Minireview

The dynamic proteome of Lyme disease *Borrelia*

Steven J Norris

Address: Departments of Pathology and Laboratory Medicine and Microbiology and Molecular Genetics, University of Texas Medical School, Houston, TX 77225-0708, USA. Email: Steven.J.Norris@uth.tmc.edu

Published: 17 March 2006

Genome ***Biology*** 2006, **7:**209 (doi:10.1186/gb-2005-7-3-209)

The electronic version of this article is the complete one and can be found online at http://genomebiology.com/2005/7/3/209

Abstract

The proteome of the spirochete bacterium *Borrelia burgdorferi*, the tick-borne agent of Lyme disease, has been characterized by two different approaches using mass spectrometry, providing a launching point for future studies on the dramatic changes in protein expression that occur during transmission of the bacterium between ticks and mammals.

We have all experienced it: the 'deer in the headlights' sensation of dumbfounded wonderment and awe when confronted with our first genome sequence. This is a particularly likely response to the genome of the spirochete *Borrelia burgdorferi* and its relatives - spiral bacteria that are transmitted by deer ticks of the *Ixodes ricinus* group and that cause the chronic disease Lyme borreliosis in humans and other animals. Although there were previous indications of its unusual characteristics, no one anticipated that the 1.5 megabase genome of *B. burgdorferi* would contain an odd mixture of approximately 12 linear and 9 circular plasmids, as well as a 0.9 megabase linear chromosome [1,2]. The plasmids, which range from 5 to 56 kilobases, are present consistently in the strains examined so far and contain genes required for the spirochete's life cycle; these replicons thus could be considered 'mini-chromosomes', although the term plasmids is typically used for simplicity. The 833 predicted plasmid-encoded open reading frames (ORFs) include 454 hypothetical genes, many of which are members of 107 paralogous families of unknown function. The plasmids are also rife with pseudogenes, leading to the conclusion that the Lyme disease spirochete genome is 'in flux': that is, it is actively evolving [2]. The unique properties of the *Borrelia* proteome are now starting to be revealed with two recent studies of global protein expression. In the first approach, Jacobs *et al.* [3] compared the protein profiles of three strains of *B. burgdorferi*, while in the second, Nowalk *et al.* [4] examined the proteins present in the soluble and membrane-associated fractions of the bacterium.

Lyme disease is a chronic disease marked by skin lesions, debilitating neurologic symptoms and arthritis. *B. burgdorferi* is the predominant cause of human Lyme borreliosis in North America, whereas *B. burgdorferi* and two related *Borrelia* species, *B. garinii* and *B. afzelii*, cause disease in Eurasia. These obligate pathogens have a precarious life cycle in which they alternate between two distinct environments: the tick intestinal tract (midgut) and mammalian (or, in some instances, avian) tissue (Figure 1). In humans, Lyme disease *Borrelia* causes a local lesion called erythema migrans at the site of the tick bite and then readily disseminates through the bloodstream to other tissues, setting up an infection that can last for months to years. The bacteria can also persist in ticks for years, but they increase greatly in numbers and migrate to the salivary glands at the time of feeding. Gene expression at the RNA level has been studied using both array and quantitative reverse transcriptase (RT)-PCR approaches, and dramatic changes in gene expression on transmission of the pathogen from ticks to humans have been found (Figure 1). For example, transcript levels for the outer surface lipoprotein OspC can increase 30- to 120-fold in 'mammalian tissue-like' conditions compared with 'unfed tick-like' conditions, whereas OspA, another surface lipoprotein that binds to the tick midgut receptor TROSPA, thus enabling the bacterium to invade its tick host, is downregulated during transmission from tick to mammal [5-8]. Some of the regulatory pathways in this adaptive process have been identified, including a pathway involving the transcription initiation factors RpoN and RpoS. Temperature, dissolved

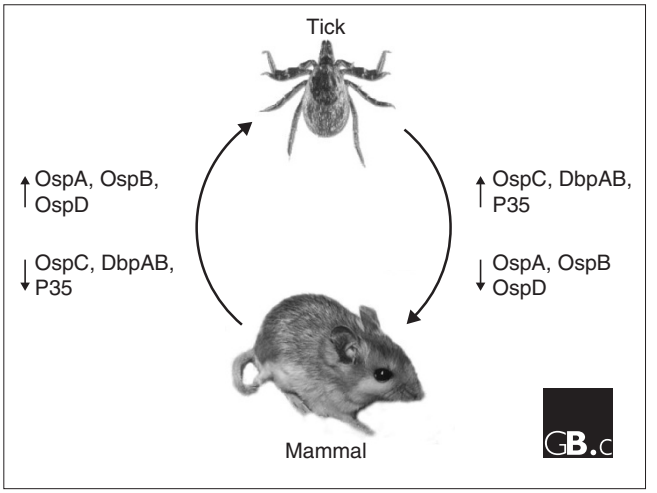

Figure 1
A simplified view of the life cycle of *Borrelia burgdorferi*. The expression of genes encoding approximately 200 proteins is dramatically altered during transmission of the bacterium from tick to mammal or mammal to tick, as exemplified by the changes in the proteins listed: ↑, upregulation; ↓, downregulation. Further details of individual proteins are in the text.

oxygen, and pH play a role in gene regulation, as well as other as-yet unidentified host factors.

It has been difficult to examine protein expression directly during mammalian or tick infection, as only a small number of spirochetes are present during most phases of infection, limiting the utility of conventional methods. A model system in which *B. burgdorferi* cultures are 'incubated' within dialysis tubing in the abdomens of rabbits or rats has been used extensively to study adaptation to the mammalian environment [9]. This set-up excludes contact with host cells and extracellular matrix, however, and some aspects of adaptation (for example, recombination in the antigenic variation gene *vlsE*, which encodes a variable lipoprotein) do not occur under these conditions. A novel approach to the direct study of bacterial protein expression in infected tissues took advantage of the fact that lipoproteins, prominent in *B. burgdorferi* and other spirochetes, are selectively partitioned to the detergent phase following solubilization in Triton X-114 [10]. By this means, VlsE, OspC, and the decorin-binding adhesin DbpA were found to be expressed at high levels in mouse joints and dermal tissue, and OspC and DbpA, but not VlsE, were found in heart tissue. These results suggest that protein expression varies between tissues; this pattern may be related to tissue tropism of the bacteria.

Global analysis of protein expression in Lyme disease *Borrelia* is now under way and will be useful not only for examining changes in gene expression, but also in understanding the biological importance of the multiple paralogous gene families and other unique properties of the predicted proteome. In one approach, Jacobs *et al.* [3] compared the protein profiles of three strains of *B. burgdorferi* using trypsin cleavage of whole-organism preparations followed by a two-step liquid chromatography separation of fragments and tandem mass spectrometry (MS/MS). In a second approach, Nowalk *et al.* [4] used two-dimensional gel electrophoresis with tryptic digestion and MS analysis of individual spots to look at the proteins in the soluble and particulate fractions of *B burgdorferi* strain B31.

Proteomic comparison of three *B. burgdorferi* strains

Jacobs *et al.* [3] compared the three strains B31, N40, and JD-1, which were chosen because they represent three genotypic groups that differ in their patterns of pathogenesis in humans and experimentally infected mice. Whole-cell lysates were treated with trypsin, and the resulting complex mixture of fragments was separated by strong cation exchange resin chromatography. Fractions of that separation were then subjected to reverse-phase capillary chromatography coupled with MS/MS analysis of the most abundant fragments in the initial MS separation. Between 13,500 and 17,300 peptides were isolated from each strain, and a total of 6,982 were confidently identified by comparison with protein predictions from the B31 genome sequence. From these data, 522, 498, and 471 proteins were detected in the B31, N40, and JD-1 preparations, respectively, which together represent 665 proteins, or roughly 38% of the predicted proteome.

The library of detected proteins from each strain exhibited a high degree of overlap (Figure 2): 52% of the proteins were detected in all three strains, and 72% were found in at least two of the strains. Between 59 and 69 proteins were identified in only one strain, which could be due to a number of factors. The most interesting possibility is that protein expression differs among the strains, correlating with the variations in pathogenesis. Some indication that this may be the case was provided by Jacobs *et al.* [3]; for example, 47% and 39% of the tryptic peptide coverage of the hypothetical protein BBH37 was identified in the B31 and JD-1 strains, respectively, whereas no BBH37 peptides were detected in N40. Two other proteins in this paralogous family (BBG01 and BBJ08) were apparently deficient in the JD-1 strain in comparison with the other two strains. Judging from the number of peptides detected, several other plasmid-encoded proteins were expressed at different levels in the three strains under the growth conditions used: OspC, OspD (BBJ09), the putative outer membrane porin Oms28 (BBA74), DbpA (BBA24), the fibronectin-binding protein BBK32, and hypothetical proteins BBI39 and BBJ34 (see [3], and in particular its Table 1 and supplementary information).

There are other possible reasons for the uneven detection of certain proteins in the three strains [3]. Only the B31 DNA sequence was used for analysis, although partial sequences

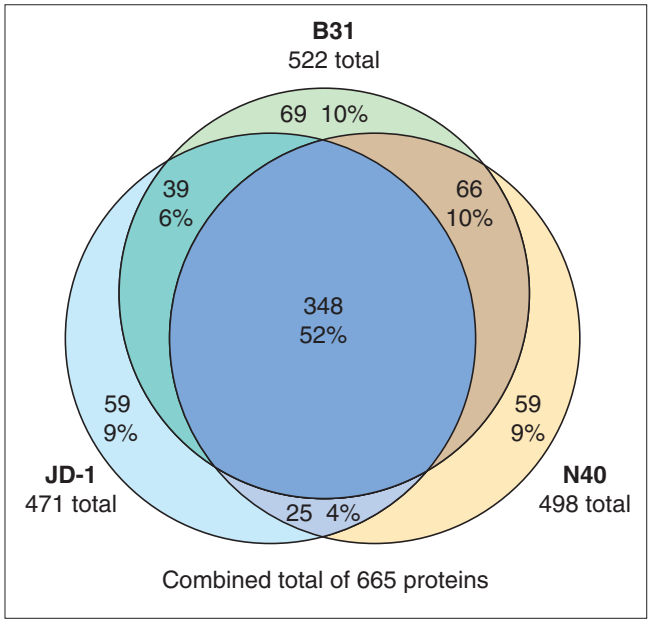

Figure 2
Correlation of the proteins detected by tandem mass spectrometry (MS/MS) of *Borrelia burgdorferi* strains B31, N40, and JD-1. Whole-cell preparations of each strain were solubilized, treated with trypsin, and then subjected to strong cation exchange chromatography and capillary reverse-phase chromatography followed by MS/MS peptide analysis. The number and percentage of proteins that were detected in three, two, or one strains are shown. Adapted from [3].

of the N40 and JD-1 genomes are now available [11]. Only 0.5% pairwise nucleotide differences were observed on average in alignments of the three DNA sequences [11], so relatively few false negatives in tryptic peptide identifications should have resulted from differences between the experimental molecular masses and those of the peptides predicted from the B31 sequence. Other reasons for the occurrence of 'unique' proteins could be low abundance or artifactual differences in abundance. The vast majority of the proteins found in only one or two of the strains were identified from fewer than three peptides (and often from only one). In some cases, the 'missing' proteins are required housekeeping proteins such as tRNA synthetases or a flagellar motor protein, and thus must actually be present in all three strains. Therefore, the uneven detection of low-abundance proteins most probably accounts for most (more than 90%) of the proteins detected in only one or two strains (Figure 2).

Characterization of the soluble and membrane-associated proteome

The other major approach to proteome characterization is two-dimensional gel electrophoresis followed by identification of individual spots on the gel by excision and MS analysis of tryptic peptides. In 1999, Jungblut *et al.* [12] analyzed the antigens of *B. garinii* using this method and the *B. burgdorferi* sequence; only a limited number of proteins were identified, however. Nowalk *et al.* [4] have now examined the proteins present in the soluble and membrane-associated (or, more accurately, the particulate) fractions of *B. burgdorferi* B31. One of the underlying problems with two-dimensional gel electrophoresis of *Borrelia* proteins is the high number and abundance of basic proteins, including the major membrane proteins OspA (pI = 8.3) and OspB (pI = 8.6). Basic proteins often migrate off the end of standard isoelectric focusing gels (even those with a broad pH range) or are poorly resolved. Nowalk *et al.* [4] addressed this issue by using either non-equilibrium pH gradient electrophoresis (NEPHGE) or immobilized pH gradients. With some modifications, NEPHGE generally out-performed the immobilized gradients in terms of both spot resolution and pH range.

A profile of the membrane-associated fraction obtained using NEPHGE followed by SDS gel electrophoresis is shown in Figure 3. Of the 160 spots detected by silver staining in this fraction, 34 proteins were identified by trypsin digestion and matrix-assisted desorption ionization time-of-flight (MALDI-TOF) MS. Similarly, the identities of 83 proteins were determined from the 185 spots in the soluble fraction; 12 proteins were detected in both fractions, so 105 proteins were identified in all. Most of the membrane-associated proteins detected were plasmid-encoded lipoproteins, consistent with the fact that the majority of the 150 predicted lipoproteins of *B. burgdorferi* are plasmid-encoded. Because two-dimensional gel electrophoresis patterns are reproducible, these identifications can be used in future studies to determine the effects of different incubation conditions, treatments, and fractionations on protein composition.

One of the surprising findings of this study was that the glycolytic pathway enzyme enolase was found in nearly equal concentrations in both the soluble and membrane-associated fractions, and aminopeptidase I and the chaperone protein GroEL were present in larger amounts in the membrane fraction than in the soluble fraction [4]. Enolase has, however, been found on the surface of staphylococci and streptococci, where it has been implicated in the adherence of the intact bacteria to plasmin, plasminogen and laminin. Because *B. burgdorferi* has both an inner and an outer membrane, however, it is not yet known whether the membrane-associated enolase is exposed on the surface. As a chaperone, GroEL may associate with membrane proteins during translocation and folding. Aminopeptidase I has also been shown to localize in the cytoplasm, periplasm, membrane and cell-wall fractions of other bacteria.

Comparison of the MS/MS- and gel-electrophoresis-derived proteomes [3,4] indicates, as expected, that the electrophoretic approach primarily identified the more abundant protein species in the MS/MS dataset (based on the number of peptides detected per protein). There were,

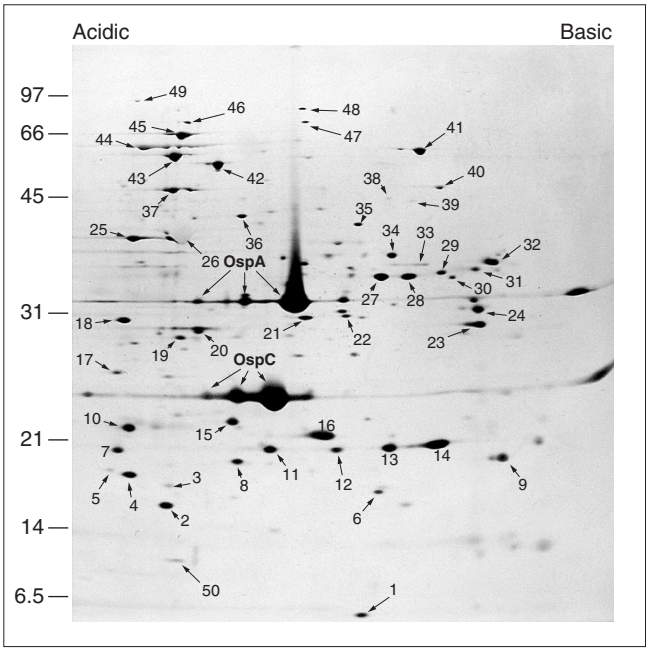

Figure 3
Separation of *Borrelia burgdorferi* membrane-associated proteins by non-equilibrium pH gradient electrophoresis (NEPHGE) followed by SDS gel electrophoresis. The spots indicated were identified by trypsin digestion followed by MALDI-TOF mass spectrometry. The locations of molecular weight markers (in kDa) are indicated on the left hand side. Reprinted with permission from [4].

however, many abundant proteins in the MS/MS group that were not present in the electrophoresis analysis, consistent with the more limited sampling in the electrophoretic approach. The ribosomal proteins were the most prominent under-represented group; others included the periplasmic serine protease DO (BB0104), the putative surface-located lipoprotein Lmp1 (BB0210), glycerol-3-phosphate dehydrogenase (BB0243), the flagellar sheath protein FlaA (BB0668), and some of the RNA polymerase subunits (BB0388, BB0389). Conversely, 11 of the 105 proteins detected by electrophoresis were not represented in the MS/MS dataset, and only one or two peptides were detected by MS/MS for several others. Thus, there are some biases in the two approaches, either in sample preparation or the detection methods themselves.

These initial proteome characterizations lay important groundwork for future studies on the expression patterns and localization of *B. burgdorferi* proteins. Both methods can be adapted to provide quantitative comparisons of protein expression under different conditions, for example, incubation at different temperatures or pH, or conceivably following host adaptation in dialysis membrane chambers implanted in animals. In the whole-cell MS/MS analysis, parallel cultures can be differentially labeled using stable isotopes (^{14}N and ^{15}N) and cysteine affinity tags to quantitate differences in expression patterns [13], as in the recent analysis of the heat-shock response in the radiation-resistant bacterium *Deinococcus radiodurans* [14]. In the electrophoretic method, scanning of gels stained with Coomassie blue or other dyes can provide some quantitation. A more elegant approach, however, is difference gel electrophoresis [15], which comprises the differential labeling of organisms from two different incubation conditions using Ettan fluorescent labeling. The two preparations are labeled separately with derivitized Cy3 and Cy5, and then mixed together before electrophoresis. Fluorescence associated with the polypeptide spots is quantitated, and the corresponding Cy3 and Cy5 signals are used to determine differences in expression patterns between the two conditions.

The initial proteomes of *B. burgdorferi* and other organisms will serve as the basis for the global analysis of protein expression in response to different environments. In turn, the interfacing of these proteome datasets with other '-omes', such as genome, transcriptome, interactome, and immunoproteome, may begin to reflect the actual complexity of bacterial physiology and pathogenesis.

References

1. Fraser CM, Casjens S, Huang WM, Sutton GG, Clayton R, Lathigra R, White O, Ketchum KA, Dodson R, Hickey EK, *et al.*: **Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*.** *Nature* 1997, **390:**580-586.
2. Casjens S, Palmer N, van Vugt R, Huang WM, Stevenson B, Rosa P, Lathigra R, Sutton G, Peterson J, Dodson RJ, *et al.*: **A bacterial genome in flux: the twelve linear and nine circular extrachromosomal DNAs in an infectious isolate of the Lyme disease spirochete *Borrelia burgdorferi*.** *Mol Microbiol* 2000, **35:**490-516.
3. Jacobs JM, Yang X, Luft BJ, Dunn JJ, Camp DG 2nd, Smith RD: **Proteomic analysis of Lyme disease: global protein comparison of three strains of *Borrelia burgdorferi*.** *Proteomics* 2005, **5:**1446-1453.
4. Nowalk AJ, Nolder C, Clifton DR, Carroll JA: **Comparative proteome analysis of subcellular fractions from *Borrelia burgdorferi* by nonequilibrium NEPHGE and IPG.** *Proteomics* 2006, doi:10.1002/pmic.200500187.
5. Rosa PA, Tilly K, Stewart PE: **The burgeoning molecular genetics of the Lyme disease spirochaete.** *Nat Rev Microbiol* 2005, **3:**129-143.
6. Revel AT, Talaat AM, Norgard MV: **DNA microarray analysis of differential gene expression in *Borrelia burgdorferi*, the Lyme disease spirochete.** *Proc Natl Acad Sci USA* 2002, **99:**1562-1567.
7. Ojaimi C, Brooks C, Casjens S, Rosa P, Elias A, Barbour A, Jasinskas A, Benach J, Katona L, Radolf J, *et al.*: **Profiling of temperature-induced changes in *Borrelia burgdorferi* gene expression by using whole genome arrays.** *Infect Immun* 2003, **71:**1689-1705.
8. Tokarz R, Anderton JM, Katona LI, Benach JL: **Combined effects of blood and temperature shift on *Borrelia burgdorferi* gene expression as determined by whole genome DNA array.** *Infect Immun* 2004, **72:**5419-5432.
9. Akins DR, Bourell KW, Caimano MJ, Norgard MV, Radolf JD: **A new animal model for studying Lyme disease spirochetes in a mammalian host-adapted state.** *J Clin Invest* 1998, **101:**2240-2250.
10. Crother TR, Champion CI, Wu XY, Blanco DR, Miller JN, Lovett MA: **Antigenic composition of *Borrelia burgdorferi* during infection of SCID mice.** *Infect Immun* 2003, **71:**3419-3428.
11. Qiu WG, Schutzer SE, Bruno JF, Attie O, Xu Y, Dunn JJ, Fraser CM, Casjens SR, Luft BJ: **Genetic exchange and plasmid transfers in *Borrelia burgdorferi* sensu stricto revealed by three-way**

genome comparisons and multilocus sequence typing. *Proc Natl Acad Sci USA* 2004, **101:**14150-14155.
12. Jungblut PR, Grabher G, Stoffler G: **Comprehensive detection of immunorelevant *Borrelia garinii* antigens by two-dimensional electrophoresis.** *Electrophoresis* 1999, **20:**3611-3622.
13. Smith RD, Anderson GA, Lipton MS, Masselon C, Pasa-Tolic L, Shen Y, Udseth HR: **The use of accurate mass tags for high-throughput microbial proteomics.** *Omics* 2002, **6:**61-90.
14. Schmid AK, Lipton MS, Mottaz H, Monroe ME, Smith RD, Lidstrom ME: **Global whole-cell FTICR mass spectrometric proteomics analysis of the heat shock response in the radioresistant bacterium *Deinococcus radiodurans*.** *J Proteome Res* 2005, **4:**709-718.
15. Marouga R, David S, Hawkins E: **The development of the DIGE system: 2D fluorescence difference gel analysis technology.** *Anal Bioanal Chem* 2005, **382:**669-678.