Journal of the Mississippi Academy of Sciences

Volume 48 July 2003 Number 3

Journal of the Mississippi Academy of Sciences

Editor
Kenneth J. Curry
University of Southern Mississippi

Associate Editors
Maria F.T. Begonia
Jackson State University

Maureen Corcoran
Waterways Experiment Station

Ibrahim O. Farah
Jackson State University

Timothy C. Lockley
USDA APHIS PPQ IFA

Todd G. Nick
Univ. of Mississippi Medical Center

Robin Rockhold
Univ. of Mississippi Medical Center

Abstracts Editor
John Boyle
Mississippi State University

The Journal of the Mississippi Academy of Sciences (ISSN 0076-9436) is published in January (annual meeting abstracts), April, July, and October, by the Mississippi Academy of Sciences. Members of the Academy receive the journal as part of their regular (non-student) membership. Inqueries regarding subscriptions, availability of back issues, and address changes should be addressed to The Mississippi Academy of Sciences, Post Office Box 55709, Jackson, MS 39296; 601-977-0627; msacad@bellsouth.net.

Manuscripts and inquiries about publication and information about advertising should be sent to the editor: Kenneth J. Curry, University of Southern Mississippi, Post Office Box 5018, Hattiesburg, MS 39406-5018.
601-266-4930 (voice & fax)
kenneth.curry@usm.edu.

Contents

General Article

132 **Life on Mars: Past, Present, and Future** (2003 Dodgen Lecture)—Christopher P. McKay

Research Articles

138 **Global Analysis of Microbial Translation Initiation Regions**— Alan P. Boyle and John A. Boyle

151 **A Simple Method of Surveying Plant Populations for Random Amplified DNA Polymorphisms**—Janet Marie Gibson McDaniel and Robert G. Hamilton

Departments

155 **Executive Officer's Column**—Charles Swann

155 **Divisional Report**—Jack Moody

158 **2004 Annual Meeting Abstract Form and Instructions**

OFFICERS OF THE MISSISSIPPI ACADEMY OF SCIENCES

President	Hamed Benghuzzi
President-Elect	Sarah Lea McGuire
Immediate Past-President	Robert C. Bateman, Jr.
Executive Officer	Charles Swann
Journal Editor	Kenneth J. Curry
Junior Academy Director	Aimée T. Lee
Directors	Dick Highfill
	Maria Begonia
	Roy Duhé
Administrative Assistant	Cynthia Huff

Election Results

We welcome Sarah Lea McGuire as our new President-elect. Roy Duhé and Maria Begonia will join our Board of Directors. Roy replaces Cliff Ochs who is cycling off; Maria replaces Sarah Lea as she assumes her responsibilities as President-elect.

The *sixty-eighth* annual meeting of the

Mississippi Academy of Sciences

will be held on
Thursday and Friday,
February 19 and 20, 2004

Biloxi, Mississippi

CONSIDERING GRADUATE SCHOOL?

Check out the

Department of Chemistry and Biochemistry

at

The University of Southern Mississippi

WWW.CHEM.USM.EDU

Our graduates get excellent jobs and postdoctoral positions.

Are you ready to turn PRO?

Ohaus introduces style and performance on a whole new level — the Scout™ *Pro*.

The new Ohaus Scout *Pro* continues in the tradition of our classic Scout balance. For quality, performance — *and value* — it's unbeatable! It lets students start weighing with almost no instruction, and it's loaded with features like parts counting with auto optimization, display hold, % weighing, totalization, and Newtons. Plus Scout *Pro* comes with a density determination kit for automatic specific gravity calculations. With an optional USB and RS232 interface kit and available software, it's simple to connect Scout *Pro* to a computer or PDA. All this, and our industry-leading service and support, makes Scout *Pro* the new star of the winning Ohaus team.

OHAUS**PRO**OHAUS**PRO**OHAUS**PRO**OHAUS**PRO**OHAUS**PRO**OHAUS**PRO**OHAUS**PRO**

OHAUS®

Get the Advantage!

www.ohaus.com

Life on Mars: Past, Present, and Future

Christopher P. McKay
Space Science Division, NASA Ames Research Center, Moffett Field CA 94035
cmckay@mail.arc.nasa.gov

Mars appears to be a cold, dry, and dead world. However there is good evidence that early in it's history it had liquid water and possibly life. The main question in the future exploration of Mars is the search for an independent origin of life on that planet. Ecosystems in cold, dry locations on Earth, such as the Antarctic dry valleys, provide examples of how life on early Mars might have survived and where to look for fossils. Fossils are not enough. We will want to determine not only if life on Mars, but if that life was a separate genesis from life on Earth. For this determination we need to access intact martian life, possibly frozen in the deep old permafrost. It is possible that at some time in the future we might recreate habitable conditions on Mars, returning it to the life-bearing state it may have enjoyed early in its history. Our studies of Mars are still in a preliminary state but everything we have learned suggests that it may be possible to restore Mars to a habitable climate.

Mars, and the possibility of life on Mars, has always held a special fascination for both scientists and the general public. This traces back at least to the advent of the telescope. When viewed through even a small telescope Mars shows features reminiscent of Earth; these include polar caps that grow in the winter and recede in the summer, and dark areas in mid latitudes that shift with the seasons. It is not surprising that some early observers assumed that these polar caps were made of ice and that the dark features were due vegetation. The idea of Earth-like conditions and life was proven incorrect by spacecraft data; the winter polar cap is CO_2 and the dark features are caused by dust. Nonetheless, the spacecraft data still show that of all the other planets Mars is the one most likely to have a biological past, present, and future.

A comparison between the conditions on the surface of Mars and Earth is shown in Table 1. The atmosphere of Mars is over 100 times thinner than the atmosphere of the Earth and it is composed predominately of CO_2. Because of the thinner atmosphere and the increased distance from the sun, the mean temperature on Mars is –60°C, compared to +15°C for Earth. The lower gravity on Mars is due to its lower mass (1/10 the mass of Earth). Another important result of this lower mass is the absence of plate tectonics. The lack of recycling by plate tectonics is thought to be ultimately responsible of the decay of the martian atmosphere from its initially think, relatively warm, state to the thin desert conditions of today (Kasting et al., 1988; McKay, 1997).

Table 1. Comparison of Mars and Earth Surface Conditions.

Parameter	Mars	Earth
Surface pressure	0.5 to 1 kPa	101.3 kPa
Temperature range	-130°C to +15°C	-60°C to +50°C
Temperature average	-60ºC	+15ºC
Composition	95% CO_2	78% N_2
	2.7% N_2	21% O_2
	1.6% Ar	1% Ar
Gravity	0.38 g	1 g
Distance from Sun	1.52 A.U.	1 A.U.
Tilt of axis	25°	23.5°
Length of mean solar day	24 hr 39.6 min	24 hr
Length of year	1.88 yr	1 yr

In 1976, the two landers of the Viking Mission reached the surface of Mars to search for evidence of life. Each lander contained three biology experiments and an instrument to detect and characterize organic material (a combination gas chromatograph mass spectrometer).

The three Viking biology experiments were: (1) the pyrolytic release experiment (PR) which sought to detect the ability of microorganisms in the soil to consume CO_2 using light (Horowitz and Hobby, 1977); (2) the gas exchange experiment (GEx) which searched for gases released by microorganisms when organic nutrients were added to the soil (Oyama and Berdahl, 1977); and (3) the labeled release experiment (LR) which sought to detect the release of CO_2 from microorganisms when radioactively labeled organic nutrients were added to the sample (Levin and Straat, 1977).

Both the GEx and the LR experiment gave interesting results. When moisture was added to the soil in the GEX experiment, $O_2$was released. The release was rapid and occurred when the soil was exposed just to water vapor. The response persisted even if the soil was heated to sterilization levels (Oyama and Berdahl, 1977). The LR experiment indicated the release of CO_2 from the added organics. This response was not present when the soil was heated to sterilization levels (Levin and Straat, 1977). The release of O_2 in the GEX experiment was not indicative of a biological response. On the other hand, the LR results were precisely what would be expected if microorganisms were present in the martian soil.

However a biological interpretation of the LR results is inconsistent with the results of the GCMS. The GCMS did not detect organics in the soil samples at the level of one part per billion (Biemann et al., 1977; Biemann, 1979). One ppb of organic material would represent a large number of individual cells existing alone in the soil (Klein, 1978; 1979). However it is not likely that there are microorganisms in the soil on Mars without associated extracellular organic material. The lack of detection of organics is the main reason for the prevailing view that non-biological factors were the cause of the reactivity of the martian soil.

It is perhaps not surprising that Mars lacks life on its surface because there is no liquid water present at any place or at any time. There are regions on Mars where the pressure and temperature is consistent with liquid water stability (Haberle et al., 2001; Lobitz et al., 2001) but at these locations no water or ice is present. At the locations were ice is present the temperature or pressure are too low to allow for liquid. All known life requires liquid water to grow or reproduce and the absence of liquid water on Mars is consistent with the absence of life there… today.

PAST HABITABILITY

Mars may be cold and dry today but there is compelling evidence that earlier in its history Mars did have liquid water. This evidence comes from the images taken from orbital spacecraft. Figure 1 shows an image of a water carved canyon on Mars and represents probably the best evidence for extended and repeated, if not continuous, flow of liquid water on Mars (Carr, 1996; Carr and Malin, 2000).

Water is the quintessence of life, and the evi-

Figure 1. Liquid water on another world. Mars Global Surveyor image showing Nanedi vallis in the Xanthe Terra region of Mars. Image covers an area 9.8 km by 18.5 km; the canyon is about 2.5 km wide. This image is the best evidence we have of liquid water anywhere outside the Earth. Photo from NASA/Malin Space Sciences.

dence that sometime in its early history Mars had liquid water is the primary motivation for the search for evidence of life (McKay, 1997). Liquid water flowed on the surface of Mars long ago and the most likely place to search for evidence of any life associated with that water is in ancient lakebeds.

Cabrol et al. (1998) have shown that Gusev Crater (Figure 2) was probably an ice-covered lake and that the material in the crater represents sediments deposited in this lake. Kempe and Kazmierczak (1997) have argued that the delta deposits which can be seen near the mouth of the river (Ma'adim Vallis) as it enters Gusev Crater may be carbonates. This large flat crater makes a promising target for a search for microfossils on Mars.

However Gusev Crater poses an apparent paradox. As seen in Figure 1, Gusev Crater was a large lake (160 km diameter) feed by a large river but the surrounding terrain appears heavily cratered and unaltered. The preservation of this cratered terrain is inconsistent with the rain that would be expected to be associated with such a large river and lake. The paradox of rivers and lakes without rain can be resolved by considering the only place on Earth where rivers and lakes exist without rain; the dry valleys of the Antarctic.

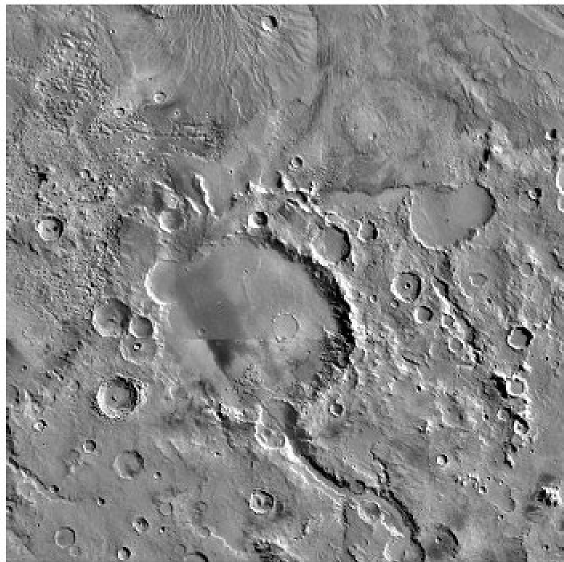

Figure 2. The search for life on Mars: Step 1 search for fossils in Gusev crater on Mars. A good site for the first step in the search for life is Gusev Crater (15°S, 185°W). Gusev is a 160 km diameter crater into which a large river once flowed producing a lake in the crater and carrying in sediments that have filled the crater floor. The resulting formation is a large flat plain easily targeted and easily navigated by a rover. It would be possible to land a rover in the center of the crater and then move toward the shoreline and up the river bed, all the while searching for fossils in the sediments. Such fossils would demonstrate that Mars once had life. NASA photo.

The dry valleys of Antarctica comprise the largest ice-free region on that continent. The valleys are a cold desert environment with mean annual temperatures of –20°C (Doran et al., 2002). Figure 3 shows a view of upper Wright Valley with Lake Vanda and the Onxy River on the valley floor. Precipitation in the dry valleys occurs only as snow and the amount is equivalent to about 1–2 cm of water. The hydrological cycle in these valleys is as follows: (1) Snow falls. (2) In the lower regions of the valleys the snow evaporates with little liquid formation or erosion. In the higher elevations snow accumulates and forms glaciers. (3)The glacial ice flows downward in the valley (as can be seen in Figure 3). (4) The lower parts of the glaciers melt in the summer providing a source of meltwater that then flows in the lakes (for Lake Vanda this is via the Onyx River). (5) The liquid water freezes to the underside of the ice cover. (6) Ablation and evaporation from the top of the ice cover returns moisture to the atmosphere and completes the cycle. This hydrological cycle has rivers in summer and lakes all year round but no rain. The minimum requirements for such a cycle are atmospheric pressures well above the triple point of water (610 Pa) and summer temperatures above freezing. McKay and Davis (1991) have shown that such conditions could have pre-

Figure 3. Aerial view of Lake Vanda, in Wright Valley Antarctica. The lake is about 5 km across. The Onyx River, which flows only for a few weeks each summer, can be seen flowing into the far end of the lake. Photo by J. Robie Vestal.

vailed on Mars for many hundreds of millions of years early in its history.

The lakes in the dry valleys of Antarctica provide an example of the physical processes that can maintain large bodies of liquid water under mean annual temperatures well below freezing. Biologically these lakes are also important analogs because of the plankton and benthic communities of microorganisms that thrive there (Parker et al., 1982; Wharton et al., 1982). Life could have existed in lakes on Mars in similar ecological conditions.

One interesting feature of the dry valley lakes visible in Figure 3 is the evidence of ancient shore levels. In these ancient lake deposits evidence of life is present and provide a basis for considering how remnants of life in martian lakes might be preserved (Wharton et al., 1995; Doran et al., 1998).

THE SEARCH FOR A SECOND GENESIS

The discovery of fossilized evidence for life in a dry lakebed on Mars would be of great interest. However it would not directly address the fundamental question of life on Mars, that is: was there a separate genesis of life on Mars. It is now known that rocks can be ejected from Mars and carried to Earth and it is assumed that the reverse is possible as well. Work on the preservation of magnetic signatures shows that the temperature in these rocks never reaches temperatures high enough to sterilize them (Weiss et al., 2000). Thus it is possible that Mars and Earth are not biologically isolated and share a common origin of life.

To determine if life on Mars is a second genesis requires more than fossils, it requires access to the organic remains of an actual martian organism (McKay, 2001). Life on Earth has a distinctive and universal signature both in its "hardware" and in its "software." The hardware of Earth-life is composed of 20 L amino acids, 5 nucleotide bases that appear in DNA and RNA, and a few D sugars the compose the polysaccharides. The common software of Earth-life is most clearly seen in the conserved sequences in the ribosomal RNA that demonstrate the phylogenetic connectivity of all life on Earth on the so-called "tree of life" (Woess, 1987). To determine if martian organism represent a second genesis of life, the hardware and software of martian life must be compared to that of Earth life.

It is unlikely that preserved organic remains of martian organisms will be found in paleo lake sediments in the equatorial regions. These locations may hold only fossils. A possible site for finding preserved organic remains of past martian life is in the polar regions. The most promising site may be at 80°S, 180°W located in the heavily cratered highlands. Locations closer to the polar cap are covered with the relatively young polar layered deposits. At this location there is also the presence of strong crustal magnetism (Connerney et al., 1999). These features were presumed to be caused by an early martian magnetic field. The apparent erasure of these magnetic features in the vicinity of the Halles and Argyre impacts suggests that such features are old, possible the oldest phenomenon detected on Mars. Feldman et al. (2002) have reported ice rich ground in the southern polar regions. Thus the ice and sediments at 80°S, 180°W may represent the oldest, coldest, undisturbed ice-rich permafrost on Mars. Here we might find remnants of martian microorganisms frozen since the early water-rich period of martian history.

FUTURE HABITABILITY

The evidence that Mars had a more clement environment early in its history has lead to the suggestion that it might be possible to restore that planet to habitable conditions (McKay et al., 1991; McKay and Marinova, 2001). Climate models for warming Mars suggest that a thick atmosphere composed primarily of carbon dioxide could be recreated on Mars is a hundred years or so (McKay and Marinova, 2001). Creating an oxygen rich atmosphere is not possible with foreseeable technologies in time scales that are comparable to human lifetimes (McKay and Marinova, 2001).

To make a biosphere on Mars requires three basic ingredients in planetary amounts: water, CO_2 and N_2. The total inventory of these key compounds on Mars is uncertain (McKay et al., 1991) but the range of plausible values include the amounts needed for terraforming. A key science goal for the exploration of Mars will be to determine the location and extent of subsurface reservoirs of these compounds.

The first step in terraforming Mars would be to simply warm its surface. Several methods have been suggested for warming Mars. Perhaps the most practical is the production of compounds that are extremely efficient greenhouse gases (Marinova et al., 2000). These include PFCs, nitrous oxide, sulfur hexafluoride and methane. When combined with water and carbon dioxide these gases can effectively block infrared radiation from leaving the surface,

creating a strong greenhouse warming. Calculations suggest (Marinova et al., 2000) that total concentrations of a few ppm or less could warm Mars to Earth-like conditions. Even at the ppm level the total mass of gases produced is much too large to transport from Earth and would have to be produced in situ. The essential elements are known to be present on the surface of Mars. As greenhouse gases are produced on Mars and the surface warms, any carbon dioxide adsorbed into the regolith or frozen onto the polar caps would evaporate into the atmosphere. This positive feedback would accelerate and amplify the warming from the greenhouse gases.

The warming of Mars due to the greenhouse effect could proceed quite rapidly. If there was enough carbon dioxide to provide for a thick atmosphere and if the super greenhouse gases were produced on Mars it would not take very long for Mars to warm appreciably. If every single photon reaching Mars from the sun were used to warm Mars it would have a warm surface in only 10 years. Another 50 years would be enough to melt a layer of water 500 meters thick over the entire planet—a very suitable ocean. However, it is not possible to trap solar energy with 100% efficiency. Even with a strong greenhouse effect a value near 10% is more probable. With 10% efficiency it would take 100 years to warm the surface of Mars and an additional 500 years to melt its ocean. The melting of the martian permafrost to release an ocean of water may take longer than the calculation above suggests. The surface warmth would have to diffuse down into the soil. To warm the subsurface to a kilometer depth simply by diffusion would take about 100,000 years. However, passive diffusion of heat would be augmented by active transport due to the flowing of the resulting liquid water. Thus Mars could have a thick warm atmosphere in about 100 years and a fully formed ocean-biosphere in 600 years. Not a long time even on human time scales.

The production of an oxygen-rich atmosphere suitable for humans to breath is much more energy intensive than just warming the planet. The conversion of an atmosphere of carbon dioxide to organic material and the concomitant production of oxygen required billions of years on the Earth and was due to photosynthesis by primitive algae. On Mars, plants would also provide a global-scale, self-replicating mechanism for the production of oxygen. Although, plants have perfected their biochemical technique over billions of years of evolution, their efficiency at incorporating the global average solar radiation into biomass is still only 0.01%. Even optimistically assuming that this can be improved by an order of magnitude by selective breeding and genetic engineering, it will take 10,000 years for an oxygen rich atmosphere to be produced. Thus, humans on Mars may doff their space suits in the relatively near future but they will require their oxygen supplies for many generations.

CONCLUSIONS

Mars presents a challenge and an opportunity. The challenge is to explore a distant planet with a complex history, first with robotic probes and eventually with human explorers. The opportunity is to learn about the nature of life, to search for a possible second type of life in our own solar system and thereby begin to understand the profound philosophical and scientific issues related to life in the universe. The ultimate challenge on Mars will be the reconstruction of a biosphere on that world. A task of restoration ecology worthy of the best efforts of a space-faring humanity.

LITERATURE CITED

Biemann, K., J. Oro, P. Toulmin III, L.E. Orgel, A.O. Nier, D.M. Anderson, P.G. Simmonds, D. Flory, A.V. Diaz, D.R. Rushneck, J.E. Biller, and A.L. LaFleur. 1977. The search for organic substances and inorganic volatile compounds in the surface of Mars. J. Geophys. Res. 82:4651–4658.

Biemann, K. 1979. The implications and limitations of the findings of the Viking Organic Analysis Experiment. J. Molec. Evol. 14: 65–70.

Cabrol, N.A., E.A. Grin, R. Landheim, R.O. Kuzmin, and R. Greeley. 1998. Duration of the Ma'adim Vallis/Gusev Crater hydrogeologic system, Mars. Icarus 133:98–108.

Carr, M.H. 1996. Water on Mars. Oxford University Press, New York.

Carr, M.H., and M.C. Malin. 2000. Meter-scale characteristics of martian channels and valleys. Icarus 146:366–386.

Connerney, J.E.P., M.H. Acuna, P. Wasilewski, N.F. Ness, H. Reme, C. Mazelle, D. Vignes, R.P. Lin, D. Mitchell, and P. Cloutier. 1999. Magnetic lineations in the ancient crust of Mars. Science 284:794–798.

Doran, P.T., R.W. Wharton, D. DesMarais, and C.P. McKay. 1998. Antarctic paleolake sediments and the search for extinct life on Mars. J. Geophys. Res. 103:28481–28493.

Doran, P.T., J.C. Priscu, W.B. Lyons, J.E. Walsh, A.G. Fountain, D.M. McKnight, D.L. Moorhead, R.A. Virginia, D.H. Wall, G.D. Clow, C.H. Fritsen, C.P. McKay, and A.N. Parsons. 2002. Antarctic climate cooling and terrestrial ecosystem response. Nature 415:517–520.

Feldman, W.C., W.V. Boynton, R.L. Tokar, T.H. Prettyman, O. Gasnault, S.W. Squyres, R.C. Elphic, D.J. Lawrence, S.L. Lawson, S. Maurice, G.W. McKinney, K.R. Moore, R.C. Reedy,. 2002. Global distribution of neutrons from Mars: Results from Mars Odyssey. Science 297:75–78.

Haberle, R.M., C.P. McKay, J. Schaeffer, N.A. Cabrol, E.A. Grin, A.P. Zent, and R. Quinn. 2001. On the possibility of liquid water on present-day Mars. J. Geophys. Res. 106:23317–23326.

Horowitz, N.H., and Hobby, G.L. 1977. Viking on Mars: The carbon assimilation experiments. J. Geophys. Res. 82:4659–4662.

Kasting, J. F., O.B. Toon, and J.B., Pollack 1988. How climate evolved on the terrestrial planets Scientific American 258(2):90–97.

Kempe, S., and J. Kazmierczak . 1997. A terrestrial model for an alkaline martian hydrosphere. Planet. Space Sci. 45:1493–1499.

Klein, H.P. 1978. The Viking biological experiments on Mars. Icarus 34:666–674.

Klein, H.P. 1979. The Viking mission and the search for life on Mars. Rev. Geophys. Space Phys. 17:1655–1662.

Levin, G.V., and Straat, P.A. 1977. Recent results from the Viking Labeled Release Experiment on Mars. J. Geophys. Res. 82: 4663–4667.

Lobitz, B., B.L. Wood, M.A. Averner, and C.P. McKay. 2001. Use of spacecraft data to derive regions on Mars where liquid water would be stable. Proc. Nat. Acad. Sci. 98:2132–2137.

Marinova, M.M., C.P. McKay, and H. Hashimoto. 2000. Warming Mars using artificial super-greenhouse gases, J. Brit. Interplanet. Soc. 53:235–240.

McKay, C.P. 1997. The search for life on Mars. Origins Life Evolution Biosph. 27:263–289.

McKay, C.P. 2001. The search for a second genesis of life in our Solar System, Pages 269–277 *in* J. Chela-Flores, ed. First Steps in the Origin of Life in the Universe.

McKay, C.P. and M.M. Marinova. 2001. The physics, biology, and environmental ethics of making Mars habitable. Astrobiology 1:89–109.

McKay, C.P., and W.L. Davis. 1991. Duration of liquid water habitats on early Mars. Icarus 90:214–221.

McKay, C.P., O.B. Toon, and J.F. Kasting. 1991. Making Mars habitable. Nature 352:489–496.

Oyama, V.I., and B.J. Berdahl. 1977. The Viking gas exchange experiment results from Chryse and Utopia surface samples. J. Geophys. Res. 82:4669–4676.

Parker, B.C., G.M. Simmons, Jr., K.G. Seaburg, D.D. Cathey, and F.T.C. Allnutt. 1982. Comparative ecology of plankton communities in seven Antarctic oasis lakes. J. Plank. Res. 4:271–286.

Weiss B.P., J.L. Kirschvink, F.J. Baudenbacher, H. Vali, N.T. Peters, F.A. Macdonald, and J.P. Wikswo. 2000. A Low Temperature Transfer of ALH84001 from Mars to Earth. Science 290:791–795.

Wharton, R.A., Jr., B.C. Parker, G.M. Simmons, Jr., K.G. Seaburg, and F.G. Love. 1982. Biogenic calcite structures forming in Lake Fryxell, Antarctica. Nature 295:403–405.

Wharton, R.A., Jr., J.M. Crosby, C.P. McKay, and J.W. Rice, Jr. 1995. Paleolakes on Mars. J. Paleolimnology 13:267–283.

Woese, C.R. 1987. Bacterial evolution. Microbiol. Rev. 51:221–271.

Global Analysis of Microbial Translation Initiation Regions

Alan P. Boyle and John A. Boyle[1]
Department of Biochemistry and Molecular Biology,
Mississippi State University, Mississippi State, MS 39762

The availability of genomic sequences from multiple bacteria has allowed global comparisons of patterns. Here we present a graphical comparison of normalized base frequencies in the vicinity of translation starts for both eubacteria and archae. The results show that most eubacterial Open Reading Frames (ORFs) are preceded by a distinctly recognizable Shine-Dalgarno (SD) sequence pattern. However, some eubacteria deviate from this arrangement and have diminished SD patterns or completely lack this sequence. On the other hand, some archae seem to use both SD sequences and leaderless transcripts in their translation initiation process. This is dependent on the position of a gene within an operon. Most archae seem to have other regular sequences located upstream from the typical SD location. Both eubacteria and archae have a surprising repetitive pattern seen within the averaged ORFs. The eubacterial and archaeal averaged patterns are slightly different from each other, and individual organisms within each domain vary from the averages. Nevertheless, the existence of such a periodicity within ORFs may allow the development of new techniques to identify actual genes from ORFs.
Keywords: translation initiation, eubacterial, archaeal, Shine-Dalgarno, alignment

Eubacteria initiate their translational process by binding mRNA to the small ribosomal subunit. This occurs because of a complementarity between a sequence at the 3' end of 16S rRNA and the Shine-Dalgarno (SD) sequence just 5' to the initiation codon (Shine and Dalgarno, 1974; Gualerzi and Pon, 1990). However, rarely, some eubacteria have been shown to lack an untranslated leader of sufficient length to contain an SD sequence (Van Etten and Janssen, 1998). In addition, there is some evidence that *Mycoplasma* species may have a high proportion of leaderless transcripts. On the other hand, archae may more routinely use heterogenous mechanisms for translation initiation (Saito and Tomita, 1999). Work on two different species, *Pyrobaculum aerophilum* (Slupska et al., 2001) and *Sulfolobus solfataricus* (Tolstrup et al., 2000), has shown that while many genes seem to have SD sequences in the proper location, a significant number of others are likely to have leaderless transcripts.

The availability of large numbers of complete genomic sequences provides the opportunity to search for patterns within and among genomes. In particular there were 56 eubacterial and 11 archaeal sequences that were published by the middle of January, 2002 (http://www.ncbi.nlm.nih.gov:80/PMGifs/Genomes/micr.html). Newly sequenced genomes are subjected to computational methods that produce a collection of open reading frames (ORFs) that presumably correspond to the genes in the organisms. This type of analysis has already led to insights into the translational process (Saito and Tomita, 1999; Sakai et al., 2001; Ma et al., 2002).

We construct a matrix using a maximum likelihood statistical approach (Hertz and Stormo, 1996) and combine it with a graphical representation of the results to show results averaged over all ORFs for all available sequenced microbial genomes. This approach can reveal common patterns and deviations from these patterns for microorganisms. We discuss the results as they relate to translation initiation mechanisms.

MATERIALS AND METHODS

Sequences for the following genomes were available as of 01-22-02 at the Entrez Genomes section of the National Center for Biotechnology Information (NCBI) Web site (http://www.ncbi.nlm.nih.gov:80/PMGifs/Genomes/micr.html) (hereafter referred to as the NCBI Microbial Genomes website):

[1]Corresponding author. (FAX 662-325-8664; email jab@ra.msstate.edu)

Agrobacterium tumefaciens Cereon circular chromosome (Goodner et al., 2001)
Agrobacterium tumefaciens Dupont circular chromosome (Wood et al., 2001)
Aeropyrum pernix K1 (Kawarabayasi et al., 1999)
Aquifex aeolicus chromosome (Deckert et al., 1998)
Archaeoglobus fulgidus (Klenk et al., 1997)
Bacillus halodurans C-125 (Takami et al., 2000)
Bacillus subtilis (Kunst et al., 1997)
Borrelia burgdorferi chromosome (Fraser et al., 1997)
Brucella melitensis chromosome I, chromosome II (DelVecchio et al., 2002)
Buchnera sp. APS (Shigenobu et al., 2000)
Campylobacter jejuni (Parkhill et al., 2000)
Caulobacter crescentus (Nierman et al., 2001)
Chlamydophila pneumoniae CWL029 (Kalman et al., 1999)
Chlamydophila pneumoniae AR39 (Read et al., 2000)
Chlamydophila pneumoniae J138 (Shirai et al., 2000)
Chlamydia trachomatis (Stephens et al., 1998)
Chlamydia muridarum chromosome (Read et al., 2000)
Clostridium acetobutylicum chromosome (Nolling et al., 2001)
Deinococcus radiodurans R1 chromosome 1, chromosome 2 (White et al., 1999)
Escherichia coli K12 (Blattner et al., 1997)
Escherichia coli O157:H7 EDL933 (Perna et al., 2001)
Escherichia coli O157:H7 (Hayashi et al., 2001)
Halobacterium sp. NRC-1 (Fleischmann et al., 1995)
Haemophilus influenzae (Ng et al., 2000)
Helicobacter pylori 26695 (Tomb et al., 1997)
Helicobacter pylori J99 (Alm et al., 1999)
Lactococcus lactis subsp. *lactis* (Bolotin et al., 2001)
Listeria monocytogenes EGD-e (Glaser et al., 2001)
Listeria innocua (Glaser et al., 2001)
Methanobacterium thermoautotrophicum (Smith et al., 1997)
Methanococcus jannaschii chromosome (Bult et al., 1996)
Mesorhizobium loti chromosome (http://www.kazusa.or.jp/rhizobase/)
Mycobacterium tuberculosis H37Rv (Cole et al., 1998)
Mycobacterium tuberculosis CDC1551 (Fleischmann, R.D., D. Alland, J.A. Eisen, L. Carpenter, O. White, J. Peterson, R. DeBoy, R. Dodson, M. Gwinn, D. Haft, et al., Whole genome comparison of *Mycobacterium tuberculosis* clinical and laboratory strains. Unpublished (listed in NCBI Microbial Genomes website).
Mycobacterium leprae (Cole et al., 2001)
Mycoplasma genitalium (Fraser et al., 1995)
Mycoplasma pneumoniae (Himmelreich et al., 1996)
Mycoplasma pulmonis (Chambaud et al., 2001)
Neisseria meningitidis MC58 (Tettelin et al., 2000)
Neisseria meningitidis Z2491 (Parkhill et al., 2000)
Nostoc sp. PCC 7120 (Kaneko et al., 2001)
Pasteurella multocida (May et al., 2001)
Pseudomonas aeruginosa (Stover et al., 2000)
Pyrococcus abyssi (Heilig, R. *Pyrococcus abyssi* genome sequence: insights into archaeal chromosome structure and evolution Unpublished (listed in NCBI Microbial Genomes website).
Pyrococus horikoshii (Kawarabayasi et al., 1998)
Rickettsia conorii Malish 7 (Ogata et al., 2001)
Rickettsia prowazekii (Ogata et al., 2001)
Ralstonia solanacearum (Salanoubat, M., S. Genin, F. Artiguenave, J. Gouzy, S. Mangenot, M. Arlat, A. Billault, P. Brottier, J.C. Camus, L. Cattolico, et al., Genome sequence of the plant pathogen *Ralstonia solanacearum* Unpublished (listed in NCBI Microbial Genomes website).
Salmonella typhimurium LT2 (McClelland et al., 2001)
Salmonella typhi (Parkhill et al., 2001)
Sinorhizobium meliloti (Capela et al., 2001)
Staphylococcus aureus N315 (Kuroda et al., 2001)
Staphylococcus aureus Mu50 (Kuroda et al., 2001)
Streptococcus pneumoniae TIGR4 (Tettelin et al., 2001)
Streptococcus pneumoniae R6 (Hoskins et al., 2001)
Streptococcus pyogenes (Ferretti et al., 2001)
Sulfolobus solfataricus (She et al., 2001)
Sulfolobus tokodaii (Kawarabayasi et al., 2001)
Synechocystis PCC6803 (Kaneko et al., 1996)
Thermoplasma acidophilum (Ruepp et al., 2000)
Thermoplasma volcanium (Kawashima et al., 1999)
Treponema pallidum (Fraser et al., 1998)
Thermotoga maritima (Nelson et al., 1999)
Ureaplasma urealyticum (Glass et al., 2000)
Vibrio cholerae chromosome I, chromosome II (Heidelberg et al., 2000)
Xylella fastidiosa chromosome (Simpson et al., 2000)
Yersinia pestis chromosome (Parkhill et al., 2001)

Software was developed using Perl and run on Sun OS 5.7. Data was read from standard formatted data files found at the Entrez-Genome site hosted by

the NCBI (NCBI Microbial Genomes website) and the information needed to align the open reading frames was extracted. Each eubacterial and archaeal DNA sequence was downloaded in FASTA format (*.fna) and the open reading frame information was selected from the same site in ProTein Table (*.ptt) format. No reformatting of the data was required. The only input required by the software was the location of the genome files and the start/stop locations for the alignment (in this case -70 to +50 where 0 is the first base in the start codon). The program first removed all the end-line characters and calculated the G-C content of the strain. Next, each ORF segment was aligned with the start codon beginning at the zero location (ORFs indicated as being in the opposite direction were computed as reverse compliments and aligned). The bases at each position were then totaled by summing over all ORFs. The sums were divided by the total number of ORFs to find the real frequency and then normalized by dividing by the expected base content at each position by using the G-C content of the organism. Next, a log probability of any given base in the region was calculated by taking the log of these normalized values (Staden, 1984; Hertz and Stormo, 1996; Stormo, 2000). The values were then output into a data file. This was repeated for each genomic file. An average for each bacterial domain was calculated from all the frequencies in that domain. A total of 58 eubacterial sequences were run requiring 1598 seconds and totaling 146,335 ORFs. A total of 11 archaeal sequences were run requiring 235 seconds and totaling 23,361 ORFs.

Subtraction of individual organismal patterns from the domain average was accomplished by direct subtraction of frequencies at each base. The result was reploted to show the differences in frequencies.

Sulfolobus solfataricus and *Halobacterium* sp. NRC-1 genomes were examined to identify ORFs that were likely to be members of operons. In order to qualify, ORFs must have been assigned an identity at the NCBI Microbial Genomes web site and be part of a recognizable cluster of related genes. The selected genes were related by function (as in collections of ribosomal protein genes or subunits of a protein like NADH dehydrogenase) or by being part of a pathway (like proteins in the cobalamin biosynthetic pathway). The gene cluster had to be closely linked with fewer than ten bases separating the genes (many overlapped slightly). The initial gene in the putative operon had to be separated by 50 to 100 bases from the preceding gene and had to have no identifiable functional relationship with it. No hypothetical protein genes were used.

A version of the program used for this work is accessible at the following web site: http://www.msstate.edu/dept/biochemistry/CBIG/. In addition the aligned genomes of all microbial species examined are available at this site.

RESULTS

Examination of individual eubacterial genomic sequence patterns and the eubacterial average over all reported ORFs (Figure 1) reveals distinctive patterns near the start codon. The averaged start codons themselves show the expected overwhelming presence of A, T, and especially G in the third position (base number 2 in the representation used). There is a general enrichment of A's both upstream and downstream from the start codon with the obvious exception of the Shine-Dalgarno region and a decline in A as the last base before the start. There is also a general decrease in G's in this start codon proximal region except for SD. The expected SD sequences upstream from start contribute to a distinctive pattern. The canonical sequence of AGGAGG may vary somewhat in its location with respect to translation start. The net result is a distribution of this sequence over a range of locations in the genomic patterns and in the eubacterial average. The high G content of SD yields a broad G peak centered around -9 to -10 bases upstream from start. There is a rise in A content upstream of this peak and a definite decline within the peak. T and C frequencies drop off dramatically in the SD region.

Individual eubacterial genomes vary from the average pattern. A few eubacteria maintain a preponderance of A's over G's throughout the SD region. This is true for *Caulobacter crescentus*, *Mycobacterium tuberculosis*, and *Xylella fastidioasa* (Figure 2). These still have the decline in T's and C's as observed for others. Some eubacteria, e.g. *Bacillus subtilis* (Figure 3) and *Staphylococcus aureus* (see Figure 1), have SD regions whose center is shifted slightly upstream from the average and have a higher frequency of G's and lower frequency of T's and C's in this region as compared to the average. This is highlighted by subtracting the organismal pattern from the eubacterial average (Figure 3).

Synechocystis PCC6803 has a clear deviation from the average pattern (Figure 4). While there is a rise in G's at SD, it is not nearly so pronounced as in the average. The decline in C's is also not conspicu-

ous. There is no obvious explanation for this anomaly. *Deinococcus radiodurans* also presents an anomalous pattern (Figure 4). Instead of a G peak in the SD region, it displays a prominent A peak and only small decreases in T and C. There is a T peak at position -7. The patterns are consistent for both chromosomes.

Other notable exceptions from the eubacterial pattern are seen in two *Mycoplasma* species, *Mycoplasma genitalium* and *Mycoplasma pneumoniae* (Figure 4). While both evidence the general rise in A's and decline in G's near the start codon, each lacks evidence of SD sequences. In contrast *Mycoplasma pulmonis* fits the standard eubacterial pattern except that a larger proportion of its ORFs start with something other than ATG. Removal of the *M. genitalium* and *M. pneumoniae* frequencies from the eubacterial average has little affect on the pattern since together they only contribute 1173 ORFs out of the almost 150,000 found in the overall average.

The average over all archaeal sequences shows a much less uniform pattern than for eubacteria (Figure 5). While this could be due to the smaller sample size of archaeal sequences, the number of ORFs involved is still over 23,000. The variations are more likely due to the diverse nature of the archaeal kingdom. It is also likely due to diverse translation initiation mechanisms for different classes of genes within individual archae (Tolstrup et al., 2000; Slupska et al., 2001).

There are some similarities between the archaeal and eubacterial averages. There is an enrichment in A's immediately upstream and downstream from the start codon. The last base before this codon is depleted in A. There is a decrease in G before the start, but, in contrast to eubacteria, there is no general depletion in G after the start codon. While there is a clear G peak and decline in A's and C's in the SD region, there is no drop in T's here. In addition the SD pattern is not so pronounced as in the eubacterial case. Individual archae lack distinct SD patterns. *Thermoplasma acidophilum, Aeropyrum pernix*, and *Halobacterium* sp. NRC-1 have none of the standard base distributions here. On the other hand, with the exception of *Aeropyrum pernix*, they and the other archae have apparent consistent structure in the region upstream from the SD location that is seen in no eubacteria (Figure 5). From -27 to -30 there is a preponderance of T's and -19 to -26 and -31 to -35 shows a preference for A in all the archae. These regions seem to show a diminished amount of C and G although this is not pronounced.

Examination of the results for *Sulfolobus solfataricus* reveals some of the complexity of archaeal genes (Figure 6). Its genomic average is very similar to the overall archaeal average. However, a sampling of 63 genes that are likely to be transcribed as internal members of polycistronic mRNAs shows a pattern that presents a clear SD signal with little upstream structure. Sampling 40 genes that are likely to be the initial sequences in polycistrons presents a noisy average but reveals no SD pattern and suggests the possible presence of the A-T rich and C-G poor region from -35 to -19. Combination of the patterns from the two gene classes would yield the sequence structures found upstream from the translation start signal in the overall genomic pattern for *Sulfolobus*.

Sampling 72 genes that appear to be internal and 29 that appear to be the first in an operon in *Halobacterium* sp. NRC-1 gives a similar result (data not shown). An A-T rich region is seen from -38 to -25 with the typical T peak surrounded by A peaks when the operon initiating genes are averaged. A G rich region from -12 to -8 and correspondingly C poor region is suggestive of an SD region for genes internal to operons. This region is obscured in the overall average.

The gene sequences downstream from the start codon show a surprising regularity in all bacteria examined. This is not an artifact of the alignment; combination of random genomic sequences (data not shown) presents only noise. A similar regularity has been observed by others for individual genes and organisms (Fickett, 1984; Tsonis et al., 1991; Suckow et al., 1998).

DISCUSSION

The statistical and graphical approach used here shows averages of the sequenced eubacterial and archaeal genomes. Patterns that relate to the translational processes are readily visualized. Comparisons with individual organisms reveal possible deviations from the standard processes. As expected, most eubacteria have a distinct indication of a Shine-Dalgarno region located just upstream from the initiation codon. There is also a general enrichment of A's near the start with the exception of a clear decrease at position -1. There is no clear evidence of a downstream box that has been suggested to be a translation enhancer in the region of +7 to +12 (Sprengart et al., 1996).

Deviations from the average results for eubacteria

could be caused by several factors. Misidentification of ORFs by various gene finding programs could produce noise in the data. An ORF may be incorrectly extended in the 5' direction beyond the actual start codon simply because there is an open reading frame available beyond that codon. This would place the real SD site within the gene when the alignment is done. Some sequences identified as ORFs may not be real genes. This would also introduce noise in the alignment. Conceivably, some genes may contain translational elements that are quite different from the average. It might be an interesting application of this technique to identify ORFs that deviate in some statistical way from the average. Do they cluster in some fashion? Do they have unusual characteristics as genes or do they look like misassigned sequences?

Some organisms deviate in definite but not critical ways from the average. A few have a G peak in the SD region that is less than the A peak but still have clear decreases in T and C. *B. subtilis* and *S. aureus* have intense but shifted SD signals. This is most clearly seen by subtracting the average results from the organismal results. These variations from the average are not likely to represent large scale differences in translation initiation.

Some organismal averages are distinct from the eubacterial average. *Synechocystis* shows an exceptionally weak SD pattern and also has reasonably prominent C frequencies on either side of the start codon. Other work has suggested that this bacterium is somewhat different in its translation initiation and may have different classes of genes (Osada et al., 1999; Sakai et al., 2001). However, examination of only the highly expressed class of genes (Mrazek et al., 2001) (data not shown) does not show much variation from the overall average. *Deinococcus radiodurans* shows the typical lower values for C and T frequencies in the SD region, but there is an A peak here for both chromosomes. G tends to increase around -8 to -9. This is unlike any other eubacterial genome. Examination of the 16S ribosomal RNA genes in both *Synechocystis* and *Deinococcus* shows the expected AGGAGG sequence (NCBI Microbial Genomes website) and so altered complementarity is not to be expected. The regularity seen within the coding region is also clearly different for *Deinococcus*. These two eubacteria are not related in any fashion with *Synechocystis* being a cyanobacterium and *Deinococcus* being one of the most unusual eubacterial extremophiles (White et al., 1999). It will be interesting to compare these results with any newly sequenced cyanobacteria to see if a unique translation initiation pattern emerges.

The results seen with *Mycoplasma genitalium* and *Mycoplasma pneumoniae* highlight their high frequency of leaderless transcripts (Weiner et al., 2000). This phenomena, whereby an mRNA is produced with few bases in front of the start codon, has been observed in every type of organism but is usually rare (Van Etten and Janssen, 1998). Translation initiation seems to be accomplished through the actions of either bacterial or eukaryotic equivalents of IF-2 (Kyrpides and Woese, 1998; Grill et al., 2000; Grill et al., 2001). These two *Mycoplasma* species show no indication at all of an SD region. While *Synechocystis* and *Deinococcus* are unusual, they still evidence a diminished G peak; *M. genitalium* and *M. pneumoniae* have no distinguishable G peak. The other species of mycoplasma in this study, *Mycoplasma pulmonis* and the related organism *Ureaplasma urealyticum*, have a readily distinguishable SD pattern and therefore are not likely to have leaderless transcripts.

Archeal translation initiation is not as well understood as eubacterial initiation. Genomic analysis has suggested that at least some archae use a combination of eubacterial and eukaryotic mechanisms (Salin et al., 1991; Saito and Tomita, 1999). Some archaeal initiation factors are homologs of the eukaryotic equivalents (Bult et al., 1996; Keeling et al., 1998; Kyrpides and Woese, 1998) yet many archaeal genes have clear SD regions, appropriately complimentary regions at the 3' end of their 16S rRNA molecules, and no 5' CAP structure. Some archaeal transcripts are leaderless (May and Dennis, 1990; Condo et al., 1999; Slupska et al., 2001). Recent work has suggested that in at least some archea, the specific translation initiation mechanism depends on whether the gene in question is located internal to an operon or is an isolated gene or the first gene in an operon (Tolstrup et al., 2000; Slupska et al. 2001).

The results here reflect both the diversity of the archaeal domain and the diversity of mechanisms within individual organisms. Not all of the archae, especially *Halobacterium* and to a lesser extent *Thermoplasma acidophilum* and *Aeropyrum pernix*, have a clearly defined SD region. However, with the exception of *Aeropyrum pernix*, all the archaeal genomes have at least some indication of common sequence elements further upstream from an expected SD site. There are A rich regions located around -34 and -24 and a T rich region centered at -28. The T region and the -24 A region could readily

correspond to an A box structure (consensus TTTA(A or T)A) (Wich et al., 1986; Reiter et al., 1988; Thomm and Wich, 1988; Reiter et al., 1990) slightly diffuse out due to heterogeneity in location. This element represents a transcription start signal yielding leaderless transcripts having no SD region and is found upstream of isolated genes and first genes in operons in *Pyrobaculum aerophilum* (Slupska et al., 2001) and *Sulfolobus solfataricus* (Tolstrup et al., 2000; She et al., 2001). It was not found upstream of internal genes in operons. The results presented here for *Sulfolobus* and *Halobacterium* support this conclusion and lead to the suggestion that each genome result and the overall archaeal average represents a superposition of at least the two classes of genes. The SD region centered at -9 is from internal members of operons; the putative A box (the combination of the T peak at -28 and the A peak at -24) is the transcription signal for leaderless products of isolated genes and first genes in operons. The A region at -34 may be another type of transcriptional signal since it is likely to be too far upstream to be involved in translation initiation. It seems that most if not all archae studied here may use at least two mechanisms for translation initiation based on these results.

The repetitive pattern seen with the averaged coding sequences is consistently present in all organisms examined. It is especially interesting that a pattern is seen even when all ORFs over both the eubacteria and archae are averaged (Figures 1 and 5). However, the observed patterns are slightly different in each case. There is a clear periodicity of three bases seen and this is likely due to codon structure. Nevertheless, there is no reason to expect such a regularity in base positions. Both eubacteria and archae have higher than expected G frequencies in the first codon position and lower than expected in the second position. This begins immediately after the start codon in archae but becomes prominent in eubacteria only after several codons. Conversely, T is elevated in the second position but diminished in the first position. The other bases show lower effects in the averages. Individual organisms can show patterns that are distinct from the averages. Figure 2 shows that *Caulobacter crescentus* has regularly higher C and lower A in the third position. Figure 5 shows that *Halobacterium* has higher C and lower T in the third position and both higher A and T in the second position. The set of all 69 organisms examined has much variety but each member always shows some regular pattern.

The pattern observed internal to the ORFs is not seen upstream nor is it seen in random sequences. The pattern is observable even if few ORFs are examined as is seen in Figure 6 when 63 and 40 ORFs are aligned. In particular, the G repeat stands out from the noisy background. Preliminary work shows that alignment of multiple codons within one gene also shows periodicity. This agrees with work done for some genes in *Pyrococcus* (Suckow et al., 1998). There has been some suggestion that examination of periodicity within ORFs could assist in gene identification (Tiwari et al., 1997). It remains to be seen whether eukaryotes will have similar patterns.

Detailed analysis should allow comparison of individual ORFs to organismal averages. This may prove useful in identifying ORFs that represent outliers in the data. Are they atypical because they are members of a particular gene class? The *Sulfolobus* and *Halobacterium* genes that begin operons show this characteristic. On the other hand, further study of some abnormal ORFs might eventually show that they have been incorrectly identified as genes.

The combination of a visual approach with a maximum likelihood statistical treatment of aligned ORFs can be powerful in revealing global patterns around translation initiation sites in both eubacteria and archae. The availability of large numbers of genomes allows formulation of reasonable averages for the two domains considered here. Since a broad range of bacteria were included, it is not likely that the results are skewed in any particular way. No bacterium contributed more than 8% of the total ORFs in question.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Carolyn Boyle for her advice on statistical validity of methods used. This work was supported by NSF Award No. EPS0082979. This is publication number J10086 of the Mississippi Agricultural and Forestry Experiment Station.

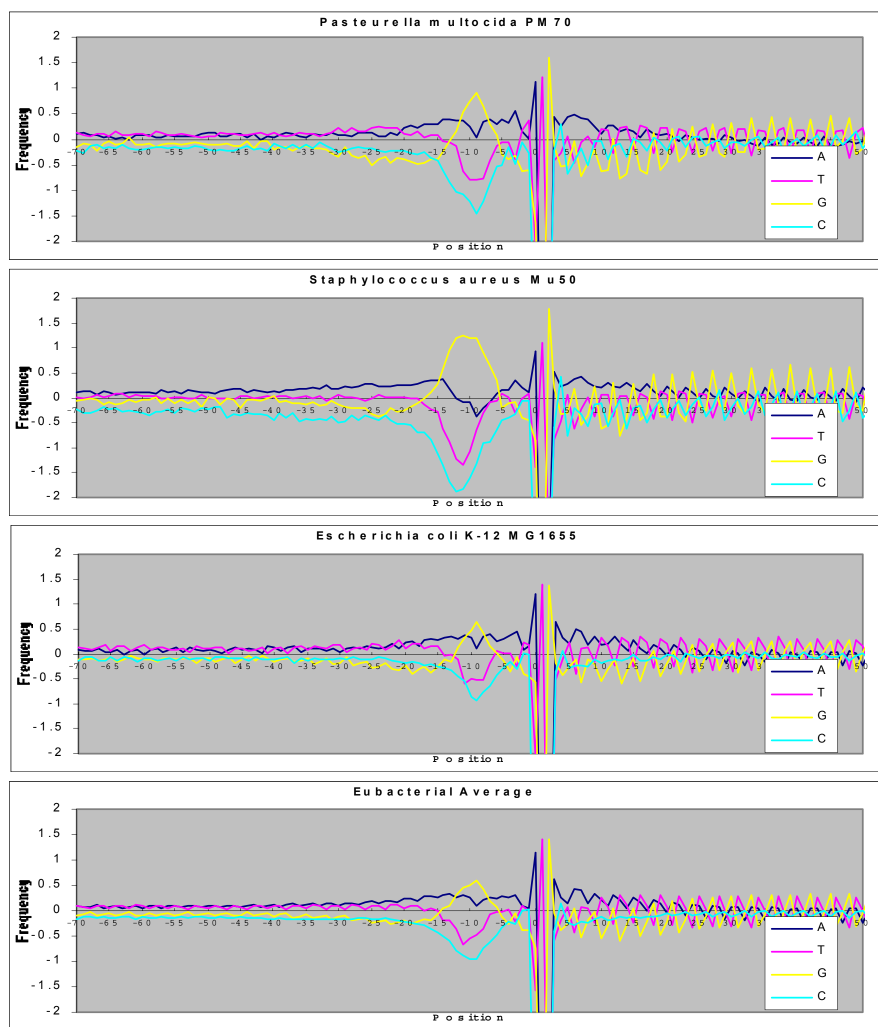

Figure 1. Aligned ORFs for three representative eubacteria, *Pasterella multocida* PM70, *Staphylococcus aureus* Mu50, and *Escherichia coli* K-12 MG1655, and for all ORFs of all eubacterial sequences available as of 01-22-02. Position represents the base position for each ORF aligned so that the first base in the start codon is aligned as zero on a number line. The base position immediately preceding the start codon is -1. The frequency represents a weighted, maximum likelihood statistical value calculated for each position as indicated in Methods and in reference 9.

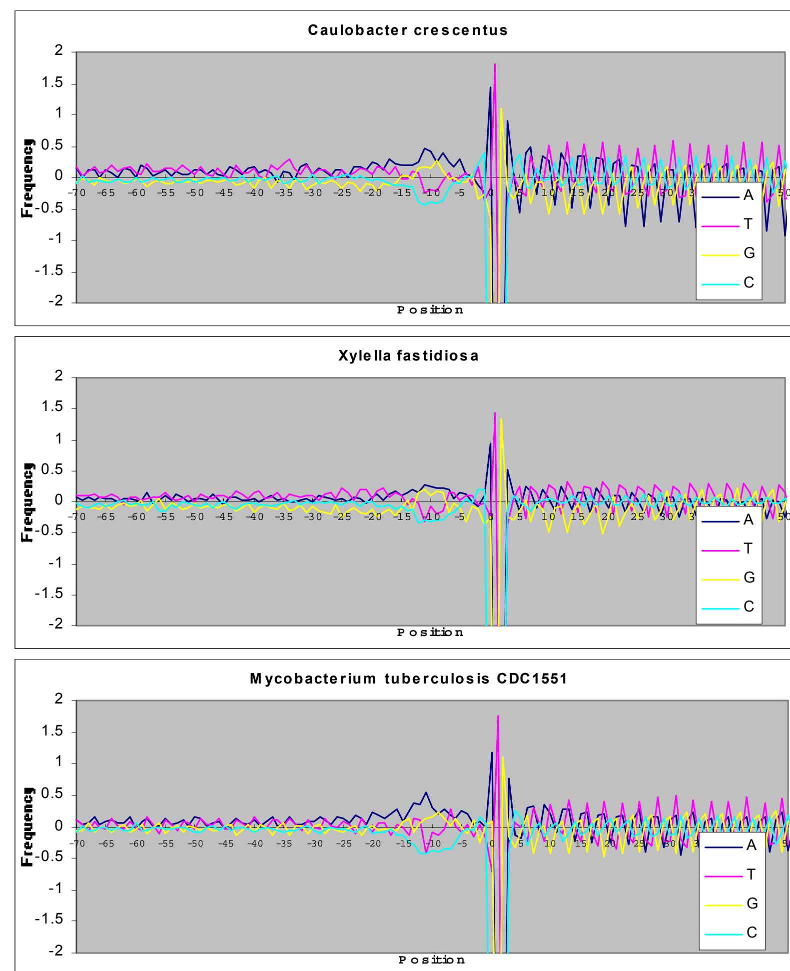

Figure 2. Aligned ORFs for *Caulobacter crescentus*, *Mycobacterium tuberculosis*, and *Xylella fastidioasa*. Position represents the base position for each ORF aligned so that the first base in the start codon is aligned as zero on a number line. The frequency represents a weighted, maximum likelihood statistical value calculated for each position as indicated in Methods and in reference 9.

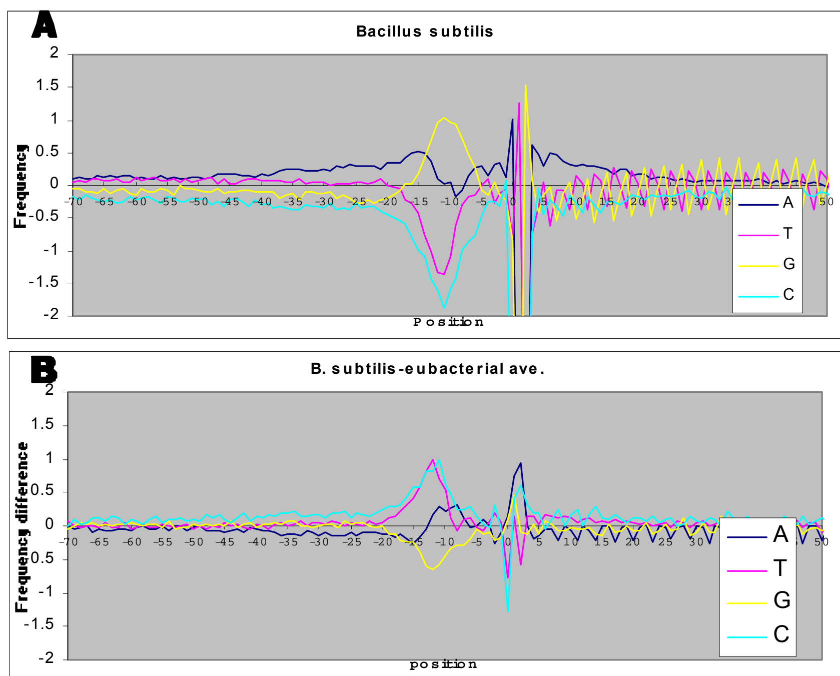

Figure 3. (A) Aligned ORFs for *Bacillus subtilis*. Position represents the base position for each ORF aligned so that the first base in the start codon is aligned as zero on a number line. The frequency represents a weighted, maximum likelihood statistical value calculated for each position as indicated in Methods and in reference 9. (B) Differences between the weighted frequencies at each position for *Bacillus subtilis* and those of the eubacterial average.

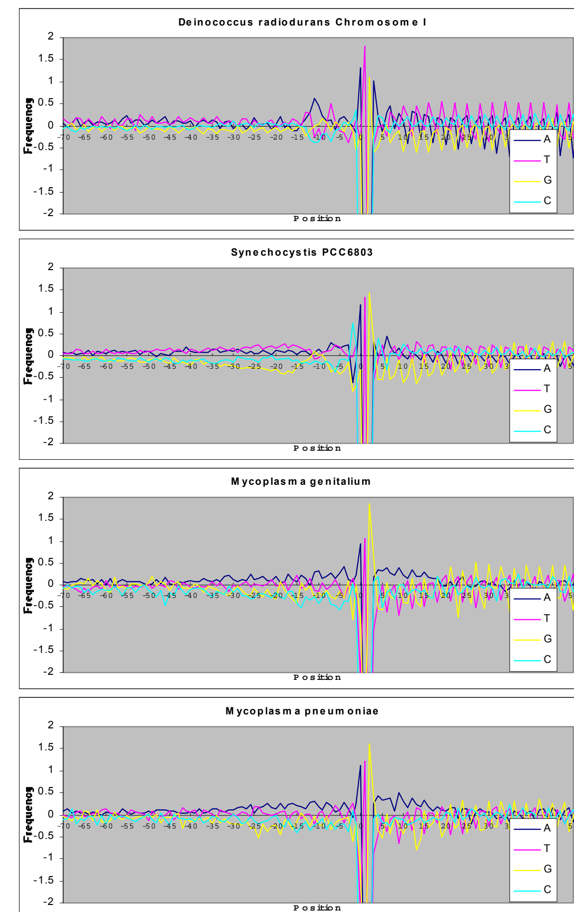

Figure 4. Aligned ORFs for *Deinococcus radiodurans, Synechocystis* PCC6803, *Mycoplasma genitalium* and *Mycoplasma pneumoniae*. Position represents the base position for each ORF aligned so that the first base in the start codon is aligned as zero on a number line. The frequency represents a weighted, maximum likelihood statistical value calculated for each position as indicated in Methods and in reference 9.

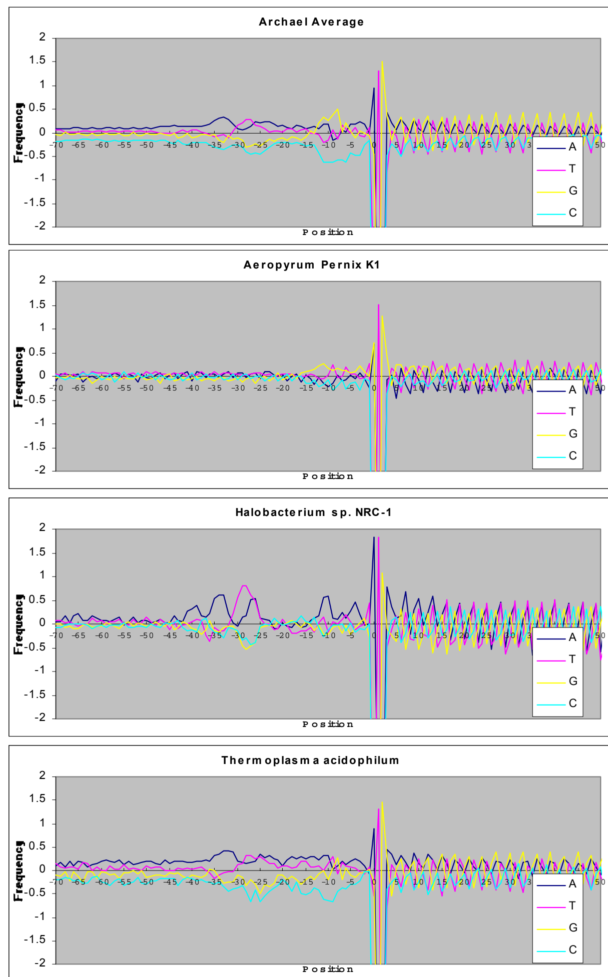

Figure 5. Aligned ORFs for all archaeal sequences available as of 01-22-02 as well as aligned ORFs for *Thermoplasma acidophilum, Aeropyrum pernix*, and *Halobacterium* sp. NRC-1. Position represents the base position for each ORF aligned so that the first base in the start codon is aligned as zero on a number line. The frequency represents a weighted, maximum likelihood statistical value calculated for each position as indicated in Methods and in reference 9.

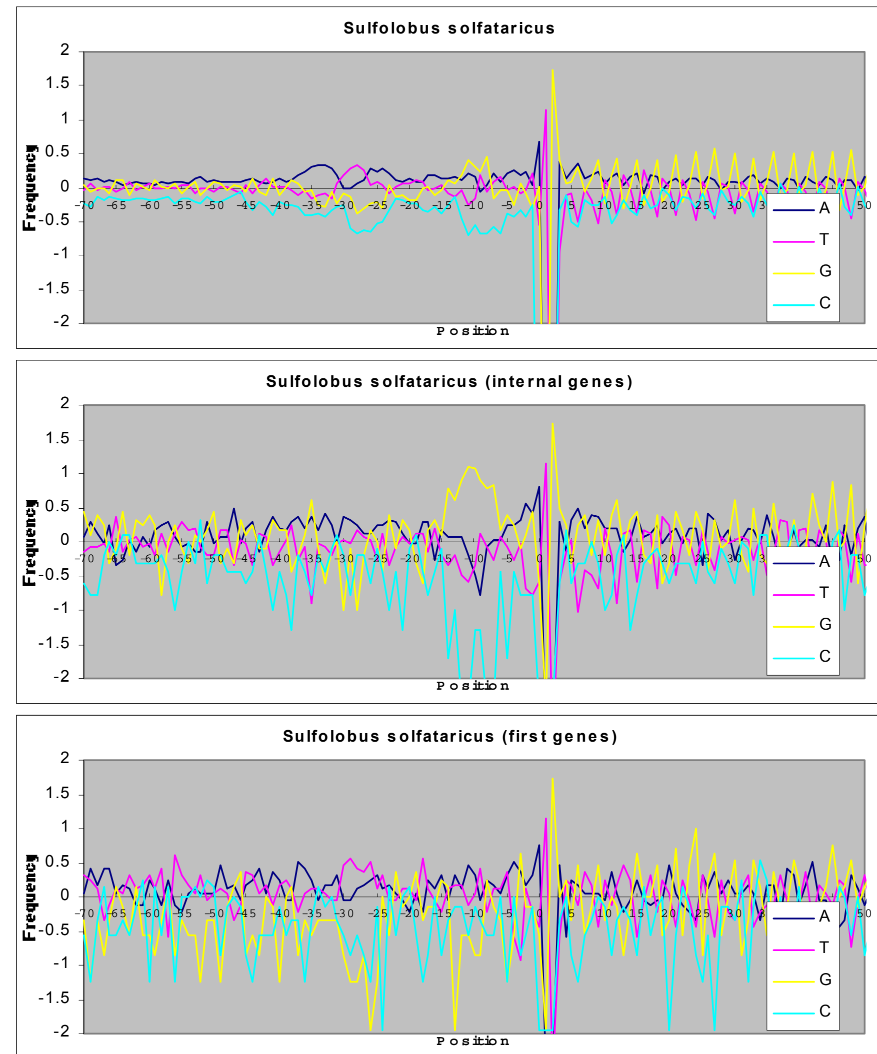

Figure 6. Aligned ORFs for *Sulfolobus solfataricus* as well as ORFs from 63 genes likely to be transcribed as internal members of operons (internal genes) or ORFs from 40 genes likely to be the first genes in operons (first genes). Position represents the base position for each ORF aligned so that the first base in the start codon is aligned as zero on a number line. The frequency represents a weighted, maximum likelihood statistical value calculated for each position as indicated in Methods and in reference 9.

LITERATURE CITED

Alm, R.A., L.-S.L. Ling, D.T. Moir, B.L. King, E.D. Brown, P.C. Doig, D.R. Smith, B. Noonan, B.C. Guild, B.L. deJonge, et al. 1999. Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. Nature 397:176–180.

Blattner, F.R., G. Plunkett, III, C.A. Bloch, N.T. Perna, V. Burland, M. Riley, J. Collado-Vides, J.D. Glasner, C.K. Rode, G.F. Mayhew, et al. 1997. The complete genome sequence of *Escherichia coli* K-12. Science 277:1453–1474.

Bolotin, A., P. Wincker, S. Mauger, O. Jaillon, K. Malarme, J. Weissenbach, S.D. Ehrlich, and A. Sorokin. 2001. The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ssp. lactis IL1403. Genome Res. 11:731–753.

Bult, C.J., O. White, G.J. Olsen, L. Zhou, R.D. Fleischmann, G.G. Sutton, J.A. Blake, L.M. FitzGerald, R.A. Clayton, J.D. Gocayne, et al. 1996. Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. Science 273:1058–1073.

Capela, D., F. Barloy-Hubler, J. Gouzy, G. Bothe, F. Ampe, J. Batut, P. Boistard, A. Becker, M. Boutry, E. Cadieu, et al. 2001. Analysis of the chromosome sequence of the legume symbiont *Sinorhizobium meliloti* strain 1021. Proc. Natl. Acad. Sci. USA 98:9877–9882.

Chambaud, I., R. Heilig, S. Ferris, V. Barbe, D. Samson, F. Galisson, I. Moszer, K. Dybvig, H. Wroblewski, A. Viari, et al. 2001. The complete genome sequence of the murine respiratory pathogen *Mycoplasma pulmonis*. Nucleic Acids Res. 29:2145–2153.

Cole, S.T., R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S.V. Gordon, K. Eiglmeier, S. Gas, C.E. Barry, III, et al. 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature 393:537–544.

Cole, S.T., K. Eiglmeier, J. Parkhill, K.D. James, N.R. Thomson, P.R. Wheeler, N. Honore, T. Ganier, C. Churcher, D. Harris, et al. 2001. Massive gene decay in the leprosy bacillus. Nature 409:1007–1011.

Condo, I., A. Ciammaruconi, D. Benelli, D. Ruggero, and P. Londei. 1999. Cis-acting signals controlling translational initiation in the thermophilic archaeon *Sulfolobus solfataricus*. Mol. Microbiol. 34:77–84.

Deckert, G., P.V. Warren, T. Gaasterland, W.G. Young, A.L. Lenox, D.E. Graham, R. Overbeek, M.A. Snead, M. Keller, M. Aujay, et al. 1998. The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*. Nature 392:353–358.

DelVecchio, V.G., V. Kapatral, R.J. Redkar, G. Patra, C. Mujer, T. Los, N. Ivanova, I. Anderson, A. Bhattacharyya, A. Lykidis, et al. 2002. The genome sequence of the facultative intracellular pathogen *Brucella melitensis*. Proc. Natl. Acad. Sci. USA 99:443–448.

Ferretti, J.J., W.M. McShan, D. Adjic, D. Savic, G. Savic, K. Lyon, C. Primeaux, S.S. Sezate, A.N. Surorov, S. Kenton, et al. 2001. Complete genome sequence of an M1 strain of *Streptococcus pyogenes*. Proc. Natl. Acad. Sci. USA 98:4658–4663.

Fickett, J.W. 1984. Fast optimal alignment. Nucleic Acids Res. 12:175–179.

Fleischmann, R.D., M.D. Adams, O. White, R.A. Clayton, E.F. Kirkness, A.R. Kerlavage, C.J. Bult, J.-F. Tomb, B.A. Dougherty, J.M. Merrick, et al. 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science 269:496–512.

Fraser, C.M., S. Casjens, W.M. Huang, G.G. Sutton, R.A. Clayton, R. Lathigra, O. White, K.A. Ketchum, R. Dodson, E.K. Hickey, M. Gwinn, B. Dougherty, et al. 1997. Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*. Nature 390:580–586.

Fraser, C.M., J.D. Gocayne, O. White, M.D. Adams, R.A. Clayton, R.D. Fleischmann, C.J. Bult, A.R. Kerlavage, G.G. Sutton, J.M. Kelley, et al. 1995. The minimal gene complement of *Mycoplasma genitalium*. Science 270:397–403.

Fraser, C.M., S.J. Norris, G.M. Weinstock, O. White, G.G. Sutton, R. Dodson, M. Gwinn, E. K. Hickey, R. Clayton, K.A. Ketchum, et al. 1998. Complete genome sequence of *Treponema pallidum*, the syphilis spirochete. Science 281:375–388.

Glaser, P., L. Frangeul, C. Buchrieser, A. Amend, F. Baquero, P. Berche, H. Bloecker, P. Brandt, T. Chakraborty, A. Charbit, et al. 2001. Comparative genomics of Listeria species. Science 294:849–852.

Glass, J.I., E.J. Lefkowitz, J.S. Glass, C.R. Heiner, E.Y. Chen, and G.H. Cassell. 2000. The complete sequence of the mucosal pathogen *Ureaplasma urealyticum*. Nature 407:757–762.

Goodner, B., G. Hinkle, S. Gattung, N. Miller, M. Blanchard, B. Qurollo, B.S. Goldman, Y. Cao, M. Askenazi, C. Halling, et al. 2001. Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tumefaciens* C58. Science 294:2323–2328.

Grill, S., C.O. Gualerzi, P. Londei, and U. Blasi. 2000. Selective stimulation of translation of leaderless mRNA by initiation factor 2: evolutionary implications for translation. EMBO J. 19(15):4101–4110.

Grill, S., I. Moll, D. Hasenohrl, C.O. Gualerziand, and U. Blasi. 2001. Modulation of ribosomal recruitment to 5'-terminal start codons by translation initiation factors IF2 and IF3. FEBS Lett. 495:167–171.

Gualerzi, C.O., and C.L. Pon. 1990. Initiation of mRNA translation in prokaryotes. Biochemistry 29:5881–5889.

Hayashi, T., K. Makino, M. Ohnishi, K. Kurokawa, K. Ishii, K. Yokoyama, C.-G. Han, E. Ohtsubo, K. Nakayama, T. Murata, et al. 2001. Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12. DNA Res. 8:11–22.

Heidelberg, J.F., J.A. Eisen, W.C. Nelson, R.A. Clayton, M.L. Gwinn, R.J. Dodson, D.H. Haft, E.K. Hickey, J.D. Peterson, L.A. Umayam, et al. 2000. DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. Nature 406:477–483.

Hertz, G.Z., and G.D. Stormo. 1996. *Escherichia coli* promoter sequences: analysis and prediction. Methods Enzymol. 273:30–42.

Himmelreich, R., H. Hilbert, H. Plagens, E. Pirkl, B.C. Li, and R. Herrmann. 1996. Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae*. Nucleic Acids Res. 24:4420–4449.

Hoskins, J.A., W. Alborn, Jr., J. Arnold, L. Blaszczak, S. Burgett, B.S. DeHoff, S. Estrem, L. Fritz, D.-J Fu, W. Fuller, et al. 2001. The Genome of the Bacterium *Streptococcus pneumoniae* strain R6. J. Bacteriol. 183:5709–5717.

Kalman, S., W. Mitchell, R. Marathe, C. Lammel, J. Fan, R.W. Hyman, L. Olinger, J. Grimwood, R.W. Davis, and R.S. Stephens. 1999. Comparative genomes of *Chlamydia pneumoniae* and *C. trachomatis*. Nat. Genet. 21:385–389.

Kaneko, T., Y. Nakamura, C.P. Wolk, T. Kuritz, S. Sasamoto, A. Watanabe, M. Iriguchi, A. Ishikawa, K. Kawashima, T. Kimura, et al. 2001. Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena sp. strain PCC 7120. DNA Res. 8:205–213.

Kaneko, T., S. Sato, H. Kotani, A. Tanaka, E. Asamizu, Y. Nakamura, N. Miyajima, M. Hirosawa, M. Sugiura, S. Sasamoto, et al. 1996. Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. DNA Res. 3:109–136.

Kawarabayasi, Y., Y. Hino, H. Horikawa, K. Jin-no, M. Takahashi, M. Sekine, S. Baba, Ankai, A., H. Kosugi, A. Hosoyama, et al. 2001. Complete genome sequence of an aerobic thermoacidophilic crenarchaeon, *Sulfolobus tokodaii* strain7. DNA Res. 8:123–140.

Kawarabayasi, Y., Y. Hino, H. Horikawa, S. Yamazaki, Y. Haikawa, K. Jin-no, M. Takahashi, M. Sekine, S. Baba, A. Ankai, et al. 1999. Complete Genome Sequence of an Aerobic Hyper-thermophilic Crenarchaeon, Aeropyrum pernix K1. DNA Res. 6:83–101.

Kawarabayasi, Y., M. Sawada, H. Horikawa, Y. Haikawa, Y. Hino, S. Yamamoto, M. Sekine, S. Baba, H. Kosugi, A. Hosoyama, et al. 1998. Complete Sequence and Gene Organization of the Genome of a Hyper-thermophilic Archaebacterium, *Pyrococcus horikoshii* OT3. DNA Research 5:55–76.

Kawashima, T., Y. Yamamoto, H. Aramaki, T. Nunoshiba, T. Kawamoto, K.Watanabe, M. Yamazaki, K. Kanehori, N. Amano, Y. Ohya, K. Makino, and M. Suzuki. 1999. Determination of the complete genomic DNA sequence of *Thermoplasma volvanium* GSS1. Proc. Jpn. Acad. 75:213–218.

Keeling, P.J., N.M. Fast, and G.I. McFadden. 1998. Evolutionary relationship between translation initiation factor eIF-2gamma and selenocysteine-specific elongation factor SELB: change of function in translation factors. J. Mol. Evol. 47:649–655.

Klenk, H.P., R.A. Clayton, J.-F. Tomb, O. White, K.E. Nelson, K.A. Ketchum, R.J. Dodson, M. Gwinn, E.K. Hickey, J.D. Peterson, et al. 1997. The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*. Nature 390:364–370.

Kunst, F., N. Ogasawara, I. Moszer, A.M. Albertini, G. Alloni, V. Azevedo, M.G. Bertero, P. Bessieres, A. Bolotin, S. Borchert, et al. 1997. The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. Nature 390:249–256.

Kuroda, M., T. Ohta, I. Uchiyama, T. Baba, H. Yuzawa,, I. Kobayashi, L. Cui, A. Oguchi, K. Aoki, Y. Nagai, et al. 2001. Whole genome sequencing of meticillin-resistant *Stapylococcus aureus*. The Lancet 357:1225–1240.

Kyrpides, N.C., and C.R. Woese. 1998. Archaeal translation initiation revisited: the initiation factor 2 and eukaryotic initiation factor 2B alpha-beta-delta subunit families. Proc. Natl. Acad. Sci. USA 95:3726–3730.

Ma, J., A. Campbell, and S. Karlin. 2002. Correlations between Shine-Dalgarno sequences and gene features such as predicted expression levels and operon structures. J. Bacteriol. 184:5733–5745.

May, B.P., and P.P. Dennis. 1990. Unusual evolution of a superoxide dismutase-like gene from the extremely halophilic archaebacterium *Halobacterium cutirubrum*. J. Bacteriol. 172:3725–3729.

May, B.J., Q. Zhang, L. Li, M.L. Paustian, T.S. Whittam, and V.S. Kapur. 2001. Complete nucleotide sequence of an avian isolate of *Pasteurella multocida*. Proc. Natl. Acad. Sci. USA 98:3460–3465.

McClelland, M., K.E. Sanderson, J. Spieth, S.W. Clifton, P. Latreille, L. Courtney, S. Powollik, J. Ali, M. Dante, F. Du, et al. 2001. The complete genome sequence of *Salmonella enterica serovar Typhimurium* LT2: features revealed by comparison to related genomes. Nature 413:852–856.

Mrazek, J., D. Bhaya, A.R. Grossman, S. Karlin. 2001. Highly expressed and alien genes of the Synechocystis genome. Nucleic Acids Res. 29:1590–1601.

Nelson, K.E., R.A. Clayton, S.R. Gill, M.L. Gwinn, R.J. Dodson, D.H. Haft, E.K. Hickey, J.D. Peterson, W.C. Nelson, K.A. Ketchum, et al. 1999. Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence of *Thermotoga maritima*. Nature 399:323–329.

Ng, W.V., S.P. Kennedy, G.G. Mahairas, B. Berquist, M. Pan, H.D. Shukla, S.R. Lasky, N.S. Baliga, V. Thorsson, J. Sbrogna, et al. 2000. Genome sequence of Halobacterium species NRC-1. Proc. Natl. Acad. Sci. USA 97:12176–12181.

Nierman, W.C., T.V. Feldblyum, I.T. Paulsen, K.E. Nelson, J. Eisen, J.F. Heidelberg, M. Alley, N. Ohta, J.R. Maddock, I. Potocka, et al. 2001. Complete Genome Sequence of *Caulobacter crescentus*. Proc. Natl. Acad. Sci. USA 98:4136–4141.

Nolling, J., G. Breton, M.V. Omelchenko, K.S. Markarova, Q. Zeng, R. Gibson, H.M. Lee, J. Dubois, D. Qiu, J. Hitti, et al. 2001. Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum. J. Bacteriol. 183:4823–4838.

Ogata, H., S. Audic, P. Renesto-Audiffren, P.-E. Fournier, V. Barbe, D. Samson, V. Roux, P. Cossart, J. Weissenbach, J.-M. Claverie, and D. Raoult. 2001. Mechanisms of Evolution in *Rickettsia conorii* and *Rickettsia prowazekii*. Science 293:2093–2098.

Osada, Y., R. Saito, and M. Tomita. 1999. Analysis of base-pairing potentials between 16S rRNA and 5' UTR for translation initiation in various prokaryotes. Bioinformatics 15:578–581.

Parkhill, J., M. Achtman, K.D. James, S.D. Bentley, C. Churcher, S.R. Klee, G. Morelli, D. Basham, D. Brown, T. Chillingworth, et al. 2000. Complete DNA sequence of a serogroup A strain of *Neisseria menigitidis* Z2491. Nature 404:502–506.

Parkhill, J., G. Dougan, K.D. James, N.R. Thomson, D. Pickard, J. Wain, C. Churcher, K.L. Mungall, S.D. Bentley, M.T.G. Holden, et al. 2001. Complete genome sequence of a multiple drug resistant *Salmonella enterica serovar Typhi* CT18. Nature 413:848–852.

Parkhill, J., B.W. Wren, K. Mungall, J.M. Ketley, C. Churcher, D. Basham, T. Chillingworth,
R.M. Davies, T. Feltwell, S. Holroyd, et al. 2000. The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. Nature 403:665–668.

Parkhill, J., B.W. Wren, N.R. Thomson, R.W. Titball, M.T.G. Holden, M.B. Prentice, M. Sebaihia, K.D. James, C. Churcher, K.L. Mungall, et al. 2001. Genome sequence of *Yersinia pestis*, the causative agent of plague. Nature 413:523–527.

Perna, N.T., G. Plunkett, III, V. Burland, B. Mau, J.D. Glasner, D.J. Rose, G.F. Mayhew, P.S. Evans, J. Gregor, H.A. Kirkpatrick, et al. 2001. Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. Nature 409:529–533.

Read, T.D., R. Brunham, C. Shen, S.R. Gill, J.F. Heidelberg, O. White, E.K. Hickey, J. Peterson, L.A. Umayam, T. Utterback, et al. 2000. Genome sequences of *Chlamydia trachomatis* MoPn and *Chlamydia pneumoniae* AR39. Nucleic Acids Res. 28:1397–1406.

Read, T.D., R.C. Brunham, C. Shen, S.R. Gill, J.F. Heidelberg, O. White, E.K. Hickey, J. Peterson, L.A. Umayam, T. Utterback, et al. 2000. Genome sequences of *Chlamydia trachomatis* MoPn and *Chlamydia pneumoniae* AR39. Nucleic Acids Res. 28:1397–1406.

Reiter, W.D, U. Hudepohl, and W. Zillig. 1990. Mutational analysis of an archaebacterial promoter: essential role of a TATA box for transcription efficiency and start-site selection *in vitro*. Proc. Natl. Acad. Sci. USA 87:9509–9513.

Reiter, W.D., P. Palm, and W. Zillig. 1988. Transcription termination in the archaebacterium Sulfolobus: signal structures and linkage to transcription initiation. Nucleic Acids Res. 16:2445–2459.

Ruepp, A., W. Graml, M.L. Santos-Martinez, K.K. Koretke, C. Volker, H.W. Mewes, D. Frishman, S. Stocker, A.N. Lupas, and W. Baumeister. 2000. The genome sequence of the thermoacidophilic scavenger *Thermoplasma acidophilum*. Nature 407:508–513.

Saito, R., and M. Tomita. 1999. Computer analyses of complete genomes suggest that some archaebacteria employ both eukaryotic and eubacterial mechanisms in translation initiation. Gene 238:79–83.

Sakai, H., C. Imamura, Y. Osada, R. Saito, T. Washio, and M. Tomita. 2001. Correlation between Shine-Dalgarno sequence conservation and codon usage of bacterial genes. J. Mol. Evol. 52:164–170.

Salin, M.L., M.V. Duke, D.P. Ma, and J.A. Boyle. 1991. *Halobacterium halobium* Mn-SOD gene: archaebacterial and eubacterial features. Free Rad. Res. Commun.12–13 Pt 1:443–449.

She, Q., R.K. Singh, F. Confalonieri, Y. Zivanovic, G. Allard, M.J. Awayez, C.C. Chan-Weiher, I.G. Clausen, B.A. Curtis, A. De Moors, et al. 2001. The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. Proc. Natl. Acad. Sci. USA 98:7835–7840.

Shigenobu, S., H. Watanabe, M. Hattori, Y. Sakaki, and H. Ishikawa. 2000. Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. APS. Nature 407:81–86.

Shine, J., and L. Dalgarno. 1974. The 3'-terminal sequence of *Escherichia coli*16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. USA 71:1342–1346.

Shirai, M., H. Hirakawa, M. Kimoto, M. Tabuchi, F. Kishi, K. Ouchi, T. Shiba, K. Ishii, M. Hattori, S. Kuhara, and T. Nakazawa. 2000. Comparison of whole genome sequences of *Chlamydia pneumoniae* J138 from Japan and CWL029 from USA. Nucleic Acids Res. 28:2311–2314.

Simpson, A.J.G., F.C. Reinach, P. Arruda, F.A. Abreu, M. Acencio, R. Alvarenga, L.M.C. Alves, J.E. Araya, G.S. Baia, C.S. Baptista, et al. 2000. The genome sequence of the plant pathogen *Xylella fastidiosa*. Nature 406:151–157.

Slupska, M.M., A.G. King, S. Fitz-Gibbon, J. Besemer, M. Borodovsky, and J.H. Miller. 2001. Leaderless transcripts of the crenarchaeal hyperthermophile *Pyrobaculum aerophilum*. J. Mol. Biol. 309:347–360.

Smith, D.R., L.A. Doucette-Stamm, C. Deloughery, H.-M. Lee, J. Dubois, T. Aldredge, R. Bashirzadeh, D. Blakely, R. Cook, K. Gilbert, et al. 1997. Complete genome sequence of *Methanobacterium thermoautotrophicum* deltaH: functional analysis and comparative genomics. J. Bacteriol. 179:7135–7155.

Sprengart, M.L., E. Fuchs, and A.G. Porter. 1996. The downstream box: an efficient and independent translation initiation signal in *Escherichia coli*. EMBO J. 15:665–674.

Staden, R. 1984. Computer methods to locate signals in nucleic acid sequences. Nucleic Acids Res.12:505–519.

Stephens, R.S., S. Kalman, C. Lammel, J. Fan, R. Marathe, L. Aravind, W. Mitchell, L. Olinger, R.L. Tatusov, Q. Zhao, et al. 1998. Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis*. Science 282:754–759.

Stormo, G.D. 2000. DNA binding sites: representation and discovery. Bioinformatics 16:16–23.

Stover, C.K., X.-Q.T. Pham, A.L. Erwin, S.D. Mizoguchi, P. Warrener, M.J. Hickey, F.S.L. Brinkman, W.O. Hufnagle, D.J. Kowalik, M. Lagrou, et al. 2000. Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen. Nature 406:959–964.

Suckow, J.M., N. Amano, Y. Ohfuku, J. Kakinuma, H. Koike, and M. Suzuki. 1998. A transcription frame-based analysis of the genomic DNA sequence of a hyper-thermophilic archaeon for the identification of genes, pseudo-genes and operon structures. FEBS Lett. 426:86–92.

Takami, H., K. Nakasone, Y. Takaki, G. Maeno, Y. Sasaki, N. Masui, F. Fuji, C. Hirama, Y. Nakamura, N. Ogasawara, et al. 2000. Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and genomic sequence comparison with *Bacillus subtilis*. Nucleic Acids Res. 28:4317–4331.

Tettelin, H., K.E. Nelson, I.T. Paulsen, J.A. Eisen, T.D. Read, S. Peterson, J. Heidelberg, R.T.
DeBoy, D.H. Haft, R.J. Dodson, et al. 2001. Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*. Science 293:498–506.

Tettelin, H., N.J. Saunders, J. Heidelberg, A.C. Jeffries, K.E. Nelson, J.A. Eisen, K.A. Ketchum, D.W. Hood, J.F. Peden, R.J. Dodson, et al. 2000. Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58. Science 287:1809–1815.

Thomm, M., and G. Wich. 1988. An archaebacterial promoter element for stable RNA genes with homology to the TATA box of higher eukaryotes. Nucleic Acids Res. 16:151–163.

Tiwari, S., S. Ramachandran, A. Bhattacharya, S. Bhattacharya, and R. Ramaswamy. 1997. Prediction of probable genes by Fourier analysis of genomic sequences. Comput. Appl. Biosci. 13:263–270.

Tolstrup, N., C.W. Sensen, R.A. Garrett, and I.G. Clausen. 2000. Two different and highly organized mechanisms of translation initiation in the archaeon *Sulfolobus solfataricus*. Extremophiles 4:175–179.

Tomb, J.-F., O. White, A.R. Kerlavage, R.A. Clayton, G.G. Sutton, R.D. Fleischmann, K.A. Ketchum, H.P. Klenk, S. Gill, B.A. Dougherty, et al. 1997. The complete genome sequence of the gastric pathogen *Helicobacter pylori*. Nature 388:539–547.

Tsonis, A.A., J.B. Elsner, and P.A. Tsonis. 1991. Periodicity in DNA coding sequences: implications in gene evolution. J. Theor. Biol.151:323–331.

Van Etten, W.J., and G.R. Janssen. 1998. An AUG initiation codon, not codon-anticodon complementarity, is required for the translation of unleadered mRNA in *Escherichia coli*. Mol. Microbiol. 27:987–1001.

Van Etten, W.J., and G.R. Janssen. 1998. An AUG initiation codon, not codon-anticodon complementarity, is required for the translation of unleadered mRNA in *Escherichia coli*. Mol. Microbiol. 27:987–1001.

Weiner, J., 3rd, R. Herrmann, and G.F. Browning. 2000. Transcription in *Mycoplasma pneumoniae*. Nucleic Acids Res. 28:4488–4496.

White, O., J.A. Eisen, J.F. Heidelberg, E.K. Hickey, J.D. Peterson, R.J. Dodson, D.H. Haft, M.L. Gwinn, W.C. Nelson, D.L. Richardson, et al. 1999. Genome Sequence of the Radioresistant Bacterium *Deinococcus radiodurans* R1. Science 286:1571–1577.

Wich, G., H. Hummel, M. Jarsch, U. Bar, and A. Bock. 1986. Transcription signals for stable RNA genes in *Methanococcus*. Nucleic Acids Res.14:2459–79.

Wood, D.W., J.C. Setubal, R. Kaul, D. Monks, L. Chen, G.E. Wood, Y. Chen, L. Woo, J.P. Kitajima, V.K. Okura, et al. 2001. The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58. Science 294:2317–2323.

Web Site References

http://www.ncbi.nlm.nih.gov:80/PMGifs/Genomes/micr.html NCBI Entrez-Genome Microbial Genomes

http://www.kazusa.or.jp/rhizobase/Rhizobase

http://www.msstate.edu/dept/biochemistry/CBIG/Mississippi State University Computational Biology and Informatics Group

A Simple Method of Surveying Plant Populations for Random Amplified DNA Polymorphisms

Janet Marie Gibson McDaniel and Robert G. Hamilton[1]
Department of Biological Sciences, Mississippi College, Clinton, MS 39058

Random amplified polymorphic DNA (RAPD) fragments can be separated on 8 cm long x 1 cm thick 5% polyacrylamide gels and silver stained to yield high quality data. This method requires far less technical expertise than other methods of separating RAPDs fragments, as gels can be purchased pre-made and complete equipment packages for such electrophoresis can be readily purchased. This method allows for analyses of RAPD variation under a much wider range of laboratory conditions than other such methods.

There is little known of the genome of many plant species. However, there are many endangered plants for which management requires knowledge of levels of genetic diversity within and among populations. A simple method of surveying plant populations is needed to allow for surveys to be completed with a minimal amount of equipment and minimal levels of expertise.

Random amplified polymorphic DNAs (RAPDs) (Williams et al., 1990) is a simple method of generating large numbers of polymorphisms in species for which the genome is largely unknown. RAPDs are generated by using 10 bp arbitrary primers, which will anneal numerous times in any given genome. RAPD fragments that anneal within 5000 bp can produce a fragment using the polymerase chain reaction (PCR). Agarose gel electrophoresis, commonly used for RAPD surveys, often lacks the sensitivity to consistently detect RAPD fragments. To increase sensitivity, we have used 43 cm 5% polyacrylamide gels to separate DNA fragments and silver staining to detect DNA fragments as described by Antolin et al. (1996) to consistently and reliably detect RAPD fragments.

However, the large 43 cm long x 0.45 mm thick gels require a level of technical expertise that may not be available for all situations. Separation of the large plates, as required by the silver staining, with the gel adhering to only one plate, often fails. We thus developed a method for using 8 cm x 1 mm thick 5% polyacrylamide gels for electrophoresis and silver staining to detect DNAs to provide a reliable, sensitive method for detection of RAPD fragments.

RAPDs have been criticized as lacking sufficient rigor for molecular analyses, as fragments cannot always be reproduced (Brown and Kresovich, 1996). While this criticism is quite valid, many RAPD fragments can be consistently reproduced (e.g., Antolin et al., 1996). RAPDs are also very powerful as a means of assessing members of the same clone in clonal plants (Hamilton et al., 2002). RAPD fragments are generally short, less than 1000 bp, and are thus quite easily cloned. Once cloned, fragments can easily be sequenced, with sequences used to develop primers specific to a given sequence and used as sequence tagged sites (STSs) (Chaffin and Hamilton, 1998). STSs representing fragments that are shown to be most revealing of differences among individuals can be further investigated. No other method of surveying populations for genetic variation allows such a useful preliminary survey while facilitating such a powerful more in depth survey. RAPDs are thus a method well suited for investigations of clonal plant populations for which there is little to no information regarding the genome.

MATERIALS AND METHODS

DNA samples were collected from daffodil leaves. DNA extractions were completed using the Qiagen DNAeasy plant kit. 20 mg of dry leaf tissue was shredded and placed into a 1.5 ml microcentrifuge tube with 106 micron glass beads added to each tube. Tissue was ground in 400 ml of Buffer A included in the kit preheated to 65°C, with 4 ml RNase A added following grinding. Instructions

[1]Author for correspondence

included with the kit were followed subsequent to the extraction step.

The "hot start" technique was used for the polymerase chain reaction. The hot start technique involves the separation of the primers and the template/polymerase until the temperature is above 36°C. We used a sterile DNase free reaction tube (Molecular Bioproducts) with bead of wax on the side of the tube. We prepared a lower reaction with primer in the reaction tube, heated the tube and added in an upper reaction mix with template/polymerase. The primer used for this example was Operon A-01 (5'-CAGGCCCTTC-3'). Primers were resuspended in 1 ml of TE buffer, pH 7.0, as per instructions from Operon, which results in about 5 pm/µl of primer.

For the Lower:

Primer	2.0 µl
dATP (10 mM)	1.0 µl
dCTP (10 mM)	1.0 µl
dGTP (10 mM)	1.0 µl
dTTP (10 mM)	1.0 µl
Stoffel Buffer (10x) (100 mM, Tris-HCL, pH 8.3, 100 mM, KCL)	2.5 µl
25 mM $MgCl_2$	10.0 µl
Sterile distilled water	6.5 µl

For the Upper:

Sterile distilled water	19.5 µl
DNA polymerase—Stoeffel fragment(5 U/ml)	1.0 µl
Template—from DNA prep	2.0 µl
Stoffel Buffer (10x) (100 mM, Tris-HCL, pH 8.3, 100 mM, KCl)	2.5 µl

The reactions were amplified using the following cycling protocol:

1. 95°C 2 minutes

   add lower, heat to 95°C for 30 seconds, remove reactions, pause program, then place reactions on ice, allow to cool, add upper.

2. 94°C 1 minute
3. 37°C 1 minute
4. 72°C 2 minutes
5. Repeat steps 2–4… a total of 45 times
6. 72°C 7 minutes

Following PCR, 10 µl of DNA sample buffer was added to samples. Polyacrylamide gel electrophoresis (PAGE) was performed using 10 well and 15 well 5% BioRad Ready Gels and a mini protean PAGE system. 1.0 x TBE (tris-borate-EDTA, diluted from 10x stock, 0.89 M tris, 0.89 M boric acid and 0.02 M EDTA, pH 8.0) was used in the lower buffer tank and 1.5x TBE was used in the upper tank as running buffers. Power was set at 20 volts and the gel was run for 11 hours for a 10 well gel, with the running buffers changed after 4 and 8 hours of electrophoresis. We ran 15 well gels at 100 volts for 1 hour and 20 volts for 6 hours, changing the running buffer after 3 hours. We used a 100 bp ladder as marker DNA (100–1000 bp, in 100 bp increments), however any marker that will allow for size determination of fragments in the range of about 100–2000 bp will work well. We have been primarily interested in analyzing fragments less than 1000 bp and thus used the 100 bp ladder described. 25 µl of PCR product was loaded onto each sample lane. More sample can be added as the wells will hold at least 50 µl.

Gels were stained using the following protocol:

1. Fix/Stop—200 ml 10% glacial acetic acid in ultrapure water.

   Fixing the gel. Agitate the gel well for 20 minutes (an orbital shaker set at 60 RPM works well) or leave overnight in the fix/stop solution. The gel can be stored in the fix/stop solution for several days.

2. Rinse the gel 3 times (2 minutes each) in ultrapure water using agitation. Drain water from gel for about 20 seconds between rinses.

3. Staining solution—0.2 g silver nitrate, 0.3 ml 37% formaldehyde in 200 ml ultrapure water.

   Stain the gel. Place the gel in the staining solution and agitate well for 30 minutes.

4. Rinse the gel for 5 seconds (no more than 10 seconds) in ultrapure water. The stain diffuses rapidly out of the gel during this step. If you rinse too long, restain. If you do not rinse at all, the excess stain will deplete the developing solution before the gel is fully developed.

5. Developing solution—6 g sodium carbonate to

200 ml in ultrapure water. Chill to 10°C. Just before use, add 40 µl of 10 mg/ml sodium thiosulfate and 0.3 ml of 37% formaldehyde.

> Develop the gel. Transfer the gel into half of the developing solution and agitate well by hand. When bands start to become visible, transfer gel to the other half of the developing solution. Continue with developing until maximal resolution is obtained. The decision of when to stop the developing reaction is somewhat subjective. You are looking for maximal contrast between the bands and the background. The reaction is stopped by adding the fix/stop solution. Staining will continue for a few seconds following the addition of the fix/stop solution.

> Note: It is critical that ultrapure water be used during silver staining, and that all glassware and plasticware be rinsed as thoroughly as possible with deionized water before use. It is also critical to use sodium carbonate of sufficient quality. We used ACS grade sodium carbonate (Sigma S-2148).

Following staining, gels were placed onto a sheet of thin acetate soaked in water. Another sheet of thin acetate was lain over the gels and air bubbles were removed from between the sheets of acetate. The gels were then dried in a BioRad gel dryer. Gels could be air dried in the absence of a gel dryer, however the acetate sheets must be fastened to each other so that they adhere stably to the gels and to each other during the drying process.

Following the drying process, gels were cut out from the acetate sheets. Copies of gels were made onto thicker acetate sheets but placing the gel on a white light box and tracing the gels onto another sheet of acetate. A line running across the bottom of the wells was also traced onto the sheets as the origin of the DNA fragments during PAGE. The distance of each fragment from the origin was measured and recorded. The 100 bp ladder marker DNA was used to create a standard graph of distance migrated vs fragment size on semi log graph paper. The standard curve was then used to determine the lengths of fragments in samples.

RESULTS

Figure 1 shows the 10 well gel run at 20 volts for 11 hours. Figure 2A shows a 15 well gel run at 20 volts for 6 hours and figure 2B shows a 15 well gel run at 100 volts for 1 hour. Data are much easier to visualize using the 10 well gel (fig. 2). Separating fragments at high voltage for a short time (fig. 2B) led to the poorest quality gel. Electrophoresis for 11 hours produced the best quality of results for the primer used. If it is necessary to observe shorter fragments, a shorter running time would be needed as fragments less than 450 or so base pairs were run off the bottom of the gel when gels were electrophoresed for 11 hours. However, electrophoresis for 6 hours did not get the 100 bp marker to the bottom of the gel (fig. 2A), and thus running times of over 6 but less than 11 hours are probably most optimal, and would have to be determined empirically for any system on a case by case basis.

CONCLUSIONS

This method can be completed easily in a minimally equipped laboratory. A microcentrifuge is necessary for DNA extractions. A thermal cycler and BioRad mini protean electrophoresis system (or a similar system) are necessary for PCR and electrophoresis. The 10 well gels produced data that were much easier to observe than the 15 well gels (compare figures 1 and 2). The 10 well gels also hold more sample, and thus are more flexible with respect to amount of sample leaded onto a gel. The quality of the data are identical to that of the 43 cm gels in the range within which fragments can be observed, with the limitation of the smaller gels being a narrowing of the range within which one can make quality observations.

This method can generate a great deal of data for assays of genetic homogeneity within populations with little preparation time (as the gels can be purchased pre-made and the DNA extraction is performed using a kit). While RAPDs are difficult to use as a means of resolving degrees of relatedness among organisms, identical or highly similar RAPD fingerprints are possibly the best indicators of genetic homogeneity available. As homogeneity is the parameter of interest to ecologists, the method outlined in this paper would be useful in molecular ecological studies; there is no need for prior information about the genome of subjects, methods are simple to perform, data can be extended by cloning

and sequencing of products and large amounts of data can be collected in a short period of time.

Data can be analyzed with a simple measure of diversity, such as the Shannon-Weaver index (Shannon and Weaver, 1949), where diversity is estimated using a measure of information content, H' where $H' = -\Sigma p_i \log p_i$, where p_i is the frequency of a fragment within a population sampled. Gustafson et al. (2002) describe more elaborate comparisons that can be made of RAPD fragment polymorphisms within and among populations.

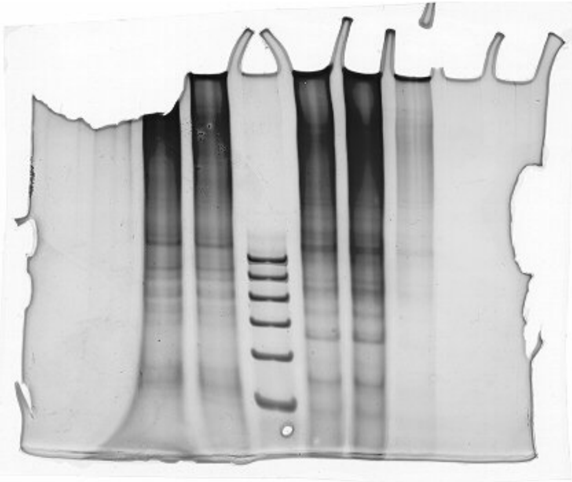

Figure 1. Silver stained 10 well 8 cm x 1 mm 5% BioRad ready gel.

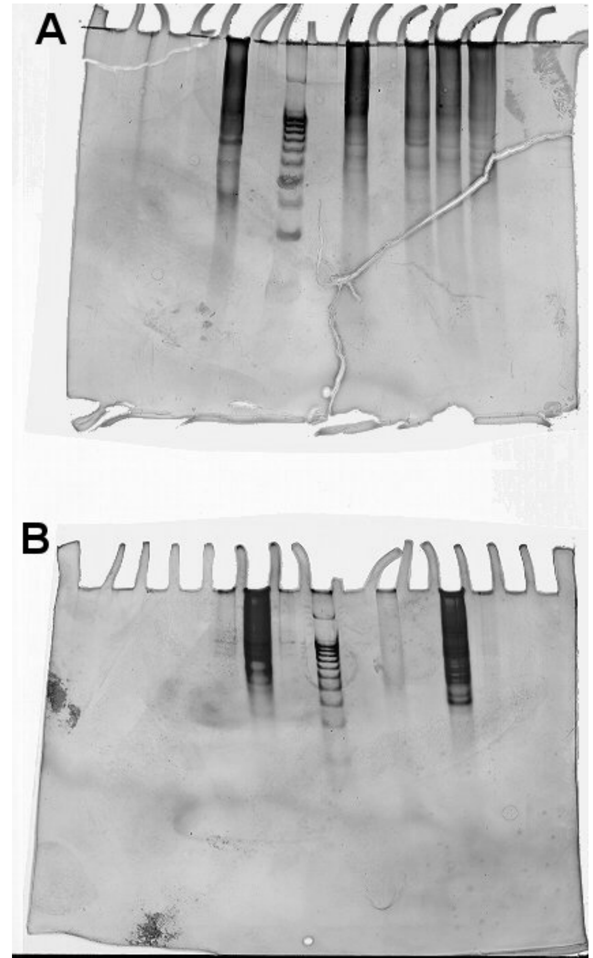

Figure 2. (A) Silver stained 5% 8 cm x 1 cm BioRad ready gel with samples run at 20 volts for 6 hours. (B) Silver stained 5% 8 cm x 1 cm BioRad ready gel with samples run at 100 volts for 1 hour.

LITERATURE CITED

Antolin, M.F., C. Bosio, J. Cotton, W. Sweeney, M. Strand, and W. Black. 1996. Intensive linkage mapping in a wasp (*Bracon hebetor*) and a mosquito (*Aedes aegypti*) with single-strand conformation polymorphism analysis of random amplified polymorphic DNA markers. Genetics 143:1727–1738.

Brown, S., and S. Kresovich. 1996. Molecular characterization for plant genetic resources conservation. Pages 86–93 *in* Patterson, ed. Genome Mapping in Plants. Academic Press. San Diego, CA.

Chaffin, C., and R. Hamilton. 1998. A method of producing STS (Sequence Tagged Site) markers from RAPDs (Random Amplified Polymorphic DNAs) in *Ceratopteris richardii*. C-fern web journal. http://cfern.bio.utk.edu/journal/journal.html.

Gustafson, D., D. Gibson, and D. Nickrent. 2002. Random amplified polymorphic DNA variation among remnant big bluestem (*Andropogon gerardii* Vitman) populations from Arkansas' grand prairie. Molecular Ecology 8:1693–1701.

Hamilton, R.T., T. Scarff, K. McGehee, and D. Morgan. 2002. Resolving variation within Narcissus cultivars. J. Miss. Acad. Sci. 47:14.

Shannon C.E., and W. Weaver. 1949. The mathematical theory of communication. University of Illinois Press, Champaign, IL.

Williams, J., A. Kubelick, K. Livak, J. Rafalski, and S. Tingey. 1990. DNA polymorphisms amplified by arbitrary primers are useful genetic markers. Nuclic Acids Res. 18: 6531–6535.

Executive Officer's Column

A changing of the guard is always an interesting time of adjustment for all concerned – those changing responsibilities, the membership and the Academy as a whole. At February's annual meeting the changes began to take place as Dr. John Boyle steps down from over 13 years of service as the MAS Executive Officer. As he adjusts to more free time and fewer responsibilities, I will be adjusting to less free time and more responsibilities. I look forward to the challenge.

As I begin the job of Executive Officer, perhaps a word about myself is in order. I am employed by the Mississippi Mineral Resources Institute, which is headquartered on the main campus of the University of Mississippi. I am a geologist by training, and I serve on the Board of the Mississippi State Board of Registered Professional Geologists. I have a broad range of research interests including geological mapping, Mississippi earthquakes, and the characteristics and concentration of naturally occurring radioactive materials. Many of my research projects have been interdisciplinary and I have enjoyed the opportunity to learn from those with related expertise.

Perhaps the membership would appreciate a map to my ideas and concerns are as they relate to the MAS? The primary concern is that of a long-term decline in membership. I hope to work with Ms. Cynthia Huff and see if there are any obvious indications as to cause. If we can identify a cause then we need to find a method to "cure the problem". We will look more closely into ways to make paying dues, abstract fees and annual meeting registration easier. The underutilization of the "Journal" is another concern. Dr. Ken Curry is the editor and has done an excellent job of keeping the publication schedules on track, but there always seems to be a shortage of papers. The "Journal" is refereed, page charges are low, and turn-around time is short. Perhaps the MAS is not promoting publication in the "Journal" to the membership enough. Another concern for us to look into.

The Geology and Geography Division sponsored a special symposium to go along with their general papers during the 1999 annual meeting. They ran out of chairs and people were standing at the back of the meeting room to hear the symposium speakers and many remained for the general presentations. Everyone agreed it very successful. I suggest other divisions may want to consider doing the same. It certainly adds a new facet to the annual meeting and with a well known speaker it can generate significant interest among other scientists and bring more into the annual meeting. Another idea is to encourage other activities at the annual meeting such as workshops, short courses and perhaps even tours and field trips.

I look forward to working with the MAS officers, and Board to plot the future course of the Academy. I also hope the membership will offer any suggestions as to how the Academy may improve its mission or make the Academy more attractive to scientists working in Mississippi.—Charles Swann

Divisional Report

Geology and Geography

The Geology and Geography Division met this year in a one-day session with 18 presentations and one poster. There was little time for relaxed conversation. Some of the broader topics discussed were:

(a) Visions of a statewide, high-quality, geographic information system (GIS) available on the Web

The 2003 Legislature passed a law that envisions the development of high-quality, digital, statewide, geographic data. This kind of information will be very useful to government, private sector, education, and citizens. The Mississippi system will integrate with the developing national system. The development of this digital information system will be as fundamental to Mississippi's digital future as literal roads were to our transportation network—each playing a vital role in Mississippi's ability to compete for a quality lifestyle for its citizens through efficient government, a thriving economy, education, recreation, etc. All of the GIS work of the future will be built on this quality foundation. As in all forms of infrastructure development, you need to build a good foundation.

(b) Licensure of geologists

Geologists whose work can impact the health, safety, and welfare of the public are required to have professional registration. This law also requires that these professionally qualified geologists be used when geology is involved in the task being performed. Mississippi has an active board of registered professional geologists.

(c) Geologic mapping

Geologic mapping is important because everything is built on the underlying geology. So, it is good to know its predictable influences on human activities such as: (1) construction sites with swelling clays, faults, boulders, or abrasive materials; (2) sanitary landfill site suitability; (3) water resources—recharge areas with susceptibility to contamination, aquitards with enhanced runoff, spring distribution, and water quality characteristics; (4) mineral resource distribution like oil and gas, sand and gravel, commercial clays, limestones, and who knows what in the future. For these reasons and many more, the surface and shallow geology has influenced the historical development, both natural and human, of the state and will continue to have influence in the future—to be forewarned is to be forearmed.

Many of the talks focused on site-specific geologic situations such as coastal geology, fossil whale excavations and findings, ground-water management techniques, seismic research, and earthquake research. Do you realize that northwest Mississippi is near the New Madrid Fault Zone? During the 1811–1812 New Madrid earthquakes, church bells were rung in Boston, the Natchez area was shaken, and islands in the Mississippi River disappeared. When the next big earthquake hits near our area, whether in New Madrid or Alabama, our geology will play a major role in the resulting effects (including liquefaction and soil amplification). That is why we need to know our geology and use that knowledge for future development. Our division session was an information-packed day.

We were so delighted to have the Mississippi Geological Society (MGS) sponsor a "Best Student Presentation Award." Seven student papers were given and Jeannie Bryson of Millsaps College won with her paper entitled "An Investigation of the Origin and Extent of a Perched Wetland, Millsaps College, Jackson, Mississippi." With the guidance of Dr. Stan Galicki, Jeannie took a multidisciplinary approach to her defined problem and resolved the study issues. She was an undergraduate who is now pursuing graduate study at the University of Arizona. Dr. Galicki is now inspiring the next class of aspiring geology students to equal levels of success and beyond. The MGS will be sponsoring best paper awards for undergraduate and graduate students at the Biloxi meeting. We greatly appreciate the support of these students and professors by such a fine professional organization as MGS.

Jeannie Bryson (right), winner of the Best Student Presentation Award, with her mentor Dr. Stan Galicki.

We are delighted to announce that the Geology and Geography Division will host a symposium on "climate change as seen in the geologic record" at the Biloxi meeting. The subject matter is of great interest to any who follow the news. Global warming—is it happening? Is the climate supposed to change or have the last 5 to 10 thousand years been the exception to the rule? The subject is fascinating and the amount of research going into it is enormous. We are so grateful for the opportunity to have this fascinating subject presented by nationally and internationally recognized experts Dr. Richard Alley of Pennsylvania State University and Dr. Joan Fitzpatrick of the United States Geological Survey. This will be an opportunity to hear the state of the knowledge presentation on a subject that will affect all our lives. Details of the symposium will be coming in future issues. We hope that professors, students and the public will take this opportunity to hear these distinguished scientists present the story. Hope to see you there.—Jack Moody

Premier Exhibitors

In addition to their participation in our annual meeting,
these exhibitors continue to support the Academy and ask for your support!

Base Pairs
Dr. Robin Rockhold
University of Mississippi Medical Center
2500 North State St.
Jackson, MS 39216-4505
601-984-1634 (phone)
rrockhold@pharmacology.umsmed.edu

Beckman Coulter
Mr. Dennis K. Dahmer
Scientific Consultants, Inc.
PO Box 82010
Baton Rouge, LA 70884
1-800-835-7941 (phone)
scilabprod@msn.com

Biomedical Solutions, Inc.
Ms Annette Gilligan
8402 Livingston Ave.
Chalmette, LA 70043
866.674.7220 x 207 (toll free)
annetteg@biomedicalsolutions.org

Department of Polymer Science
Marek W. Urban, Ph.D.
Director, NSF Materials Research Sciences and Engineering
Box 10076
Hattiesburg, MS 39406
(601) 266-6868 (phone)
(601) 266-5504 (fax)
Marek.Urban@usm.edu

Fisher Scientific Company LLC
Ms Susie Emfinger
Account Representative
9999 Veterans Memorial Drive
Houston, TX 77038-2499
800-955-999 X6135 (phone)
601-825-0711 (fax)
jhemfinger@aol.com

Micro-Performance
Mr. Jimmie Smith
Regional Manager
165 Mill Chase
Covington, GA 30016
770-385-8700 (phone)
770-385-8701 (fax)
georgia-optics@mail.com

Mississippi Center for Supercomputing Research
Mr. David G. Roach
Director
Office of Information Technology, The University of Mississippi
303 Powers Hall
University, MS 38677
ccdavid@olemiss.edu

Scimetrics, Inc.
Mr. Jeff Case
20 Parlange Dr.
Destrehan, LA 70047
800-231-2065 x120 (phone)
504-764-0564 (fax)
jeff.case@scimetricsinc.com

Swift Instruments, Inc.
Mr. Jim Vandygriff
614 Green Hills Rd.
Duncanville, TX 75137
972-780-1065 (phone)
jim.sharon@juno.com

MISSISSIPPI ACADEMY OF SCIENCES ABSTRACT FORM/MEMBERSHIP FORM

ABSTRACT INFORMATION

Abstract title ______________________________

Name of presenting author(s) ______________________________

(Presenter must be a current (i.e., 2004 membership dues must be paid) student member, regular member, or life member of the MAS)

Telephone ______________ Email ______________

Check the division in which you are presenting

__ Agriculture and Plant Science
__ Cellular, Molecular and Dev. Biol.
__ Chemistry and Chem. Engineering
__ Ecology and Evolutionary Biology
__ Geology and Geography
__ Health Sciences
__ History and Philosophy of Science
__ Math., Computer Sci. and Statistics
__ Marine and Atmospheric Sciences
__ Physics and Engineering
__ Psychology and Social Sciences
__ Science Education
__ Zoology and Entomology

Type of presentation

__ Poster presentation __ Workshop
__ Lecture presentation __ Invited symposium

If the presenting author for this paper is also presenting in another division, please list the other division: ______________

Audio-visual equipment needs

__ 2" x 2" slide projector
__ Overhead projector

Other audio-visual equipment including computers and computer projection equipment must be provided by the speaker.

MEMBERSHIP INFORMATION

New __ Renewal __
Mr. Ms Dr. ______________________________
Address ______________________________
City, State, Zip ______________________________
School or Firm ______________________________
Telephone ______________ Email address ______________
PLEASE INDICATE DIVISION WITH WHICH YOU WISH TO BE AFFILIATED ______________
Regular member $25 Student member $5 Life member $ 250
Educational $150 Corporate Patron $1000 Corporate Donor $500

CHECKLIST

The following MUST be DONE:

__ Enclose copy of abstract (even if abstract has been submitted electronically)
__ Complete and enclose abstract form /membership form(this form)
__ Enclose the following payments (make check payable to Mississippi Academy of Sciences):
  __ $25 per abstract
  __ $25 regular membership fee OR $5 student membership fee (2003 membership must be paid for abstract to be accepted)
__ You must supply a check # ______________ or P.O. # ______________ (credit cards are not accepted)

In addition you MAY preregister at this time:

__ Enclose the following payments:
  __ $20 regular member (after 15 Jan.) __ $12 regular member (Preregistration before Jan. 15, 2004)
  __ $10 student member (after 15 Jan.) __ $ 5 student member (Preregistration before Jan. 15, 2004)
  __ $50 nonmember (after 15 Jan.) __ $40 nonmember (Preregistration before Jan. 15, 2004)

NOTE: **Abstracts that are resubmitted for changes will incur a $10 resubmission fee. Late abstracts will be accepted with a $10 late fee during November increased to $25 after that. Late abstracts will be accepted only if there is room in the appropriate division.** They will be published in the April issue of the MAS JOURNAL.

MISSISSIPPI ACADEMY OF SCIENCES—ABSTRACT INSTRUCTIONS
PLEASE READ ALL INSTRUCTIONS BEFORE YOU SUBMIT YOUR ABSTRACT

- Your paper may be presented orally or as a poster. Oral presentations are generally 15 minutes although some divisions allow more time. The speaker should limit a 15 minute presentation to 10–12 minutes to allow time for discussion; longer presentations should be limited accordingly. Instructions for poster presentations are given on the reverse side of this sheet.
- Enclose a personal check, money order, institutional check, or purchase order for $25 publication charge for each abstract to be published, payable to the Mississippi Academy of Sciences. The publication charge will be refunded if the abstract is not accepted.
- The presenting author must be a member of the Academy at the time the paper/poster is presented. Payment for membership of the presenting author must accompany the abstract.
- Attendance and participation at all sessions requires payment of registration.
- Note that three separate fees are associated with submitting and presenting a paper at the annual meeting of the Mississippi Academy of Sciences. (1) An abstract fee is assessed to defray the cost of publishing abstracts and (2) a membership fee is assessed to defray the costs of running the Academy. (3) Preregistration payment ($12 regular; $5 student) may accompany the abstract, or you may elect to pay this fee before January 15th, or pay full registration fees at the meeting.
- Abstracts may be submitted by e-mail or entered directly through the MAS website. The URL is http://www.msacad.org. This abstract submission form and the appropriate fees should be sent by US mail even if the abstract has been submitted electronically.
- Abstracts may be submitted as a WordPerfect, Word, ASCII, ANSI, or .RTF file on a PC readable diskette. *Formatting should be minimal*. This abstract submission form and the appropriate fees should be sent by US mail even if a diskette is used for the abstract.
- Abstracts may be submitted typed or printed on clean white paper. Abstracts received in this form will be scanned into a computer. Leave ample margins and use a sanserif type font to help minimize errors in scanning.
- **Abstracts that are resubmitted for changes will incur a $10 resubmission fee.**
- **Late abstracts will be accepted with a $10 late fee during November increased to $25 after that. Late abstracts will be accepted only if there is room in the appropriate division.** They will be published in the April issue of the MAS JOURNAL.
- Submit your abstract and appropriate fees to the Abstracts' Editor, John Boyle, **TO BE RECEIVED NO LATER THAN OCTOBER 31, 2003**.
- Late abstracts will be accepted with a $10 late fee and only if there is room in the appropriate division. They will be published in the April issue of the MAS journal.

Dr. John Boyle
Mississippi State University
Dept. of Biochemistry
P.O. Drawer 9650
Mississippi State, MS 39762

FORMAT FOR ABSTRACT

- Your abstract should be informative, containing: (a) a sentence statement of the study's specific objectives, unless this is given in the title; (b) brief statement of methods, if pertinent; (c) summary of the results obtained; (d) statement of the conclusions. It is not satisfactory to state, "The results will be discussed."
- Your abstract, including a concise, descriptive title, author(s), location where work was done, text and acknowledgment, may not exceed 250 words. **Excessively long abstracts will be truncated**.
- The title should be all capital letters. Use significant words descriptive of subject content.
- Authors' names start a new line.

- The institution where your research was done should include city, state, and zip code. Do not include institutional subdivisions such as department.
- The abstract should be one paragraph, single spaced, starting with a 3-space indentation.
- Use standard abbreviations for common units of measure. Other words to be abbreviated, such as chemical names, should be spelled out in full for the first use, followed by the abbreviation in parenthesis. Do not abbreviate in the abstract title.
- Special symbols not on your printer or typewriter must be in black ink.
- Use italics for scientific names of organisms.
- Begin authors' names on a new line. Place an asterisk (*) after the presenter(s), if there are multiple authors.
- Use superscripts for institutional affiliations where necessary to avoid ambiguity.
- Refer to these examples as guides.

EXAMPLES OF TITLES AND AUTHORS:

[single author, no ambiguity about designated speaker or affiliation]
AN EXPERIMENTAL MODEL FOR CHEMOTHERAPY ON DORMANT TUBERCULOUS INFECTION WITH PARTICULAR REFERENCE TO RIFAMPICIN
Joe E. Jones, Mississippi State University, Mississippi State, MS 39762
   Abstract body starts here . . .

[two authors, both designated as speakers, different affiliations, but no ambiguity]
AN EXPERIMENTAL MODEL FOR CHEMOTHERAPY ON DORMANT TUBERCULOUS INFECTION WITH PARTICULAR REFERENCE TO RIFAMPICIN
Joe E. Jones* and Ralph A. Smith*, Mississippi State University, Mississippi State, MS 39762, and University of Mississippi Medical Center, Jackson, MS 39216
   Abstract body starts here . . .

[two authors, one designated speaker, different affiliations, but no ambiguity]
AN EXPERIMENTAL MODEL FOR CHEMOTHERAPY ON DORMANT TUBERCULOUS INFECTION WITH PARTICULAR REFERENCE TO RIFAMPICIN
Joe E. Jones and Ralph A. Smith*, Mississippi State University, Mississippi State, MS 39762, and University of Mississippi Medical Center, Jackson, MS 39216
   Abstract body starts here . . .

[three authors, one designated speaker, different affiliations]
AN EXPERIMENTAL MODEL FOR CHEMOTHERAPY ON DORMANT TUBERCULOUS INFECTION WITH PARTICULAR REFERENCE TO RIFAMPICIN
Joe E. Jones[1], Ralph A. Smith[1]*, and Alice D. Doe[2], [1]Mississippi State University, Mississippi State, MS 39762, and [2]University of Mississippi Medical Center, Jackson, MS 39216
   Abstract body starts here . . .

GUIDELINES FOR POSTER PRESENTATIONS

- The Academy provides poster backboards. Each backboard is 34" high by 5' wide. Mount the poster on the board assigned to you by your Division Chairperson. Please do not draw, write, or use adhesive material on the boards. You must provide your own thumb tacks.
- Lettering for your poster title should be at least 1" high and follow the format for your abstract. Lettering for your poster text should be at least 3/8" high.
- Posters should be on display during the entire day during which their divisional poster session is scheduled. They must be removed at the end of that day.
- Authors must be present with their poster to discuss their work at the time indicated in the program.