
Mary F. Lopez, PhD,* David A. Sarracino,* Maryann Vogelsang,* Jennifer N. Sutton,* Michael Athanas,* Bryan Krastins,* Alejandro Garces,* Amol Prakash,* Scott Peterman,* Zareh Demirjian,† Ignacio Inglessis-Azuaje I,† Kathleen Feeney,† Mikaela Elia,† David McMullin,† G. William Dec,† Igor Palacios,† Eng H. Lo,† Ferdinand Buonanno,† and MingMing Ning, MD†

Abstract: Patent foramen ovale (PFO) is highly prevalent and associated with more than 150,000 strokes per year. Traditionally, it is thought that PFOs facilitate strokes by allowing venous clots to travel directly to the brain. However, only a small portion of PFO stroke patients have a known tendency to form blood clots, and the best treatment for this multiorgan disease is unclear. Therefore, mapping the changes in systemic circulation of PFO-related stroke is crucial in understanding the pathophysiology to individualize the best clinical treatment for each patient. We initiated a study using a novel quantitative, 2-pass discovery workflow using high-resolution liquid chromatography–mass spectrometry/mass spectrometry coupled with label-free analysis to track protein expression in PFO patients before and after endovascular closure of the PFO. Using this approach, we were able to demonstrate quantitative differences in protein expression between both PFO-related and non–PFO-related ischemic stroke groups as well as before and after PFO closure. As an initial step in understanding the molecular landscape of PFO-related physiology, our methods have yielded biologically relevant information on the synergistic and functional redundancy of various cell-signaling molecules with respect to PFO circulatory physiology. The resulting protein expression patterns were related to canonical pathways including prothrombin activation, atherosclerosis signaling, acute-phase response, LXR/RXR activation, and coagulation signaling. These findings demonstrate the feasibility and robustness of using a proteomic approach for biomarker discovery to help gauge therapeutic efficacy in stroke.

Key Words: proteomics, biomarker, discovery, stroke, cerebrovascular disease, ischemic stroke, patent foramen ovale, PFO, mass spectrometry

From the *Thermo Fisher Scientific BRIMS, Cambridge; and †Clinical Proteomics Research Center and Cardio-Neurology Clinic, Department of Neurology, Massachusetts General Hospital, Harvard Medical School, Boston, MA.

Received September 28, 2012.

Reprints: Mary F. Lopez, PhD, BRIMS, Thermo Fisher Scientific, 790 Memorial Dr, Cambridge, MA 02139. E-mail: mary.lopez@thermofisher.com; or MingMing Ning, MD, Clinical Proteomics Research Center, Massachusetts General Hospital/Harvard Medical School, 15 Parkman St, ACC 720; Boston, MA 02114. E-mail: ning@hms.harvard.edu.

This work was supported by the NIH/NINDS, R01 NS067139 (M.M.N.) and P01-NS55104 (E.H.L.).

Supplemental digital content is available for this article. Direct URL citation appears in the print text and is provided in the HTML and PDF versions of this article on the journal’s Web site (www.investigativemed.com).

Copyright © 2012 by The American Federation for Medical Research
ISSN: 1081-5589
DOI: 10.231/JIM.0b013e318276ede0

Patent foramen ovale (PFO) is an independent stroke risk factor, but its best treatment is not clear (Fig. 1).1,2 Highly prevalent (25%-30% of the general population), PFO is often discovered only after a stroke and is associated with more than 150,000 strokes per year.3–5 Traditionally, it is thought that PFOs facilitate paradoxical embolism by allowing venous clots to travel directly to the brain. However, there is a significant disconnect between this simple mechanism and clinical data, as less than approximately 15% of PFO stroke patients have a known tendency to form venous clots.3,6 Clinical trials to investigate treatment options are ongoing, but because individual risks vary, and preferred treatment is likely not one-size-fits-all, controversies regarding PFO remain unresolved.6,7 In part, this is due to a poor understanding of the molecular landscape of PFO-related neurovascular injury.8–10

Because PFO is a complex, multiorgan disease involving the brain, heart, and circulation, we initiated a study using mass spectrometry (MS) to follow blood protein expression in PFO patients before and after endovascular closure of the PFO. Clinical endovascular closure of PFO provides a rare bedside model in which to study the effects of a specific mechanical intervention on circulatory protein signaling, both immediately and over time.11

Although MS has been applied to biomarker discovery for at least a decade, one of the most difficult problems has been the interpretation and ranking of putative biomarkers derived from differential expression liquid chromatography–MS/MS experiments.38 Limitations of the classic approaches that depend on data-dependent MS acquisition from complex peptide mixtures include lack of rigorous quantification and independent parameters for evaluating the “usefulness” of a particular biomarker. In addition, the high dynamic range of clinical samples, such as plasma and serum, effectively limits protein identification of low-abundance proteins when data-dependent MS acquisition methods are used, because only the highest abundance proteins are identified repeatedly.39 Previous approaches have incorporated physical fractionation of protein samples to dig deeper into the proteome.40 However, albumin depletion and offline fractionation (such as cation exchange) often result in protein losses and greatly increased instrument run time and expense.41 In addition, rigorous quantification becomes increasingly difficult when sample preparation is so complex. As an improvement to these approaches, we have developed a 2-pass, quantitative discovery workflow that includes application of expression trend
ratios and receiver operating characteristic (ROC) analyses for the efficient evaluation and scoring of putative biomarkers (Fig. 2).22 The 2-pass approach couples very accurate, full-scan quantification with MS/MS acquisition driven by an inclusion list generated from analysis of the full-scan data. Application of the inclusion list for MS/MS acquisition essentially uses the mass spectrometer to "fractionate" the sample and results in the increased identification of lower abundance and clinically useful biomarkers. In addition, because multiple physical fractions do not need to be analyzed by MS, more clinical samples can be analyzed in the same amount of time, allowing for better statistics and evaluation of biological variability. Using this approach, we were able to demonstrate clear, quantitative differences in protein expression across the various sample groups and relate the expression patterns to relevant biological pathways. We utilize this 2-pass, quantitative discovery workflow to explore circulatory protein expression in stroke patients before and after endovascular closure of the PFO.

MATERIALS AND METHODS

Clinical Serum Samples

A total of 64 plasma samples were obtained from stroke patients and healthy controls with similar risk factors from the Cardio-Neurology Clinic of Massachusetts General Hospital in accordance with institutional review board approval. Patent foramen ovale–related "cryptogenic ischemic strokes" were identified by 2 vascular board-certified neurologists. Rigorous
inclusion/exclusion criteria were applied to ensure proper diagnosis in each group, and all patients underwent the following testing to rule out other etiology related to ischemic embolic infarct: (1) physical examination by a vascular neurologist to document clinical syndrome consistent with ischemic infarct; (2) magnetic resonance imaging/magnetic resonance angiography and computed tomographic angiography to document ischemic infarct and to rule out other reasons of infarct such as intracranial stenosis, large vessel occlusion, and dissection; (3) transthoracic and/or transesophageal echocardiogram to assess and document presence of PFO and rule out atrial/ventricular thrombus or any valvular lesions that may be related to embolism; and (4) extended cardiac monitoring to rule out atrial fibrillation or other cardiac arrhythmia. In addition, other ischemic stroke subtypes such as lacunar infarct related to hypertension, vasculitis, endocarditis, or venous infarction were excluded. Also excluded were all patients with active infection, active pregnancy, or renal/liver failure, as these conditions may alter protein signaling profiles. Moreover, patients without baseline imaging such as computed tomography or magnetic resonance imaging needed to confirm clinical stroke syndromes were excluded. Controls were recruited from subjects with similar baseline risk factors to match the study population. Subject demographics (n = 14 PFO-related ischemic stroke, n = 7 non-PFO-related cryptogenic ischemic stroke, and n = 18 control subjects) are listed in Table 1. Patients were consecutively, prospectively enrolled post–ischemic stroke in accordance with the approval of the institutional review board.

Standard operating procedures were strictly observed for all samples. All personnel were trained with standard operating procedure to process samples in the same fashion, as follows: blood from venous or atrial source was collected into EDTA-coated tubes and immediately processed (within 5 minutes) to obtain plasma by centrifugation at 3400 revolutions/min for 15 minutes at 20°C (to avoid platelet activation), removing the plasma supernatant without disturbing the clot, aliquoted immediately, and frozen at −80°C to ensure minimal protein degradation.

Samples were carefully transported frozen to be processed at the same time, in random order, by investigators blinded to the clinical data to avoid bias and batch variations.

**Trypsin Digestion, Reduction/Alkylation, and Desalting**

Plasma samples (25 μL) were thawed on ice and processed as previously described.43

**Liquid Chromatography and High-Resolution MS**

As shown in Figure 1, the 2-pass workflow strategy consists of the separate optimization of MS parameters and configuration for protein quantification and identification.

**Pass 1.** Plasma samples (500 ng of human samples) were prepared as described above and injected onto a Thermo Scientific Easy nLC system configured with a 10 cm × 100-μm trap column and a 25 cm × 100-μm internal diameter resolving column. The sample load was optimized for optimum quantification (i.e., full-scan data). Buffer A was 98% water, 2% methanol, and 0.2% formic acid. Buffer B was 10% water, 10% isopropanol, 80% acetonitrile, and 0.2% formic acid. Samples were loaded at 4 μL/min for 10 minutes, and a gradient from 0% to 45% B at 375 nL/min was run over 130 minutes, for a total run time of 150 minutes (including regeneration and sample loading). The Thermo Scientific LTQ Orbitrap Velos mass spectrometer was run in a standard top-10 data-dependent configuration, except that a higher trigger-threshold (20K) was used to ensure that the MS2 did not interfere with the full-scan duty cycle. This ensured optimal full-scan data for quantification. MS2 fragmentation and analysis were performed in the ion trap mass analyzer.

**Data Analysis and Pass 2.** Proteomics data analysis was performed using Thermo Scientific SIEVE software version 2.0 that features chromatographic alignment, framing, differential ROC, ratio, and trend analyses. Both top-10 data-dependent scans and full-scan data were analyzed with the SIEVE software using chromatographic alignment followed by feature extraction using unsupervised statistical techniques including isotope deconvolution. Based on various criteria including ROC area under the curve, low ratios, high ratios, high abundance, low abundance, or trend ratios, an inclusion list was created for the best differentially expressed candidates. This inclusion list was used exclusively for MS2 acquisition in pass 2. A larger sample load was used in the pass 2 runs (600–800 ng), allowing for higher-quality MS2 spectra. Because these full-scan spectra would not be used for quantification, peak shape and intensity reproducibility were not crucial. Fragmentation scans from pass 2 were analyzed using SEQUEST and FDR analysis to make identifications. The fragmentation search results from pass 2, and the quantitative information obtained from pass 1, were combined and further analyzed using the SIEVE software.

**Ingenuity Pathways Data Analysis**

The differential expression results were uploaded and analyzed using Ingenuity Pathways Analysis (IPA), http://www.ingenuity.com/products/pathways_analysis.html, according to the manufacturer’s instructions.

**RESULTS**

**Two-Pass Workflow**

Figure 2 shows a graphical representation of the workflow that was applied to the clinical samples. This approach was recently developed in our laboratory to improve quantification by decoupling it from protein/peptide identification.42 During pass

---

**TABLE 1.** Patient Clinical Characteristics and Sample Collection: PFO-Related Versus Non–PFO-Related Versus Normal Experiment

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>PFO-Related (n = 14)</th>
<th>Non–PFO-Related (n = 7)</th>
<th>Normal (n = 18)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean (range), y</td>
<td>46 (26–55)</td>
<td>45 (25–53)</td>
<td>44 (22–56)</td>
</tr>
<tr>
<td>Sex</td>
<td>50% Male</td>
<td>57% Male</td>
<td>50% Male</td>
</tr>
<tr>
<td>Race</td>
<td>100% White</td>
<td>100% White</td>
<td>94% White</td>
</tr>
<tr>
<td>Hyperlipidemia</td>
<td>14%</td>
<td>14%</td>
<td>11%</td>
</tr>
<tr>
<td>Hypertension</td>
<td>7%</td>
<td>0</td>
<td>5%</td>
</tr>
<tr>
<td>Diabetes</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Residual right-to-left shunting after PFO closure on cardiac echo</td>
<td>0</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>
1, parameters are optimized for acquisition of full-scan data with an emphasis on chromatographic reproducibility and robust spray. MS2 spectra are acquired only in a “top 5 or 10” data-dependent manner so full-scan measurements are not compromised. This approach makes high-quality relative quantification possible, even in complex samples such as plasma, but also provides a limited list of MS2 spectra for identification of high-abundance proteins. The full-scan data are analyzed, and differentially expressed features fulfilling a desired pattern filter are compiled into a target inclusion list for acquisition in pass 2. Because the mass spectrometer acquires only data from masses that are on the target list (as opposed to default data-dependent acquisition), less intense (and therefore less abundant) masses may be triggered to produce fragmentation spectra for identification. Using this approach, no fractionation or depletion was necessary to identify proteins at intensities in the picomolar range. Figure 3 shows the distribution of protein IDs in passes 1 and 2 that passed the filter criteria. As is evident in the figure, the majority of proteins, 95%, were uniquely identified in pass 2, with 0.4% uniquely identified in pass 1 data. Only 1 protein fulfilling the filtering criteria was identified in both passes 1 and 2 data.

Differential Protein Expression in PFO-Related Stroke Samples Versus Non-PFO-Related Stroke Samples

When the 2-pass workflow was applied to the PFO-related stroke, non-PFO-related stroke, and normal (nonstroke) samples in Table 1 in a trend analysis, a striking differential expression pattern was evident (Fig. 4). To identify which proteins were differentially expressed between PFO-related and non-PFO-related stroke, we applied the following expression filter to the complete data set at the protein level: normalized ratio PFO-related/normal > 1.2 AND normalized ratio PFO-related/normal < 0.6. The resulting data were further analyzed, and only proteins with at least 2 peptides and the 1% FDR and $P < 0.01$ stringency parameters were retained. The resulting collection of 57 proteins, shown in Figure 4, provides an initial list of biomarker candidates differentiating the PFO-related and non-PFO-related sample groups. The proteins with the highest expression ratios in PFO-related versus non-PFO-related samples (ratio > 2.0) were $\alpha_2$-macroglobulin, serum albumin, immunoglobulin $\lambda$, and complement C3 precursor. Conversely, the proteins with the highest expression ratios in non-PFO-related versus PFO-related samples were serine/threonine protein kinase and SrC homology 3 domain containing guanine nucleotide exchange factor. We further analyzed this protein data set with IPA to determine canonical pathways that were significantly associated with the data set. Figure 5 shows that several pathways including extrinsic and intrinsic prothrombin activation, atherosclerosis signaling, aldosterone signaling, renin-angiotensin signaling, cardiac $\beta$-adrenergic signaling, and thrombin signaling pathways had highly significant ($P < 0.001$) overlap with the PFO-related and non-PFO-related putative marker data set. The highest overlap was with extrinsic prothrombin activation where 2 of 20 proteins had decreased and 4 of 20 proteins had increased expression versus the experimental data set (Figure 5; and Supplemental Table 1, Supplemental Digital Content 1, http://links.lww.com/JIM/A13).

Differential Protein Expression Before, During, and After PFO Endovascular Closure In PFO-Related Stroke Samples

To investigate the protein expression pattern related to endovascular closure of PFO in stroke, we analyzed the complete set of longitudinal samples from 14 patients described in Table 2. The matched samples were obtained from patient baseline preendovascular PFO closure venous blood (Preop), from the left (PRLA) and right (PRRA) atria immediately before endovascular closure, from the left (PSLA) and right (PSRA) atria immediately after endovascular closure, and venous blood at 3 months’ follow-up. We applied the following filter to the protein data obtained from the SIEVE trend analysis:

\[ \text{Ratio of 3 month follow---up/Preop > 1.5 OR < 0.6} \]

Only proteins represented by at least 2 peptides were allowed. The resulting protein expression pattern is shown in Figure 6. There was considerable variability in $P$ value across the matched samples, but several proteins were significantly ($P < 0.05$) differentially expressed, including $\beta$-globin, serum albumin, $\alpha_2$-macroglobulin, serotransferrin, similar to immunoglobulin $\lambda$–like polypeptide 1, apolipoprotein A1, and $\alpha_1$-antitrypsin precursor. When we applied IPA canonical pathway analysis to the PFO closure data set, several pathways were
correlated with high significance ($P < 0.0001$) (Fig. 7). The significant pathways included acute-phase response signaling, LXR/RXR activation, and extrinsic and intrinsic prothrombin activation pathways. Contrasting with the results from the previous (PFO-related stroke vs. non-PFO-related stroke) IPA analysis, numerous proteins in the PFO closure data set demonstrated reduced...
expression after PFO closure with respect to the canonical pathways such as acute inflammatory response, cholesterol, and coagulation signaling (Supplementary Table 2, Supplemental Digital Content 2, http://links.lww.com/JIM/A14).

**DISCUSSION**

As an initial step in understanding the molecular landscape of PFO-related physiology, our methods have yielded biologically relevant information on the synergistic and functional redundancy of various cell-signaling molecules with respect to PFO circulatory physiology. In particular, after PFO closure, numerous proteins demonstrated reduced expression in stroke-related canonical pathways such as acute inflammatory response, LXR/RXR, and coagulation signaling. These findings demonstrate the feasibility and robustness of using the 2-pass approach for biomarker discovery in a clinical model with patients as their own controls to gauge therapeutic efficacy of PFO endovascular closure.

Patent foramen ovale-related neurovascular injury certainly has a major impact on public health, and the high prevalence of PFO emphasizes the importance of proper prevention, treatment, and management in at-risk individuals.1,5,9,11,12,20,25,36,44 In addition to paradoxical embolic strokes, PFOs are present in more than 60% of all migraine-with-aura patients and 20% to 30% of the general population.45–51 But who should be screened to prevent future strokes? What therapeutic targets and options are best for stroke and migraine patients? This is likely a multiorgan disease dependent on interaction between the brain, heart, and lungs—but with potential connection through circulatory signaling in blood. For example, we found that serotonin (5-HT), a vasoactive prothrombotic substance that induces cardiac oxidative stress and a neurotransmitter carefully regulated and inactivated by the lung, is decreased after percutaneous closure of PFO—supporting the hypothesis that chronic right-to-left shunting may allow harmful vasoactive mediators to escape deactivation by the pulmonary vasculature.37,52,53 This increase in the brain vasculature’s direct exposure to various vasoactive mediators may elevate the risk of cerebrovascular injury.37 Results of the current study support our earlier findings. In this study, the plasma signal change before and after PFO closure suggests that PFO with large venous-to-arterial shunting may allow for harmful circulatory factors to travel directly from the venous (right atrial, deoxygenated) to the arterial (left atrial, oxygenated) side, bypassing important deactivation by the lung—potentially triggering a proinflammatory, procoagulant state. Conversely, obliterating this shunting by PFO closure seems to alter circulatory phenotype into a less procoagulant and proinflammatory state.

These data are hypothesis generating, and further studies in larger cohorts are needed to investigate and confirm individual factors. Although this is a very small sample of patients, we made every effort to include patients with similar age, risk factor, and comorbidities. Patent foramen ovale–related stroke patients are known to have fewer conventional vascular risk factors, which make them ideal for proteomic studies. The study design of obtaining preprocedure and postprocedure samples helps to minimize confounders by utilizing each individual patient’s
preoperative baseline as their own control. In addition, all patients in this cohort are on the same dose of 325 mg of aspirin throughout the study. And all baseline blood and endovascular closure samples were remote from the time of stroke (9-3 months), so that the changes we found in acute inflammatory factors over time were not related to decrease in acute inflammatory response over time. Migraine patients were excluded from this study so that the molecular mechanism of migraine would not confound findings in this study for PFO-related stroke. However, we are actively investigating this subpopulation, because migraine subtypes in specific patient populations are related to PFO and stroke. There were also no changes in any other medications throughout the longitudinal follow-up, such that no new confounders were introduced except for the PFO

* P-value < 0.05
** P-value < 0.001

FIGURE 6. Protein expression pattern before, during, and after PFO endovascular closure.

Proteomic profiling studies with respect to long-term clinical outcome, such as recurrent TIA or stroke, and comparison of various antiplatelet and anticoagulation treatments and of patients with concurrent migraines are important and underway. In summary, we find that proteomic exploration can help to identify targets of PFO-related neurovascular disease in the context of a largely unexplored territory of circulatory signaling with the potential to triage and monitor therapeutic efficacy.

REFERENCES


