Whole Genome Expression of Cellular Response to Stroke

Frank R. Sharp, MD; Glen C. Jickling, MD, MAS

Many attempts have been made at developing biomarkers for stroke. Although successful to some degree,\textsuperscript{1-3} none have been sufficiently robust to be used in clinical practice. Thus, there is still a great need for more in-depth studies of the biology of human stroke to understand its pathogenesis better. This should make it possible to develop blood tests for stroke and transient ischemic attacks (TIAs) to guide treatment and ultimately improve outcomes.

The traditional approach to develop stroke biomarkers has been to select candidate markers based on known pathobiology. The majority of markers that have been evaluated are proteins that are measured in patients at various times before or after stroke. The rationale for this approach is that brain injury releases molecules into blood that can be measured as evidence of brain injury; or that other cells and organs release molecules that either cause or contribute to a stroke, or are a response to the stroke. These approaches have been handicapped because they require a guess at the most reliable biomarker.

Our group took a different approach to the problem by assessing the immune system after stroke.\textsuperscript{4,5} The rationale for this approach is shown in Figure 1. The expression of genes in leukocytes is influenced by many factors associated with ischemic stroke. Leukocytes interact with blood clots, platelets, atherosclerotic plaque, and injured brain endothelial cells via adhesion molecules.\textsuperscript{2} In addition, leukocytes detect circulating cytokines, chemokines, and hormones. Each has the potential to modulate RNA expression in leukocytes.

To provide proof-of-principle that circulating leukocytes could provide insight into the pathogenesis of stroke, we performed a series of animal studies. For the first study, blood of rats was obtained 24 hours after ischemic stroke, intracerebral hemorrhage, status epilepticus, hypoxia, and hypoglycemia.\textsuperscript{6} RNA was isolated from the blood and processed on whole genome microarrays. Whole genome microarrays were used because they permit an unbiased selection of molecular markers for each type of brain injury, rather than making a guess at what those markers might be. The following results demonstrated several key principles that have guided our subsequent studies:\textsuperscript{5} (1) A large number of RNAs changed expression in leukocytes after brain injuries. (2) No single RNA was sufficient to distinguish different types of brain injury. (3) A panel of RNAs (or so-called gene profile) did characterize a given type of injury, and different gene profiles could distinguish each brain injury. This led us to hypothesize that there would be a distinct gene profile for different brain injuries in humans, and that these would be useful for diagnosis and prognosis. We have gone on to show that brief global cerebral ischemia produces unique gene profiles in blood of rats depending on whether there is neuronal cell death in the hippocampus or not.\textsuperscript{7}

In addition, brief periods of focal ischemia that mimic TIAs in humans produce unique gene profiles that differ based on the duration of focal ischemia.\textsuperscript{8,9}

The first proof-of-principle studies in humans were reported by Moore et al\textsuperscript{10} in 2005. They showed an RNA expression profile from peripheral blood mononuclear cells could distinguish ischemic stroke from control patients.\textsuperscript{11} We confirmed these studies in humans and showed that genes specifically expressed in whole blood before 3 hours, at 5 hours, and at 24 hours after ischemic stroke could distinguish ischemic stroke from controls with >85% sensitivity and specificity.\textsuperscript{12} One of the surprises that came from these first whole genome studies of human stroke was that the genes expressed in human blood after ischemic stroke were very different than those expressed in rodent blood. For these reasons, almost all of our subsequent studies have been in humans.

With this promising human data, we then evaluated these results in a second validation cohort. This is a crucial, but often under-appreciated because biomarkers often fail to replicate in a second population. The reasons for this are many but can include small samples on which the predictor was identified, nonrepresentative samples for derivation or validation cohort, and lack of robust biological effects that are swamped by interindividual and technical variability. Stamova et al\textsuperscript{13} and Barr et al\textsuperscript{14} were able to confirm our initial study. Profiles from our original study predicted ischemic stroke with >85% sensitivity and specificity, and new profiles based on a larger cohort were able to distinguish ischemic stroke from disease controls with >85% sensitivity and specificity.\textsuperscript{13} Further study, however, is required to compare stroke to diseases that mimic stroke.

Studies evaluating RNA expression differences in stroke pathogenesis have also been performed. In the first study, large vessel stroke was compared with cardioembolic stroke. A total of 77 genes differed between the 2; and 23 genes could distinguish the 2 types of stroke.\textsuperscript{15} A subsequent larger study...
by Jickling et al.\textsuperscript{16} evaluated 194 samples from 76 patients with acute ischemic stroke. RNA was isolated from blood and run on Affymetrix microarrays. Genes that distinguished large vessel from cardioembolic stroke were determined at 3, 5, and 24 hours after stroke. A 40-gene profile differentiated cardioembolic stroke from large vessel stroke with >95% sensitivity and specificity. A separate 37-gene profile distinguished cardioembolic stroke attributable to atrial fibrillation from nonatrial fibrillation causes with >90% sensitivity and specificity.\textsuperscript{16} Finally, our most recent study shows specific profiles for lacunar stroke compared with large vessel and cardioembolic ischemic strokes.\textsuperscript{17}

Having developed cardioembolic, large vessel, and lacunar profiles for cause of stroke, we applied them to cryptogenic stroke.\textsuperscript{18} RNA was isolated from peripheral blood of 131 cryptogenic strokes and compared with profiles derived from 149 strokes of known cause. Each sample was run on Affymetrix microarrays. Cause of cryptogenic stroke was predicted using gene expression in blood and infarct location. Cryptogenic strokes were predicted to be 58% cardioembolic, 18% arterial, 12% lacunar, and 12% unclear pathogenesis (Figure 2).\textsuperscript{18} Cryptogenic stroke of predicted cardioembolic pathogenesis had more prior myocardial infarction and higher CHA(2)DS(2)-VASc scores compared with stroke of predicted arterial pathogenesis. Predicted lacunar strokes had higher systolic and diastolic blood pressures and lower National Institutes of Health Stroke Scale compared with predicted arterial and cardioembolic strokes. Cryptogenic strokes of unclear predicted pathogenesis were less likely to have a prior TIA or ischemic stroke.\textsuperscript{18} These data provide the first proof-of-principle that gene expression profiles in blood could be used to predict a probable cause in cryptogenic strokes.

Given these promising results, the more difficult problem of TIAs has been tackled. A challenge in TIA is the lack of a gold standard for a transient neurological event attributable to cerebral ischemia. We first performed an animal study showing that 5 and 10 minutes of focal ischemia that might simulate TIAs seen in humans produced characteristic gene expression profiles.\textsuperscript{9} This work was translated to humans, where RNA expression in blood of patients with TIA (n=26) was compared with vascular risk factor control subjects without symptomatic cardiovascular disease (n=26).\textsuperscript{10} There were 449 genes differentially expressed between TIA and controls. Hierarchical cluster analysis of the identified genes suggested the presence of 2 patterns of RNA expression in patients with TIA, with 1 group possibly being associated with those with high risk of stroke, given positive diffusion-weighted imaging–magnetic resonance imaging in 2 patients, and another going on to have a stroke.\textsuperscript{19}

On further study, 74 genes expressed in TIA were found to be common to those in ischemic stroke.\textsuperscript{20} Functional pathways common to TIA and stroke involved granulocytes and B cells. A prediction model using 26 of the 74 ischemia genes distinguished TIA and stroke subjects from control subjects with 89% sensitivity and specificity. In the validation cohort, 17 of 17 TIA diffusion-weighted imaging–positive/minor strokes were predicted to be ischemic, and 10 of 13 nonischemic transient neurological events were predicted to be nonischemic. In transient neurological events of unclear pathogenesis, 71% were predicted to be ischemic. Thus, this study identified a common molecular response to ischemia in TIA and stroke.\textsuperscript{20} We have used a similar approach to show that genes associated with white matter hyperintensities in brain are not associated with those found in ischemic stroke or TIAs.\textsuperscript{21}

Despite the successes, there remain several hurdles to translate these proof-of-concept studies to practice. For example, for prospective prediction problems with data normalization need to be solved to eliminate batch effects that occur over time. If accomplished, large studies will be needed to confirm the results.

**Acknowledgments**

We thank recent collaborators of this article.

**Sources of Funding**

These studies were supported by grants from the National Institutes of Health and the American Heart Association.
Disclosures

None.

References


Key Words: arrays ■ cryptogenic ■ gene expression ■ leukocytes ■ stroke

by guest on June 1, 2017 http://stroke.ahajournals.org/ Downloaded from
Whole Genome Expression of Cellular Response to Stroke
Frank R. Sharp and Glen C. Jickling

doi: 10.1161/STROKEAHA.112.679357

The online version of this article, along with updated information and services, is located on the World Wide Web at:

http://stroke.ahajournals.org/content/44/6_suppl_1/S23