WATER CHEMISTRY DATA

Chemical	2002 (mg/L)	2003 (mg/L)
Calcium Hardness	2.84	4.40
Total Chlorine	0.19	0.00
Hydrogen Sulfide	0.01	0.00
Total Iron	0.09	0.00
Nitrate	1.70	0.06
Silica	20.20	126.00
Sulfate	8.00	16.00
Zinc	0.05	0.19

SITE: Fairy Geyser / Site Image (2003)

WATER CHEMISTRY DATA

Chemical	2002 (mg/L)	2003 (mg/L)
Calcium Hardness	3.36	3.77
Total Chlorine	0.39	0.01
Hydrogen Sulfide	0.01	0.00
Total Iron	0.03	0.07
Nitrate	1.40	0.50
Silica	28.00	137.00
Sulfate	15.00	25.00
Zinc	0.08	0.40

no address information and made no specific reference to online database development. Upon performing Google™ searches with author names and the keywords "database" and "Yellowstone," we located only one additional IAR database. All IAR databases we located will now be briefly described. The first project is C. Kaag's "Scientific Research in Yellowstone National Park 1872-1997: a Bibliographic Guide and Database", written using Procite® (Thomson ISI Researchsoft™). This Washington State University-supported site featured a search engine, although "quirks" were noted by the authors. The second project was C. Skinner's Obsidian Database, a simple HTML list of bibliographic references pertaining to Yellowstone and beyond. The third project was D. Stoner's "Integrated Biogeochemical Database" written using Microsoft Access™ and linked to Arc Internet Map Server™ (ESRI). Although published goals of this database project include microbiological data, query functions, and

Table 2. **Yellowstone IAR Database Projects**. To compile this table, we reviewed the National Park Service Investigator's Annual Report (NPS IAR) database, searching under title and objective options using the keyword "database."

Investigator	Reporting Year(s)	Topic	Database
S. Boomer	1999	Microbiology	wilbur.wou.edu/pls/wou/nsboomer.public3.main
K. Cannon	1999-2002	Obsidian Geology	none located online
S. Consolo-Murphy	1994, 1998-1999	Beaver Ecology	none located online
M. Glascock	1997-2001	Obsidian Geology	none located online
J. Harris	2001-2002	Mine Impact	none located online
C. Kaag	1997-1999, 2000	Bibliographies	www.wsulibs.wsu.edu/Yellowstone/
K. Keurs	1995-1996, 2000	Fire Ecology	none located online
R. Lidstrom	1992-1997	Microbiology	none located online
C. Noble	1998-2000	Wetland Geology	none located online
E. Schrader	2000	Mine Impact	none located online
C. Skinner	2002	Obsidian Geology	www.peak.org/obsidian/bibliography.html
D. Stoner	1999-2000	Biogeochemistry	gis.inel.gov/ynp/ynphtml
B. Wall	1999-2000	Paleontology	none located online

online data entry forms, this site currently only contains spatial and physical data and source photographs (Stoner et al. 2001).

To rectify possible search limitations during our IAR review, we performed a general Google™ search using the keywords "Yellowstone" and "database." Reviewing the approximately 57,000 hits (most commercial or travel-related) is beyond the scope of this review; thus, our presentation should be considered non-exhaustive. We noted several other bibliographic databases, the most comprehensive of which was the "Greater Yellowstone Bibliography" (**http://www-lib.uwyo.edu/db/ynp/**), supported by The University of Wyoming Libraries at Laramie (the authors did not cite what software they used). We noted several geospatial data clearinghouses that provided Yellowstone-specific GIS datasets, most derived from or linked to the comprehensive searchable National Park Service GIS Data Clearinghouse (**http://www.nps.gov/gis/**). We located few microbial-oriented databases about Yellowstone, the most useful being G. Chan's "Microbial Geochemistry" (**http://mysite.verizon.net/gwchan/Research/research.html**). In addition to providing links for Yellowstone research pages, "Microbial Geochemistry" provides simple stand-alone spreadsheet tables that catalogued cultured thermophiles from Yellowstone. Similar link-providing pages were also found on some federal grant agency-related sites (e.g.,
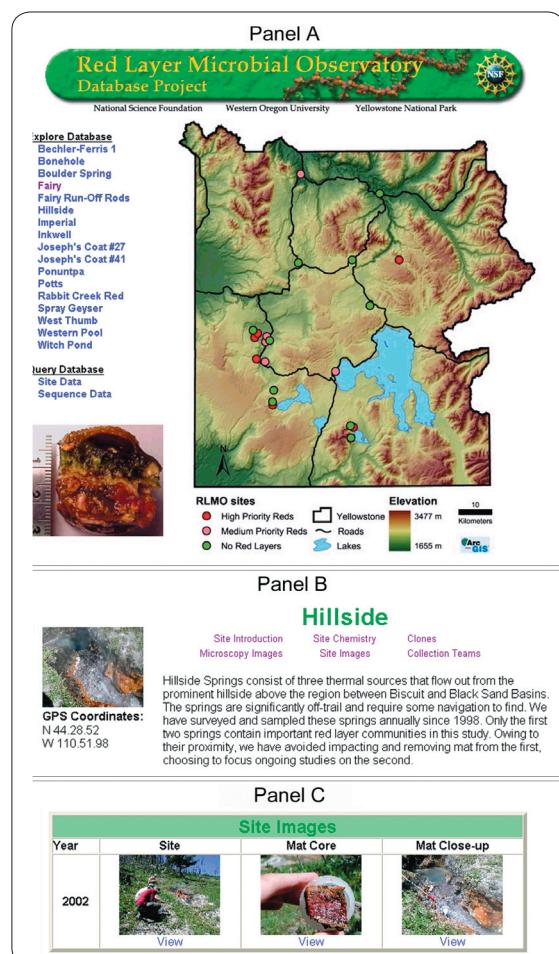
the NASA-supported Astrobiology Web site, **http://astrobiology.arc.nasa.gov**).

3.0 THE RLMO DATABASE

Undergraduate teams from Western Oregon University began surveying Yellowstone for new red layer communities in 1998. In the field, we used global positioning system equipment, assessed habitats using pH and temperature sensors, photo-documented new red layer sites and mats, and coordinated the labeling of collected samples with adjoining field data forms. Immediately following survey trips, we assessed red layer community pigments and microscopic composition. Samples that contained both red filaments and Bchl *a* were frozen (-80°C) for DNA analysis. The latter methods provide the backbone for research-driven microbiology and molecular laboratory courses, and outreach activities that have had a combined impact on 223 high school students, undergraduates, and science teachers.

While formulating our long-term research goals and hypotheses (1998-2001), we described, archived, and/or published more than 150 red layer-derived 16S sequences and 10 RLMO sites—the latter adjoined to multiple digital photographs (microscopic and macroscopic) and GPS, temperature, pH, and pigment data. Because this project was performed by undergraduates and supported

via Research at Undergraduate Institute (RUI) category funding, we also archived information about field collection teams and course-based research participants. At the time, our database was limited to HTML tables on our long-term RLMO Project Web site (**http://www.wou.edu/~boomers/research/allresearch.html**, online since 2000).
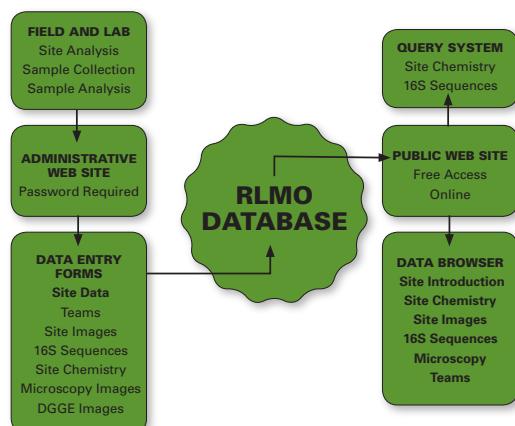

Figure 1. Public RLMO Database Images. Panel A shows the current (November, 2003) cover of the RLMO Database. The map and legend graphics, an ArcGIS™ map of Yellowstone National Park (data obtained from Yellowstone Spatial Analysis Center Web site, **www.nps.gov/yell/technical/gis/**), indicate locations of all RLMO sites in the database. Panel B shows a sample of the site introduction page for Hillside Springs RLMO site. Panel C displays one year of sample site images taken for Hillside Springs.

With the proposed addition of more data dimensions (chemistry and population assessment using denaturing gradient gel electrophoresis, DGGE) in 2001-2002, the idea of continuing to utilize HTML tables to manage a five-year monitoring study became unwieldy, and we turned to Western Oregon University's Computer Services for database assistance. Given campus expertise with Oracle™, we committed to this software using PL/SQL for online access.

The publicly accessible RLMO database cover is shown in **Figure 1, Panel A**; its overall structure summarized in **Figure 2**. After reviewing target data collected or to be collected, we developed the following data entry fields, each of which is dependently linked to a single Site Data Form that archives site name, location, and introductory remarks (**Figure 1, Panel B**):

1. Site Images, consisting of mat and source, mat close-up, and mat core sample (**Figure 1, Panel C**)

2. Site Chemistry, consisting of 15 common inorganic salts or metals, pH, temperature, and pigments (**Figure 3, Panel C, next page**)

3. Microscopic Images, consisting of light and fluorescent techniques

Figure 2. RLMO Database Structure. This flow chart schematically illustrates the conceptual design of the database, following the process from field collection and analysis through data tables for administrative entry and public viewing.

4. 16S rDNA sequences, consisting of clone name, active link to respective GenBank flatfile, and BLAST-based phylum identify

5. DGGE Images

Panel A

Admin Menu

Edit Existing Site
Select...   ▾
Get Record

Add New Site

Create New Team
Edit Existing Team
Create New Collector

Panel B

New Site Chemistry (- 245 -)

Year: ▢

Water

Acid/Base: ▢
Temperature (°C): ▢
Ferrous Iron (mg/L): ▢
Zinc (mg/L): ▢
Molybdenum (mg/L): ▢
Total Chlorine (mg/L): ▢
Copper (mg/L): ▢
Ca Hardness (mg/L): ▢

Manganese (mg/L): ▢
Hydrogen Sulfide (mg/L): ▢
Sulfate (mg/L): ▢
Aluminum (mg/L): ▢
Silica (mg/L): ▢
Total Iron (mg/L): ▢
Nitrate (mg/L): ▢
Reactive Phosphorus (mg/L): ▢
Mg Hardness (mg/L): ▢

Notes

Panel C

Hillside

Site Introduction   Site Chemistry   Clones
Microscopy Images   Site Images   Collection Teams

Site Chemistry - 2002

Water

Acid/Base:	6.50	Manganese (mg/L):	0.02
Temperature (°C):	83.50	Hydrogen Sulfide (mg/L):	0.01
Ferrous Iron (mg/L):	0.00	Sulfate (mg/L):	8.00
Zinc (mg/L):	0.05	Aluminum (mg/L):	0.00
Molybdenum (mg/L):	0.00	Silica (mg/L):	20.20
Total Chlorine (mg/L):	0.19	Total Iron (mg/L):	0.09
Copper (mg/L):	0.00	Nitrate (mg/L):	1.70
Ca Hardness (mg/L):	2.84	Reactive Phosphorus (mg/L):	1.11
		Mg Hardness (mg/L):	1.36

↑ **Figure 3.** **RLMO Database Data Entry and Display Features.** Panel A indicates the password-protected gateway page for administrative data entry and access. Panel B shows a sample site chemistry data entry table. Data entry and "notes" boxes typify all data entry forms, allowing for both precise data entry and flexible note spaces for recording special methods, equipment, or collection circumstances. Panel C illustrates how site chemistry data is displayed for the public access portion of the RLMO database.

For each specific datum entered in all fields, we have linked "notes" boxes that allow us to provide relevant accessory information (e.g., notable equipment deviations from standard procedures and references). Although each Site Data Form is static in that it contains no reference to time, each dependent data entry field can be updated on an annual basis and linked to specific collection times. Thus, there will be as many site-specific data entry fields as there are collection times for each RLMO site in the study. Given the RUI-focus of this project, we also created tables for each field team collection member (name, current position) that could be assembled to build teams and linked to given site collection times (**Figure 4, next page**).

All database entry forms are accessed online using an administrative, password-protected Web site. Upon entry of access information, administrators may choose to edit existing sites using a pull-down menu, add new sites, or enter team/collection data (**Figure 3, Panel A**). Upon the initial creation of a given Site Data Form, administrators may further select to add or edit data from dependent data entry field tables. As with typical online databases, administrators never directly enter data using raw Oracle™ tables. For example, **Figure 3, Panel B** shows a representative administrative site chemistry data entry table. During database development or modification, four members of the lab (PI, research assistant, and two undergraduates) served to troubleshoot data entry issues. In the summer of 2004, we actively taught undergraduates involved in summer research experiences to perform data entry using raw data they gather in the field and lab. This important skill improves specific training in database entry and reiterates the need for excellent lab notebook maintenance in the lab and field.

The public RLMO database Web site is accessible without the use of a password. Users can explore the database via two formats: a Data Browser and a Query System. The former allows users to navigate RLMO site data using the data entry field hierarchy described above, viewing all or some collection time-points for each field of data. The current data Query System allows users to analyze discrete Site Chemistry or 16S sequence information. The former allows users to define year(s) using pull-down boxes, and

specify site(s) and geochemical characteristics by checking specific boxes that reflect database content (**Figure 5, Panel A**). Data output consists of a series of tables that summarize each geochemical characteristic selected as a function of site and time (**Figure 5, Panel B**). The Sequence Query system allows users to specify site(s) and Inferred Identity(ies) by checking specific boxes that reflect database content (**Figure 5, Panel C**). Alternatively, users may simply type in an accession number that links to additional clone and site information. Data output consists of a series of tables

Hillside

Site Introduction   Site Chemistry   Clones
Microscopy Images   Site Images   Collection Teams

Collection Teams - All Years

1998
June 1998 (50)

Sarah Boomer	RLMO PI, Associate Professor of Biology
Robin Leitch	Health Care Management, Oregon
Joseph Hayes	Business/Sales, Salem, Oregon
Kody Phillis	Research and Diagnostics Technician, Rock Creek Labs, Oregon
Ben Stern	Paramedic, Sheridan, Oregon
Chandell Terwilliger	Lab Manager, GeneTools LLC, Corvallis, Oregon
Daniel Russell	State Government Legislative Assistant, Salem, Oregon
Eric Hayes	Unknown in 2003

1999
August 1999 (70)

Sarah Boomer	RLMO PI, Associate Professor of Biology
Rex Addis	Youth Corrections Officer, Oregon
Chris Coverdill	Research Technician, OHSU, Oregon
Brian Hedlund	Assistant Professor, UNLV, Nevada
Ben Stern	Paramedic, Sheridan, Oregon
Kevin Larson	Dental School Student, OHSU, Oregon

2000
September 2000 (71)

Sarah Boomer	RLMO PI, Associate Professor of Biology
Jessica Cameron	Public Health Department, Colorado
Jeannine Earnest	Biotechnology Production, GeneTools LLC, Corvallis, Oregon
Michelle Hase	Biotechnology Research, AVI BioPharma, Corvallis, Oregon
Danny Lodge	Former undergraduate, current research assistant, WOU, Oregon
Nicole Mullins	Continuing Biology Education Degree, WOU, Oregon
Eric Stroup	Biotechnology Productions, Bend Resaerch, Bend, Oregon
Mandi Ziglinski	Middle School Science Teacher, Mt. Angel, Oregon

↑ Figure 4. Sample Public Team Collection Site. Because the RLMO Database project is supported by RUI category funding, we archive information (names and current location) about undergraduate collection team members involved in field collection. Individual team collector information is archived and administratively assembled into team tables, linked to specific years and RLMO sites. For this figure, team collection information for Hillside Springs is shown for 1998-2000.

that summarize clone information (inferred identity and/or accession numbers) as a function of site (**Figure 5, Panel D**). In both query formats, all site names contain active links back to individual Site Data Forms.

4.0 CONCLUSIONS AND FUTURE DIRECTIONS

The RLMO Database Project represents the product of two undergraduates, both supported via modest NSF Microbial Observatory-derived summer stipends. At the inception of this project, the primary investigator, used to dealing only with simple Excel™ and HTML tables, did not

↑ **Figure 5.** **Sample Query and Data Output**. Panels A and B show sample site query and data. In this query, we retrieved three site parameters (acid/base, temperature, and aluminum) from two RLMO sites (Fairy and Hillside Springs) for the years 2002-2003. Panels C and D illustrate sample sequence query and data. In this query, we retrieved all Gram Positive, Chloroflexi, and Proteobacteria sequences from two RLMO sites (Fairy and Hillside Springs). Panel D shows the query results for the Panel C query. Data tables display accession numbers, each of which is linked directly to NCBI GenBank.

truly grasp the sheer utility of what such an online database could bring to the teaching and research lab setting. Three years later, the RLMO database provides the centerpiece of a long-term geochemical and molecular monitoring project that seeks to examine how site-specific chemistry affects genetic variation and selection of Red Chloroflexi community members. At this writing, basic site information (location, pH, temperature, and site images) for 17 RLMO communities has been entered into this database. Nearly 500 16S sequences from 11 different RLMO sites have been archived, alongside chemical assessments for four sites over two years.

The RLMO Database was not modeled after any particular scientific database because, when we began this project in 2000, we found no online database examples that sought to archive what we wanted to develop. Like many other molecular ecology-oriented researchers, we admired the immense power of the NCBI database but were frustrated with shortcomings of GenBank flatfiles in linking site- or collection-dependent information with retrieved molecular sequences (well stated in Sheldon et al. 2002). Shortly after the first version of the RLMO Database (September 2000), we learned of the exemplary Sapelo Island Microbial Observatory (SIMO) Database (**http://simo.marsci. uga.edu/mainWeb/pages/database.htm**), written using Microsoft SQL Server™, and made similar upgrades to our database. Although universally accessible databases seem an ideal answer to sharing and analyzing a variety of data from common places, a simple comparison of methods, equipment, kinds of site data, microenvironments, etc., quickly demonstrates that attaining common researcher standards is highly challenging. Moreover, the cost of setting up and maintaining meaningful small project-driven databases is such that project-specific database development should not be prohibitive for any research lab to undertake.

The future of the RLMO database includes annual data entry by researchers and undergraduates. While major structural elements of the database must remain constant, we envision modifications and improvements that will likely include the following features. First, as new information is published or reported, we anticipate adding new collection parameters (e.g., other chemicals to our current survey). Second, we plan to develop a query-based reporting system that will compile site and sequence data reports into simple flatfiles for easy viewing. Although such display options would provide convenient access to information, we believe that it is inappropriate to assume that general site parameters (e.g., pH, temperature, chemical signatures) are precisely localized to the microenvironment where a given organism and its 16S gene were located at the time of collection. Third, we plan to include more sequence-linked data entry fields (e.g., DNA isolation method, primers). Finally, we plan to develop a separate data-entry access site in order for other researchers to add single Red Chloroflexi sequences and simple site data to this collection. We anticipate this modification would be limited yet enable us to integrate important Red Chloroflexi data from Yellowstone and around the world.

REFERENCES

Baxevanis, A.D., and B.F.F. Ouellette. 1998. *Bioinformatics: A Practical Guide to the Analysis of Genes & Proteins*. John Wiley & Sons.

Bonheyo, G.T., B.W. Fouke, J. Frias-Lopez, and B. Sanzenbacher. 2001. Microbial 16S rRNA diversity within travertine depositional facies from 73 degrees to 25 degrees Celsius at Angel Terrace, Mammoth Hot Springs, Yellowstone National Park, USA. unpublished GenBank submissions.

Boomer, S.M., D.P. Lodge, B.E. Dutton, and B. Pierson. 2002. Molecular characterization of novel red green nonsulfur bacteria from five distinct hot spring communities in Yellowstone National Park. *Appl Environ Microbiol* 68:346-55.

Boomer, S.M., B.K. Pierson, R. Austinhirst, and R.W. Castenholz. 2000. Characterization of novel bacteriochlorophyll-a-containing red filaments from alkaline hot springs in Yellowstone National Park. *Arch Microbiol* 174:152-61.

Castenholz, R.W. 1984. *Microbial Mats: Stromatolites*. New York: Alan R. Liss, Inc.

Hanada, S., S. Takaichi, K. Matsuura, and K. Nakamura. 2002. *Roseiflexus castenholzii* gen. nov., sp. nov., a thermophilic, filamentous, photosynthetic bacterium which lacks chlorosomes. *Int J Syst Bacteriol* 52:187-193.

Nubel, U., M.M. Bateson, V. Vandieken, A. Wieland, M. Kuhl, and D.M. Ward. 2002. Microscopic examination of distribution and phenotypic properties of phylogenetically diverse *Chloroflexaceae*-related bacteria in hot spring microbial mats. *Appl Environ Microbiol* 68:4593-603.

Oyaizu, H., B.A. Debrunner-Vossbrinck, L. Mandelco, J.A. Studier, and C.R. Woese. 1987. The green non-sulfur bacteria: a deep branching in the eubacterial line of descent. *Appl Microbiol* 9:47-53.

Pierson, B.K. and R.W. Castenholz. 1992. The family Chloroflexacae. In *The Prokaryotes*, 2nd ed.. A. Balows, H. Truper, M. Dworkin, W. Harder, and K.H. Schleifer. New York: Springer, Berlin Heidelberg.

Pierson, B.K., S.J. Giovannoni, and R.W. Castenholz. 1984. Physiological ecology of a gliding bacterium containing bacteriochlorophyll a. *Appl Environ Microbiol* 47:576-84.

Pierson, B.K., S.J. Giovannoni, D.A. Stahl, and R.W. Castenholz. 1985. *Heliothrix oregonensis*, gen. nov., sp. nov., a phototrophic filamentous gliding bacterium containing bacteriochlorophyll a. *Arch Microbiol* 142:164-7.

Schopf, J.W., and B.M. Packer. 1987. Early Archean (3.3-billion to 3.5-billion-year-old) microfossils from Warrawoona Group, Australia. *Science* 237:70-3.

Sheldon, W.M., Moran, M.A., and Hollibaugh, J.T. 2002. Efforts to link ecological metadata with bacterial gene sequences at the Sapelo Island Microbial Observatory. Proceedings of the 6th World Multiconference on Systematics, Cybernetics, and Informatics.

Stoner, D.L., M.C. Geary, L.J. White, R.D. Lee, J.A. Brizzee, A.C. Rodman, and R.C. Rope. 2001. Mapping microbial biodiversity. *Appl Environ Microbiol* 67:4324-8.

Walsh, M.M., and D.R. Lowe. 1985. Filamentous microfossils from the 3500-myr-old Onverwacht Group, Barberton Mountain Land, South Africa. *Nature* 341:530-32.

Ward, D.M., M.J. Ferris, S.C. Nold, and M.M. Bateson. 1998. A natural view of microbial biodiversity within hot spring cyanobacterial mat communities. *Microbiol Mol Biol Rev* 62:1353-70.

Weller, R., M.M. Bateson, B.K. Heimbuch, E.D. Kopczynski, and D.M. Ward. 1992. Uncultivated cyanobacteria, *Chloroflexus*-like inhabitants, and spirochete-like inhabitants of a hot spring microbial mat. *Appl Environ Microbiol* 58:3964-9.

Woese, C.R. 1987. Bacterial evolution. *Microbiol Rev* 51:221-71.