RESEARCH ARTICLE

$CD15^+$ granulocyte and $CD8^+$ T lymphocyte based gene expression clusters for ischemic stroke detection

Mateusz G. Adamski, MD, PhD[1, 2], Yan Li[1], Erin Wagner, MPA–C[1], Chloe Seales-Bailey, RPA-C[1], Nikia Bennett, MD[1], Hua Yu, MD[1], Michael Murphy, PhD[3], Steven A. Soper, PhD[4], Alison E. Baird, FRACP, PhD[1]

Affiliations:

[1] Department of Neurology, SUNY Downstate Medical Center, Brooklyn, NY, USA;
[2] Jagiellonian Centre for Experimental Therapeutics (JCET), Jagiellonian University, Krakow, Poland
[3] Louisiana State University, Baton Rouge, LA;
[4] University of Kansas, KA

*** Corresponding Author:** Alison E. Baird, Department of Neurology, SUNY Downstate Medical Center, 450 Clarkson Ave, Box 1274, Brooklyn, NY 11203, USA, Email: alison.baird@downstate.edu; Tel: 718/221-5188, Fax: 718/221-5761

Abstract

Using whole blood and peripheral blood mononuclear cells, microarray-derived gene expression profiles have shown promise for the detection of acute ischemic stroke. Circulating leukocytes contain multiple cellular subsets of highly specific functions that may provide more powerful and more specific stroke detection than whole blood based expression profiles. The objectives of this study were to determine the cellular sources of gene expression changes in whole blood in ischemic stroke and the utility of leukocyte subset profiles for stroke detection. Using high-throughput reverse transcription real time PCR, the absolute expression of 41 stroke-related transcripts identified in whole blood and peripheral blood mononuclear cells was quantified in 6 major leukocyte subsets – $CD15^+$ granulocytes, $CD14^+$ monocytes, $CD4^+$ T lymphocytes, $CD8^+$ T lymphocytes, $\gamma\delta TCR^+$ cells and $CD20^+$ B lymphocytes. Hierarchical cluster analyses were used to identify clusters of cell subset specific gene expression patterns for ischemic stroke detection. $CD15^+$ granulocytes and $CD8^+$ T lymphocytes were found to be the major sources of the expression changes in ischemic stroke, with 14 and 16 genes up-regulated respectively. None of the genes were significantly altered in $CD14^+$ monocytes or $CD20^+$ B lymphocytes. Multiple clusters of transcripts were identified that discriminated between ischemic stroke and control, most notably in $CD15^+$ granulocytes (p=2.88e-5) and $CD8^+$T lymphocytes (p=1.71e-5). A $CD15^+$ granulocyte-derived 3 gene cluster (*CA4*, *MMP9*, *NAIP*) showed high accuracy for ischemic stroke detection (AUC=0.813) and was 100% sensitive in a validation cohort. We conclude that transcripts identified in microarray studies in circulating leukocytes in stroke are predominantly expressed in $CD15^+$ granulocytes and $CD8^+$ T lymphocytes. Leukocyte subset specific gene expression clusters show promise for ischemic stroke detection.

Word count: 4421

Key words: qPCR, gene expression, transcript, ischemic stroke, granulocyte, lymphocyte

Running title: Adamski: Cellular sources of gene expression in stroke

Non-standard Abbreviations:

HT RT-qPCR	High throughput reverse transcription quantitative polymerase chain reaction
IS	Ischemic stroke
PBMC	Peripheral blood mononuclear cells
WB	Whole blood

1. Introduction

Stroke is a leading cause of death and disability in the United States and there is a pressing need for new diagnostic methods. Using whole blood (WB) and peripheral blood mononuclear cells (PBMC) microarray-derived gene expression profiles have proven to be accurate for the detection of acute ischemic stroke (IS).[1-3] These microarray results have recently been quantitatively validated in WB samples using high throughput quantitative real time polymerase chain reactions (HT RT-qPCR) and a number of WB-based expression clusters for IS detection were identified[4] and point of care technologies are under active development.

Circulating leukocytes comprise multiple cell subsets of highly specific and often opposing functions. Hence leukocyte subset specific gene expression changes may provide more powerful and specific stroke detection that may be masked in WB and PBMC based expression profiles. To date, leukocyte subset specific gene expression profiles have not been reported for ischemic stroke. As the innate immune response is known to be the predominant system involved in the acute phase of ischemic stroke we hypothesized that the cells of the nonspecific innate immune response - neutrophils and monocytes - would be the predominant leukocyte subsets showing altered gene expression.[5-7] We studied 6 major leukocyte subsets and sought to determine the utility of leukocyte subset profiles for stroke detection.

2. Methods

Where applicable, the conduct and reporting of the study are in accordance with the Minimum Information for Publication of Quantitative Real-Time PCR Experiments criteria.[8]

2.1. Study Subjects

Peripheral blood samples were obtained from 25 ischemic stroke patients (including 7 IS patients who comprised the validation cohort) admitted to the University Hospital of Brooklyn at SUNY Downstate Medical Center and at Long Island College Hospital and 15 gender and race matched control subjects recruited from the local community. The subjects' characteristics are provided in **Table I**. The median (interquartile range) time of blood draw was 36.0 (23.0, 48.0) hours post stroke onset. Stroke was diagnosed according to World Health Organization stroke criteria.[9] The Institutional Review Board at the State University of New York (SUNY) approved the study and all study participants or their authorized representatives gave full and signed informed consent.

The study inclusion criteria were: over 18 years of age and acute ischemic stroke. The exclusion criteria were: current immunological diseases, taking steroid or

immunosuppressive therapies, severe allergies, acute infection and severe anemia. The following clinical data were recorded: age, gender, self-reported race/ethnicity, self–reported risk factors, medications at the time of stroke onset or, in the control subjects, medications at the time of blood draw, National Institutes of Health Stroke Scale (NIHSS) score in the stroke subjects and complete blood counts (CBC), including total white blood cell count and white cell differential counts. Hypertension was defined as a prior (at any time in the past) diagnosis of hypertension by the subject's physician or currently receiving treatment for hypertension. Diabetes was defined as a past medical history of known diabetes mellitus. Coronary artery disease was defined as a physician-diagnosed past history of ischemic heart disease or angina. Hyperlipidemia was defined as a past history of documented elevation in total cholesterol (>200mg/dl). Smoking was defined as current or prior smoking. Atrial fibrillation was defined as a past or current history of physician-diagnosed atrial fibrillation. The ischemic stroke subtypes established according to the TOAST criteria[10] were: lacune 3, cryptogenic 10, cardiac embolic 4, large artery atherosclerosis 7 and unavailable 1.

2.2. Leukocyte Separation

The detailed protocol for leukocyte separation has been published previously.[11] Thirty milliliters (ml) of whole blood (WB) were drawn into ethylenediaminetetraacetic acid tubes and a complete blood count (CBC) with differential was performed on each subject. Density gradient centrifugation with Histopaque 1077 and 1119 (Sigma Aldrich) were used to separate peripheral blood mononuclear cells (PBMC) and granulocytes from the WB. Cells were separated using positive magnetic bead separation (Miltenyi Biotec). The WB granulocyte fraction was purified to $CD15^+$ cells. PBMC's were used to subsequently isolate the positive fractions of the $CD14^+$, $CD4^+$, $\gamma\delta TCR^+$, $CD20^+$ and $CD8^+$ leukocyte subsets. The purity of the separated cell populations was assessed using a fluorescence activated cell analyzer (FACA, Epics XL flow cytometer, Beckman Coulter) and analyzed using "FlowJo" software (ver. 9.5.2). Up to 2 million cells were used for RNA extraction from granulocytes ($CD15^+$), monocytes ($CD14^+$), T lymphocytes ($CD4^+$), $\gamma\delta TCR^+$ T lymphocytes, B lymphocytes ($CD20^+$), and T lymphocytes ($CD8^+$). The cellular viability and number were assessed using a hemocytometer and Trypan-blue dye staining.

2.3. Primer Selection and Development

40 transcripts identified in 3 previously published studies were selected for analysis (**Supplemental Table I**).[1-3] The 3 studies had identified 9, 18 and 22 genes within panels with some overlap among the studies. Hox 1.11, transcript identified in Tang's et al study,[2] was not studied because it is a non-coding RNA sequence. Hypothetical protein FLJ22662 Laminin A motif from the Moore et al list[4] is now termed phospholipase B domain containing 1 (*PLBD1*) according to current nomenclature. Two variants of *CD14* were studied to give a total of 41 transcripts that were tested. The complete primer characteristics were published earlier.[11] The RT-qPCR primers were self-designed, commercially synthesized by Invitrogen and wet tested using regular RT-qPCR (StepOnePlus Real-Time PCR Systems; Applied Biosystems).

2.4. HT RT-qPCR

cDNA was synthetized using the High Capacity cDNA Reverse Transcription Kit (Life Technologies, Carlsbad, CA), based on random hexamers, according to the manufacturer's protocol. In addition to study samples two commercial cDNA samples (Universal cDNA Reverse Transcribed by

Random Hexamer: Human Normal Tissues; Biochain, Newark, CA) were run on each plate to perform normalization. HT RT-qPCR was run on the BioMark HD System, using 96 × 96 Fluidigm Dynamic Arrays (Fluidigm, South San Francisco, CA). Four plates were used for this study. The percent present calls were over 90%. Detailed methods have been published previously.[11]

Table I. Clinical and laboratory characteristics of IS patients and controls

Factor	**All (n=40)**	**Stroke Primary Cohort (n=18)**	**Control (n=15)**	**Stroke Validation Cohort (n=7)**	**p**
Age	68.5±13.7	71.6±13.0	58.1±12.3	73.7±8.2	0.001
Gender– male	17 (42)	7 (39)	7 (47)	3 (43)	0.93
Race– black	36 (90)	16 (89)	13 (87)	7 (100)	1.0
Risk factors					
Hypertension	35 (87)	17 (94)	11 (73)	7 (100)	0.11
Diabetes	16 (40)	8 (39)	7 (53)	1 (14)	0.74
Coronary artery disease	12 (30)	5 (28)	3 (20)	4 (57)	0.48
Smoking history	8 (20)	6 (33)	2 (13)	0 (0)	0.68
Atrial fibrillation	5 (12)	4 (22)	0 (0)	1 (14)	0.17
Hyperlipidemia	19 (47)	8 (44)	8 (53)	3 (43)	0.81
Medications					
Diuretics	10 (25)	6 (33)	2 (13)	2 (28)	0.34
ACEIs/ARBs	15 (37)	7 (39)	3 (20)	5 (71)	0.15
Beta blockers	24 (60)	14 (78)	7 (47)	3 (43)	0.32
Calcium channel blockers	10 (25)	5 (28)	3 (20)	2 (28)	0.85
Antithrombotics	23 (57)	10 (55)	8 (53)	5 (71)	0.93
Statins	17 (42)	7 (39)	7 (47)	3 (43)	0.93
Stroke-Related					
NIHSS score	7.0 (5.0, 10.0)	7.5 (4.2, 10.0)		6.0 (5.0, 7.5)	

Results are mean (standard deviation) or median (interquartile range) for continuous factors and numbers (%) for categorical factors. ACEI – angiotensin converting enzyme inhibitor, ARB - angiotensin receptor blocker, N/A – not applicable, NIHSS – National Institutes of Health Stroke Scale. Student's t-tests were used to compare the ages between the stroke patients and the control subjects. Chi–square and Fishers' exact tests were used to compare the remaining demographic, risk factors and medications between the stroke patients and the control subjects. P values are for comparisons of all stroke patients versus the control subjects.

2.5. Gene Expression Data Analyses

The absolute gene expression for each sample was measured using the input sample quantity method[12] – that is independent of control genes - after adjusting for the input cell count and normalizing to a standard volume of a standard cDNA sample

(Universal cDNA Reverse Transcribed by Random Hexamer: Human Normal Tissues; Biochain, Newark, CA). A measure of the copy number per cell for each sample was obtained according to the equation:

$$X_C = \frac{(1+E)^{(nCq,cDNA - nCq,X)}}{ccX}$$

where X_c is the transcript number per cell, E is the efficiency of target cDNA amplification, *nCq,cDNA* and *nCq,X* are the cycle number at which amplification crosses the threshold respectively for the standard cDNA sample and for sample X, *cc* is the number of cells used for RNA extraction based on CBC result. The results for the stroke patients and control subjects were then compared.

2.6. Development and Validation of the Gene Classifier

Based on gene expression results from 18 IS patients and 15 referent control subjects and employing hierarchical cluster analysis, clusters of transcripts were identified for each studied cell subset. Stroke classification accuracy was determined for a selected and highly significant cluster using the Area Under Curve (AUC) test. The diagnostic utility of selected clusters was tested based on a validation cohort (n=7).

2.7. Statistical Analyses

The data were analyzed using R version 2.15.1. Shapiro's tests were used to assess for normality of the data. For grouped data Student's t tests and Wilcoxon rank sum tests were used to compare groups. Chi-square tests and Fisher's exact tests were used to compare categorical values. The hierarchical cluster analyses - non-supervised techniques to detect hidden associations in the data - used Ward's method and log-transformed data. Corrections for multiple comparisons used the Benjamini and Hochberg (false discovery rate, [FDR]) and Bonferroni algorithms. Receiver operating curve analysis and sensitivity and specificity analyses were used to test diagnostic value of the 3 transcript cluster. p-values <0.05 were considered statistically significant.

3. Results

3.1. Expression of 41 transcripts in 6 leukocyte subsets

Individual genes were significantly up-regulated in IS patients in 4 leukocyte populations: $CD15^+$ granulocytes (14 transcripts), $CD8^+$ T lymphocytes (16 transcripts), $\gamma\delta TCR^+$ cells (2 transcripts) and $CD4^+$ T lymphocytes (1 transcript). None of the genes were significantly altered in $CD14^+$ monocytes or $CD20^+$ B lymphocytes. **Figure 1** demonstrates the stroke-related fold changes for the 41 transcripts in the 6 cellular subsets. It also reveals that upregulation of transcripts predominated in $CD8^+$ T and $CD15^+$ cells and that individual transcripts were differentially expressed among the different leukocyte subsets.

3.2. Patterns of altered gene expression in leukocyte subsets

Genes significantly altered in IS patients came from both WB and PBMC-derived transcript panels. The expression of *NAIP*, *MMP9*, *ADM*, *LTA4H*, *PYGL*, *FCGR1A and IL13RA1* was exclusively upregulated in $CD15^+$ granulocytes. The expression of *ENTPD1*, *PILRA*, *PLBD1*, *F5*, *NPL and FPR1* was specifically altered in $CD8^+$ T cells. The expression of 6 genes was altered both in granulocytes and $CD8^+$ T lymphocytes. *VCAN* was significantly up-regulated in $CD8^+$ and $CD4^+$T lymphocytes and was down-regulated in $CD15^+$ granulocytes (**Table II**).

Figure 1. Hierarchical cluster analysis and heatmap of fold changes in expression of 41 genes, in 6 leukocyte subsets, between IS (n=18) and control subjects (n=15).

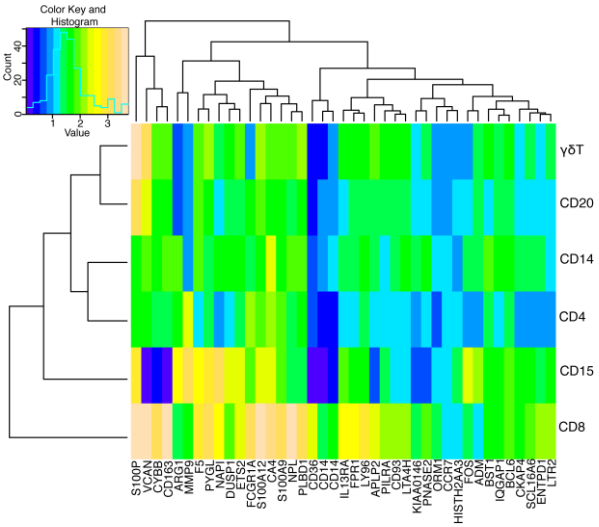


Table II. Patterns of Altered Gene Expression

Pattern	Transcripts	Cellular sources original study
Elevated in CD15 cells	*NAIP* *MMP9* *ADM* *LTA4H* *PYGL* *FCGR1A* *IL13RA1*	PBMC WB PBMC PBMC WB PBMC PBMC
Elevated in both CD15 and CD8 cells	*FOS* *LY96* *S100A9* *S100A12* *CA4* *ETS2*	PBMC PBMC WB WB WB WB and PBMC
Elevated in CD15 cells and/or >1 T cell subset (CD8 and γδT)	*S100P* *DUSP1*	WB PBMC
Elevated in CD8 cells	*ENTPD1* *PILRA* *PLBD1* *F5* *NPL* *FPR1*	PBMC PBMC PBMC WB PBMC WB
Elevated in CD8 and CD4 (decreased in CD15) cells	*VCAN*	WB and PBMC

3.3. Clusters of altered gene expression in leukocyte subsets

In hierarchical cluster analyses 43 clusters of transcripts were identified, specific to the six subsets studied, that significantly discriminated between stroke and control. All clusters contained transcripts from both WB and PBMC panels. After correction for multiple comparisons 14 clusters from 5 cell subsets (CD15+ granulocytes, CD8+ T lymphocytes, CD14+ monocytes, γδTCR+ cells and CD4+ T lymphocytes) remained significant, p values ranged between 9.87e-3 and 1.71e-5 **(Table III)**.

Table III. Gene expression clusters significantly characteristic for IS identified in hierarchical cluster analyses in 4 leukocyte subsets

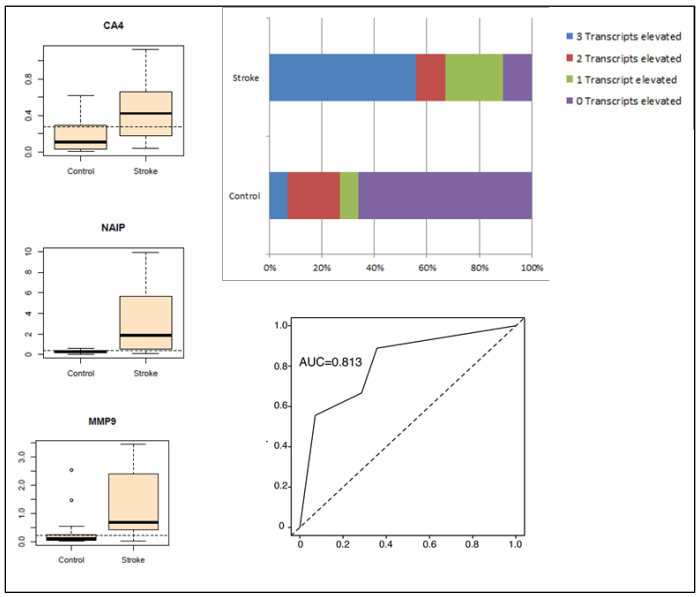
Transcripts	P value, of cluster, stroke versus control	Adjusted p value*	Adjusted p value**
CD15- Cluster 1			
IQGAP1, SLC16A6, NPL, CD93, PYGL, PLBD1	9.4e-6	8.84e-5	4.41e-4
CD15- Cluster 2			
ADM, CKAP4, FOS, BST1	2.94e-5	1.73e-4	1.38e-3
CD15- Cluster 3			
ENTPD1, IL13RA1, LTA4H, S100P	9.7e-5	3.80e-4	4.56e-3
CD15- Cluster 5			
DUSP1, HIST2H2AA3, BCL6, PILRA, FCGR1A, TLR2	7.70e-5	3.29e-4	3.62e-3
CD15- Cluster 7			
LY96, S100A9, FPR1, S100A12, RNASE2, CCR7	0.0012	3.01e-3	5.73e-3
CD15- Cluster 8			
CA4, MMP9, NAIP	6.14e-7	9.62e-6	2.88e-5
CD14- Cluster 4			
PLBD1, BST1, LTA4H, CYBB, SCL16, BCL6, VCAN, FCGR1A	0.00019	6.38e-4	8.93e-3
CD4- Cluster 3			
IQGAP1, NPL, FOS, PLBD1, BST1, VCAN	0.000146	5.28e-4	6.86e-3
CD8- Cluster 1			
IL13, APLP2, ENTPD1, ETS2, PYGL, DUSP1, KIAA, ADM, S100P, CD36,	3.64e-7	8.55e-6	1.71e-5
CD8- Cluster 3			
CYBB, BST1, CD93, NPL, IQGAP1	0.00021	6.58e-4	9.87e-3
CD8- Cluster 4			
FOS, VCAN, PLBD1, MMP9, CA4	1.42e-5	1.11e-4	6.67e-4
CD8- Cluster 5			
BCL6, SLC16, LTA4H, CKAP4, FPR1, FCGR1A	2.58e-5	1.73e-4	1.21e-3
γδT- Cluster 1			
IQGAP1, NPL, FOS, DUSP1, CD93, CKAP4, PLBD1, BST1, VCAN	7.52e-6	8.84e-5	3.53e-4
γδT- Cluster 4			
ETS2, IL13, ENTPD1, PYGL, ADM, KIAA, APLP2, MMP9, CA4	5.19e-5	2.44e-4	2.44e-3

Wilcoxon rank sum tests used for analyses, *-FDR, **-Bonferroni; CD15- granulocytes; CD14- monocytes; CD4, CD8, γδT- respectively $CD4^+$, $CD8^+$, γδTCR T lymphocytes

3.4 . Performance and validation of 3-gene expression cluster

The diagnostic utility of a 3-gene cluster (*CA4*, *MMP9*, *NAIP*) from $CD15^+$ granulocytes was assessed in the primary cohort (PC) of 18 IS patients and 15 control subjects and validated in an independent cohort of 7 IS patients. The upper threshold levels for each of the transcripts were based on the third quartile in the control subjects (**Figure 2a**). Absent calls were noted in the PC for *NAIP* in 7 controls and 6 stroke subjects; this could reflect low or absent transcript expression. The proportions of patients in the PC with elevated expression of between 0 and 3 transcripts are shown in **Figure 2b**. In the PC elevated $CD15^+$ granulocyte expression of at least one transcript in this 3-gene cluster classified stroke with a sensitivity of 89% and a specificity of 67%. The overall accuracy of the 3-gene cluster for PC was high (AUC=0.813, **Figure 2c**). The number of subjects with elevated expression of each transcript in the PC and validation cohort is shown in **Table IV**. The three-gene cluster accurately classified all 7 IS stroke patients in the validation cohort.

Figure 2. Characteristics of a 3-transcript classifier for ischemic stroke detection



(A) Boxplots demonstrating the threshold values for defining elevated expression of each of the transcripts (*CA4*, *NAIP*, *MMP9*). The threshold was set at above the third quartile value in the control group (dashed line on each boxplot). The threshold value was the normalized transcript copy number. (B) Bar graphs depicting the number of transcripts elevated in the stroke patients and the control subjects. In the stroke and control bars the value for the 3 transcript elevation represents the proportions of subjects who had all 3 transcripts elevated, the value for 2 transcripts represents the proportions of subjects who had 2 transcripts elevated, the value for 1 transcript elevated represents the proportion of subjects who had 1 transcript elevated, and the value for 0 transcripts represents the proportions of subjects who had 0 transcripts elevated. In this cluster 89% (16/18) of the stroke patients had 1 or more of the 3 transcripts elevated while 33% (5/15) of the control group showed elevation of 1 or more of the 3 transcripts. Hence the sensitivity was 89% and the specificity was 67%. (C) ROC analysis for Cluster 1 for stroke classification revealed that the AUC was 0.813. Elevation of 3 or more transcripts gave the greatest sensitivity and specificity**.**

Table IV. Performance of a 3-gene cluster in CD15$^+$ granulocyte in primary and validation cohorts

Transcript	Threshold	Primary cohort IS*	Validation cohort IS*	Control subjects*
CA4	>0.273	11/18 (61%)	3/7(43%)	4/15 (27%)
MMP9	>0.229	16/18 (89%)	7/7 (100%)	4/15 (27%)
NAIP	>0.386	10/12 (84%)	5/7 (71%)	2/8 (25%)
3 transcripts in Cluster 8	1 or more transcripts elevated	16/18 (89%)	7/7 (100%)	5/15 (33%)

*- indicates number of subjects with elevated transcript copy number

4. Discussion

In this study, we found that gene expression changes in circulating leukocytes in acute IS are leukocyte subset specific, with CD15$^+$ granulocytes and CD8$^+$ T lymphocytes being the major sources of gene expression changes. Contrary to our hypothesis we did not find significantly altered gene expression in monocytes. Leukocyte-subset-specific gene expression clusters for IS detection were also identified and subset-specific transcript clusters showed diagnostic utility.

4.1. Cell subset specific gene expression in IS

The diversity of various immune cell functions has recently been documented in clinical and experimental IS studies.[13] These results highlighted that the immune response is not only specific to cell subset but may change radically within a given cell subset; for example, T regulatory lymphocytes in the acute phase are strong mediators of IS while in the delayed phase are key cerebroprotective modulators of post IS damage.[13]

Hence, the importance of leukocyte subset specific gene expression changes in the clinical setting is increasingly being recognized, but until now has only been studied indirectly; suggesting that neutrophils, monocytes and platelets were the predominant cell types expressing stroke-related genes in control subjects.[14] Selected pro-inflammatory, pro-apoptotic and adhesion genes have been studied in monocytes, macrophages, and T and B lymphocytes.[15] In contrast to our hypothesis, in the current study CD15$^+$ granulocytes and CD8$^+$ T lymphocytes (cytotoxic cells from the adaptive immune system) were found to play a major role.

Neutrophils are essential players of the innate immune system. Secondary to local and systemic mechanisms they infiltrate the ischemic area causing brain tissue damage mediated by MMP-9 - an enzyme involved in extracellular matrix degradation.[16] Yilmaz et al showed in experimental stroke a pivotal role of CD8$^+$ T-lymphocytes (along with CD4$^+$ T lymphocytes) in the cerebral microvascular dysfunction, inflammation, and tissue injury associated with ischemia/reperfusion injury.[17] At the same time, in a clinical study significant changes in the cytotoxic function of CD8$^+$T lymphocytes in IS have been found.[18] In aging, immunosurveillant CD8 T cells are primed to potentiate inflammation and leukocyte recruitment following ischemic injury.[19]

4.2. Clinical application of a 3-gene cluster

The 3-gene granulocyte cluster (*CA4*, *MMP9* and *NAIP*), having high sensitivity for IS

detection and a low gene number, was the most promising for potential clinical application. This cluster also has the potential to be applied to a point-of-care testing (POCT) technology that is in development.[20] The 3-gene cluster uniquely addresses the major requirements of POCT technology: employs an easily accessible blood sample, shows high accuracy for IS detection and is based on quantitative gene expression changes. The 3-gene cluster, once applied to the POCT technology employing Single-Pair Fluorescence Resonance Energy Transfer could allow for rapid turn-around-time, and hence clinical diagnosis at the bed side.[20]

Sources of Funding

This work was supported by project funding from the NIBIB at the NIH to A.E.B., M.M. and S.A.S. (R01 EB010087).

Disclosures

None

References

1. Moore DF, Li H, Jeffries N, Wright V, Cooper RA Jr, Elkahloun A, Gelderman MP, Zudaire E, Blevins G, Yu H, Goldin E, Baird AE. Using peripheral blood mononuclear cells to determine a gene expression profile of acute ischemic stroke: a pilot investigation. *Circulation*. 2005;111(2):212–21. doi:10.1161/01.CIR.0000152105.79665.C6.

2. Tang Y, Xu H, Du X, et al. Gene expression in blood changes rapidly in neutrophils and monocytes after ischemic stroke in humans: a microarray study. *J Cereb Blood Flow Metab*. 2006;26(8):1089–102. doi:10.1038/sj.jcbfm.9600264.

3. Barr TL, Conley Y, Ding J, Dillman A, Warach S, Singleton A, Matarin M. Genomic biomarkers and cellular pathways of ischemic stroke by RNA gene expression profiling. *Neurology*. 2010;75(11):1009–14. doi:10.1212/WNL.0b013e3181f2b37f.

4. Adamski MG, Li Y, Wagner E, Yu H, Seales-Bailey C, Soper SA, Murphy M, Baird AE. Expression profile based gene clusters for ischemic stroke detection. *Genomics*. 2014;104(3):163–9. doi:10.1016/j.ygeno.2014.08.004.

5. Dirnagl U, Iadecola C, Moskowitz MA. Pathobiology of ischaemic stroke: an integrated view. Trends Neurosci. 1999 Sep;22(9):391-7. PMID: 10441299.

6. Barone FC, Hillegass LM, Price WJ, White RF, Lee EV, Feuerstein GZ, Sarau HM, Clark RK, Griswold DE. Polymorphonuclear leukocyte infiltration into cerebral focal ischemic tissue: myeloperoxidase activity assay and histologic verification. J Neurosci Res. 1991 Jul;29(3):336-45. PMID: 1656059.

7. Herz J, Sabellek P, Lane TE, Gunzer M, Hermann DM, Doeppner TR. Role of Neutrophils in Exacerbation of Brain Injury After Focal Cerebral Ischemia in Hyperlipidemic Mice. Stroke. 2015 Oct;46(10):2916-25. doi: 10.1161/STROKEAHA.115.010620. PMID:2633796.

8. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem* [Internet]. 2009 [cited 2012 Oct 26];55:611–22.

Available from: http://www.ncbi.nlm.nih.gov/pubmed/19246619.

9. Hatano S. Experience from a multicentre stroke register: a preliminary report. Bulletin of the World Health Organisation. 1976;54(5):541–553.

10. Adams HP, Bendixen BH, Kappelle LJ, Biller J, Love BB, Gordon DL, Marsh EE. Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in Acute Stroke Treatment. *Stroke* [Internet]. 1993;24:35–41. Available from: http://www.ncbi.nlm.nih.gov/pubmed/7678184.

11. Adamski MG, Li Y, Wagner E, Yu H, Seales-Bailey C, Soper SA, Murphy M, Baird AE. Next-generation qPCR for the high-throughput measurement of gene expression in multiple leukocyte subsets. *J Biomol Screen* [Internet]. 2013 [cited 2014 Feb 26];18:1008–17. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23690294.

12. Adamski MG, Gumann P, Baird AE. A Method for Quantitative Analysis of Standard and High-Throughput qPCR Expression Data Based on Input Sample Quantity. *PLoS One* [Internet]. 2014 [cited 2014 Aug 5];9:e103917. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25090612.

13. Kleinschnitz C, Kraft P, Dreykluft A, et al. Regulatory T cells are strong promoters of acute ischemic stroke in mice by inducing dysfunction of the cerebral microvasculature. *Blood*. 2013;121(4):679–91. doi:10.1182/blood-2012-04-426734.

14. Du X, Tang Y, Xu H, . Genomic profiles for human peripheral blood T cells, B cells, natural killer cells, monLit L, Walker W, Ashwood P, Gregg JP, Sharp FRocytes, and polymorphonuclear cells: comparisons to ischemic stroke, migraine, and Tourette syndrome. *Genomics*. 2006;87(6):693–703. doi:10.1016/j.ygeno.2006.02.003.

15. Kassner SS, Kollmar R, Bonaterra GA, Hildebrandt W, Schwab S, Kinscherf R. The early immunological response to acute ischemic stroke: differential gene expression in subpopulations of mononuclear cells. *Neuroscience*. 2009;160(2):394–401. doi:10.1016/j.neuroscience.2009.02.050.

16. Justicia C, Panés J, Solé S, Cervera A, Deulofeu R, Chamorro A, Planas AM. Neutrophil infiltration increases matrix metalloproteinase-9 in the ischemic brain after occlusion/reperfusion of the middle cerebral artery in rats. *J Cereb Blood Flow Metab*. 2003;23(12):1430–40. doi:10.1097/01.WCB.0000090680.07515.C8.

17. Yilmaz G, Arumugam T V, Stokes KY, Granger DN. Role of T lymphocytes and interferon-gamma in ischemic stroke. *Circulation.* 2006; 113(17):2105–12.

18. Li G, Wang X, Huang LWang Y, Hao JJ, Ge X, Xu XY, Cytotoxic function of CD8+ T lymphocytes isolated from patients with acute severe cerebral infarction: an assessment of stroke-induced immunosuppression. *BMC Immunol*. 2013;14:1. doi:10.1186/1471-2172-14-1.

19. Ritzel RM, Crapser J, Patel AR, Verma R, Grenier JM, Chauhan A, Jellison ER,

McCullough LD. Age-Associated Resident Memory CD8 T Cells in the Central Nervous System Are Primed To Potentiate Inflammation after Ischemic Brain Injury. J Immunol. 2016 Apr 15;196(8):3318-30. doi: 10.4049/jimmunol.1502021. Epub 2016 Mar 9. PMID:26962232.

20. Peng Z, Young B, Baird AE, Soper SA. Single-pair fluorescence resonance energy transfer analysis of mRNA transcripts for highly sensitive gene expression profiling in near real time. *Anal Chem*. 2013;85(16):7851–8. doi:10.1021/ac400729q.

SUPPLEMENTAL MATERIAL

Supplemental Table I Stroke related transcript panels identified in microarray gene expression studies

Transcript And Study	**Up or down-regulation**		**Transcript And Study**	**Up or down-regulation**
PBMCs			**Whole Blood**	
MOORE et al 2005[1]			**TANG et al, 2006[2]**	
CD163	UP		*Hox 1.11*	UP
PLBD1	UP		*CKAP4*	UP
ADM	UP		*S100A9*	UP
KIAA0146	UP		*MMP9*	UP
APLP2	UP		*S100P*	UP
NPL	UP		*F5-1*	UP
FOS	UP		*FPR1*	UP
TLR2	UP		*S100A12*	UP
NAIP	UP		*RNASE2*	UP
CD36	UP		*ARG1*	DOWN
DUSP1	UP		*CA4*	UP
ENTPD1	UP		*LY96*	UP
VCAN	UP		*SLC16A6*	UP
CYBB	UP		*HIST2H2AA3*	UP
IL13RA1	UP		*ETS2*	UP
LTA4H	UP		*BCL6*	UP
ETS2	UP		*PYGL*	UP
CD14-1	UP		*NPL*	UP
CD14-2	UP			
BST1	UP		**BARR et al, 2010[3]**	
CD93	UP		*ARG1*	DOWN
PILRA	UP		*CA4*	UP
FCGR1A	UP		*CCR7*	UP
			VCAN	UP
			IQGAP1	UP
			LY96	UP
			MMP9	UP
			ORM1	UP
			S100A12	UP

