

Tabled Document

Inquiry: Lyme-like illness
 Date: 20/4/16
 Witness: Dr Gary Lum
 Organisation: D.H.

RESEARCH ARTICLE



Printed by Dr Gary Lum



mBio®

Longitudinal Transcriptome Analysis Reveals a Sustained Differential Gene Expression Signature in Patients Treated for Acute Lyme Disease

Jerome Bouquet,^a Mark J. Soloski,^b Andrea Swei,^c Chris Cheadle,^b Scot Federman,^a Jean-Noel Billaud,^d Alison W. Rebman,^b Beniwende Kabre,^a Richard Halpert,^d Meher Boorgula,^b John N. Aucott,^b Charles Y. Chiu^a

Department of Laboratory Medicine, University of California, San Francisco, San Francisco, California, USA^a; Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA^b; Department of Biology, San Francisco State University, San Francisco, California, USA^c; Qiagen Bioinformatics, Redwood City, California, USA^d

J.B., M.J.S., J.A., and C.Y.C. conceived the project and designed the research. J.B., M.J.S., A.S., and B.K. performed the experiments. J.A. and A.R. enrolled patients in the SLICE study and collected clinical and laboratory data. J.B., M.J.S., S.F., J.-N.B., R.H., M.B., C.C., and C.Y.C. analyzed next-generation sequencing and microarray data. J.B. and C.Y.C. wrote the manuscript.

ABSTRACT Lyme disease is a tick-borne illness caused by the bacterium *Borrelia burgdorferi*, and approximately 10 to 20% of patients report persistent symptoms lasting months to years despite appropriate treatment with antibiotics. To gain insights into the molecular basis of acute Lyme disease and the ensuing development of post-treatment symptoms, we conducted a longitudinal transcriptome study of 29 Lyme disease patients (and 13 matched controls) enrolled at the time of diagnosis and followed for up to 6 months. The differential gene expression signature of Lyme disease following the acute phase of infection persisted for at least 3 weeks and had fewer than 44% differentially expressed genes (DEGs) in common with other infectious or noninfectious syndromes. Early Lyme disease prior to antibiotic therapy was characterized by marked upregulation of Toll-like receptor signaling but lack of activation of the inflammatory T-cell apoptotic and B-cell developmental pathways seen in other acute infectious syndromes. Six months after completion of therapy, Lyme disease patients were found to have 31 to 60% of their pathways in common with three different immune-mediated chronic diseases. No differential gene expression signature was observed between Lyme disease patients with resolved illness to those with persistent symptoms at 6 months post-treatment. The identification of a sustained differential gene expression signature in Lyme disease suggests that a panel of selected human host-based biomarkers may address the need for sensitive clinical diagnostics during the “window period” of infection prior to the appearance of a detectable antibody response and may also inform the development of new therapeutic targets.

IMPORTANCE Lyme disease is the most common tick-borne infection in the United States, and some patients report lingering symptoms lasting months to years despite antibiotic treatment. To better understand the role of the human host response in acute Lyme disease and the development of post-treatment symptoms, we conducted the first longitudinal gene expression (transcriptome) study of patients enrolled at the time of diagnosis and followed up for up to 6 months after treatment. Importantly, we found that the gene expression signature of early Lyme disease is distinct from that of other acute infectious diseases and persists for at least 3 weeks following infection. This study also uncovered multiple previously undescribed pathways and genes that may be useful in the future as human host biomarkers for diagnosis and that constitute potential targets for the development of new therapies.

Received 20 January 2016 Accepted 26 January 2016 Published 12 February 2016

Citation Bouquet J, Soloski MJ, Swei A, Cheadle C, Federman S, Billaud J-N, Rebman AW, Kabre B, Halpert R, Boorgula M, Aucott JN, Chiu CY. 2016. Longitudinal transcriptome analysis reveals a sustained differential gene expression signature in patients treated for acute Lyme disease. mBio 7(1):e00100-16. doi:10.1128/mBio.00100-16.

Editor Alan G. Barbour, University of California Irvine

Copyright © 2016 Bouquet et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

Address correspondence to John Aucott, jaucott@jhmi.edu, or Charles Y. Chiu, charles.chiu@ucsf.edu.

Lyme disease, a systemic tick-borne infection caused by the bacterial spirochete *Borrelia burgdorferi*, is the most common vector-borne disease in the United States and Europe (1). Over 30,000 cases in the United States are reported annually to the Centers for Disease Control and Prevention (CDC) (<http://www.cdc.gov/lyme/stats/humancases.html>). However, actual prevalence estimates are at least 10 times as high because of underreporting of cases and overreliance on insensitive diagnostic tests in the acute phase of infection (2). Lyme disease has been associated with arthritis, meningitis, facial palsy, and (rarely) myocarditis resulting in sudden death (3). Most patients treated with appropriate antibiotics recover rapidly and completely, but a minority

of patients develop persistent symptoms correlating with disseminated disease, a greater severity of illness at presentation, and delayed antibiotic therapy (4). The proportion of Lyme disease patients with persistent symptoms varies greatly, from 0 to 50%, depending on the cohort of interest and the case definition used (4, 5). When lingering or recurrent symptoms are associated with a functional decline and persist for greater than 6 months, patients are considered to meet clinical criteria for post-treatment Lyme disease syndrome (PTLDS) (6), although the exact molecular mechanisms underlying this condition remain unknown.

Control of *B. burgdorferi* infection in early Lyme disease requires both innate and adaptive immune responses (7). Phago-

cytes constitute the first line of defense, engulfing the spirochete and producing Th1-type proinflammatory cytokines. Spirochetal lipoproteins can directly stimulate the B-cell response, and both lipidated and nonlipidated proteins trigger T-cell-dependent humoral responses. Decreased Th1 and increased Th17 responses have also been shown to play a role in the development of post-treatment Lyme disease symptoms during the chronic phase of the illness (8, 9). However, with the exception of antibiotic-refractory Lyme arthritis, very few studies have looked at the molecular mechanisms underlying persistent symptomatology in treated Lyme disease patients, and all to date have used targeted approaches assaying specific cytokine levels (8–10). The overall global and temporal pathways involved in human clinical infection with *B. burgdorferi* remain to be elucidated.

In this study, we applied next-generation sequencing of peripheral blood mononuclear cells (PBMCs) to investigate the transcriptomes of 29 patients with acute Lyme disease longitudinally from the time of diagnosis to 6 months post-treatment and those of 13 matched controls. We performed network and pathway analyses in order to gain insights into the molecular mechanisms underpinning acute Lyme disease and post-treatment symptoms and to discover potential diagnostic biomarkers.

RESULTS

Patient enrollment, sample collection, and transcriptome analysis. This study included a cohort of 29 patients with acute Lyme disease and 13 matched controls without acute illness. Transcriptome profiling by RNA sequencing (RNA-Seq) and pathway analysis were performed with PBMC samples collected at three time points, V1 (time of acute Lyme disease diagnosis and prior to starting antibiotic therapy), V2 (immediately after the completion of a 3-week course of doxycycline treatment), and V5 (6 months after the completion of therapy) (Fig. 1). Approximately 73 (± 43 [standard deviation]) million reads were generated per sample, and on average, 64.9% of the genes had nonzero counts (see Fig. S1 in the supplemental material).

No significant differences in age, sex, ethnicity, or preexisting comorbidities were noted between Lyme disease patients and controls (Table 1). Two-tiered antibody testing for Lyme disease with whole-cell lysates was positive in 20 (71.4%) of 28 patients tested, with 14/28 (50%) patients testing positive at the pretreatment visit and an additional 6/28 (21.4%) seroconverting during treatment (Table 1). The 29 Lyme disease patients were enrolled in a single season at the same geographic location, an outpatient clinic in suburban Maryland. At the 6-month follow-up visit (V5), 15 patients had fully recovered from the infection while 13 experienced persistent symptoms post-treatment, defined as new-onset fatigue, widespread musculoskeletal pain involving ≥ 3 joints, and/or cognitive dysfunction (11); 1 patient was lost to follow-up. Of the 13 patients with persistent symptoms, 4 were diagnosed with PTLDS on the basis of a recently proposed standardized case definition that included a documented functional decline at 6 months as a key criterion (6).

Six (40%) of the 15 patients with resolved illness and 6 (46%) of the 13 with persistent symptoms presented with early disseminated disease consisting of multiple erythema migrans (EM) lesions at the time of diagnosis (see Table S1 in the supplemental material). The average duration of acute illness, defined as the time from onset of EM rash and/or influenza-like symptoms to study enrollment and initiation of doxycycline therapy, was sig-

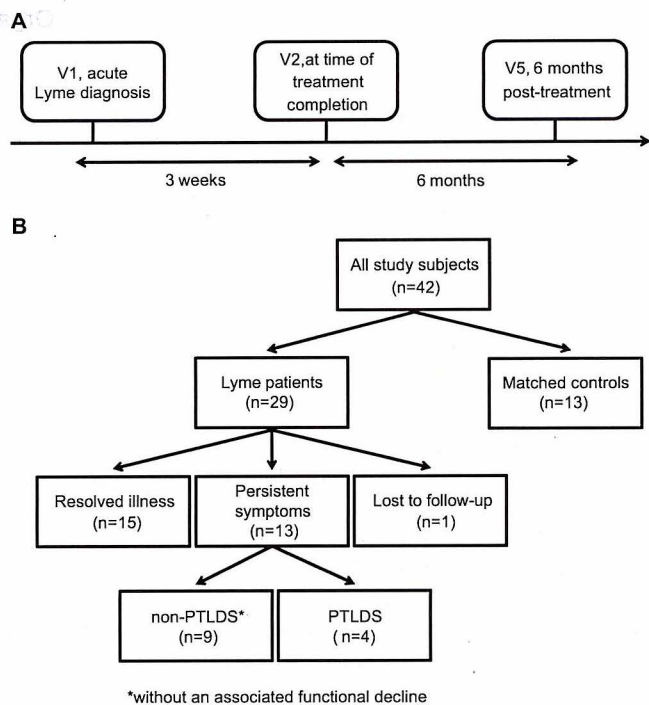


FIG 1 Schematic description of study design. (A) Timeline of clinical evaluation and PBMC sampling. (B) Flowchart of the number of patients with resolved illness or persistent symptoms. Abbreviations: non-PTLDS, post-treatment Lyme disease symptoms and no functional decline; PTLDS, PTLDS with a functional decline.

nificantly longer in patients developing persistent symptoms (9.7 days for non-PTLDS and 19.3 days for PTLDS) than in patients with resolved illness (5.2 days) ($P < 0.036$) (see Table S1 in the supplemental material). In addition, the number of symptoms was significantly higher at all time points in patients with persistent symptoms than in those with resolved illness ($P < 0.04$) (see Table S1 in the supplemental material).

Lyme disease gene expression signature. We initially compared the transcriptomes of 29 Lyme disease patients at the time of diagnosis (V1) with those of 13 matched controls. This analysis revealed a total of 1,235 differentially expressed genes (DEGs) (Fig. 2A; Table 2). Approximately 69% ($n = 847$) of the DEGs were upregulated, and 31% ($n = 388$) were downregulated (Fig. 2A). Three weeks after diagnosis (V2), at the time of completion of a standard course of antibiotic treatment, 1,060 DEGs were found in both Lyme disease patients and controls, with 63% ($n = 670$) upregulated and 37% ($n = 390$) downregulated. Sixty-two percent of the DEGs occurred at both the V1 and V2 time points (Fig. 2B). At 6 months after treatment completion (V5), the Lyme disease transcriptome did not fully return to the baseline relative to controls, with 686 DEGs, 54% ($n = 373$) upregulated and 46% ($n = 313$) downregulated. Partially overlapping clusters were observed for each sample category (V1, V2, V5, and controls) by principal component analysis (PCA) (Fig. 2C).

We then calculated differential gene expression between subjects with single versus multiple disseminated EM lesions and between seropositive and seronegative subjects (Table 2). While no DEGs were identified on the basis of single versus multiple lesions,

TABLE 1 Demographic and clinical characteristics of 29 patients with early Lyme disease and 13 matched controls^a

Variable	Lyme disease patients ^e	Controls ^f	P value ^g
Avg age (yr)	52 (36–61) [20–71]	50 (42–61) [22–70]	0.62
Females	10/29 (34.5)	8/13 (61.5)	0.22
Non-Hispanic Caucasians	27/29 (93.1)	12/13 (92.3)	0.67
≥1 comorbidities	11/29 (38.9)	6/13 (46.2)	0.62
Carpal tunnel syndrome, mo	2/29 (6.9)	1/13 (7.7)	0.41
Diabetes, mo	0/29 (0.0)	2/13 (14.8)	
Heart disease, mo	4/29 (13.8)	1/13 (7.7)	
Lung disease, mo	1/29 (3.4)	0/13 (0.0)	
Migraines, mo	4/29 (13.8)	3/13 (23.1)	
Thyroid disease, mo	3/29 (10.3)	1/13 (7.7)	
Two-tier serology ^b			NA ^h
Pretreatment positive	14/28 (50.0)	0/13 (0.0)	
Seroconverted during treatment	6/28 (21.4)	0/13 (0.0)	
Negative	8/28 (28.6)	13/13 (100)	
Recovery status at V5			NA
Resolved	15/28 (53.6)	NA	
Persistent symptoms	13/28 (46.4)	NA	
Non-PTLDS ^c	9/28 (32.1)	NA	
PTLDS ^d	4/28 (14.3)	NA	
Lost to follow-up	1/29 (2.8)	NA	
Sampling season			
V1			<0.00001
Spring	3/29 (10.3)	3/13 (23.1)	
Summer	24/29 (82.8)	1/13 (7.7)	
Autumn	2/29 (6.9)	1/13 (7.7)	
Winter	0/29 (0.0)	8/13 (61.5)	
V2			<0.00001
Spring	1/28 (3.6)	3/13 (23.1)	
Summer	22/28 (78.6)	1/13 (7.7)	
Autumn	5/28 (17.9)	1/13 (7.7)	
Winter	0/28 (0.0)	8/13 (61.5)	
V5			0.023
Spring	7/28 (25.0)	3/13 (23.1)	
Summer	0/28 (0.0)	1/13 (7.7)	
Autumn	1/28 (3.6)	1/13 (7.7)	
Winter	20/28 (71.4)	8/13 (61.5)	

^a Number/total (%) reported for categorical variables and median, IQR interquartile range (in parentheses), and range (in brackets) presented for continuous variables.^b One patient missing two-tier serology data.^c Non-PTLDS (persistent symptoms with no functional decline).^d PTLDS (persistent symptoms with functional decline).^e n = 29.^f n = 13.^g Lyme disease patients versus controls.^h NA, not applicable.

four DEGs were found to be upregulated in seronegative Lyme disease patients relative to those who were seropositive, namely, *HLA-DQA1*, *HLA-DQB1*, *HLA-DRB5*, and *NSA2* (see Fig. S3 in the supplemental material).

Pathway analyses of the Lyme disease transcriptome. Pathway analysis of Lyme disease DEGs revealed predicted activation of inflammatory response, immune cell trafficking, and hematologic system pathways at V1, as expected in the setting of the acute phase of an infection such as Lyme disease (Fig. 2D and 3). However, the same categories also remained activated following the completion of antibiotic treatment and the clinical resolution of symptoms (V2 and V5), with the general pattern of gene expression more inhibitory at V5 (Fig. 3).

Eight, 10, and 4 of the top 10 predicted canonical pathways at V1, V2, and V5, respectively, were directly related to the host immune response (Fig. 2E). The eIF2 signaling pathway, modulating translational initiation and elongation, was found to be sig-

nificantly downregulated at all three time points (Fig. 2E; see Fig. S4 and Table S2 in the supplemental material). In contrast, TREM1-mediated activation of a Th1 proinflammatory response through upregulation of the factors DAP12, interleukin-6 (IL-6), and IL-12 was prominent at only V1 and V2 (Fig. 2E; see Fig. S4 in the supplemental material). Multiple Toll-like receptors (TLRs) associated with inflammation and apoptosis were also found to be significantly upregulated at V1 and V2 (TLR1, -2, -4, -7, and -8) (see Fig. S4 in the supplemental material).

The most important upstream regulators in Lyme disease at V1 and V2 were found to be proinflammatory cytokines and markers (CSF2, gamma interferon [IFN- γ], IL-1 β , IL-6, tumor necrosis factor alpha [TNF- α]), anti-inflammatory cytokines (IL-6, IL-10), the cell surface marker CD40L, transforming growth factor β 1, the signal transduction mediator TICAM, the transcriptional regulator NF- κ B, and the immunoglobulin complex (Fig. 2F), with TNF- α shown to be a master regulator of eIF2 signaling,

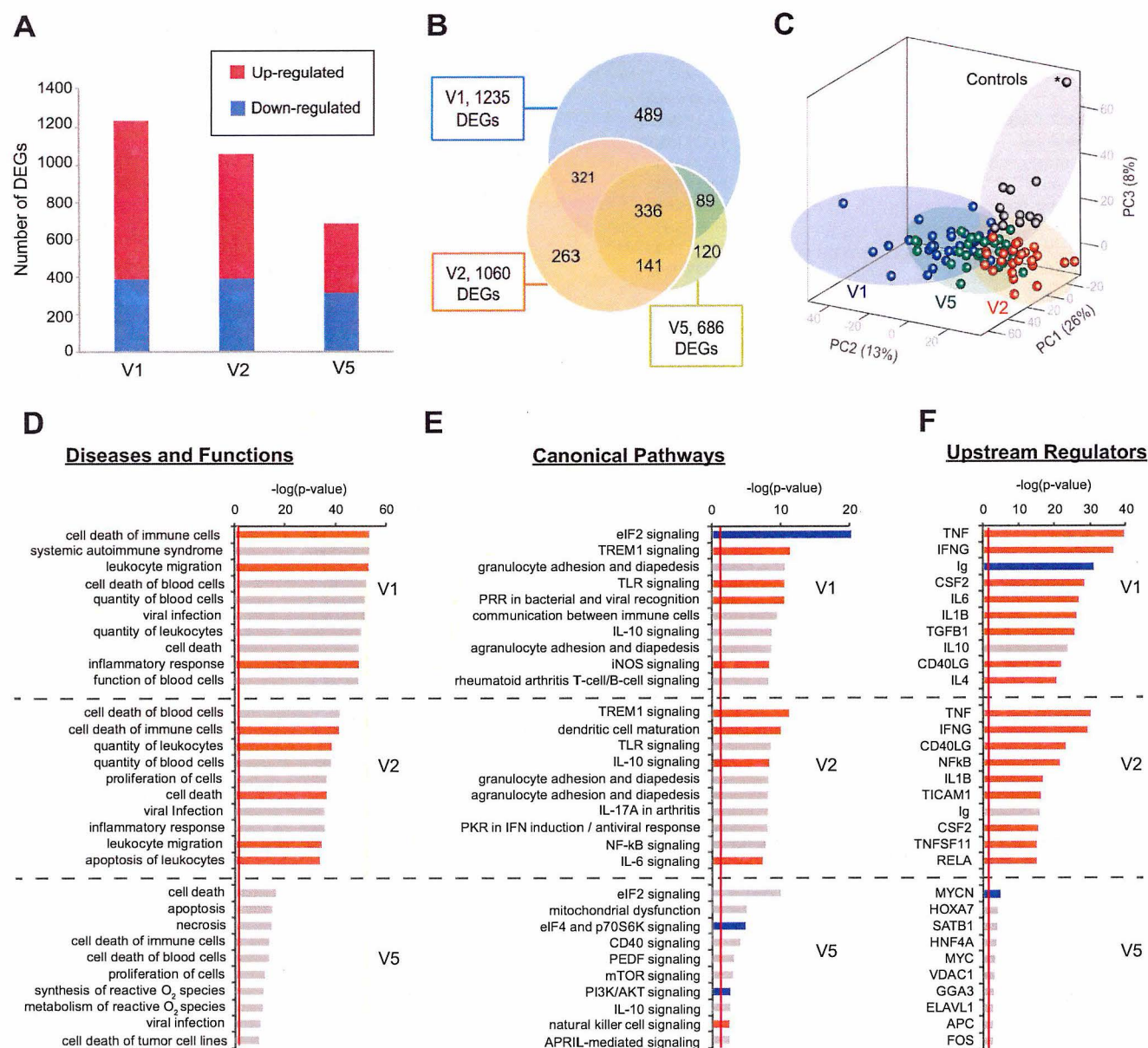


FIG 2 Longitudinal differential gene expression and pathway analysis of Lyme disease. (A) Bar chart of the numbers of genes found to be upregulated or downregulated at Lyme disease diagnosis (V1), 3 weeks post-treatment (V2, after a standard course of antibiotics), and 6 months post-treatment (V5). (B) Venn diagram representing the number of DEGs between Lyme disease patients and controls at three time points. (C) Principal component analysis (PCA) of Lyme disease patients and controls at three time points on the basis of 1,759 unique DEGs identified at V1, V2, and V5. The asterisk represents a subject in the control group who looks like an outlier in the PCA plot but is not shown to be an outlier by PCA analysis of the control samples (see Fig. S2 in the supplemental material). Note that the PC3 axis in the PCA plot accounts for only 8% of the variance in the data set. (D to F) Top 10 disease and functional categories (D), top 10 canonical pathways (E), and top 10 upstream regulators (excluding drug categories) (F) predicted to be involved in Lyme disease at (V1, V2, and V5) with categories, pathways, and genes ranked by the negative log of the P value of the enrichment score. The color scheme is based on Z scores, with activation in orange, inhibition in blue, and undetermined directionality in gray. The red line represents the designated significance threshold ($P < 0.05$).

TREM1, and TLR pathways (see Fig. S4 in the supplemental material). At V5, the top upstream regulators were predominantly involved in the regulation of gene expression (*MYCN*, *HOXA7*, *SATB1*, *HNF4A*, *MYC*, *FOS*, and *ELAVL1*) (Fig. 2F).

Comparison of acute Lyme disease with other infections. We compared our V1 RNA-Seq data, derived from patients with acute Lyme disease, to 12 available, previously published transcriptome data sets from cell culture models of *in vitro* infection or from

patients with acute viral and bacterial infections other than Lyme disease (Fig. 4). Unsurprisingly, the greatest overlap in shared DEGs (44%) was observed with *in vitro* *B. burgdorferi* infection of human PBMCs (44%), followed by infection of human endothelial cells (29%), human neuroblastoma cells (13%), or primary monkey glial cells (11%) (Fig. 4A). We also compared DEGs from acute Lyme disease patients with those corresponding to human PBMCs stimulated *in vitro* by lipopolysaccharides (LPS), infected

TABLE 2 Number of DEGs with a change of greater than ± 1.5 -fold, a P value of < 0.05 , and an FDR of $< 0.1\%$

Condition 1 (no. of subjects)	Condition 2 (no. of subjects)	No. of DEGs at:			All time points
		V1 ^k	V2 ^l	V5 ^m	
All Lyme disease (29)	Control (13)	1,235	1,060	686	644
Resolved Lyme disease (15)	Control (13)	1,021	1,090	238	524
Persistent symptoms ^a (13)	Control (13)	1,358	576	181	641
Persistent symptoms ^a (13)	Resolved Lyme disease (15)	0	0	0	1 ^d
Non-PTLDS ^b (9)	Resolved Lyme disease (15)	0	0	0	1 ^e
PTLDS ^c (4)	Resolved Lyme disease (15)	1 ^f	0	0	3 ^g
PTLDS ^c (4)	Resolved Lyme disease + non-PTLDS ^b (24)	0	0	0	2 ^h
Disseminated EM (12)	Single EM (17)	0	0	0	0
Seronegative (8)	Seropositive (20)	1 ⁱ	0	0	4 ^j
Control (8)	Control (5)	NA ⁿ	NA	NA	0

^a All patients with persistent symptoms following treatment completion.^b Non-PTLDS (persistent symptoms with no functional decline).^c PTLDS (persistent symptoms with functional decline).^d GPR15.^e MIAT.^f GPR15.^g CCDC163P, GRP15, ZNF266.^h GPR15, ZNF266.ⁱ HLA-DQB1.^j HLA-DQA1, HLA-DQB1, HLA-DRB5, NSA2.^k Acute Lyme disease diagnosis, pretreatment.^l After 3-week antibiotic treatment.^m At 6 months post-treatment.ⁿ NA, not applicable.

by the fungal pathogen *Candida albicans*, or infected by two tick-borne bacterial agents, i.e., *Francisella tularensis* (tularemia) and *Anaplasma phagocytophilum* (anaplasmosis). Interestingly, stimulation of PBMCs by LPS (39%) resulted in a greater overlap of shared DEGs than *in vitro* infection with *C. albicans* (27%), *F. tularensis* (28%), or *A. phagocytophilum* (15%) (Fig. 4A). Next, we compared the acute Lyme disease transcriptome at V1 to transcriptomes corresponding to other acute infectious syndromes (Fig. 4A and B). Patients with acute influenza had 35% of their DEGs in common with Lyme disease patients, while they had only 28, 26, and 21% of their DEGs in common with patients with bacteremia due to *Staphylococcus aureus*, *Streptococcus pneumoniae*, and *Escherichia coli*, respectively.

To determine if these infectious diseases have pathways in common, we then compared the top canonical pathways significantly involved in acute Lyme disease at V1 ($P < 0.05$) with those corresponding to other clinical infectious diseases or *in vitro* cell culture models of infection. The greatest number of shared pathways was observed with *S. pneumoniae* bacteremia (82%) and *in vitro* *F. tularensis* (80%)- and *C. albicans* (78%)-infected cell cultures. Strikingly, downregulation of eIF2 signaling was restricted to Lyme disease and not found in other infectious diseases or *in vitro* cell culture infection models (Fig. 4B). In contrast, TREM1 and TLR pathways were activated in all five infections, whereas upregulation of IFN signaling pathways was predicted in only Lyme disease and influenza. B-cell development and downregulation of calcium-induced T-cell apoptosis, prominent in other viral and bacterial infections, were not found to be significant pathways in acute Lyme disease (Fig. 4B).

Pathway analysis of Lyme disease patients at 6 months post-treatment and comparison with chronic disease syndromes. We next compared the PBMC transcriptome from all treated Lyme

disease patients at V5 (both resolved illness and persistent symptoms) to publicly available transcriptome data sets from patients diagnosed with chronic illnesses that have some symptoms in common with PTLDS, including chronic fatigue syndrome (CFS), systemic lupus erythematosus (SLE), and rheumatoid arthritis (RA). The percentages of shared DEGs and pathways ranged from 9 to 18% and from 31 to 60%, respectively (Fig. 4C). Inhibition of eIF2 signaling was common to Lyme disease, SLE, and (to a lesser extent) RA (Fig. 4D; see Fig. S5 in the supplemental material). Glutathione-mediated detoxification and IL-6 signaling pathways were found to be specific to Lyme disease patients. No additional differences were revealed by subset analyses of Lyme disease patients manifesting persistent symptoms (non-PTLDS and PTLDS) and patients with these chronic illnesses.

Comparison of differential gene expression between Lyme disease patients with resolved versus persistent disease. No DEGs were significantly expressed at any single time point between patients with resolved Lyme disease ($n = 15$) and patients with persistent symptoms, including both non-PTLDS and PTLDS patients ($n = 13$) (Table 2). A single DEG (GPR15) was found at V1 when patients with resolved Lyme disease ($n = 15$) were compared with patients with PTLDS ($n = 4$). When all of the time points were combined, a total of four different DEGs overall were identified in Lyme disease patients with persistent symptoms (non-PTLDS and/or PTLDS) compared to those with resolved disease, i.e., MIAT, CCDC163P, ZNF266, and GPR15.

DISCUSSION

We provide the first transcriptome analysis of *B. burgdorferi* infection in Lyme disease patients, revealing a gene signature that persisted for at least 3 weeks following the acute phase of infection and had fewer than 44% DEGs in common with other infectious

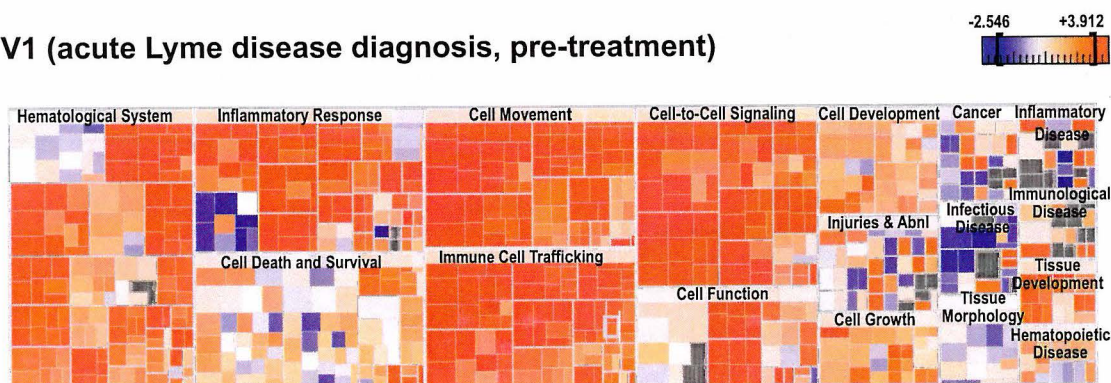
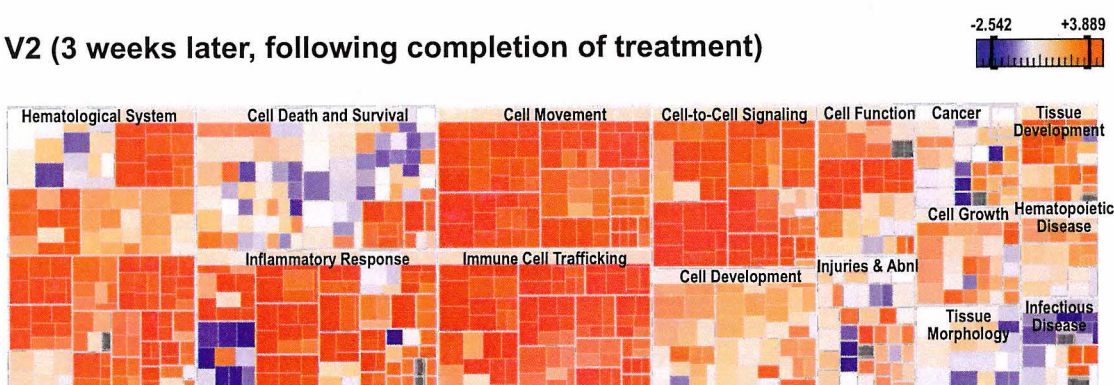
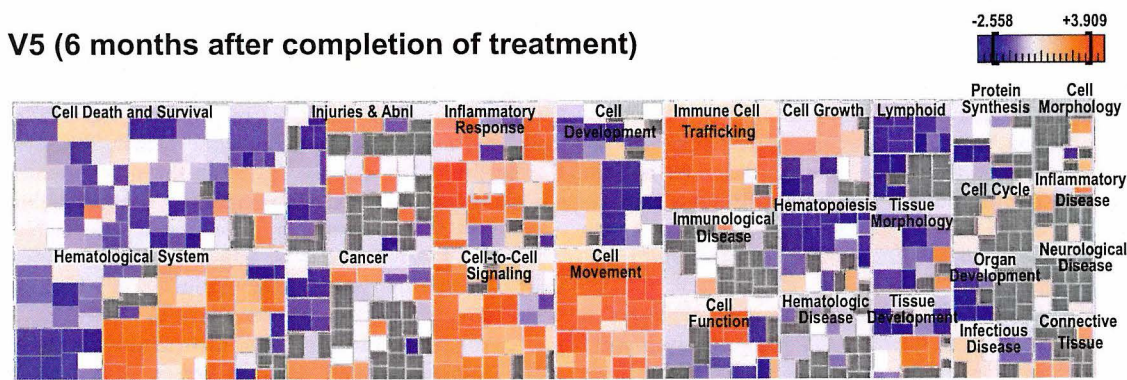
V1 (acute Lyme disease diagnosis, pre-treatment)**V2 (3 weeks later, following completion of treatment)****V5 (6 months after completion of treatment)**

FIG 3 Heat maps of disease and functional categories predicted to be involved in Lyme disease. Shown are heat maps at V1, V2, and V5. The color scheme shown is based on Z scores, with activation in orange, inhibition in blue, and undetermined directionality in gray.

or noninfectious syndromes. Notably, no differences in gene expression were observed between Lyme disease patients with resolved illness and those with persistent symptoms at 6 months, although larger cohort studies are needed to confirm this finding. The identification of a distinct and sustained transcriptome signature in early Lyme disease may facilitate the development and validation of human gene expression biomarker panels to improve diagnostic testing in the future, in parallel with other published studies investigating cytokine (12) or metabolic (13) bio-signatures.

To define the longitudinal transcriptome profile of patients with acute Lyme disease at 0 weeks, 3 weeks, and >6 months,

unbiased RNA-Seq analysis was employed with the goal of investigating the molecular basis of early and convalescent-phase Lyme disease. Potential advantages of RNA-Seq relative to microarrays include detection of low-abundance transcripts, a broader dynamic range in detecting fold changes in gene expression, unbiased detection of novel isoforms and transcripts, and elimination of hybridization-based limitations such as background noise, saturation, and probe redundancy (14). However, the utility of RNA-Seq data is dependent on a number of factors, including the number and quality of samples, sequencing depth, and designated thresholds for gene expression and differential analyses. In this study, we estimated the statistical power as 98% when analyzing

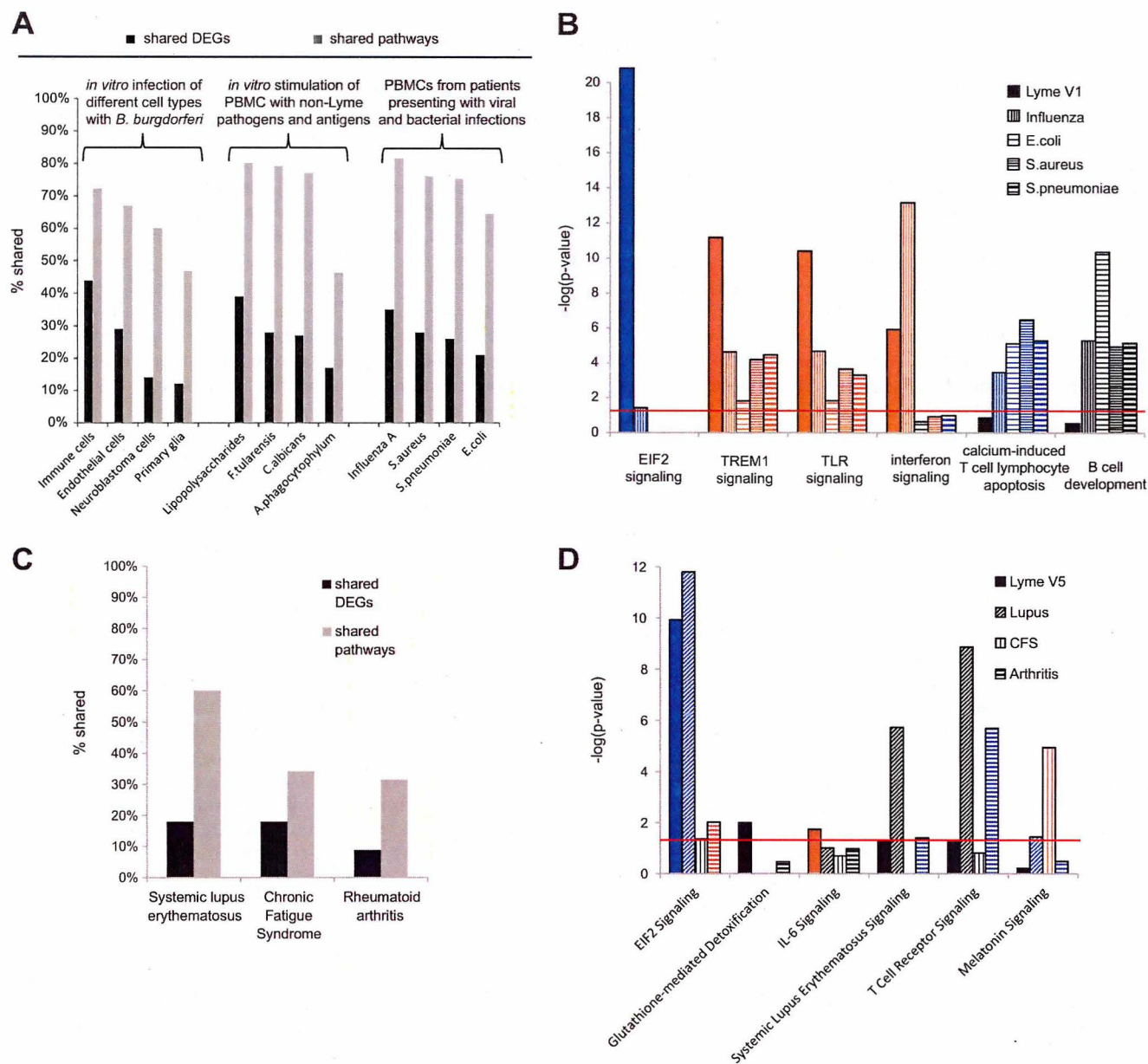


FIG 4 Comparison of the Lyme disease transcriptome with transcriptomes corresponding to other clinical illnesses or cell models. (A) Bar chart of the percentages of overlapping DEGs and canonical pathways between acute Lyme disease at V1 and other infectious diseases or *in vitro* cell culture models of infection. (B) Bar chart of the negative log of the *P* value of the enrichment score corresponding to selected pathways predicted to be involved in acute Lyme disease at V1 compared to other bacterial and viral infections. (C) Bar chart of the percentage of overlapping DEGs and canonical pathways between Lyme disease patients at V5 and three chronic and/or autoimmune syndromes (SLE, CFS, and RA). (D) Bar chart of the negative log of the *P* value of the enrichment score from selected pathways predicted to be involved in Lyme disease 6 months post-treatment (V5) compared to those from other chronic and/or autoimmune syndromes. The color scheme depicted is based on *Z* scores, with activation in orange, inhibition in blue, and undetermined directionality in gray. The red line represents a *P* value significance threshold of 0.05.

samples at all three time points combined, 78% for samples collected at a single time point, and only 62% when considering a stratification of the Lyme disease cohort according to PTLDS status, serology, or the presence of disseminated lesions (15).

The finding of a profound and sustained change in the transcriptome of acute Lyme disease patients refutes the idea that treatment and resolution of the infection result in a prompt return to a transcriptional baseline, as typically seen in the acute phase of

other infections (16). In addition, failure to return to a transcriptome baseline cannot be accounted for solely by patients with persistent symptoms, given that no DEGs were found comparing Lyme disease patients with resolved illness to those with persistent symptoms (Table 2). Persistent transcriptional changes may be characteristic of not only Lyme disease but also a number of other infections. For example, viral clearance in hepatitis C patients did not result in normalization of the baseline transcriptome (17). To

our knowledge, this is first time that sustained changes in the human host transcriptome have been reported for a bacterial infection after treatment with appropriate antibiotics. Persistence of such a signature for at least 3 weeks following infection suggests that a clinical diagnostic test for acute Lyme disease based on host gene expression is feasible. Such a test would also directly address the current diagnostic gap created by the “window period” between acute Lyme disease infection and the subsequent appearance of detectable antibody.

Infection by *B. burgdorferi* drives a complex immune response with robust inflammation and overt clinical signs and symptoms in early stages of the disease (7). The eIF2 signaling pathway, found to be downregulated here during all stages of Lyme disease, plays a central role in protein synthesis in response to cellular stress (18). Intracellular bacterial pathogens such as *Legionella pneumophila* encode effectors that actively disrupt and downregulate the eIF2 signaling pathway (19). However, *Borrelia* spirochetes are not known to enter cells during infection *in vivo*, nor are they thought to express such effectors (20). Consistent with a previous report (21), the eIF2 pathway in this study was also found to be downregulated in SLE as well as PTLDS patients. Inhibitors of the eIF2 pathway have been reported as potential therapeutic drugs for inflammatory bowel disease, and further studies are needed to assess whether eIF2 inhibitors may constitute potential targets for inflammatory sequelae of Lyme disease (22). Nevertheless, it remains to be determined whether inhibition of the eIF2 pathway in Lyme disease patients is caused directly by *Borrelia*-mediated immune dysregulation or is strictly a host response mechanism to limit tissue injury.

The prominent TREM1 signaling in acute Lyme disease observed here is consistent with previously published *in vitro* gene expression data of *B. burgdorferi* infection of human neural and primary monkey glial cell lines (23). TREM1 acts as an amplifier of the immune and inflammatory response *in vivo* (24), and modulation of TREM1 has been shown to impact a number of inflammatory conditions, including septic shock, and acute dengue virus infection (25, 26). Our data also showed upregulation of more TLRs (TLR2, TLR4, TLR7, and TLR8) in acute Lyme disease than previously described (27). This broad upregulation is likely to be indirect, reflecting a general increase in TLR regulatory activity rather than direct association of TLRs with *B. burgdorferi* proteins. In the present study, TNF- α was predicted to be a common upstream regulator of the eIF2, TREM1, and TLR signaling pathways. Notably, anti-TNF- α therapy has been proposed to reduce inflammation in the Jarisch-Herxheimer response to *Borrelia recurrentis* infection (28), and treatment was previously reported to be clinically efficacious in 4/4 patients with antibiotic-refractory Lyme arthritis (29).

Comparisons with 15 previously published transcriptome data sets found that the greatest overlap of DEGs (44%) was with the transcriptome of PBMCs stimulated with *B. burgdorferi* *in vitro*. Although this observation is to be expected, given the same infectious agent and cell type, the only partial overlap likely reflects differences between *in vivo* or *in vitro* *B. burgdorferi* infections and underscores the critical importance of analyzing “real-life” clinical samples from patients in studies of disease pathogenesis. Given the lymphocytic infiltrates characteristic of Lyme disease, in contrast with the suppurative lesions common to many bacterial infections (1), it is perhaps not surprising that the percentage of DEGs in Lyme disease patients also found in patients with acute

influenza was greater than that of DEGs also found in patients with other bacterial infections. Among bacterial infections, infection with *S. pneumoniae* had the highest number of top canonical pathways in common with acute Lyme disease, consistent with similarities in virulence factors shared by *S. pneumoniae* and *B. burgdorferi*, such as lipoproteins, that produce shared IgM-mediated immunological responses (30).

Importantly, Lyme disease patients did not show any changes in the calcium-dependent T-cell apoptosis pathway, in contrast to the marked downregulation observed in other bacterial and viral diseases (Fig. 4B). In addition, an absence of significant DEGs linked to B-cell development in Lyme disease relative to other infections was observed. These findings suggest that Lyme disease may be associated with a smaller proportion of B and T cells in peripheral blood than other diseases. Interestingly, suppression of long-lived humoral responses has been observed in a mouse model of *Borrelia* infection (31). The absence of DEGs corresponding to B-cell maturation may also potentially explain why prior infection with *B. burgdorferi* is associated with a serological response yet does not appear to confer immunity to reinfection. Certain alleles of HLA genes have been previously reported to be associated with serological responses to Lyme disease infection (32). Here we found that upregulation of certain HLA genes (*HLA-DQA1*, *HLA-DQB1*, *HLA-DRB5*) is associated with seronegativity in Lyme disease and may thus constitute potential diagnostic biomarkers for seronegative patients.

Following the acute phase of infection, recent treatment trials among patients with EM have estimated that approximately 10 to 20% of patients treated for Lyme disease experience lingering symptoms that may progress to PTLDS, although the incidence can be as high at 50% (4). The pathogenetic mechanisms of PTLDS remain unknown, but autoantigens and/or central nervous system sensitization have been postulated to play a role (10, 33–35). In our study, the relatively large proportion of post-treatment Lyme disease patients with persistent symptoms of fatigue, widespread musculoskeletal pain, and/or cognitive dysfunction (13 [46.4%] of 28) can be potentially accounted for by more stringent enrollment criteria at the time of presentation (requiring the presence of EM and concurrent influenza-like symptoms rather than EM alone). This may have resulted in the selection of patients with more severe disease and thus with an increased likelihood of persistent symptoms (36). Of note, according to the proposed formal case definition for PTLDS, which requires a functional decline in patients in addition to lingering symptoms, only 4 (14.3%) of our 28 patients met all of the criteria, within the range of the 10 to 20% frequency reported in the literature (4).

Notably, Lyme disease at 6 months post-treatment (V5) had 60 and 31% of their predicted pathways overall in common with SLE and RA, respectively. Circulating immune complexes have been identified as features common to all three conditions (37, 38). Symptoms of fatigue and cognitive impairment occur in a variety of chronic syndromes, including SLE, CFS, and PTLDS. Although some pathways were common to Lyme disease at V5 and CFS, melatonin signaling, prominent in CFS, was not predicted to be involved in Lyme disease (Fig. 4D). As melatonin is a hormone that regulates the circadian rhythms of the sleep-wake cycle and thus is strongly linked to fatigue, the absence of increased melatonin signaling suggests that the fatigue in Lyme disease patients with persistent symptoms is related to a different mechanism.

Overall, our results, showing only 18% of the DEGs and 34% of the pathways common to CFS and Lyme disease, are consistent with a proteomic study of cerebrospinal fluid that clearly discriminates between the two conditions (39).

Transcriptome analysis of Lyme disease patients with persistent symptoms (non-PTLDS and/or PTLDS) versus those with resolved illness revealed an absence of DEGs at each of the three time points, with the sole exception of a single gene (*GPR15*), which was upregulated at V1 in PTLDS patients relative to controls. Possible explanations for the overall lack of observed differences include (i) lack of statistical power from low sample numbers, (ii) sampling at designated time points instead of during periods of peak symptomatology, and (iii) that transcriptome profiling of PBMCs in blood is insufficient to discriminate between Lyme disease patients with persistent symptoms and those with resolved illness. Larger studies with increased sampling resolution are needed to establish whether there are indeed any detectable differences in gene expression between these two groups.

MATERIALS AND METHODS

Patient information. Patient enrollment, collection of clinical data and biological samples, and analysis of clinical samples by transcriptome profiling were done under protocols approved by the Institutional Review Boards of Johns Hopkins University and the University of California, San Francisco. Written informed consent was received from all participants prior to inclusion in this study.

All 29 participants with Lyme disease included in this study presented with a physician-documented EM rash of ≥ 5 cm and concurrent influenza-like symptoms that included at least one of the following: fever, chills, fatigue, headache, and/or new muscle or joint pains. At the time of enrollment, all of the participants with Lyme disease were treatment naive and subsequently underwent 3 weeks of doxycycline therapy between the first and second follow-up visits. All 29 subjects with Lyme disease were enrolled at the same geographic location (an outpatient clinic in Maryland) in a single season, from 1 May to 23 November 2009, with follow-up visits 3 weeks and 6 months after the first visit. Controls were matched by age and gender and enrolled from the same physician practice as case participants and across different seasons to account for seasonal variations in the transcriptome. Two-tier antibody testing for Lyme disease by whole-cell sonicate enzyme immunoassay, followed by IgM/IgG Western immunoblot assays, was performed for all patients and controls by a clinical reference laboratory (Quest Diagnostics). Seropositivity was assessed according to established CDC criteria (40) by the investigators (A.R. and J.A.) who enrolled and provided clinical care for the patients enrolled in this study. All control subjects were required to have a negative Lyme disease antibody test in order to be enrolled in this study. We screened both patients and controls prior to enrollment for a history of chronic fatigue, fibromyalgia, autoimmune, immunodeficiency, neurologic, psychiatric, and malignancy disorders, in which case they were excluded from the study. Prospective case patients and controls were also excluded if they had a prior documented history of Lyme disease and/or if they had previously received the Lyme disease vaccine.

Controls were enrolled primarily during the winter and spring seasons, while most Lyme disease patients were enrolled in summer during the peak season for tick bites, a difference that was statistically significant ($P < 0.03$) (Table 1). Nonetheless, the differences in seasonal sampling did not result in gene expression bias, as shown by the absence of seasonal clustering by PCA of the overall gene expression of the 13 controls (see Fig. S2b in the supplemental material). In addition, an intragroup comparison of eight controls sampled during the winter and five controls sampled during other seasons did not yield any significant DEGs (Table 2).

PBMCs from whole-blood samples at V1 (the acute phase of infection, prior to initiation of antibiotic treatment), V2 (3 weeks later, at the time of treatment completion), and V5 (6 months following treatment comple-

tion) were analyzed in this study (Fig. 1A). V2 and V5 were specifically chosen for analysis because fever and rash from acute Lyme disease typically resolve by completion of treatment (V2), while chronic persistent symptoms are clinically apparent after 6 months (V5).

The presence of persistent symptoms in Lyme disease patients at V5 was assessed by using a standardized case definition proposed by the Infectious Diseases Society of America (6, 11) that incorporates the presence of at least one of the following: new-onset fatigue, widespread musculoskeletal pain, or cognitive dysfunction. For a diagnosis of PTLDS, patients were also required to have a composite score of ≤ 45.00 on four subscales of the SF-36, a measurement of health-related quality of life (6) (Fig. 1B). The chi-square test was used to evaluate the statistical significance of differences between independent samples in one or more categorical variables, while Welch's *t* test was used for continuous variables.

Sample processing. PBMCs were isolated from fresh whole blood with Ficoll (Ficoll-Paque Plus; GE Healthcare), and total RNA was extracted from 10^7 PBMCs with TRIzol reagent (Life Technologies). mRNA was isolated with the Oligotex mRNA minikit (Qiagen). The ScriptSeq RNA-Seq library preparation kit (Epicentre) was used to generate RNA-Seq libraries according to the manufacturer's protocol. Libraries were sequenced as 100-bp paired-end runs on a HiSeq 2500 (Illumina). One hundred samples from the first cohort (29 patients at three time points and 13 control subjects, matched by age, sex, and geography) were mixed and blindly processed in three batches. Three samples, 01-36_V2, 01-42_V2, and 01-51_V1, were not included in the pooled analysis because of insufficient read counts and transcriptome coverage (see Fig. S1 in the supplemental material). No batch effect was observed by PCA of the global expression of all 25,278 genes (see Fig. S6 in the supplemental material).

Next-generation sequencing data analysis. Paired-end reads were mapped to the human genome (hg19), followed by annotation of exons and calculation of FPKM (fragments per kilobase of exon per million fragments mapped) values for all 25,278 expressed genes with version 2 of the TopHat-Cufflinks pipeline (41). Differential expression of genes was calculated by using the variance modeling at the observational level transformation (42), which applies precision weights to the matrix count, followed by linear modeling with the Limma package (43). Genes were considered to be differentially expressed when the change was greater than ± 1.5 -fold, the *P* value was < 0.05 , and the adjusted *P* value (or false-discovery rate, FDR) was $< 0.1\%$ (44). Pathway and network analyses of the transcriptome data were performed with Ingenuity Pathway Analysis (IPA) software (Qiagen) (45). The molecule activity predictor tool in the IPA software was used to predict the upstream and/or downstream activation or inhibition of a given pathway. The *P* value of the enrichment score was used to evaluate the significance of the overlap between observed and predicted gene sets, while the activation *Z* score was used to assess the match between observed and predicted patterns of upregulation and downregulation. The statistical significance of the difference in gene expression levels was determined with Welch's *t* test for independent samples by two-group comparisons. The statistical power for the transcriptome study was determined according to the algorithm developed by Hart et al. (15), with the use of a generalized linear model on normalized FPKM data instead of a negative binomial distribution on raw gene count data. The generalized linear model has been reported to be more reliable for differential analysis of data sets with small sample sizes (41, 43).

Comparison of RNA-Seq and microarray data. Microarray transcriptome data were downloaded from public servers (<http://www.ncbi.nlm.nih.gov/geo>) and include expression sets GSE12108, GSE2405, GSE42606, GSE8650, GSE6269, GSE6092, GSE14577, and GSE15573 (46–52). Raw data were extracted and preprocessed by using the Robust Multichip Average algorithm (53). Differential expression was calculated with the Limma package (43), which is applicable for analysis of both RNA-Seq and microarray data (42, 43). Genes were considered to be differentially expressed when the change was greater than ± 1.5 -fold, the *P* value was < 0.05 , and the FDR was $< 0.1\%$, in accordance with conven-

tional thresholds (44). Microarray data were not available for one study of *in vitro* *B. burgdorferi* infection (23), so tables of DEGs were used as provided instead, incorporating a change of greater than ± 1.5 -fold as a threshold cutoff for differential expression.

Data availability. All of the transcriptome data obtained in this study have been submitted to the Gene Expression Omnibus data repository under accession number GSE63085.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <http://mbio.asm.org/lookup/suppl/doi:10.1128/mBio.00100-16/-/DCSupplemental>.

Table S1, XLSX file, 0.01 MB.
Table S2, XLSX file, 0.04 MB.
Figure S1, AI file, 1.2 MB.
Figure S2, AI file, 0.1 MB.
Figure S3, AI file, 0.2 MB.
Figure S4, PDF file, 1.2 MB.
Figure S5, PDF file, 4.9 MB.
Figure S6, AI file, 2.3 MB.

ACKNOWLEDGMENTS

This work was supported in part by awards from the National Institutes of Health (R01-HL105704 and P30-AR05350), the Swartz Foundation, the Bay Area Lyme Foundation, the Stabler Foundation, the Global Lyme Alliance, the Lyme Disease Research Foundation, Inc., and Abbott Laboratories.

FUNDING INFORMATION

Swartz Foundation provided funding to Charles Y. Chiu. Bay Area Lyme Foundation provided funding to Jerome Bouquet and Charles Y. Chiu. Stabler Foundation provided funding to Mark J. Soloski. Lyme Research Alliance provided funding to Jerome Bouquet, John N. Aucott, and Charles Y. Chiu. Lyme Disease Research Foundation, Inc. provided funding to John N. Aucott. Office of Extramural Research, National Institutes of Health (OER) provided funding to Charles Y. Chiu under grant number HL105704. Office of Extramural Research, National Institutes of Health (OER) provided funding to Mark J. Soloski under grant number P30-AR05350. Abbott Laboratories provided funding to Charles Y. Chiu.

The funders of this publication had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

REFERENCES

1. Stanek G, Wormser GP, Gray J, Strle F. 2012. Lyme borreliosis. *Lancet* 379:461–473. [http://dx.doi.org/10.1016/S0140-6736\(11\)60103-7](http://dx.doi.org/10.1016/S0140-6736(11)60103-7).
2. Hinckley AF, Connally NP, Meek JI, Johnson BJ, Kemperman MM, Feldman KA, White JL, Mead PS. 2014. Lyme disease testing by large commercial laboratories in the United States. *Clin Infect Dis* 59:676–681. <http://dx.doi.org/10.1093/cid/ciu397>.
3. Centers for Disease C, Prevention. 2013. Three sudden cardiac deaths associated with Lyme carditis—United States, November 2012–July 2013. *MMWR Morb Mortal Wkly Rep* 62:993–996. <http://www.cdc.gov/mmwr/preview/mmwrhtml/mm6249a1.htm>.
4. Marques A. 2008. Chronic Lyme disease: a review. *Infect Dis Clin North Am* 22:341–360, vii–viii. <http://dx.doi.org/10.1016/j.idc.2007.12.011>.
5. Aucott J, Morrison C, Munoz B, Rowe PC, Schwarzwald A, West SK. 2009. Diagnostic challenges of early Lyme disease: lessons from a community case series. *BMC Infect Dis* 9:79. <http://dx.doi.org/10.1186/1471-2334-9-79>.
6. Aucott JN, Crowder LA, Kortte KB. 2013. Development of a foundation for a case definition of post-treatment Lyme disease syndrome. *Int J Infect Dis* 17:e443–e449. <http://dx.doi.org/10.1016/j.ijid.2013.01.008>.
7. Steere AC, Coburn J, Glickstein L. 2004. The emergence of Lyme disease. *J Clin Invest* 113:1093–1101. <http://dx.doi.org/10.1172/JCI21681>.
8. Sjöwall J, Fryland L, Nordberg M, Sjögren F, Garpmo U, Jansson C, Carlsson SA, Bergström S, Ernerudh J, Nyman D, Forsberg P, Ekerfelt C. 2011. Decreased Th1-type inflammatory cytokine expression in the skin is associated with persisting symptoms after treatment of erythema migrans. *PLoS One* 6:e18220. <http://dx.doi.org/10.1371/journal.pone.0018220>.
9. Strle K, Stupica D, Drouin EE, Steere AC, Strle F. 2014. Elevated levels of IL-23 in a subset of patients with post-Lyme disease symptoms following erythema migrans. *Clin Infect Dis* 58:372–380. <http://dx.doi.org/10.1093/cid/cit735>.
10. Moniuszko A, Czupryna P, Pancewicz S, Rutkowski K, Zajkowska O, Swierzbinska R, Grygorczuk S, Kondrusik M, Owłasiuk P, Zajkowska J. 2014. Evaluation of CXCL8, CXCL10, CXCL11, CXCL12 and CXCL13 in serum and cerebrospinal fluid of patients with neuroborreliosis. *Immunol Lett* 157:45–50. <http://dx.doi.org/10.1016/j.imlet.2013.11.002>.
11. Wormser GP, Dattwyler RJ, Shapiro ED, Halperin JJ, Steere AC, Klemperner MS, Krause PJ, Bakken JS, Strle F, Stanek G, Bockenstedt L, Fish D, Dumler JS, Nadelman RB. 2006. The clinical assessment, treatment, and prevention of Lyme disease, human granulocytic anaplasmosis, and babesiosis: clinical practice guidelines by the Infectious Diseases Society of America. *Clin Infect Dis* 43:1089–1134. <http://dx.doi.org/10.1086/508667>.
12. Soloski MJ, Crowder LA, Lahey LJ, Wagner CA, Robinson WH, Aucott JN. 2014. Serum inflammatory mediators as markers of human Lyme disease activity. *PLoS One* 9:e93243. <http://dx.doi.org/10.1371/journal.pone.0093243>.
13. Molins CR, Ashton LV, Wormser GP, Hess AM, Delorey MJ, Mahapatra S, Schrieffer ME, Belisle JT. 2015. Development of a metabolic biosignature for detection of early Lyme disease. *Clin Infect Dis* 60:1767–1775. <http://dx.doi.org/10.1093/cid/civ185>.
14. Guo Y, Sheng Q, Li J, Ye F, Samuels DC, Shyr Y. 2013. Large scale comparison of gene expression levels by microarrays and RNAseq using TCGA data. *PLoS One* 8:e71462. <http://dx.doi.org/10.1371/journal.pone.0071462>.
15. Hart SN, Therneau TM, Zhang Y, Poland GA, Kocher JP. 2013. Calculating sample size estimates for RNA sequencing data. *J Comput Biol* 20:970–978. <http://dx.doi.org/10.1089/cmb.2012.0283>.
16. Zhai Y, Franco LM, Atmar RL, Quarles JM, Arden N, Bucsas KL, Wells JM, Niño D, Wang X, Zapata GE, Shaw CA, Belmont JW, Couch RB. 2015. Host transcriptional response to influenza and other acute respiratory viral infections—a prospective cohort study. *PLoS Pathog* 11:e1004869. <http://dx.doi.org/10.1371/journal.ppat.1004869>.
17. Hou J, van Oord G, Groothuisink ZM, Claassen MA, Kreeft K, Zaaraoui-Boutahar F, van IJcken WF, Osterhaus AD, Janssen HL, Andeweg AC, de Knecht RJ, Boonstra A. 2014. Gene expression profiling to predict and assess the consequences of therapy-induced virus eradication in chronic hepatitis C virus infection. *J Virol* 88:12254–12264. <http://dx.doi.org/10.1128/JVI.00775-14>.
18. Park SH, Moon Y. 2013. Integrated stress response-altered pro-inflammatory signals in mucosal immune-related cells. *Immunopharmacol Immunotoxicol* 35:205–214. <http://dx.doi.org/10.3109/08923973.2012.742535>.
19. Fontana MF, Banga S, Barry KC, Shen X, Tan Y, Luo ZQ, Vance RE. 2011. Secreted bacterial effectors that inhibit host protein synthesis are critical for induction of the innate immune response to virulent *Legionella pneumophila*. *PLoS Pathog* 7:e1001289. <http://dx.doi.org/10.1371/journal.ppat.1001289>.
20. Embers ME, Ramamoorthy R, Philipp MT. 2004. Survival strategies of *Borrelia burgdorferi*, the etiologic agent of Lyme disease. *Microbes Infect* 6:312–318. <http://dx.doi.org/10.1016/j.micinf.2003.11.014>.
21. Grolleau A, Kaplan MJ, Hanash SM, Beretta L, Richardson B. 2000. Impaired translational response and increased protein kinase PKR expression in T cells from lupus patients. *J Clin Invest* 106:1561–1568. <http://dx.doi.org/10.1172/JCI9352>.
22. Okazaki T, Nishio A, Takeo M, Sakaguchi Y, Fukui T, Uchida K, Okazaki K. 2014. Inhibition of the dephosphorylation of eukaryotic initiation factor 2 α ameliorates murine experimental colitis. *Digestion* 90:167–178. <http://dx.doi.org/10.1159/000366414>.
23. Myers TA, Kaushal D, Philipp MT. 2009. Microglia are mediators of *Borrelia burgdorferi*-induced apoptosis in SH-SY5Y neuronal cells. *PLoS Pathog* 5:e1000659. <http://dx.doi.org/10.1371/journal.ppat.1000659>.
24. Sharif O, Knapp S. 2008. From expression to signaling: roles of TREM-1 and TREM-2 in innate immunity and bacterial infection. *Immunobiology* 213:701–713. <http://dx.doi.org/10.1016/j.imbio.2008.07.008>.
25. Gibot S, Massin F, Marcou M, Taylor V, Stidwill R, Wilson P, Singer M, Bellingan G. 2007. TREM-1 promotes survival during septic shock in

- mice. *Eur J Immunol* 37:456–466. <http://dx.doi.org/10.1002/eji.200636387>.
26. Ruiz-Pacheco JA, Vivanco-Cid H, Izaguirre-Hernández IY, Estrada-García I, Arriaga-Pizano L, Chacón-Salinas R, Fonseca-Coronado S, Vaughan G, Tovar KR, Rivera-Osorio MP, Escobar-Gutiérrez A. 2014. TREM-1 modulation during early stages of dengue virus infection. *Immunol Lett* 158:183–188. <http://dx.doi.org/10.1016/j.imlet.2014.01.003>.
 27. Oosting M, Buffen K, van der Meer JWM, Netea MG, Joosten LAB. 25 June 2014. Innate immunity networks during infection with *Borrelia burgdorferi*. *Crit Rev Microbiol* <http://dx.doi.org/10.3109/1040841X.2014.929563>.
 28. McSweeney E. 2007. Anti-tumor necrosis factor- α treatment for Lyme borreliosis. *J Infect Dis* 196:1866–1867. <http://dx.doi.org/10.1086/523828>.
 29. Steere AC, Angelis SM. 2006. Therapy for Lyme arthritis: strategies for the treatment of antibiotic-refractory arthritis. *Arthritis Rheum* 54:3079–3086. <http://dx.doi.org/10.1002/art.22131>.
 30. Racine R, Winslow GM. 2009. IgM in microbial infections: taken for granted? *Immunol Lett* 125:79–85. <http://dx.doi.org/10.1016/j.imlet.2009.06.003>.
 31. Elsner RA, Hastey CJ, Olsen KJ, Baumgarth N. 2015. Suppression of long-lived humoral immunity following *Borrelia burgdorferi* infection. *PLoS Pathog* 11:e1004976. <http://dx.doi.org/10.1371/journal.ppat.1004976>.
 32. Wang P, Hilton E. 2001. Contribution of HLA alleles in the regulation of antibody production in Lyme disease. *Front Biosci* 6:B10–B16. <http://dx.doi.org/10.2741/Wang>.
 33. Crowley JT, Drouin EE, Pianta A, Strle K, Wang Q, Costello CE, Steere AC. 2015. A highly expressed human protein, apolipoprotein B-100, serves as an autoantigen in a subgroup of patients with Lyme disease. *J Infect Dis* 212:1841–1850. <http://dx.doi.org/10.1093/infdis/jiv310>.
 34. Drouin EE, Seward RJ, Strle K, McHugh G, Katchar K, Londoño D, Yao C, Costello CE, Steere AC. 2013. A novel human autoantigen, endothelial cell growth factor, is a target of T and B cell responses in patients with Lyme disease. *Arthritis Rheum* 65:186–196. <http://dx.doi.org/10.1002/art.37732>.
 35. Pianta A, Drouin EE, Crowley JT, Arvikar S, Strle K, Costello CE, Steere AC. 2015. Annexin A2 is a target of autoimmune T and B cell responses associated with synovial fibroblast proliferation in patients with antibiotic-refractory Lyme arthritis. *Clin Immunol* 160:336–341. <http://dx.doi.org/10.1016/j.clim.2015.07.005>.
 36. Nowakowski J, Nadelman RB, Sell R, McKenna D, Cavaliere LF, Holmgren D, Gaidici A, Wormser GP. 2003. Long-term follow-up of patients with culture-confirmed Lyme disease. *Am J Med* 115:91–96. [http://dx.doi.org/10.1016/S0002-9343\(03\)00308-5](http://dx.doi.org/10.1016/S0002-9343(03)00308-5).
 37. Marques AR, Hornung RL, Dally L, Philipp MT. 2005. Detection of immune complexes is not independent of detection of antibodies in Lyme disease patients and does not confirm active infection with *Borrelia burgdorferi*. *Clin Diagn Lab Immunol* 12:1036–1040. <http://dx.doi.org/10.1128/CDLI.12.9.1036-1040.2005>.
 38. Weissmann G. 2009. Rheumatoid arthritis and systemic lupus erythematosus as immune complex diseases. *Bull NYU Hosp Jt Dis* 67:251–253.
 39. Schutzer SE, Angel TE, Liu T, Schepmoes AA, Clauss TR, Adkins JN, Camp DG, Holland BK, Bergquist J, Coyle PK, Smith RD, Fallon BA, Natelson BH. 2011. Distinct cerebrospinal fluid proteomes differentiate post-treatment Lyme disease from chronic fatigue syndrome. *PLoS One* 6:e17287. <http://dx.doi.org/10.1371/journal.pone.0017287>.
 40. Centers for Disease Control and Prevention. 1995. Recommendations for test performance and interpretation from the Second National Conference on Serologic Diagnosis of Lyme Disease. *MMWR Morb Mortal Wkly Rep* 44:590–591. <http://www.cdc.gov/mmwr/preview/mmwrhtml/00038469.htm>.
 41. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. 2012. Differential gene and transcript expression analysis of RNA-Seq experiments with TopHat and cufflinks. *Nat Protoc* 7:562–578. <http://dx.doi.org/10.1038/nprot.2012.016>.
 42. Law CW, Chen Y, Shi W, Smyth GK. 2014. voom: precision weights unlock linear model analysis tools for RNA-Seq read counts. *Genome Biol* 15:R29. <http://dx.doi.org/10.1186/gb-2014-15-2-r29>.
 43. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. 2015. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 43:e47. <http://dx.doi.org/10.1093/nar/gkv007>.
 44. Dalman MR, Deeter A, Nimishakavi G, Duan ZH. 2012. Fold change and *P* value cutoffs significantly alter microarray interpretations. *BMC Bioinformatics* 13(Suppl 2):S11. <http://dx.doi.org/10.1186/1471-2105-13-S2-S11>.
 45. Krämer A, Green J, Pollard J, Jr., Tugendreich S. 2014. Causal analysis approaches in ingenuity pathway analysis. *Bioinformatics* 30:523–530. <http://dx.doi.org/10.1093/bioinformatics/btt703>.
 46. Borjesson DL, Kobayashi SD, Whitney AR, Voyich JM, Argue CM, Deleo FR. 2005. Insights into pathogen immune evasion mechanisms: *Anaplasma phagocytophilum* fails to induce an apoptosis differentiation program in human neutrophils. *J Immunol* 174:6364–6372. <http://dx.doi.org/10.4049/jimmunol.174.10.6364>.
 47. Butchar JP, Cremer TJ, Clay CD, Gavrilin MA, Wewers MD, Marsh CB, Schlesinger LS, Tridandapani S. 2008. Microarray analysis of human monocytes infected with *Francisella tularensis* identifies new targets of host response subversion. *PLoS One* 3:e2924. <http://dx.doi.org/10.1371/journal.pone.0002924>.
 48. Dame TM, Orenzoff BL, Palmer LE, Furie MB. 2007. IFN- γ alters the response of *Borrelia burgdorferi*-activated endothelium to favor chronic inflammation. *J Immunol* 178:1172–1179. <http://dx.doi.org/10.4049/jimmunol.178.2.1172>.
 49. Gow JW, Hagan S, Herzyk P, Cannon C, Behan PO, Chaudhuri A. 2009. A gene signature for post-infectious chronic fatigue syndrome. *BMC Med Genomics* 2:38. <http://dx.doi.org/10.1186/1755-8794-2-38>.
 50. Ramilo O, Allman W, Chung W, Mejias A, Ardura M, Glaser C, Wittkowski KM, Piqueras B, Banchereau J, Palucka AK, Chaussabel D. 2007. Gene expression patterns in blood leukocytes discriminate patients with acute infections. *Blood* 109:2066–2077. <http://dx.doi.org/10.1182/blood-2006-02-002477>.
 51. Smeekens SP, Ng A, Kumar V, Johnson MD, Plantinga TS, van Diemen C, Arts P, Verwiel ET, Gresnigt MS, Fransen K, van Sommeren S, Oosting M, Cheng SC, Joosten LA, Hoischen A, Kullberg BJ, Scott WK, Perfect JR, van der Meer JW, Wijmenga C, Netea MG, Xavier RJ. 2013. Functional genomics identifies type I interferon pathway as central for host defense against *Candida albicans*. *Nat Commun* 4:1342. <http://dx.doi.org/10.1038/ncomms2343>.
 52. Teixeira VH, Olaso R, Martin-Magniette ML, Lasbleiz S, Jacq L, Oliveira CR, Hilliquin P, Gut I, Cornelis F, Petit-Teixeira E. 2009. Transcriptome analysis describing new immunity and defense genes in peripheral blood mononuclear cells of rheumatoid arthritis patients. *PLoS One* 4:e6803. <http://dx.doi.org/10.1371/journal.pone.0006803>.
 53. Bolstad BM, Irizarry RA, Astrand M, Speed TP. 2003. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* 19:185–193. <http://dx.doi.org/10.1093/bioinformatics/19.2.185>.

Quantification of *Borrelia burgdorferi* Membrane Proteins in Human Serum: A New Concept for Detection of Bacterial Infection

Crystal S. F. Cheung,^{†,‡} Kyle W. Anderson,^{†,‡} Kenia Y. Villatoro Benitez,[‡] Mark J. Soloski,[§] John N. Aucott,[§] Karen W. Phinney,[†] and Illarion V. Turko^{*,†,‡}

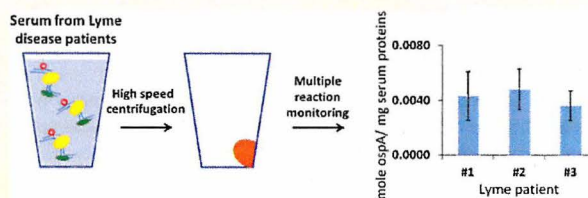
[†]Biomolecular Measurement Division, National Institute of Standards and Technology, Gaithersburg, Maryland 20899, United States

[‡]Institute for Bioscience and Biotechnology Research, Rockville, Maryland 20850, United States

[§]Department of Medicine, Johns Hopkins School of Medicine, Baltimore, Maryland 21224, United States

Supporting Information

ABSTRACT: The *Borrelia burgdorferi* spirochete is the causative agent of Lyme disease, the most common tick-borne disease in the United States. The low abundance of bacterial proteins in human serum during infection imposes a challenge for early proteomic detection of Lyme disease. To address this challenge, we propose to detect membrane proteins released from bacteria due to disruption of their plasma membrane triggered by the innate immune system. These membrane proteins can be separated from the bulk of serum proteins by high-speed centrifugation causing substantial sample enrichment prior to targeted protein quantification using multiple reaction monitoring mass spectrometry. This new approach was first applied to detection of *B. burgdorferi* membrane proteins supplemented in human serum. Our results indicated that detection of *B. burgdorferi* membrane proteins, which are $\approx 10^7$ lower in abundance than major serum proteins, is feasible. Therefore, quantitative analysis was also carried out for serum samples from three patients with acute Lyme disease. We were able to demonstrate the detection of ospA, the major *B. burgdorferi* lipoprotein at the level of 4.0 fmol of ospA/mg of serum protein. The results confirm the concept and suggest that the proposed approach can be expanded to detect other bacterial infections in humans, particularly where existing diagnostics are unreliable.



Lyme disease is a multiorgan tick-borne disease caused by spirochetes of the genus *Borrelia*. *Borrelia burgdorferi* is the prototypical Lyme disease spirochete in North America.^{1,2} If left untreated, Lyme disease may lead to neurological and rheumatic manifestations that may last for years and adversely affect health-related functioning. Various post-treatment Lyme disease symptoms that may be severe and chronic have been described as well.³ Overall, Lyme disease and post-treatment Lyme disease symptoms are associated with significant health care costs.⁴

Unfortunately, the diagnosis of Lyme disease at its earliest stage is most often based upon clinical manifestations only including presence of the primary skin lesion, called erythema migrans. Confirmatory laboratory testing is limited to serological tests for the presence of antibodies that react to *B. burgdorferi* antigens. However, serology is hampered by the long time of analysis. It generally takes 3–6 weeks before *Borrelia*-specific antibodies can be detected. In addition, various species-specific factors likely lower the sensitivity and specificity of serological tests, which may be misinterpreted and have false negatives or positives.³ Molecular assays to detect *B. burgdorferi* based on polymerase chain reaction (PCR) on DNA extracted from tissue or fluid specimens have also been described^{3,5–7} but are currently performed only for confirmation and research purposes. A drawback of using PCR is that *B. burgdorferi* DNA can be detected in samples long after spirochetes are no longer

viable.⁸ Thus, a positive PCR result can be valuable for early detection but needs to be interpreted with caution when efficiency of treatment and post-treatment symptoms are evaluated.

In addition to DNA, unique *B. burgdorferi* proteins can be targets of molecular assays. We have previously used a well-established mass spectrometry-based assay, multiple reaction monitoring (MRM), to detect and quantify target proteins.^{9–12} MRM assay relies on stable isotope-labeled internal standards added to the biological sample and is typically performed on a triple quadrupole mass spectrometer. Current instrumentation allows for the measurement of many proteins in a single sample, making MRM an ideal assay to perform high-throughput measurements on a panel of target proteins.^{13–15} Successful application of the MRM assay for detection of *Borrelia* proteins in human skin biopsies has been recently reported;¹⁶ however, direct MRM assay in human blood or serum for early detection of Lyme disease poses additional challenges due to the extremely low abundance of total circulating *Borrelia* proteins.

In the present study, we capitalize on the fact that human serum is a cell-free substance which does not have membrane vesicles under normal conditions. Bacterial infection triggers

Received: July 24, 2015

Accepted: October 22, 2015

Published: October 22, 2015



various defense mechanisms including those of the innate immune system, which can disrupt the integrity of the bacterial wall.¹⁷ Broken *B. burgdorferi* cells^{18,19} can release membrane vesicles populated with membrane proteins into the serum. These membrane vesicles can be separated from the bulk of soluble serum proteins by high-speed centrifugation and can be a ready source of unique membrane proteins for MRM detection of *B. burgdorferi* infection. As an initial proof of concept, we have been able to quantify the *B. burgdorferi* membrane protein ospA in the serum of Lyme disease patients. We believe this approach can be expanded to detection of many bacterial infections in humans, particularly where existing diagnostics are unreliable.

■ EXPERIMENTAL SECTION

Borrelia Extracts. Low-passage-number *B. burgdorferi* strain B31 was grown in modified Barbour–Stoenner–Kelly II medium at 34 °C. Bacteria are harvested from log phase cultures, washed in PBS, sonicated, and filtered as described.²⁰ Protein concentration was adjusted to 1 mg/mL, and the protein was frozen at –80 °C.

Lyme Disease Patient Serum. Serum samples were obtained from three patients with early, untreated Lyme disease at the time of diagnosis (first visit) and 2 months later on the third visit. All patients had a physician-observed erythema migrans rash at time of the first visit, as well as a recent history of flulike symptoms, including fever, fatigue, and/or musculoskeletal pain. Two of the three patients were positive on standard, commercial two-tier serologic testing at the time of enrollment, and the third seroconverted during treatment and was positive at repeat testing 3 weeks later. Control samples were obtained from nonhospitalized individuals with no prior history of Lyme disease. Handling of serum for sample processing conformed to University of Maryland regulations. All data were analyzed anonymously. Demographic information on the deidentified donors is summarized in the Supporting Information Table S1.

Two-Dimensional Polyacrylamide Gel Electrophoresis (PAGE). The first dimension of separation was performed on 7 cm strips with 3–10 immobilized pH gradient using a PROTEAN IEF cell (Bio-Rad Laboratories, Hercules, CA). The strips were rehydrated with 125 μ L of a protein solution in 2 mol/L thiourea, 7 mol/L urea, 4% CHAPS, 0.5% ASB-14, 0.2% of (3–10) BioLytes, and bromophenol blue. Isoelectric focusing was conducted at 50 V for 12 h, linearly increased over 2 h to a maximum of 4000 V, and then run to accumulate a total of 20 000 V/h. For the second dimension, the immobilized pH gradient strips were equilibrated for 15 min in 50 mmol/L Tris–HCl (pH 8.8), 6 mol/L urea, 30% glycerol, 2% sodium dodecyl sulfate (SDS), and bromophenol blue. The strips were then embedded in 0.7% (w/v) agarose on the top of 9% home-made polyacrylamide gels and proteins were separated by SDS–PAGE. In some experiments, 8–16% mini-PROTEAN TGX gels from Bio-Rad were used as well. All gels were stained with a Pierce silver stain kit for mass spectrometry.

In-Gel Protein Identification. Silver-stained gel pieces were excised and destained in accordance with the manufacturer's protocol. In-gel digestion was then carried out with a sequencing grade modified trypsin (Promega, Madison, WI) in 25 mmol/L NH_4HCO_3 (pH 7.9) for 15 h at 37 °C. After digestion, samples were dried, and dissolved in 5 mg/mL α -cyano-4-hydroxycinnamic acid in 50% acetonitrile containing 0.1% trifluoroacetic acid (TFA). After spotting onto an ABI

01-192-6-AB target plate, the MS analysis was performed using an AB4700 proteomics analyzer (Applied Biosystems, Framingham, MA). All matrix-assisted laser desorption/ionization (MALDI) mass spectra were internally calibrated using the monoisotopic masses of the autolysis peptides of trypsin at 842.51 and 2211.10. Automated combined acquisition of MS and MS/MS data was controlled with 4000 Series Explorer software 3.0. Data analysis was performed with GPS Explorer software 3.5 utilizing Mascot 2.0 (MatrixScience, London, U.K.) as the search engine. During searching, the mass tolerance was 0.08 Da for the precursor ions and 0.2 Da for the fragment ions. A protein was listed as identified protein when the MOWSE score was higher than a MOWSE score at which statistical significance ($p < 0.05$) occurred for that particular search.

¹⁵N-Labeled QconCAT Expression, Purification, and Characterization. The quantification concatamer (QconCAT) is an artificial protein composed of concatenated tryptic peptides from targeted proteins and used as an internal standard for quantification of these targeted proteins by MRM. The amino acid sequence of a QconCAT designed for quantification of *B. burgdorferi* proteins was coded into the corresponding DNA sequence and incorporated into the pET21a expression vector, with codon optimization for *Escherichia coli* (Biomatik, Cambridge, Ontario). The plasmid was transformed into One Shot BL21 (DE3) competent *E. coli* cells (Invitrogen, Grand Island, NY) and grown in M9 minimal media with 1 g/L $^{15}\text{NH}_4\text{Cl}$ (Cambridge Isotope Laboratories, Andover, MA) as the sole nitrogen source at 37 °C until the optical density reached 0.6–0.8 at 600 nm. Protein expression was induced by 0.5 mmol/L isopropyl β -D-1-thiogalactopyranoside. After 3 h of growth, the cells were harvested by centrifugation at 5000g for 10 min and resuspended in 0.1 mmol/L dithiothreitol (DTT) and sonicated. Following centrifugation at 35 000g for 30 min, the supernatant was discarded and the pellet which contained the QconCAT was resuspended in 100 mmol/L Na_2HPO_4 /10 mmol/L Tris–HCl (pH 8.0) containing 8 mol/L urea and 10 mmol/L imidazole. The suspension was centrifuged at 20 000g for 10 min, and the supernatant was used for QconCAT purification.

Purification of 6xHis-tagged QconCAT was performed using Ni–NTA agarose resin by the gravity-flow method (Qiagen, Valencia, CA). The binding, washing, and eluting buffers were 100 mmol/L Na_2HPO_4 /10 mmol/L Tris–HCl (pH 8.0) containing 8 mol/L urea and 10, 30, and 100 mmol/L imidazole, respectively. The eluted fraction was concentrated and buffer-exchanged to 100 mmol/L Na_2HPO_4 /10 mmol/L Tris–HCl (pH 8.0) containing 6 mol/L urea using an Amicon filter (30 kDa MWCO, Millipore, Billerica, MA). The QconCAT concentration was subsequently determined by a BCA protein assay with bovine serum albumin as a standard (Thermo Scientific, Waltham, MA). QconCAT expression and purification were evaluated with SDS–PAGE and mass spectrometry analysis on a 4700 proteomics analyzer. The isotope incorporation was determined at the peptide level after digestion of the purified QconCAT with trypsin. MALDI spectra of three representative peptides were imported to Isotopic Enrichment Calculator (<http://www.nist.gov/mml/bmd/bioanalytical/isoenrichcalc.cfm>),²¹ and the mean value was higher than 99% of ^{15}N incorporation. This was accepted as a complete labeling and no correction was applied to the data.

Processing of Human Serum Samples. To prove the concept, 1 mL samples of normal human serum (cat. no. S-7023, Sigma-Aldrich, St. Louis, MO) were supplemented with

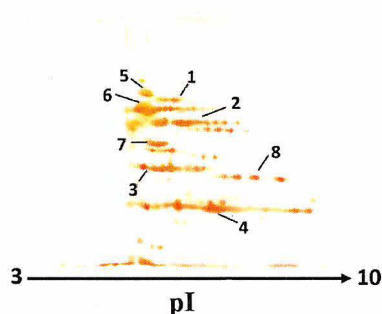


Figure 1. Two-dimensional PAGE pattern of *B. burgdorferi* proteins. The first and second dimensions were performed on 7 cm pH 3–10 immobilized pH gradient strips and 9% polyacrylamide slab gel, respectively. After separation, proteins were detected by silver staining. The proteins identified in the numbered spots are integral outer membrane protein P66 (no. 1), aminopeptidase 1 (no. 2), basic membrane protein A (no. 3), outer surface protein A (no. 4), chaperone protein Dnak (no. 5), 60 kDa chaperonin (no. 6), enolase (no. 7), and glyceraldehyde 3-phosphate dehydrogenase (no. 8).

3.0, 1.0, 0.3, 0.1, 0.03, or 0.01 μg of total *B. burgdorferi* protein. The membrane proteins were pulled down by high-speed centrifugation at 106 000g for 60 min at 4 °C. The membrane pellet was then resuspended in 1 mL of 0.1 mol/L Na_2CO_3 and centrifuged at 179 000g for 60 min at 4 °C. The Na_2CO_3 washed membrane pellet was resuspended in 150 μL of 25 mmol/L NH_4HCO_3 /1% SDS/10 mmol/L DTT and supplemented with 15 pmol of QconCAT. The mixture was incubated at room temperature for 60 min to allow reduction of cysteines and then treated with 55 mmol/L iodoacetamide for another 60 min to alkylate the reduced cysteines. Alkylated samples were precipitated with chloroform/methanol²² to deplete salts, urea, and SDS from the samples. Protein pellets were then sonicated in 100 μL of 25 mmol/L NH_4HCO_3 /0.1% RapiGest SF surfactant (Waters, Milford, MA) and treated with trypsin for 15 h at 37 °C. The substrate/trypsin ratio was 10:1 (w/w). After trypsin digestion, the samples were treated with 0.5% TFA for 30 min at 37 °C to break down acid-cleavable RapiGest. The insoluble byproduct of RapiGest was then

removed by centrifugation at 106 000g for 30 min. After centrifugation, the supernatants were dried using a vacuum centrifuge (Vacufuge, Eppendorf AG, Hamburg, Germany).

Human serum samples from control and Lyme disease patients (1 mL each) were processed as described above without supplementation with spirochete proteins, but supplemented with 3 pmol of QconCAT.

Liquid Chromatography–Tandem Mass Spectrometry (LC–MS/MS) Analysis. Dried peptides were reconstituted in 30 μL of 3% acetonitrile/97% water containing 0.1% formic acid and 5 μL were used for each LC–MS/MS run. Instrumental analyses were performed on an Agilent Zorbax Eclipse Plus C18 RRHD column (2.1 mm $\times$ 50 mm, 1.8 μm particle) coupled to an Agilent 6490 triple quadrupole LC/MS system (Santa Clara, CA). Peptides were eluted over a 35 min gradient from 5% to 80% acetonitrile containing 0.1% formic acid at a flow rate of 200 $\mu\text{L}/\text{min}$. The gradient settings were 5–10% solvent B in 5 min, 10–30% solvent B in 25 min, 30–80% solvent B in 5 min, then returned to 5% solvent B in 5 min. Solvent A was water containing 0.1% formic acid, and solvent B was 100% acetonitrile with 0.1% formic acid. The acquisition method used the following parameters in positive mode: fragmentor 380 V, electron multiplier 500 V, and capillary voltage 3500 V. Collision energy was optimized for each peptide using the default equation from Agilent, $\text{CE} = 0.036 m/z - 4.8$. Dwell time for all transitions was set at 120 ms.

Data Analysis. MRM peak area integration was performed using Agilent MassHunter Qualitative Analysis B.06. Excel was used to calculate peak area ratios. Peak integration was manually inspected and adjusted if necessary. The peak ratios from transitions were averaged to yield the peptide ratios. All experiments were performed in duplicate with three replicate injections to assess error and reproducibility. Data is represented as the mean $\pm$ SD.

RESULTS AND DISCUSSION

Selection of Target *B. burgdorferi* Proteins and Design of a QconCAT. Our primary focus was on those *B. burgdorferi* proteins that demonstrate the highest abundance and possess unique sequence. We have used a cultured *B. burgdorferi* B31 isolate A3. Figure 1 shows a two-dimensional (2D) gel pattern

Table 1. Quantification of *B. burgdorferi* Proteins Supplemented in the 1 mL of Human Serum

protein/peptide	supplementation, μg of <i>B. burgdorferi</i> protein/mL of human serum						av concn
	3.0	1.0	0.3	0.1	0.03	0.01	
	concn, pmol of a protein/mg of <i>B. burgdorferi</i> protein ^a						
ospA							4220 $\pm$ 349
YDLLATVDK	4073 $\pm$ 437	4974 $\pm$ 1040	3874 $\pm$ 559	3862 $\pm$ 1548	nd ^b	nd	
EGTVTLTK	4093 $\pm$ 258	4749 $\pm$ 82	3994 $\pm$ 278	nd	nd	nd	
GYVLEGLTAEK	4345 $\pm$ 187	4419 $\pm$ 514	3983 $\pm$ 492	4014 $\pm$ 483	4254 $\pm$ 1166	nd	
fla							987 $\pm$ 125
NNGINAANLSK	843 $\pm$ 128	1056 $\pm$ 49	1062 $\pm$ 136	nd	nd	nd	
p66							302 $\pm$ 60
STYYGFPSNDR	293 $\pm$ 53	311 $\pm$ 79	237 $\pm$ 58	nd	nd	nd	
LDLTFAGGTGTGGR	270 $\pm$ 17	296 $\pm$ 13	354 $\pm$ 22	nd	nd	nd	
NLLDQNETK	237 $\pm$ 14	294 $\pm$ 20	430 $\pm$ 53	nd	nd	nd	
ospC							83 $\pm$ 6
EVEALLSSIDEIAAK	77 $\pm$ 10	85 $\pm$ 21	88 $\pm$ 16	nd	nd	nd	
bmpA							79 $\pm$ 7
ALNIFTSNHLK	81 $\pm$ 39	77 $\pm$ 21	nd	nd	nd	nd	

^aAll experiments were performed in duplicate with three replicate injections. Data is presented as the mean $\pm$ SD. ^bnd means not detected.

for the whole homogenate of this prototypical *B. burgdorferi*. On the basis of silver staining, the list of most abundant proteins includes integral outer membrane protein P66 (p66), aminopeptidase 1 (apeA), basic membrane protein A (bmpA), outer surface protein A (ospA), chaperone protein DnaK (dnaK), 60 kDa chaperonin (groL), enolase (eno), and glyceraldehyde 3-phosphate dehydrogenase (gap). The identification statistic of these proteins is summarized in the [Supporting Information Table S2](#). For all of these identifications, the MOWSE score from Mascot software (http://www.matrixscience.com/search_form_select.html) search was higher than the score at which statistical significance ($p < 0.05$) occurred for that particular search. Identified proteins also correspond well to the expected molecular mass and PI values.

To be selected as a target protein for *B. burgdorferi* detection in human serum, the protein amino acid sequence has to be distinguishable from any human protein(s). BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) and LALIGN (http://www.ch.embnet.org/software/LALIGN_form.html) searches revealed that dnaK, groL, eno, and gap proteins from *B. burgdorferi* have high level of homology with corresponding human proteins and cannot be used for selective *B. burgdorferi* detection. Four other proteins, p66, apeA, bmpA, and ospA, have amino acid sequences that can generate multiple tryptic peptides, which are unique and will identify *B. burgdorferi* unambiguously. Although not observed in our two-dimensional gel, three more *B. burgdorferi* proteins, namely, outer surface protein C (ospC), flagellar filament 41 kDa core protein (fla), and DNA-binding protein HU (hup), have also attracted our attention. These proteins were previously reported as abundant *B. burgdorferi* proteins^{16,23–26} and have unique tryptic peptides for selective detection of *B. burgdorferi*. Consequently, the list of tryptic peptides for the seven *B. burgdorferi* proteins was generated in silico. On the basis of published rules,⁹ peptides acceptable for MRM analysis were selected and compiled into a QconCAT sequence ([Supporting Information Figure S1](#)). The ¹⁵N-labeled QconCAT was expressed, purified, and characterized ([Supporting Information Figure S2](#)).

While this QconCAT was designed to quantify *B. burgdorferi* only, it is important to note that the target proteins possess many tryptic peptides which are not only species-selective, but are also strain-selective. This means it is possible to design QconCATs, which will allow selective quantification of specific *Borrelia* species (such as *B. burgdorferi*, *B. garinii*, *B. afzelii*, etc.) in one LC–MS/MS run. Selective quantification of specific strains of *B. burgdorferi* is possible as well.

Detecting *B. burgdorferi* Membrane Proteins Supplemented in the Human Serum. We hypothesized that, upon bacterial infection, the broken bacterial cells would generate outer membrane vesicles populated with membrane proteins and become the source of membrane proteins in the serum. Five *B. burgdorferi* proteins included in the QconCAT are membrane proteins: p66 is a single pass transmembrane protein; bmpA, ospA, and ospC have an N-terminal lipid anchor; apeA was recovered from a membrane fraction.²⁴ Previously, we have demonstrated that washing the membrane pellet with 0.1 mol/L Na₂CO₃ can enrich the sample with those membrane proteins which were not detectable in membrane pellet after high-speed centrifugation.¹⁰ However, prior to applying 0.1 mol/L Na₂CO₃ washing to our samples, we confirmed that this washing step does not cause loss of the target proteins. Visual comparison of p66, apeA, bmpA, and ospA staining on the two-dimensional gels shows no changes in the amount of these proteins before and

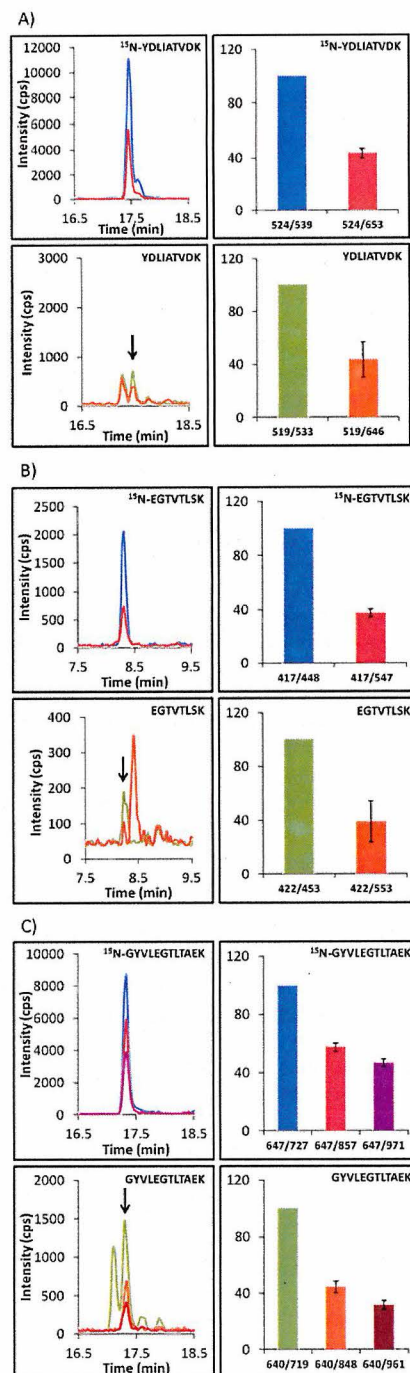


Figure 2. Extracted ion chromatograms and MRM transitions ratio monitored for ospA in the serum from Lyme disease patients, collected at the time of first visit. Data are presented for the YDLIATVDK (A), EGTVTLSK (B), and GYVLEGLTAEK (C) peptides. Heavy and light versions of peptides represent QconCAT and endogenous ospA, respectively. Overlaid extracted ion chromatograms and bars for transitions ratio are color-coordinated. For transitions ratio, the most intensive transition was taken as a 100 and intensities of other transitions were plotted as a portion of a 100. Measurements were performed in duplicate for three Lyme disease patients with three analytical replicates.

Table 2. Quantification of ospA in the Serum from Lyme Disease Patients Collected on Their First Visit

	$^{15}\text{N}/^{14}\text{N}$ ratio ^b	fmol of ospA/mg of serum protein ^a				
		patient no. 1	patient no. 2	patient no. 3	av	consensus
YDLIATVDK	16.3 $\pm$ 4.0	2.4 $\pm$ 0.5	3.3 $\pm$ 0.6	3.0 $\pm$ 0.6	2.9 $\pm$ 0.6	3.8 $\pm$ 0.4
EGTVTLTK	14.7 $\pm$ 5.4	3.7 $\pm$ 1.4	4.1 $\pm$ 1.6	2.7 $\pm$ 1.0	3.5 $\pm$ 1.4	
GYVLEGTLTAEK	9.4 $\pm$ 2.5	5.4 $\pm$ 1.4	5.8 $\pm$ 0.8	4.2 $\pm$ 1.0	5.1 $\pm$ 1.3	

^aAll experiments for three patients were performed in duplicate with three replicate injections. Data is presented as the mean $\pm$ SD. ^bAverage ratio shows the proportion between ^{15}N -labeled internal standard and ^{14}N -analyte signal intensities.

after 0.1 mol/L Na_2CO_3 washing (Supporting Information Figure S3). At the same time, measurement of the total protein in these samples shows that washing with 0.1 mol/L Na_2CO_3 removed 57% of the total protein, resulting in approximately 2-fold sample enrichment with target proteins. Such a step is important for detection of low-abundance proteins.

To prove the proposed concept, we supplemented 1 mL of normal human serum with various amount of total bacterial protein obtained from *B. burgdorferi* B31 isolate A3. The membrane proteins were pulled down by high-speed centrifugation and washed with 0.1 mol/L Na_2CO_3 . The washed membrane pellet was supplemented with ^{15}N -labeled QconCAT, and mass spectrometry analysis was focused on detecting *B. burgdorferi* proteins included in the QconCAT. The quantitative data are summarized in Table 1. Two proteins, *apeA* and *hup*, were not detected in these experiments and are not included in Table 1. Quantification of five other proteins for different added amounts of total *B. burgdorferi* protein was consistently reproducible with the SD generally under 20% of the mean value. The lowest detection for these proteins concurs well with their relative abundance in *B. burgdorferi*. For example, *bmpA* with an average concentration 79 ± 7 pmol/mg of total *B. burgdorferi* protein was detected at 1 μg supplementation while *ospA* with an average concentration 4220 ± 349 pmol/mg of total *B. burgdorferi* protein was detected at 0.03 μg supplementation. None of the proteins was detected for 0.01 μg supplementation of total *B. burgdorferi* protein. The data in Table 1 allow two simple calculations. First, the molecular mass of *ospA* is 29.4 kDa and the concentration 4220 pmol of *ospA* per mg of total *B. burgdorferi* protein means that *ospA* represents approximately 12% of total protein in *B. burgdorferi* B31 isolate A3. Second, 0.03 μg supplementation represents approximately 4 ng of *ospA*. This amount was supplemented into 1 mL of human serum, which has 60 mg of total protein. The resulting dynamic range between 60 mg/mL and 4 ng/mL is 1.5×10^7 . We believe this would be the largest dynamic range reported for quantitative measurements of a target protein in the human serum. Overall, the developed protocol for quantitative analysis of bacterial membrane proteins in human serum encouraged us to proceed to measurements of serum samples from control and Lyme disease-diagnosed patients.

Detecting *B. burgdorferi* Membrane Proteins in the Serum of Lyme Disease-Diagnosed Patients. There is a well-documented diversity of the spirochetes protein expression pattern in response to changing environmental factors.^{23,25–29} In the previous section, we have used a clonal and low-passage infectious *B. burgdorferi* B31 isolate A3 to optimize the serum processing protocol and detecting of spirochetes proteins supplemented in human serum. The average protein concentrations presented in Table 1 are correct for this specific isolate and its cultured conditions. For example, average concentration for *ospC* at 83 pmol/mg of total *B. burgdorferi* protein is 50-fold lower than concentration for *ospA*, which is 4220 pmol of *ospA*

per mg of total *B. burgdorferi* protein. This explains why we do not see *ospC* in our 2D-PAGE images (Figure 1 and Supporting Information Figure S3). However, previously published 2D-PAGE images of a different isolate of *B. burgdorferi* B31²⁴ show approximately equal levels of expression for both proteins. Therefore, when approaching measurements in human serum from naturally infected patients, we decided to track all proteins included in QconCAT no matter how abundant they are in the *B. burgdorferi* B31 isolate A3.

In our measurements, we used 1 mL size serum samples from control patients and from Lyme disease patients, collected at the time of their first and third visits (Supporting Information Table S1). We did not detect *B. burgdorferi* proteins in the control and third visit samples, but we were able to detect *ospA* in the Lyme disease samples, collected at the time of first visit. Detection has been made based on three *ospA* peptides (Figure 2): YDLIATVDK (two transitions), EGTVTLSK (two transitions), and GYVLEGTLTAEK (three transitions). It is important to emphasize that the relative ratios of these transitions for every pair of labeled and nonlabeled peptides were almost identical (Figure 2). This confirms that *ospA* quantification is not affected by nonspecific interference from the biological sample. In the contrast to supplementation experiments, we cannot normalize these measurements to the amount of the total *B. burgdorferi* protein. However, taking a value of 60 mg of serum protein/mL, we can normalize *ospA* measurements to a mg of total serum protein. On the basis of YDLIATVDK, EGTVTLSK, and GYVLEGTLTAEK peptides, we have detected an average concentration of *ospA* as 2.9 ± 0.6 , 3.5 ± 1.4 , and 5.1 ± 1.3 fmol of *ospA*/mg of serum protein for three Lyme disease patients, respectively (Table 2). As a consensus, it brings us to the detection limit on the level of approximately 4.0 fmol of *ospA*/mg of serum protein.

CONCLUSIONS

In the summary, broken bacterial cells may be a source of membrane proteins in human serum during early bacterial infection with *B. burgdorferi*. Targeting these proteins is strongly supported by the fact that a simple experimental step such as high-speed centrifugation allows substantial enrichment of the sample before LC-MS/MS analysis. Therefore, detection of proteins, whose abundance is $\approx 10^7$ lower than abundance of major serum proteins, became feasible. We report here a proof of concept based on detecting *ospA* protein in the serum from patients diagnosed with Lyme disease, and we believe that this approach may be universally applicable to detection of other bacterial infections in human serum. We anticipate that future investigations with additional serum samples will further substantiate this approach and expand its potential range of applications.

■ ASSOCIATED CONTENT

● Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.analchem.5b02803.

Donor information, identification of proteins in Figure 1, design of QconCAT sequence, characterization of ¹⁵N-labeled QconCAT, and 2D-PAGE pattern of *B. burgdorferi* membrane proteins (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: iturko@umd.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported in part by the Lyme Disease Research Foundation, Inc. (www.lymemd.org), the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health under award number P30AR05350, and the Stabler Foundation (www.stablerfoundation.org). Certain commercial materials, instruments, and equipment are identified in this manuscript in order to specify the experimental procedure as completely as possible. In no case does such identification imply a recommendation or endorsement by the National Institute of Standards and Technology nor does it imply that the materials, instruments, or equipment identified are necessarily the best available for the purpose.

■ REFERENCES

- (1) Centers for Disease Control and Prevention Reported Cases of Lyme Disease by Year, United States 1996–2010. <http://www.cdc.gov/lyme/stats/chartstables/casesbyyear.html>.
- (2) Hinckley, A. F.; Connally, N. P.; Meek, J. I.; Johnson, B. J.; Kemperman, M. M.; Feldman, K. A.; White, J. L.; Mead, P. S. *Clin. Infect. Dis.* **2014**, *59*, 676–681.
- (3) Borchers, A. T.; Keen, C. L.; Huntley, A. C.; Gershwin, M. E. *J. Autoimmun.* **2015**, *57*, 82–115.
- (4) Adrion, E. R.; Aucott, J.; Lemke, K. W.; Weiner, J. P. *PLoS One* **2015**, *10*, e0116767.
- (5) Nocton, J. J.; Dressler, F.; Rutledge, B. J.; Rys, P. N.; Persing, D. H.; Steere, A. C. *N. Engl. J. Med.* **1994**, *330*, 229–234.
- (6) Cerar, T.; Ogrinc, K.; Cimperman, J.; Lotric-Furlan, S.; Strle, F.; Ruzic-Sabljic, E. *J. Clin. Microbiol.* **2008**, *46*, 3375–3379.
- (7) Eshoo, M. W.; Crowder, C. C.; Rebman, A. W.; Rounds, M. A.; Matthews, H. E.; Picuri, J. M.; Soloski, M. J.; Ecker, D. J.; Schutzer, S. E.; Aucott, J. N. *PLoS One* **2012**, *7*, e36825.
- (8) Iyer, R.; Mukherjee, P.; Wang, K.; Simons, J.; Wormser, G. P.; Schwartz, I. J. *Clin. Microbiol.* **2013**, *51*, 857–862.
- (9) Liao, W.-L.; Heo, G.-Y.; Dodder, N. G.; Pikuleva, I. A.; Turko, I. V. *Anal. Chem.* **2010**, *82*, 5760–5767.
- (10) Wang, M.; Heo, G.-Y.; Omarova, S.; Pikuleva, I. A.; Turko, I. V. *Anal. Chem.* **2012**, *84*, 5186–5191.
- (11) Chen, J.; Huang, R. Y.-C.; Turko, I. V. *Anal. Chem.* **2013**, *85*, 6011–6017.
- (12) Cheung, C. S. F.; Anderson, K. W.; Wang, M.; Turko, I. V. *Anal. Chem.* **2015**, *87*, 1097–1102.
- (13) Liebler, D. C.; Zimmerman, L. J. *Biochemistry* **2013**, *52*, 3797–3806.
- (14) Gillette, M. A.; Carr, S. A. *Nat. Methods* **2013**, *10*, 28–34.
- (15) Picotti, P.; Aebersold, R. *Nat. Methods* **2012**, *9*, 555–566.
- (16) Schnell, G.; Boeuf, A.; Westermann, B.; Jaulhac, B.; Lipsker, D.; Carapito, C.; Boulanger, N.; Ehret-Sabatier, L. *Mol. Cell. Proteomics* **2015**, *14*, 1254–1264.
- (17) Rus, H.; Cudrici, C.; Niculescu, F. *Immunol. Res.* **2005**, *33*, 103–112.
- (18) Kochi, S. K.; Johnson, R. C.; Dalmasso, A. P. *J. Immunol.* **1991**, *146*, 3964–3970.
- (19) Breitner-Ruddock, S.; Würzner, R.; Schulze, J.; Brade, V. *Med. Microbiol. Immunol.* **1997**, *185*, 253–260.
- (20) Craft, J. E.; Grodzicki, R. L.; Steere, A. C. *J. Infect. Dis.* **1984**, *149*, 789–795.
- (21) Kilpatrick, E. L.; Liao, W.-L.; Camara, J. E.; Turko, I. V.; Bunk, D. M. *Protein Expression Purif.* **2012**, *85*, 94–99.
- (22) Liao, W.-L.; Turko, I. V. *Anal. Biochem.* **2008**, *377*, 55–61.
- (23) Jacobs, J. M.; Yang, X.; Luft, B. J.; Dunn, J. J.; Camp, D. G., II; Smith, R. D. *Proteomics* **2005**, *5*, 1446–1453.
- (24) Nowalk, A. J.; Nolder, C.; Clifton, D. R.; Carroll, J. A. *Proteomics* **2006**, *6*, 2121–2134.
- (25) Yang, X.; Promnares, K.; Qin, J.; He, M.; Shroder, D. Y.; Kariu, T.; Wang, Y.; Pal, U. J. *Proteome Res.* **2011**, *10*, 4556–4566.
- (26) Gesslbauer, B.; Poljak, A.; Handwerker, C.; Schuler, W.; Schwendenwein, D.; Weber, C.; Lundberg, U.; Meinke, A.; Kungl, A. J. *Proteomics* **2012**, *12*, 845–858.
- (27) Revel, A. T.; Talaat, A. M.; Norgard, M. V. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 1562–1567.
- (28) Angel, T. E.; Luft, B. J.; Yang, X.; Nicora, C. D.; Camp, D. G., II; Jacobs, J. M.; Smith, R. D. *PLoS One* **2010**, *5*, e13800.
- (29) Schnell, G.; Boeuf, A.; Jaulhac, B.; Boulanger, N.; Collin, E.; Barthel, C.; De Martino, S.; Ehret-Sabatier, L. *Proteomics* **2015**, *15*, 1280–1290.