References and Notes

1. C. Parmesan, G. Yohe, *Nature* **421**, 37 (2003).
2. C. D. Thomas *et al.*, *Nature* **427**, 145 (2004).
3. C. M. Wood, D. G. McDonald, Eds., *Global Warming: Implications for Freshwater and Marine Fish* (Cambridge Univ. Press, Cambridge, 1997).
4. K. Brander *et al.*, *Int. Counc. Explor. Sea Mar. Sci. Symp.* **219**, 261 (2003).
5. G. Beaugrand, K. M. Brander, J. A. Lindley, S. Souissi, P. C. Reid, *Nature* **426**, 661 (2003).
6. G. Beaugrand, P. C. Reid, F. Ibañez, J. A. Lindley, M. Edwards, *Science* **296**, 1692 (2002).
7. Materials and methods are available as supporting material on *Science* Online.
8. R. J. Knijn, T. W. Boon, H. J. L. Heessen, J. F. G. Hislop, *Atlas of North Sea Fishes* [International Council for the Exploration of the Sea (ICES) Cooperative Research Report 194, ICES, Copenhagen, 1993].
9. A. N. Strahler, A. H. Strahler, *Physical Geography: Science and Systems of the Human Environment* (Wiley, New York, 1997).
10. R. M. Thomas, in *A Rehabilitated Estuarine Ecosystem: The Environment and Ecology of the Thames Estuary*, M. J. Attrill, Ed. (Kluwer, Dordrecht, Netherlands, 1998), pp. 115–139.
11. S. Jennings *et al.*, *Fish. Res.* **40**, 125 (1999).
12. ICES Advisory Committee on Fishery Management (ACFM), *Report of the Northern Pelagic and Blue Whiting Fisheries Working Group* (ICES CM 2004/ACFM:24, ICES, Copenhagen, 2004).
13. S. I. Rogers, J. R. Ellis, *Int. Counc. Explor. Sea J. Mar. Sci.* **57**, 866 (2000).
14. N. Daan, H. Gislason, J. Pope, J. Rice, *Changes in the North Sea Fish Community: Evidence of Indirect Effects of Fishing?* (ICES CM N:10, ICES, Copenhagen, 2003).
15. U.K. Climate Impacts Programme, *UKCIP02 Scenarios Gateway—Maps Gateway*; available at www.ukcip.org.uk/scenarios/marine/marine.html.
16. N. K. Dulvy, Y. Sadovy, J. D. Reynolds, *Fish Fish.* **4**, 25 (2003).
17. S. Jennings, J. D. Reynolds, S. C. Mills, *Proc. R. Soc. London Ser. B* **265**, 333 (1998).
18. J. D. Reynolds, S. Jennings, N. K. Dulvy, in *Conservation of Exploited Species*, J. D. Reynolds, G. M. Mace, K. H. Redford, J. G. Robinson, Eds. (Cambridge Univ. Press, Cambridge, 2001), pp. 145–168.
19. M. Edwards, A. J. Richardson, *Nature* **430**, 881 (2004).
20. S. A. Murawski, *Trans. Am. Fish. Soc.* **122**, 647 (1993).
21. We thank A. Taylor, P. C. Reid, T. Osborn, P. Jones, the ICES Oceanographic Database, and the UK Climate Impacts Programme for providing data; J. Gill, W. Sutherland, and A. Watkinson for commenting on earlier drafts; and M. Attrill, K. Brander, R. Clark, C. Fox, J. Gill, and S. Jennings for valuable discussion. This research was undertaken under the Defra-funded project MF0730, with additional support from a Commonwealth Scholarship to A.L.P. and an NERC grant to J.D.R.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1111322/DC1
Materials and Methods
Tables S1 to S4
References

22 February 2005; accepted 28 April 2005
Published online 12 May 2005;
10.1126/science.1111322
Include this information when citing this paper.

Community Proteomics of a Natural Microbial Biofilm

Rachna J. Ram,[1] Nathan C. VerBerkmoes,[3,4] Michael P. Thelen,[1,6] Gene W. Tyson,[1] Brett J. Baker,[2] Robert C. Blake II,[7] Manesh Shah,[5] Robert L. Hettich,[4] Jillian F. Banfield[1,2]*

Using genomic and mass spectrometry–based proteomic methods, we evaluated gene expression, identified key activities, and examined partitioning of metabolic functions in a natural acid mine drainage (AMD) microbial biofilm community. We detected 2033 proteins from the five most abundant species in the biofilm, including 48% of the predicted proteins from the dominant biofilm organism, *Leptospirillum* group II. Proteins involved in protein refolding and response to oxidative stress appeared to be highly expressed, which suggests that damage to biomolecules is a key challenge for survival. We validated and estimated the relative abundance and cellular localization of 357 unique and 215 conserved novel proteins and determined that one abundant novel protein is a cytochrome central to iron oxidation and AMD formation.

Microbial communities play key roles in the Earth's biogeochemical cycles. Our knowledge of the structure and activities in these communities is limited, because analyses of microbial physiology and genetics have been largely confined to studies of organisms from the few lineages for which cultivation conditions have been determined (*1*). An additional limitation of pure culture–based studies is that potentially critical community and environmental interactions are not sampled. Recent acquisition of genomic data directly from natural samples has begun to reveal the gene content of communities (*2*) and environments (*3*). Here we combined "shotgun" mass spectrometry (MS)–based proteomics (*4–6*) with community genomic analysis to evaluate in situ microbial activity of a low-complexity natural microbial biofilm.

The biofilm samples used in this study and prior work (*2*) were collected from underground sites in the Richmond mine at Iron Mountain, near Redding, California (USA). These pink biofilms grew on the surface of sulfuric acid–rich (pH ~0.8), ~42°C solutions that contain near-molar concentrations of Fe and millimolar concentrations of Zn, Cu, and As (*7*) (Fig. 1). We used oligonucleotide probes (*8*) to demonstrate that *Leptospirillum* group II dominated the sample, but it also contained *Leptospirillum* group III, *Sulfobacillus*, and Archaea related to *Ferroplasma acidarmanus* and "G-plasma" (Fig. 2). This was similar in structure and composition to the community previously used as a source of genomic sequence (*2*).

In general, proteins could be assigned to organisms, because the genes that encode them are on DNA fragments (scaffolds) that have been assigned to different organism types (*2*). From the genomic dataset (*2*), we created a database of 12,148 proteins (Biofilm_db1) that was used to identify two-dimensional (2D) nano–liquid chromatography (nano-LC) (200 to 300 nl/min) tandem mass spectrometry (MS/MS) spectra (*8–13*) from different biofilm fractions. One or more peptides were assigned to ~5994 proteins (Table 1). This corresponds to ~49% of all proteins encoded by the genomes of the five most abundant organisms. We estimated the likelihood of false-positive protein identification using a variety of detection criteria and databases derived from organisms not present in this environment (*8*). Because of these results, for all subsequent analyses, we required matching of two or more peptides per protein for confident detection (*8*). After removal of duplicates, we detected 2003 different proteins (table S1). An additional 30 proteins were found that were encoded by alternative or overlapping open reading frames (*8*).

We detected 48% of the predicted proteins (i.e., 1362 of 2862) from *Leptospirillum* group II (table S2). This percentage exceeded those of most prior proteomic studies of microbial isolates (*10*, *12*, *13*). In part, this may reflect the presence of cells in many different growth stages, as well as microniches within the biofilm (*14*). We also detected 270 *Leptospirillum* group III, 84 *Ferroplasma* type I, 99 *Ferroplasma* type II, and 122 "G-plasma" proteins. In addition, we found 30 proteins on unassigned archaeal scaffolds and 36 on unassigned bacterial scaffolds. The proportion of proteins detected from each organism type was similar to the proportion of cells from each organism type in the biofilm (*8*). Most proteins from low-abundance members were probably in concentrations too low to be detected by the presence of two peptides.

[1]Department of Environmental Science, Policy, and Management, [2]Department of Earth and Planetary Science, University of California at Berkeley, Berkeley, CA 94720, USA. [3]Graduate School of Genome Science and Technology University of Tennessee–Oak Ridge National Laboratory, 1060 Commerce Park, Oak Ridge, TN 37830, USA. [4]Chemical Sciences Division, [5]Life Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA. [6]Biosciences Directorate, Lawrence Livermore National Laboratory, Livermore, CA 94551, USA. [7]College of Pharmacy, Xavier University, New Orleans, LA, 70125, USA.

*To whom correspondence should be addressed. E-mail: jill@seismo.berkeley.edu

Furthermore, we were unable to identify proteins from organisms such as *Sulfobacillus*, for which there is little genome sequence (*2*).

Using the MS data, we estimated the relative abundance of individual proteins in different biofilm fractions (*15*). Overall, the biofilm was dominated by novel proteins that are the products of genes previously annotated as "hypothetical" (42% of genes in the original genomic dataset). On the basis of the BLASTP algorithm (*16*), these proteins lack significant homology (*17*) to proteins with functional assignments. Of the abundant proteins, 15% were "unique" (not significantly similar to any known protein), and 2% were "conserved" (similar to other predicted proteins, but not significantly similar to characterized proteins). Other functional groups prominent in the group of most abundant biofilm proteins were ribosomal proteins (13%), chaperones (11%), thioredoxins (9%), and proteins involved in radical defense (8%). Ten abundant disulfide isomerases were detected, at least four of which were present in the extracellular fraction. Together these findings suggest that protein stability in pH < 1 solutions is achieved in part by refolding, catalyzed by abundant enzymes that were tolerant of low pH. Peroxiredoxin and some other abundant proteins (e.g., rubrerythrin and catalase) are involved in defense against oxidative stress, which indicates an important challenge in the AMD environment.

On the basis of MS sequence coverage, the extracellular fraction was enriched (*18*) in unique proteins (52%) and contained ~14% conserved novel proteins. These proteins presumably function in AMD solution and may be important for adaptation to the acidic, metal-rich environment. Among extracellular proteins, the three with the highest sequence coverage in this fraction are encoded by hypothetical genes. One showed weak similarity to c-type cytochromes and Fe/Pb permeases (table S2), and originated from *Leptospirillum* group II.

Although 67% of the predicted amino acid sequence of the protein that resembles cytochrome c could be reconstructed from multiple overlapping peptides, there were three gaps (Fig. 3). Further analysis revealed that one gap contains a signal peptide (*19*); the two others contain single amino acid differences in the cytochrome variant found in the biofilm used for proteomic analysis versus the biofilm used to generate the genomic data (*8*). After taking into account these factors, 100% of the protein was recovered by MS. Therefore, combining community genomics with proteomics data allows detection of protein variants and signal sequence cleavage in natural samples.

This extracellular protein also contained a heme-binding consensus sequence, indicating that it may play a role in electron transport. We verified by gel electrophoresis analysis and N-terminal sequencing that this is a heme-containing protein that is abundant in the extracellular fraction (*8*) (fig. S3). Interestingly, the peptide sequence differed from the predicted sequence, because it contained leucine in place of isoleucine at a position encoded by ATC, which suggests atypical codon usage. The proteomic analysis is blind to this substitution because isoleucine and leucine share the same mass. Abundant iron-oxidizing cytochromes with absorption peaks around 579 nm have been detected in *Leptospirillum ferrooxidans* (a member of *Leptospirillum* group I) and *Leptospirillum ferriphilum* (a member of *Leptospirillum* group II) isolates (*20*, *21*) (fig. S4). Amino acid sequences for these cytochromes have not been reported. Because the absorption at 579 nm is unique to

Table 1. Number of proteins detected through triplicate analysis of biofilm fractions using different filtering levels. The LCQ and LTQ datasets were derived from redundant protein counts from global contrast files using two different MS techniques (*8*). Liberal filters require at least one peptide per gene; conservative filters require at least two peptides per gene; ultraconservative filters require at least three peptides per gene. Xcorrs of at least 1.8 (+1), 2.5 (+2), 3.5 (+3) were used in all cases.

Filtering level	LCQ data set	LTQ data set	Combined data sets
Liberal filters	3127	5534	5994
Conservative filters	1160	2077	2146
Ultraconservative filters	837	1419	1435

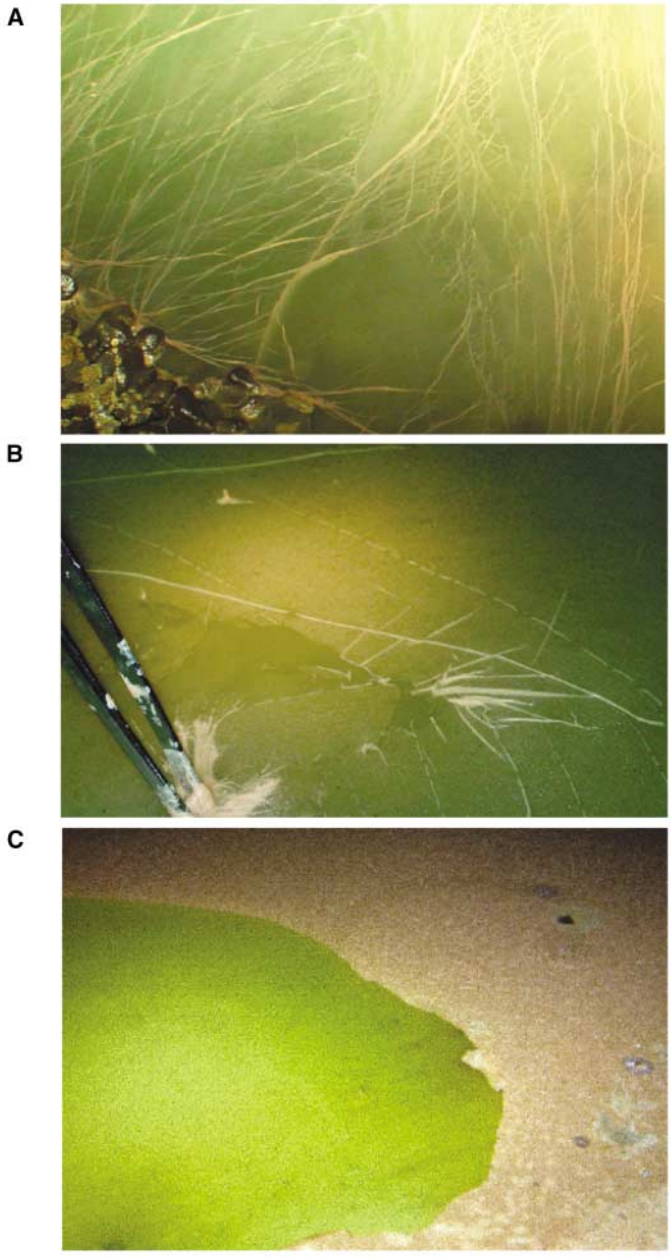


Fig. 1. (**A**) Photograph of the biofilm during collection in January 2004. The biofilm occurs as a continuous sheet over the surface of the AMD pool; wrinkles form because of movement of the solution. [Photograph taken from the AB end location (fig. S1).] (**B**) Close-up photograph during sample collection showing that the biofilm is thin and apparently homogeneous. (**C**) A thicker biofilm in the same location 5 months later, which suggests that the initial biofilm was actively growing when sampled. [Photographs by T. Johnson]

the abundant *Leptospirillum* cytochrome type, we henceforth refer to these proteins as cyt_{579}.

The first 40 amino acids of cyt_{579} purified from the periplasm of *L. ferriphilum* (*8*) matched the predicted sequence of the extracellular cytochrome in *Leptospirillum* group II derived from the biofilm after taking into account loss of the 23 amino acid signal peptide and the substitution of leucine for isoleucine. The cleavage of a leader sequence from the N terminus indicates that the mature protein is exported across the cytoplasmic membrane. The rate constant for the ferrous iron–dependent reduction of cyt_{579} purified from *L. ferriphilum* was greater than or equal to the overall turnover number for the transfer of electrons from ferrous iron to molecular oxygen by whole cells (*8*). Furthermore, the apparent standard reduction potential of cyt_{579} was ~670 mV (*8*). This relatively high potential would be expected for an electron carrier in a transport chain where the electron donor (ferrous ions complexed with sulfate in acidic solution) has a reduction potential of ≥660 mV. Thus, cyt_{579} has both kinetic and thermodynamic properties consistent with a central role in iron oxidation by *Leptospirillum* group II. The high abundance, localization to the extracellular fraction, and enzymatic activity of cyt_{579} imply that it is the primary iron oxidant in the electron transport chain (fig. S5). Eight c-type cytochromes and three other novel proteins with heme-binding motifs were also detected.

As the supply of ferric iron is the rate-limiting step for pyrite oxidation in this environment (*22*), the metabolic activity of iron-oxidizing microorganisms largely determines the rate of AMD formation. *Leptospirillum* group II dominates most biofilms from the Richmond mine and is frequently detected at other mining sites and bioleaching plants (*23*). Thus, cyt_{579} is likely one of the key enzymes that connects the biology and geochemistry of metal-rich acidic environments.

Overall a probable function was assigned to 69% of the detected *Leptospirillum* group II proteins on the basis of sequence similarity (table S2). We assigned all detected *Leptospirillum* group II proteins to functional categories on the basis of clusters of orthologous genes (COGs) (*24*) to evaluate the degree of expression of novel proteins and to estimate how biochemical resources were partitioned to different metabolic activities by this organism. Most commonly detected were unique and conserved novel proteins (Fig. 4). Proteins involved in amino acid metabolism, translation, and energy production and conversion were the next most commonly detected, followed by cell envelope biogenesis, coenzyme metabolism, and protein folding and modification.

Many proteins involved in cobalamin and heme biosynthesis were detected. The reason for a high cobalamin demand by *Leptospirillum* group II is unclear. Additional analysis is needed to determine whether other community members manufacture this vitamin or obtain it from *Leptospirillum* group II. Heme is essential for biosynthesis of cytochromes such as cyt_{579}. A high demand for cyt_{579} is consistent with the relatively low energy yield associated with iron oxidation (*25*). Heme is also likely incorporated into the abundant catalase and peroxidase proteins, which are important for peroxide and radical detoxification. Similarly, the detection of many enzymes involved in protein refolding may reflect the challenge associated with maintaining protein integrity in the hot, acid environment. The apparently abundant thioredoxins may also construct and maintain the conformation of the abundant acid-exposed heme-based proteins localized in the periplasm.

Proteins from COG families involved in secondary metabolite biosynthesis, transport, and catabolism; cell division and chromosome partitioning; and inorganic ion transport and metabolism made up the smallest numbers of detected proteins. In part, this may reflect our inability to assign these metabolic roles to novel environment- and lineage-specific proteins.

Despite the predominance of novel proteins, it is noteworthy that only 38% of the proteins encoded by conserved hypothetical genes and 35% of the proteins encoded by unique hypothetical genes were detected. In contrast, we detected 86% of proteins predicted to be involved in amino acid metabolism and 86% of those involved in translation. This suggests that many hypothetical genes are nonfunctional, encode proteins required at low abundance, or are expressed under conditions different from those at the time of sampling. We compared the fraction of genes in the genome associated with each function with the fraction of proteins in the proteome associated with that function. Amino acid metabolism, translation, nucleotide metabolism, protein refolding and modification were all more highly represented in the proteome than in the genome. In all other categories (except transposases), we observed that representation in the proteome was similar to that in the genome (Fig. 4).

Proteomic data can provide direct insights into how essential functions are carried out and partitioned among members of natural communities. For example, although we detected a RuBisCO (ribulose-1,5-bisphosphate carboxylase-oxygenase)–like protein in *Leptospirillum* group II (*2*) (47% MS sequence

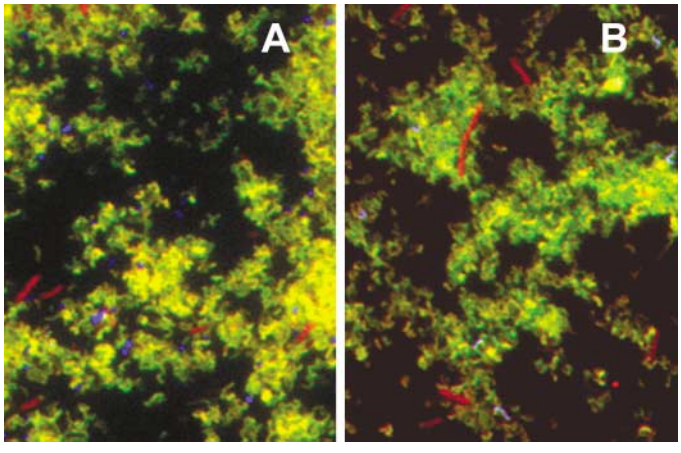


Fig. 2. Fluorescence in situ hybridization analysis (*8*) of the biofilm collected from the same site as Fig. 1 (AB end) in January 2004. In both images, *Leptospirillum* group II is yellow, and other bacteria (*Sulfobacillus* spp.) are red. Archaea are probed in (**A**) and appear blue; *Leptospirillum* group III are probed in (**B**) and appear white.

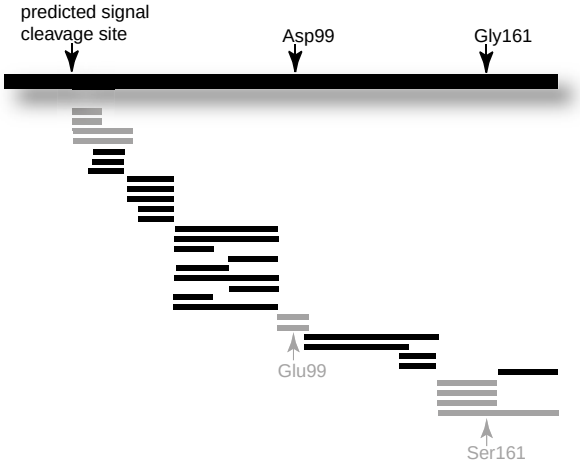


Fig. 3. Recovery of peptides spanning the entire sequence of a natural variant of cyt_{579}. The predicted sequence for cyt_{579}, on the basis of the community genomic data (*2*), is represented by a large black bar. Below, the smaller black bars represent tryptic peptides identified through proteomic analysis. Gray bars represent regions of the mature protein recovered by MS after consideration of two amino acid differences due to strain variation and cleavage of an N-terminal signal peptide.

coverage), further examination suggests that this protein plays another role, possibly in methylthioadenosine recycling (*26*). The high MS detection of *Por* genes and carbon monoxide dehydrogenase with acetyl coenzyme A (acetyl-CoA) synthase in *Leptospirillum* group II may reflect carbon fixation via the acetyl-CoA pathway. NifH, encoded by *Leptospirillum* group III was detected, which is consistent with the inference that this relatively low abundance organism is central to nitrogen fixation in the AMD system (*2*). We also detected proteins involved in nitrogen regulation and ammonia uptake in *Leptospirillum* group II, as would be expected if this organism depends on nitrogen fixed by *Leptospirillum* group III. However, failure to detect other *Leptospirillum* Nif proteins (*27*) suggests a relatively low level of nitrogen fixation at the time of sampling. Both nitrogen fixation and carbon fixation via the acetyl-CoA pathway suggest that some metabolic activities occur in microenvironments protected from molecular oxygen.

Each of the genomes from the biofilm organisms encodes genes that are potentially used to synthesize extracellular polymers. Genes involved in production of cellulose, a likely biofilm constituent (*28*), are expressed by *Leptospirillum* group II; however, many of the polymer production proteins of *Leptospirillum* group II were not detected. Despite the much lower degree of sampling of the *Leptospirillum* group III and *Ferroplasma* type II genomes, we detected proteins from these organisms that may be involved in creating biofilm architecture. Overall the proportion of detected proteins involved in carbohydrate metabolism is considerably greater in *Leptospirillum* group III than in *Leptospirillum* group II (Fig. 4), which implies that *Leptospirillum* group III plays a key role in this biofilm-essential function.

For 15 putative operons composed only of hypothetical genes, we detected all the predicted proteins (table S2). For example, one operon encodes five *Leptospirillum*-specific proteins, and another encoded three *Leptospirillum* group II–specific proteins, all of which were found in the membrane and extracellular fractions (*8*). These operons of lineage-specific proteins may provide functions that are central to adaptation to the AMD environment. Clues to the functions of other novel proteins were inferred from operon structure (Fig. 5). For *Leptospirillum* group II, we detected 280 novel proteins encoded within 212 operons, almost all of which encoded one or more genes with a functional annotation (table S2). For example, one operon that encoded two novel membrane-associated proteins also encoded three proteasome subunits. Two unique proteins were encoded in a four-gene operon with two putative nitrogen regulatory proteins, which suggests roles in nitrogen metabolism. Four other novel proteins, possibly involved in motility, were encoded in an operon of 15 genes that included at least eight flagellar genes.

Both the *Leptospirillum* group II proteome and the community proteome display a bimodal distribution of isoelectric points around ~6 to 6.9 and ~9 to 9.9 (figs. S2 and S6). In contrast, the distribution of isoelectric points for proteins enriched in the extracellular fraction is predominantly in the range ~9 to 10.9 (fig. S6). Thus, separation based on isoelectric points may be useful during purification of acid-stable proteins for functional characterization.

Often the hypothetical proteins not detected were encoded in blocks. In *Leptospirillum* group II, more than 15 genome fragments including at least 273 genes (50% of them novel) have been identified as of probable plasmid origin on the basis of the presence of typical plasmid genes and the absence or low abundance of core metabolic genes and the absence of tRNAs (table S2). On average, products of only 14% of genes encoded on the putative plasmid fragments were detected, in contrast to 51% of genes on chromosome-like fragments. This low incidence of protein detection (table S2) indicates that many laterally acquired genes serve no function, are rarely important, or are expressed only at low levels.

The *Leptospirillum* group II genome encodes more than 100 transposases or transposase fragments, and products of 8 of at least 17 distinct transposase groups were detected. It is noteworthy that the products of transposase genes that are highly duplicated in the *Leptospirillum* group II genome (e.g., genes grouped as T5, table S2) were not detected. However, we did find the products of transposase genes that occur only once or twice in the genome (e.g., T1 and T2 in table S2). One detected transposase is a strain-specific type found in some *Ferroplasma* type I variants.

More than 20 partial or complete integrases or recombinases were encoded in the *Leptospirillum* type II genome, and three distinct integrases or recombinases were found associated with genomic regions inferred to be acquired by lateral transfer. This finding indicates substantial plasmid and/or phage activity in the community at the time of sampling. One detected integrase identical in *Ferroplasma* type I and *Ferroplasma* type II has an unusually low GC content (29%, compared with ~38% in *Ferroplasma* typically) and may be derived from a prophage that recently infected both species.

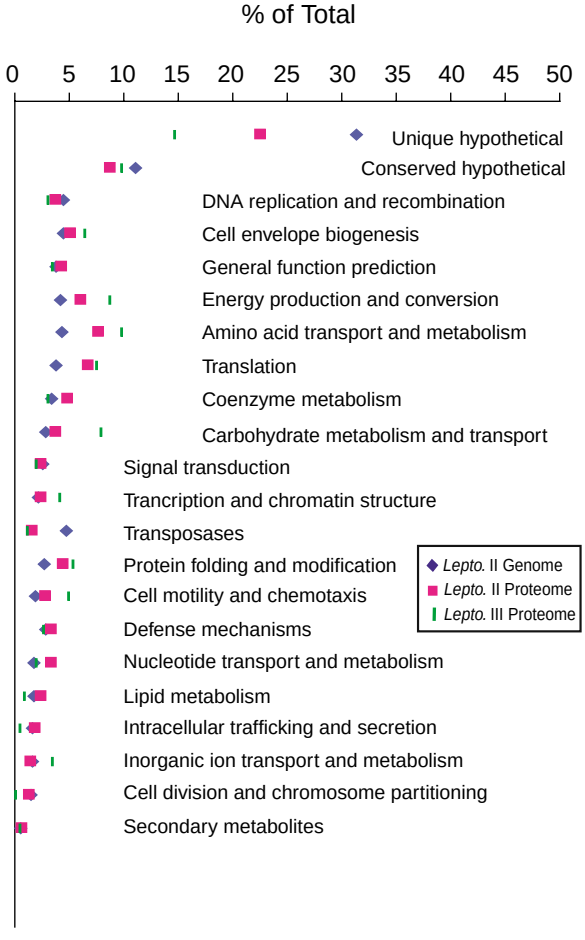


Fig. 4. Functional categories of *Leptospirillum* group II proteins predicted from the genome dataset and *Leptospirillum* groups II and III detected in the proteome. The percentage of total proteins or genes in each category are depicted.

Although it is well documented that microorganisms promote AMD formation (*22*, *23*), little was known about how these organisms function in their natural environments or as consortia. Our "proteogenomic" methods revealed that, in addition to extreme acidity and metal toxicity, reactive oxygen species are a significant challenge in this environment. Moreover, biofilm polymer production and nitrogen fixation appeared to be partitioned among community members, and many novel environment- and/or lineage-specific proteins were expressed that are presumably important. One of these was an abundant acid-stable protein capable of iron oxidation, a process central to AMD generation. However, many novel genes associated with phage- and plasmid-like insertions were only weakly expressed or not expressed.

MS–based de novo sequencing approaches (*29*) and MS^3 analysis (*30*) should reduce the requirement for exact gene sequence data, which would broaden the applicability of genome sequence information for environmental studies. As more environmental genomic sequences become available and MS methods improve, proteogenomics may be applied to other natural communities of environmental, medical, or industrial importance.

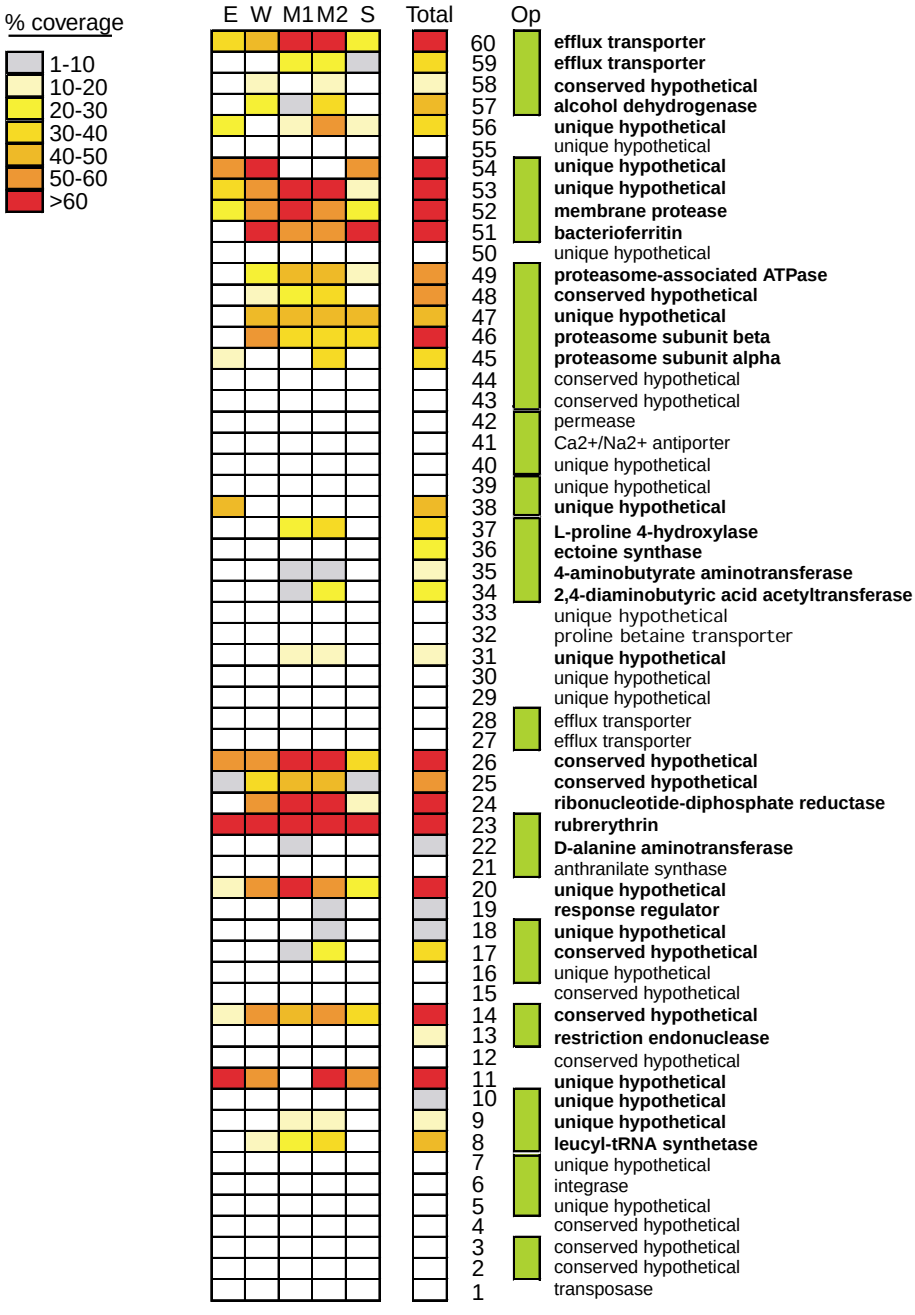


Fig. 5. Characterization of a genome fragment using the proteome dataset. The diagram shows the annotation, putative operon (Op) structure, and gene number on *Leptospirillum* group II scaffold 21. If the protein encoded by a gene was confidently detected (i.e., matching of two or more peptides), its annotation is in bold type. Colored boxes convey the percentage of each protein detected via MS in extracellular (E), whole-cell (W), membrane (M1 and M2), and cytoplasmic (S) fractions, as well as in the combined biofilm fractions (T). Membrane fractions were prepared by using two different protocols (*8*).

References and Notes

1. E. F. Delong, N. R. Pace, *Syst. Biol.* **50**, 470 (2001).
2. G. W. Tyson *et al.*, *Nature* **428**, 37 (2004).
3. J. C. Venter *et al.*, *Science* **304**, 66 (2004).
4. M. S. Lipton *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 11049 (2002).
5. J. Peng, J. E. Elias, C. C. Thoreen, L. J. Licklider, S. P. Gygi, *J. Proteome Res.* **2**, 43 (2003).
6. W. Zhu, C. I. Reich, G. J. Olsen, C. S. Giometti, J. R. Yates III, *J. Proteome Res.* **3**, 538 (2004).
7. G. K. Druschel, B. J. Baker, T. H. Gihring, J. F. Banfield, *Geochem. Trans.* **5**, 13 (2004).
8. For details, supplementary online materials are available at *Science* Online.
9. A. S. Essader, B. J. Cargile, J. L. Bundy, J. L. Stephenson Jr., *Proteomics* **5**, 24 (2005).
10. N. C. Verberkmoes *et al.*, *J. Proteome Res.* **1**, 239 (2002).
11. W. H. McDonald *et al.*, *Int. J. Mass Spectrom.* **219**, 245 (2002).
12. M. P. Washburn, D. Wolters, J. R. Yates III, *Nat. Biotechnol.* **19**, 242 (2001).
13. R. W. Corbin *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 9232 (2003).
14. P. Watnick, R. Kolter, *J. Bacteriol.* **182**, 2675 (2000).
15. We used spectral count (number of spectra detected per protein), peptide count (number of times any unique peptide is detected from a protein), and sequence coverage (percentage of protein sequence detected) to infer the 10 most abundant proteins in each fraction. All three measures gave generally similar results.
16. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, *J. Mol. Biol.* **215**, 403 (1990).
17. Significant homology required an EXPECT (E) value of $<e^{-10}$ by BLASTP analysis when compared with nonredundant proteins in the National Center for Biotechnology Information (NIH) database.
18. Proteins enriched in the extracellular fraction were defined as having an average MS sequence coverage in this fraction that is >10% and more than twice the average sequence coverage in any other fraction.
19. J. D. Bendtsen, H. Nielsen, G. von Heijne, S. Brunak, *J. Mol. Biol.* **340**, 783 (2004).
20. A. Hart, J. C. Murrell, R. K. Poole, P. R. Norris, *FEMS Microbiol. Lett.* **81**, 89 (1991).
21. R. C. Blake II *et al.*, *FEMS Microbiol. Rev.* **11**, 9 (1993).
22. G. W. Luther *et al.*, *Geochim. Cosmochim. Acta* **51**, 3193 (1987).
23. B. J. Baker, J. F. Banfield, *FEMS Microbiol. Rev.* **44**, 139 (2003).
24. R. L. Tatusov, E. V. Koonin, D. J. Lipman, *Science* **278**, 631 (1997).
25. R. C. Blake II, B. Johnson, *Environmental Microbe-Metal Interactions*, D. R. Lovley, Ed. (ASM Press, Washington, DC, 2000), pp. 53–78.
26. T. E. Hanson, F. R. Tabita, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 4397 (2001).
27. V. Parro, M. Moreno-Paz, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 7883 (2003).
28. R. J. Ram *et al.*, unpublished observations.
29. K. G. Standing, *Curr. Opin. Struct. Biol.* **13**, 595 (2003).
30. J. V. Olsen, M. Mann, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 13417 (2004).
31. http://compbio.ornl.gov/biofilm_amd/

32. We thank T. Arman (owner, Iron Mountain mines), R. Sugarek, D. Dodds, and R. Carver for mine access and on-site assistance. We thank D. Tabb, the Yates Laboratory, the Institute for Systems Biology for software, F. Larimer for computer access, U. Kappler for heme-staining protocols, W. Sand for strain P3A, E. Allen for help with genomic analysis, J. Flanagan

for genomic DNA, and R. Whitaker and G. Fox for their comments on the manuscript. Research was supported by grants from the U.S. Department of Energy (DOE) Microbial Genome Program (J.F.B.), NSF Biocomplexity Program (J.F.B.), NASA Astrobiology Institute (J.F.B.), DOE Energy Biosciences Program (R.C.B.), and DOE Genomes to Life Program (R.H.). DDBJ/EMBL/GenBank accession numbers for the community genomic data are AADL01000000 and AADL00000000. All data, databases, and resulting identifications are available online; see (*31*).

Supporting Online Material
www.sciencemag.org/cgi/content/full/1109070/DC1
Materials and Methods
SOM Text
Figs. S1 to S6
Tables S1 and S2
References and Notes

23 December 2004; accepted 27 April 2005
Published online 5 May 2005;
10.1126/science.1109070
Include this information when citing this paper.

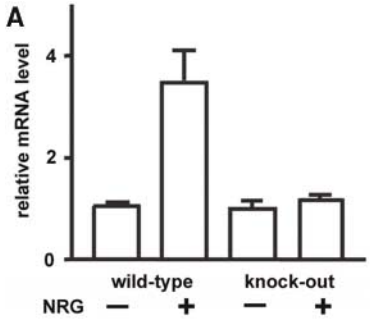
Synapses Form in Skeletal Muscles Lacking Neuregulin Receptors

P. Escher,[1]* E. Lacazette,[1]* M. Courtet,[1] A. Blindenbacher,[1] L. Landmann,[2] G. Bezakova,[3] K. C. Lloyd,[4] U. Mueller,[5] H. R. Brenner[1]†

The formation of the neuromuscular junction (NMJ) is directed by reciprocal interactions between motor neurons and muscle fibers. Neuregulin (NRG) and Agrin from motor nerve terminals are both implicated. Here, we demonstrate that NMJs can form in the absence of the NRG receptors ErbB2 and ErbB4 in mouse muscle. Postsynaptic differentiation is, however, induced by Agrin. We therefore conclude that NRG signaling to muscle is not required for NMJ formation. The effects of NRG signaling to muscle may be mediated indirectly through Schwann cells.

Motor neurons, terminal Schwann cells, and muscle fibers express two neuregulin genes, *nrg-1* and *nrg-2* (*1*–*3*). NRGs activate receptor tyrosine kinases of the ErbB family (*4*) that are expressed in the same cells. Motor neuron–derived NRG1β, originally isolated from brain as acetylcholine receptor (AChR)–inducing activity (ARIA) in cultured myotubes (*1*), is widely accepted to induce *AChR* gene transcription in subsynaptic myonuclei by activating ErbB receptors; the receptors stimulate the same regulatory elements in *AChR* gene promoters as the motor nerve (*5*–*7*). Agrin, another nerve-derived signal for NMJ formation (*8*), activates its receptor MuSK in muscle (*9*) and thus recruits AChRs, muscle-derived NRGs, ErbBs, and other synaptic components to the subsynaptic muscle membrane. Whereas the importance of Agrin in NMJ formation has been demonstrated by genetic means (*10*), the analysis of NRG/ErbB signaling has been less obvious. Mice lacking *nrg1*, *erbB2*, or *erbB4* die during development because of heart malformation, which prevents analysis of their function in synapse development. Transgene rescue experiments or muscle-specific gene ablation have been used to circumvent embryonic lethality. However, because Schwann cells depend on NRG signals from motor neurons for their survival and differentiation (*11*) and signals from Schwann cells regulate motor neuron function (*12*), the effect of NRG on subsynaptic muscle differentiation could not be resolved in the genetically rescued mice. Moreover, whereas ErbB2 is required for muscle spindle development (*13*), neither ErbB2 nor ErbB4 receptors alone are essential for NMJ formation (*14*). However, they may have redundant functions on synapse-specific gene expression (*3*).

To abolish all NRG signaling to muscle, we inactivated both the ErbB2 and ErbB4 receptors selectively in muscle without affecting their expression in Schwann cells. To this end, we crossed mice double-homozygous for loxP-flanked alleles of the *erbB2* (*13*) and *erbB4* (*15*) genes with a mouse line that begins to express Cre recombinase under the control of the human skeletal actin promoter in muscle precursors at embryonic day 9 (E9) (*16*), i.e., before neuromuscular synapse formation at E13.



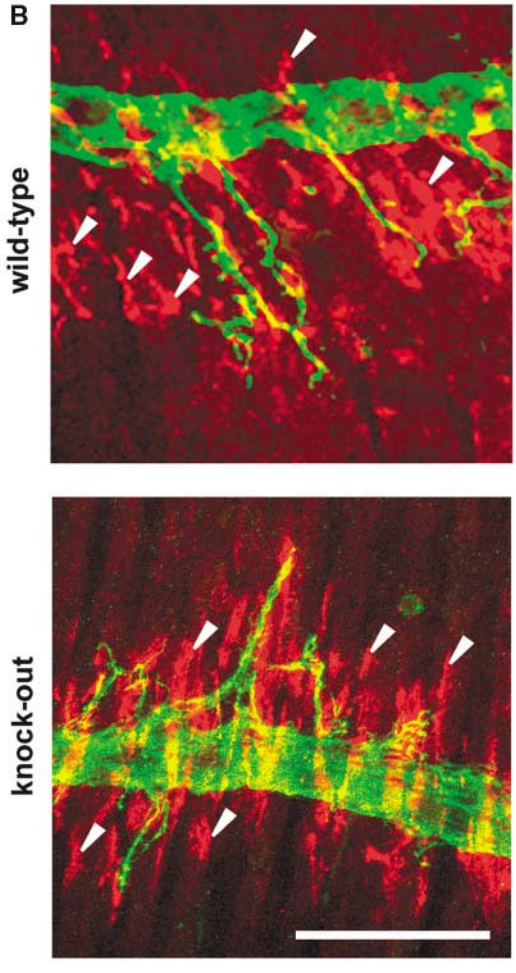


Fig. 1. NRG signaling is blocked in mutant myotubes; fetal neuromuscular junctions in *HSA-Cre;erbB2$^{fl/-}$;erbB4$^{fl/fl}$* mice develop normally. (**A**) Wild-type, but not ErbB-deficient myotubes express AChRε mRNA in response to 5 nM NRG1β (means ± SEM, from three mutant and four wild-type cultures). *P* values in two-tailed *t* tests are <0.01 for wild-type and >0.2 for mutant cultures. (**B**) Wild-type and mutant diaphragm muscles stained for AChR (red), neurofilament, and synaptophysin (green) at E14.5. Arrowheads denote nerve-free AChR clusters. Scale bar, 50 μm.

[1]Institute of Physiology and [2]Institute of Anatomy, Department of Clinical-Biological Sciences, [3]Division of Pharmacology, Biozentrum, University of Basel, 4056 Basel, Switzerland. [4]Center for Comparative Medicine, School of Veterinary Medicine, University of California, Davis, CA 95616, USA. [5]Department of Cell Biology and Institute for Childhood and Neglected Disease, The Scripps Research Institute, La Jolla, CA 92037, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: Hans-Rudolf.Brenner@unibas.ch