



Published in final edited form as:

*Stroke*. 2016 December ; 47(12): 2896–2903. doi:10.1161/STROKEAHA.116.013869.

## Altered Expression of Long Noncoding RNAs in Blood Following Ischemic Stroke and Proximity to Putative Stroke Risk Loci

Cheryl Dykstra-Aiello, BS, Glen C Jickling, MD, Bradley P Ander, PhD, Natasha Shroff, BS, Xinhua Zhan, MD/PhD, DaZhi Liu, PhD, Heather Hull, BS, Miles Orantia, BS, Boryana S Stamova, PhD\*, and Frank R Sharp, MD\*

Department of Neurology, University of California at Davis, Sacramento, CA 95817

### Abstract

**Background and Purpose**—Though peripheral blood mRNA and microRNA change following ischemic stroke, any role for long noncoding RNA (lncRNA), which comprise most of the genome and have been implicated in various diseases, is unknown. Thus, we hypothesized that lncRNA expression also changes following stroke.

**Methods**—lncRNA expression was assessed in 266 whole-blood RNA samples drawn once per individual from ischemic stroke patients and matched vascular risk factor controls. Differential lncRNA expression was assessed by Analysis of Covariance (ANCOVA,  $p$ -value  $< 0.005$ ; fold change  $> |1.2|$ ), principal components analysis and hierarchical clustering on a derivation set ( $n=176$ ) and confirmed on a validation set ( $n=90$ ). Post-stroke temporal lncRNA expression changes were assessed using ANCOVA with confounding factor correction ( $p < 0.005$ ; partial correlation with time since event  $> |0.4|$ ). Because sexual dimorphism exists in stroke, analyses were performed for each sex separately.

**Results**—299 lncRNAs were differentially expressed between stroke and control males, whereas 97 lncRNAs were differentially expressed between stroke and control females. Significant changes of lncRNA expression with time after stroke were detected for 49 lncRNAs in males and 31 lncRNAs in females. Some differentially expressed lncRNAs mapped close to genomic locations of previously identified putative stroke-risk genes, including Lipoprotein, Lp(A)-Like 2, ABO blood group, Prostaglandin 12 Synthase, and  $\alpha$ -Adducins.

**Conclusions**—This study provides evidence of altered and sexually dimorphic lncRNA expression in peripheral blood of stroke patients compared to controls and suggests lncRNAs have potential for stroke biomarker development. Some regulated lncRNA could regulate some previously identified putative stroke-risk genes.

Correspondence to: Boryana Stamova, Department of Neurology, UC Davis School of Medicine; MIND Institute Biosciences Building, Rm. 2417, 2805 50<sup>th</sup> Street, Sacramento, CA. 95817 bsstamova@ucdavis.edu Phone: 916-703-0384 FAX: 916-703-0369.  
\*Co-senior authors

**Supplementary Material:** Please see <http://stroke.ahajournals.org>.

**Disclosures:** None

## Keywords

lncRNA; long noncoding RNA; noncoding RNA; stroke; ischemic stroke; gene expression; sexual dimorphism; transcriptional profiling

---

## INTRODUCTION

Previous studies show that coding messenger RNA (mRNA) expression changes in peripheral human blood following stroke and that this expression is sexually dimorphic and temporally dependent<sup>1-9</sup>. Noncoding RNAs (ncRNAs) comprise a majority of the genome and play important developmental and physiological roles, as well as roles in disease mainly by regulating gene expression<sup>10, 11</sup>. This regulation by ncRNAs can be on proximal genes or distant genes via many mechanisms, including modifying chromatin, controlling transcription, degrading mRNA, and splicing<sup>12</sup>. Although both long and short ncRNAs have regulatory roles, small noncoding microRNAs have been most studied in ischemic stroke (IS)<sup>9, 13-15</sup>.

Long noncoding RNAs (lncRNAs) are traditionally defined as having greater than 200 nucleotides but are inconsistently annotated and subclassified by function and/or by genomic location<sup>10, 16-19</sup>. LncRNAs are often cell and tissue specific and can be found between gene coding regions (long intergenic noncoding RNAs (lincRNAs)), overlapping coding genes in either sense or antisense directions, as pseudogenes that have sequence similarity to coding genes, or as mRNAlike lncRNAs which, like mRNAs, can undergo transcription, 5'-capping, splicing and polyadenylation<sup>20-22</sup>.

Several lncRNAs have been shown to play roles in rodent models of stroke<sup>23</sup> and human stroke vascular risk factors like atherosclerosis<sup>24-26</sup>. One group found stroke changes expression of lncRNAs in rat brain and some regulated lncRNAs bind to corepressors of REST, a transcription factor that plays a role in stroke-mediated cell death<sup>23</sup>. Thus, to begin to determine whether lncRNAs play a role in human ischemic stroke, we investigated lncRNA expression in whole blood of ischemic stroke compared to control subjects and examined changes of lncRNA expression over time following ischemic stroke.

## METHODS

### Study subjects

Male and female stroke patients and control subjects were recruited between 2005 and 2013 from medical centers at the Universities of California at Davis and San Francisco and at the University of Alberta, Canada. The protocol was approved by the Institutional Review Boards at all participating Universities and written informed consent was provided by the participants or their proxy. Patient histories, examinations, and evaluations were performed and stroke diagnoses made by board-certified neurologists. Stroke assessments included: electrocardiogram and Holter monitor, carotid Doppler, computed tomography (CT) brain scan and/or magnetic resonance imaging (MRI) and magnetic resonance arteriogram of extra- and intracranial vessels or CT angiography<sup>27</sup>.

## Derivation and validation sets

In order to validate differential lncRNA expression, ischemic stroke and control patients of both sexes were chosen from all recruited patients and assigned randomly to either a derivation set or a validation set. All samples were matched for age, race, sex and vascular risk factors (VRFs), including hypertension, diabetes mellitus and hyperlipidemia.

## Criteria for exclusion included

previous stroke (for control subjects); pre-blood draw treatments with thrombolytics (e.g. rt-PA), or anticoagulants (e.g. heparin or Coumadin); subarachnoid hemorrhage or hemorrhagic infarction; pre- or post-stroke infection; abuse of recreational drugs; dialysis; cancer; blood abnormalities (e.g. anemia), or steroid and other immune-suppressive treatment. The derivation set was used to determine significant differentially expressed lncRNAs, which were then used to filter the validation set for visualization with principal components analysis.

## Blood collection and RNA isolation

PAXgene tubes (Pre-AnalytiX, GmbH, Switzerland), which stabilize blood RNA, were used for venipuncture to obtain whole blood. Blood was collected only once per patient after symptom onset with times varying from a few hours to many days. For control samples, this time point was recorded as zero. Blood draw timing was included as a covariate in the statistical analyses<sup>4, 5, 9, 28-30</sup>. Blood was frozen and stored at  $-80^{\circ}\text{C}$ . RNA was isolated with PAXgene blood RNA isolation kits according to the manufacturer's protocol. Quantification of RNA was determined with a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA) and quality assessment made by an Agilent 2100 Bioanalyzer (Agilent, Palo Alto, CA, USA).

## Array hybridization

A two-cycle target labeling protocol was used for RNA before hybridizing cDNA to Affymetrix GeneChip® Human Transcriptome 2.0 arrays (Affymetrix, Santa Clara, CA), which cover > 40,000 noncoding transcripts (Affymetrix GeneChip® Human Transcriptome Array Data Sheet). Arrays were washed and processed on a Fluidics Station 450 and scanned on a GeneChip Scanner 3000 7G (Affymetrix). Raw expression values (probe level data) for each gene were saved in Affymetrix.CEL and Affymetrix.DAT files. Variation in microarray experiments can be affected by such factors as processing batches on different days (technical variation)<sup>31</sup>. In order to decrease these batch effects, each microarray batch contained randomly allocated control and stroke samples, and batch effect was factored into the analysis.

## Data normalization and statistical analysis

Partek Genomics Suite (PGS) software (Partek Inc., St. Louis, MO, USA) version 6.6 was used to import probes. Raw gene expression data was quantile normalized and  $\log_2$ -transformed using Robust Multi-array Average (RMA) with pre-background adjustment for GC content. After import, 925,032 probeset regions (PSRs) were annotated using Affymetrix specified library files HTA-2\_0.r3 version hg19 and then filtered to include only

lncRNA PSRs. To minimize noise, the remaining 103,351 lncRNA PSRs were then filtered to exclude any with maximum  $\log_2$  expression signal less than three. Thus 102,937 and 102,783 lncRNA PSRs corresponding to 22784 and 22783 unique lncRNAs were analyzed in male and female samples, respectively. Analysis of Covariance (ANCOVA), including diagnosis (stroke, control), age, time since stroke, sex, batch scan date, and VRFs (hypertension, diabetes mellitus and hypercholesterolemia) was first performed on both male and female samples. In order to decrease bias related to hormonal differences and dosage effect of sex chromosomes between the sexes due to X-chromosome inactivation escapes in females and because sexual dimorphism exists in stroke<sup>4, 28</sup>, each sex was analyzed separately. The results were then overlapped to identify post-stroke sex-specific and common lncRNA expression. To eliminate those genes with very low expression and reduce noise, differentially expressed lncRNAs were considered significant with absolute fold change (FC) $>1.2$  and  $p<0.005$ <sup>3, 32, 33</sup>. Prior to visualization on principal components analysis (PCA) and hierarchical clustering, effects due to batch scan date and time since stroke were removed from both the derivation and validation sets.

### Changes of lncRNA expression over time

In order to investigate differential lncRNA expression that changes with time, an analysis of covariance (ANCOVA) was performed on all ischemic stroke samples within each sex with factors time since stroke, batch scan date, age and VRFs. A lncRNA list ( $p<0.005$ ) was derived and then filtered on a partial correlation for time since stroke symptom onset ( $r > |0.4|$ )<sup>34</sup>. Partial correlation coefficients,  $r$ , describe the linear relation between two variables (gene expression versus time following stroke onset) while controlling for the effects of one or more additional variables<sup>35</sup>.

### Noncoding RNA

Differentially expressed lncRNAs that were identified using the PGS software were further annotated with the NetAffx™ Analysis Center's Exon/Gene Expression Batch Query tool ([www.Affymetrix.com/analysis](http://www.Affymetrix.com/analysis)). Biological interpretation of the differentially expressed lncRNAs was examined manually by literature review and with the UCSC Genome Browser<sup>36</sup>, NONCODE<sup>37</sup> and OMIM®-Online Mendelian Inheritance in Man® databases<sup>38</sup>.

## RESULTS

### Patient Characteristics

Of the 266 samples (133 IS, 133 controls), 176 were assigned as derivation set samples (male,  $n=88$ ; female,  $n=88$ ) and 90 assigned to the validation set (male,  $n=40$ ; female,  $n=50$ ). Each set was comprised of equal numbers of ischemic stroke and vascular risk factor matched control (VRFC) samples. The majority of the samples were Caucasian. A sub-analysis that included race as a co-factor revealed that race did not contribute a large source of variation. Thus race was not considered in the statistical model of the main analyses. Patient and clinical characteristics are presented in Supplementary Tables I and II. There were no significant statistical differences in patient characteristics (sex, race (%Caucasian), time since event, age, and vascular risk factors) between groups. Average age (in years  $\pm$

standard deviation (s.d.) of the male (n=44) and the female (n=44) stroke derivation sets was  $61.9 \pm 13.4$  and  $65.1 \pm 12.5$ , respectively. Average age of the male (n=44) and female (n=44) control derivation sets were  $58.7 \pm 13.5$  and  $61.67 \pm 11.64$ , respectively. Average time since stroke (in hours  $\pm$  s.d.) of the male and female stroke derivation sets were  $49.0 \pm 28.1$  and  $54.6 \pm 31.1$ , respectively. Time since event ranged from 4.4 to 156.0 hours after symptom onset allowing for investigation of a large range of early to late post-stroke biology.

### Differentially Expressed lncRNAs

There were 404 (0.4%) lncRNA PSRs (Supplementary Table III,  $p < 0.005$ ,  $FC > |1.2|$ ) differentially expressed between stroke and control males, which were associated with 299 (1.3%) differentially expressed lncRNAs (Table 1). Of these lncRNAs, 146 (48.8%) had increased expression levels and 153 (51.2%) had decreased expression levels (Figure 1). Most (n=161, 53.9%) of the differentially expressed lncRNAs were lincRNAs (Table 1).

There were 113 (0.1%) lncRNA PSRs (Supplementary Table III,  $p < 0.005$ ,  $FC > |1.2|$ ) differentially expressed between stroke and control females which were associated with 97 (0.4%) differentially expressed lncRNAs (Table 1). Of these lncRNAs, 76 (78.4%) had increased expression levels and 21 (21.7%) had decreased expression levels (Figure 2) and the majority were lincRNAs (Table 1). There were significantly more lncRNAs with increased expression levels than with decreased expression levels in the female samples ( $p = 0.0001$ ).

Six lncRNAs were common between the sexes. However, two shared decreased expression in both sexes whereas the remaining four showed increased expression in the female samples and decreased expression in the male samples (Supplementary Table IV,  $p < 0.005$  ( $FC > |1.2|$ )).

### Changes of lncRNA Expression Over Time

In the male ischemic stroke samples, there were 299 (0.3%) lncRNA PSRs that significantly changed expression over time ( $p < 0.005$ ) with partial correlation r-values ranging from  $-0.50$  to  $+0.44$ . Using a partial correlation  $r = |0.4|$  as a cut off, 49 lncRNAs remained that changed over time (Table 2, Supplementary Table V;  $p < 0.005$ ,  $r = |0.4|$ ). Of these 49 lncRNAs, three were also differentially expressed between stroke and controls (PSR01064543, PSR04026704 and PSR22016723; Supplementary Table VI). The most positively and negatively correlated lncRNAs with time in ischemic stroke males are presented in Figure 3.

In the female stroke samples, there were 300 (0.3%) lncRNA PSRs that significantly changed expression over time ( $p < 0.005$ ) with partial correlation r-values ranging from  $-0.50$  to  $+0.41$ . Using a partial correlation  $r = |0.4|$  as a cut off, 31 lncRNAs remained that changed over time (Table 2, Supplementary Table V;  $p < 0.005$ ,  $r = |0.4|$ ). Of these 31 lncRNAs, one was also significantly differentially expressed between stroke and controls (PSR05035106; Supplementary Table VI). The most positively and negatively correlated lncRNAs with time in ischemic stroke females were mRNAlike lncRNAs (Figure 3). In the female samples, one of the lncRNAs significant for changes over time, linc-CEP120-1, was also differentially upregulated in stroke vs control females (Supplementary Table VI). In

comparison to ischemic stroke males, ischemic stroke females had significantly more downregulated lncRNAs over time ( $p < 0.001$ ).

### Biological Interpretation – Proximity to Putative Stroke- and VRF-Risk Genes

In the male samples, PSR06037329 is differentially expressed and is annotated to linc-SLC22A2 with genomic coordinates that map to the putative ischemic stroke risk loci, SLC22A3/LPAL2/LPA, at cytoband 6q25.3 (Table 3). Similarly, PSR09027609 is annotated to linc-OBP2B-1, another lncRNA that has genomic coordinates 48.1 kilo base pairs (kb) upstream from rs579459, an ischemic stroke-risk associated SNP at 9q34.2 (Table 3). Among the female differentially expressed lncRNAs, annotation of PSR20015349 identifies OTTHUMT00000079682, an mRNAlike lncRNA that maps to within 132.7 kb of stroke-implicated gene Prostaglandin 12 Synthase (PTGIS) at 20q13.13, while linc-luo-1172 (PSR04025299) maps to within 332.6 kb of another stroke associated gene,  $\alpha$ -Adducins (ADD1), at 4p16.3 (Table 3). Additionally, lncRNAs significant for changes over time were mapped to VRF-associated genes (Table 3). Among the male ischemic stroke patient lncRNAs, ENST00000507441 (PSR19028831) at 19q13.32 maps to within 219.6 kb of Apolipoprotein E (APOE), an atherosclerosis-associated gene and ENST0000079667 (PSR20017082) is within 29.0 kb of Prostaglandin I2 (Prostacyclin) Synthase (PTGIS), associated with essential hypertension. The female ischemic stroke lncRNA, NR\_036641 (PSR04029574) was within the atherosclerosis-associated gene, Platelet Derived Growth Factor C (PDGFC).

## DISCUSSION

This is the first study to demonstrate changes of expression of long noncoding RNA (lncRNA) in blood of humans following ischemic stroke. There were a separate set of lncRNA that change expression over time, and both the expression changes and changes over time are sexually dimorphic. Notably, selected lncRNA were in proximity to or within known or putative ischemic stroke genetic risk loci and could play a role in their regulation. Finally, we suggest lncRNA might now be included with mRNA and miRNA as possible biomarkers **to be explored further** for stroke diagnosis and its causes.

Ischemic stroke can be difficult to differentiate from intracerebral hemorrhage and stroke mimics particularly in acute settings where brain imaging is not feasible or available<sup>39</sup>. Because treatment of ischemic stroke must be done quickly, development of a rapid, inexpensive and accurate blood test would be valuable. Previous work aimed at developing a stroke blood test focused either on proteins or on coding mRNAs<sup>6, 29, 40, 41</sup>. The current findings that lncRNA are differentially expressed in stroke vs controls, and that expression of selected lncRNA change over time following ischemic stroke, could be considered in development of future biomarker panels to differentiate ischemic stroke from intracerebral hemorrhage and stroke mimics.

Long noncoding RNAs have been studied in neurological disorders<sup>42</sup>, cardiovascular disease<sup>19</sup> and rodent stroke models<sup>43</sup>. lncRNA are of interest because they affect mRNA expression in many ways including DNA methylation<sup>44</sup>. One study reported differences in DNA methylation in blood of healthy male and female donors<sup>45</sup>, suggesting sexual

dimorphism in lncRNA expression. Similarly, the current study found that leukocyte lncRNA expression differed between male and female samples supporting previous evidence of sexual dimorphism in post-stroke mRNA expression<sup>4, 9, 28</sup>.

The regulated lncRNA described here have the potential for regulating mRNA expression. Colorectal cancer studies have shown that lncRNA, CCAT1-L, is 515 kb away from the MYC gene it transcriptionally regulates<sup>46</sup>. Therefore, it is plausible that lncRNAs may regulate ischemic stroke genes from similar distances. Indeed, albeit different for each sex, significant differentially upregulated lncRNAs were discovered that mapped to genomic locations previously suggested to be associated with ischemic stroke risk (Table 3).

In male samples, PSR06037329 was differentially expressed between stroke and controls and mapped to the Lipoprotein, Lp(A)-Like 2, Pseudogene (LPAL2). LPAL2 is part of the risk loci, SLC22A3/LPAL2/LPA, that has been suggested to be associated with both ischemic stroke and coronary artery disease<sup>47, 48</sup>. A single nucleotide polymorphism (SNP), rs10455872, in this risk locus, predicts increased plasma levels of lipoprotein(a) (LPA) which is correlated with increased risk of stroke<sup>47</sup>. Although the exact function of linc-SLC22A2 remains to be determined, as a pseudogene, it may regulate gene expression in many ways, including coding for truncated proteins or mRNA interference<sup>21, 44</sup>.

Another significant differentially expressed lncRNA in male samples, linc-OBP2B-1, mapped to within 48.1 kb of SNP rs579459 at 9q34.2 that is associated with ischemic stroke<sup>48</sup> and within 44.3 kb of the ABO blood group gene (transferase A, alpha 1-3-N-acetylgalactosaminyltransferase; transferase B, alpha 1-3-galactosyltransferase) (ABO). Polymorphisms in this gene locus are associated with increased risk of ischemic stroke<sup>49</sup>.

In female samples, the PTGIS gene was less than 150 kb away from OTTHUMT00000079682, one of the significant differentially upregulated mRNAlike lncRNAs in ischemic stroke patients. PTGIS catalyzes synthesis of prostacyclin which inhibits aggregation of platelets and vasoconstriction and mutations of this gene are correlated with ischemic stroke<sup>50</sup>. Additionally, a lncRNA with increased expression levels, linc-luo-1172, was within 350 kb of the ADD1 gene. A genetic variant of ADD1 conveyed risk of ischemic stroke particularly in hypertensive patients<sup>51, 52</sup>.

Changes over time in post-stroke RNA expression are important for understanding stroke pathophysiology given the short time window for acute treatment and long time window for recovery. Accordingly, and in keeping with previous research that demonstrated changes of mRNA expression over time following ischemic stroke<sup>4, 5, 28, 29</sup>, this study showed that lncRNA expression changes over time after ischemic stroke and some of these lncRNA are in the proximity of putative vascular risk factor-associated genes<sup>53-55</sup>.

Functional analysis of lncRNAs with respect to stroke pathophysiology are just beginning. An *in vitro* study of an atherosclerotic pathway identified an upregulated lncRNA NFIA-A1 that was associated with increased microRNA-382-5p expression and a downregulated nuclear factor IA (NFIA), a gene that was both overlapped by the lncRNA and a target of the microRNA<sup>26</sup>. We have shown changes of microRNA in blood following stroke<sup>15</sup>. Future

studies will be needed to determine if microRNA-lncRNA interactions are one of the ways in which gene expression in stroke is also regulated<sup>56</sup>.

### Limitations and Future Directions

Although we report very little lncRNA overlap between the sexes, this is most likely due to some extent to computational stringency ( $p < 0.005$ ) rather than biology because, at  $p < 0.05$  ( $FC > |1.2|$ ), there were 95 differentially expressed lncRNAs common between 3120 differentially expressed lncRNAs in male samples and 1223 in female samples (Supplementary Table VII, hypergeometric  $p = 2.2E-16$ ).

While outside the scope of this study, ischemic stroke subtype should be included in future studies to further refine differential lncRNA expression results. We performed a sub-analysis on all the male ( $n=64$ ) and female ( $n=69$ ) ischemic stroke samples (separately per sex) using subtype (Supplementary Table II) as a diagnosis factor in our model in place of ischemic stroke, and found that each stroke subtype presents with a unique lncRNA expression profile (Venn Diagram, Supplementary Figure I).

In addition, stroke severity needs to be included in analyses when investigating lncRNA expression. Therefore, to determine whether lncRNA expression levels might be correlated to stroke severity, we performed a sub-analysis of lncRNA expression on National Institutes of Health Stroke Scale scores (NIHSS) at hospital admission on a small subset ( $n=22$ , 41% male, Supplementary Table VIII) of ischemic stroke patients with blood draw time since event  $< 24$ h. We found 330 long noncoding PSRs that significantly correlated with NIHSS, of which 13 mapped near to 8 putative stroke risk loci, including the recently reported FOXF2 risk locus<sup>57</sup> (Supplementary Table IX).

This study used samples with a wide range of blood draw times since event which allowed for study of a range of post-stroke biology. However, future studies should refine these time frames in order to better understand how expression of lncRNAs changes over time after stroke and in each sex.

Future validation of individual lncRNAs in a larger cohort is needed.

### Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

### Acknowledgments and Sources of Funding

These studies were supported by an AHA Fellow to Faculty award to GCJ, American Heart Association grants to DL and BS, and NIH/NINDS grants to FRS, GCJ and BS (NS079153, NS075035, NS097000).

### REFERENCES

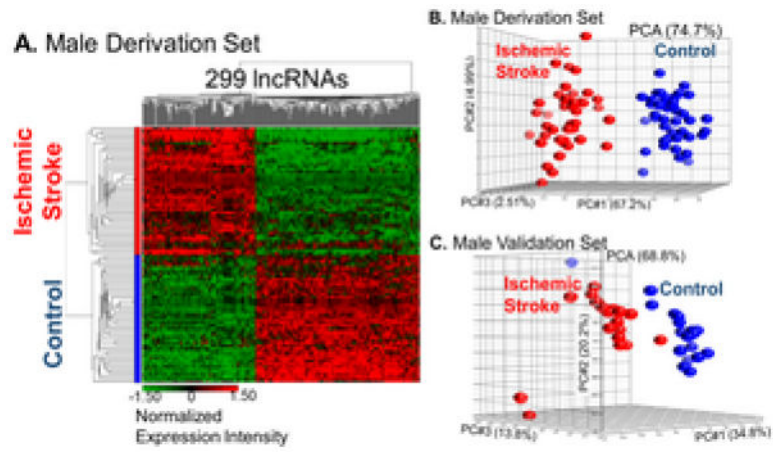
1. Jickling GC, Stamova B, Ander BP, Zhan X, Liu D, Sison SM, et al. Prediction of cardioembolic, arterial, and lacunar causes of cryptogenic stroke by gene expression and infarct location. *Stroke; a journal of cerebral circulation*. 2012; 43:2036–2041. [PubMed: 22627989]
2. Jickling GC, Stamova B, Ander BP, Zhan X, Tian Y, Liu D, et al. Profiles of lacunar and nonlacunar stroke. *Ann Neurol*. 2011; 70:477–485. [PubMed: 21796664]



3. Jickling GC, Xu H, Stamova B, Ander BP, Zhan X, Tian Y, et al. Signatures of cardioembolic and large-vessel ischemic stroke. *Annals of neurology*. 2010; 68:681–692. [PubMed: 21031583]
4. Stamova B, Tian Y, Jickling G, Bushnell C, Zhan X, Liu D, et al. The x-chromosome has a different pattern of gene expression in women compared with men with ischemic stroke. *Stroke; a journal of cerebral circulation*. 2012; 43:326–334. [PubMed: 22052522]
5. Stamova B, Xu H, Jickling G, Bushnell C, Tian Y, Ander BP, et al. Gene expression profiling of blood for the prediction of ischemic stroke. *Stroke; a journal of cerebral circulation*. 2010; 41:2171–2177. [PubMed: 20798371]
6. Stamova BS, Apperson M, Walker WL, Tian Y, Xu H, Adamczyk P, et al. Identification and validation of suitable endogenous reference genes for gene expression studies in human peripheral blood. *BMC medical genomics*. 2009; 2:49. [PubMed: 19656400]
7. Tian Y, Stamova B, Jickling GC, Xu H, Liu D, Ander BP, et al. Y chromosome gene expression in the blood of male patients with ischemic stroke compared with male controls. *Gender medicine*. 2012; 9:68–75. e63. [PubMed: 22365286]
8. Moore DF, Li H, Jeffries N, Wright V, Cooper RA Jr, Elkahoulou A, et al. Using peripheral blood mononuclear cells to determine a gene expression profile of acute ischemic stroke: A pilot investigation. *Circulation*. 2005; 111:212–221. [PubMed: 15630028]
9. Stamova B, Jickling GC, Ander BP, Zhan X, Liu D, Turner R, et al. Gene expression in peripheral immune cells following cardioembolic stroke is sexually dimorphic. *PLoS One*. 2014; 9:e102550. [PubMed: 25036109]
10. Clark MB, Choudhary A, Smith MA, Taft RJ, Mattick JS. The dark matter rises: The expanding world of regulatory rnas. *Essays in biochemistry*. 2013; 54:1–16. [PubMed: 23829523]
11. Kitagawa M, Kitagawa K, Kotake Y, Niida H, Ohhata T. Cell cycle regulation by long non-coding rnas. *Cellular and molecular life sciences : CMLS*. 2013; 70:4785–4794. [PubMed: 23880895]
12. Amaral PP, Dinger ME, Mercer TR, Mattick JS. The eukaryotic genome as an rna machine. *Science (New York, N.Y.)*. 2008; 319:1787–1789.
13. Dharap A, Bowen K, Place R, Li LC, Vemuganti R. Transient focal ischemia induces extensive temporal changes in rat cerebral micrornaome. *Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism*. 2009; 29:675–687.
14. Saugstad JA. Non-coding rnas in stroke and neuroprotection. *Frontiers in neurology*. 2015; 6:50. [PubMed: 25821444]
15. Jickling GC, Ander BP, Zhan X, Noblett D, Stamova B, Liu D. Microrna expression in peripheral blood cells following acute ischemic stroke and their predicted gene targets. *PloS one*. 2014; 9:e99283. [PubMed: 24911610]
16. Esteller M. Non-coding rnas in human disease. *Nature reviews. Genetics*. 2011; 12:861–874.
17. Taft RJ, Pang KC, Mercer TR, Dinger M, Mattick JS. Non-coding rnas: Regulators of disease. *The Journal of pathology*. 2010; 220:126–139. [PubMed: 19882673]
18. St Laurent G, Wahlestedt C, Kapranov P. The landscape of long noncoding rna classification. *Trends in genetics : TIG*. 2015; 31:239–251. [PubMed: 25869999]
19. Devaux Y, Zangrando J, Schroen B, Creemers EE, Pedrazzini T, Chang CP, et al. Long noncoding rnas in cardiac development and ageing. *Nature reviews. Cardiology*. 2015; 12:415–425. [PubMed: 25855606]
20. Wilusz JE, Sunwoo H, Spector DL. Long noncoding rnas: Functional surprises from the rna world. *Genes & development*. 2009; 23:1494–1504. [PubMed: 19571179]
21. Pink RC, Carter DR. Pseudogenes as regulators of biological function. *Essays in biochemistry*. 2013; 54:103–112. [PubMed: 23829530]
22. Sun M, Kraus WL. From discovery to function: The expanding roles of long noncoding rnas in physiology and disease. *Endocrine reviews*. 2015; 36:25–64. [PubMed: 25426780]
23. Dharap A, Pokrzywa C, Vemuganti R. Increased binding of stroke-induced long non-coding rnas to the transcriptional corepressors sin3a and corest. *ASN neuro*. 2013; 5:283–289. [PubMed: 24063527]
24. Yin KJ, Hamblin M, Chen YE. Non-coding rnas in cerebral endothelial pathophysiology: Emerging roles in stroke. *Neurochemistry international*. 2014; 77:9–16. [PubMed: 24704794]

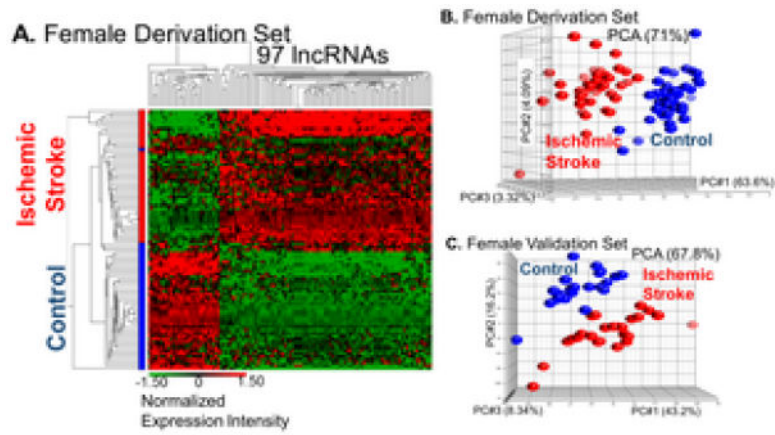
25. Vemuganti R. All's well that transcribes well: Non-coding rnas and post-stroke brain damage. *Neurochemistry international*. 2013; 63:438–449. [PubMed: 23954844]
26. Hu YW, Zhao JY, Li SF, Huang JL, Qiu YR, Ma X, et al. Rpl5-833a20.1/mir-382-5p/nfia-dependent signal transduction pathway contributes to the regulation of cholesterol homeostasis and inflammatory reaction. *Arteriosclerosis, thrombosis, and vascular biology*. 2015; 35:87–101.
27. Yew KS, Cheng E. Acute stroke diagnosis. *Am Fam Physician*. 2009; 80:33–40. [PubMed: 19621844]
28. Tian Y, Stamova B, Jickling GC, Liu D, Ander BP, Bushnell C, et al. Effects of gender on gene expression in the blood of ischemic stroke patients. *Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism*. 2012; 32:780–791.
29. Tang Y, Xu H, Du X, Lit L, Walker W, Lu A, et al. Gene expression in blood changes rapidly in neutrophils and monocytes after ischemic stroke in humans: A microarray study. *Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism*. 2006; 26:1089–1102.
30. Yilmaz G, Granger DN. Leukocyte recruitment and ischemic brain injury. *Neuromolecular medicine*. 2010; 12:193–204. [PubMed: 19579016]
31. Lander ES. Array of hope. *Nature genetics*. 1999; 21:3–4. [PubMed: 9915492]
32. Ander BP, Barger N, Stamova B, Sharp FR, Schumann CM. Atypical mirna expression in temporal cortex associated with dysregulation of immune, cell cycle, and other pathways in autism spectrum disorders. *Molecular autism*. 2015; 6:37. [PubMed: 26146533]
33. Stamova B, Ander BP, Barger N, Sharp FR, Schumann CM. Specific regional and age-related small noncoding rna expression patterns within superior temporal gyrus of typical human brains are less distinct in autism brains. *Journal of child neurology*. 2015; 30:1930–1946. [PubMed: 26350727]
34. Tian Y, Stamova B, Ander BP, Jickling GC, Gunther JR, Corbett BA, et al. Correlations of gene expression with ratings of inattention and hyperactivity/impulsivity in tourette syndrome: A pilot study. *BMC medical genomics*. 2012; 5:49. [PubMed: 23110997]
35. Morrison, DF. *Multivariate statistical methods*. McGraw Hill; New York: 1976.
36. Speir ML, Zweig AS, Rosenbloom KR, Raney BJ, Paten B, Nejad P, et al. The ucsc genome browser database: 2016 update. *Nucleic acids research*. 2015; 44:D717–D725. [PubMed: 26590259]
37. Zhao Y, Yuan J, Chen R. Noncodev4: Annotation of noncoding rnas with emphasis on long noncoding rnas. *Methods in molecular biology (Clifton, N.J.)*. 2016; 1402:243–254.
38. McKusick VA. Mendelian inheritance in man and its online version, omim. *American journal of human genetics*. 2007; 80:588–604. [PubMed: 17357067]
39. Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, Blaha MJ, Cushman M, et al. Heart disease and stroke statistics-2016 update: A report from the american heart association. *Circulation*. 2015; 133:e38–e360. [PubMed: 26673558]
40. Baird AE, Soper SA, Pullagurla SR, Adamski MG. Recent and near-future advances in nucleic acid-based diagnosis of stroke. *Expert review of molecular diagnostics*. 2015; 15:665–679. [PubMed: 25837776]
41. Jickling GC, Sharp FR. Biomarker panels in ischemic stroke. *Stroke; a journal of cerebral circulation*. 2015; 46:915–920. [PubMed: 25657186]
42. Fenoglio C, Ridolfi E, Galimberti D, Scarpini E. An emerging role for long non-coding rna dysregulation in neurological disorders. *International journal of molecular sciences*. 2013; 14:20427–20442. [PubMed: 24129177]
43. Dharap A, Nakka VP, Vemuganti R. Effect of focal ischemia on long noncoding rnas. *Stroke; a journal of cerebral circulation*. 2012; 43:2800–2802. [PubMed: 22949471]
44. Batista PJ, Chang HY. Long noncoding rnas: Cellular address codes in development and disease. *Cell*. 2013; 152:1298–1307. [PubMed: 23498938]
45. El-Maarri O, Becker T, Junen J, Manzoor SS, Diaz-Lacava A, Schwaab R, et al. Gender specific differences in levels of DNA methylation at selected loci from human total blood: A tendency toward higher methylation levels in males. *Human genetics*. 2007; 122:505–514. [PubMed: 17851693]

46. Kim T, Cui R, Jeon YJ, Lee JH, Lee JH, Sim H, et al. Long-range interaction and correlation between myc enhancer and oncogenic long noncoding rna carlo-5. *Proceedings of the National Academy of Sciences of the United States of America*. 2014; 111:4173–4178. [PubMed: 24594601]
47. Ronald J, Rajagopalan R, Cerrato F, Nord AS, Hatsukami T, Kohler T, et al. Genetic variation in lpa2, lpa, and plg predicts plasma lipoprotein(a) level and carotid artery disease risk. *Stroke; a journal of cerebral circulation*. 2011; 42:2–9. [PubMed: 21127300]
48. Dichgans M, Malik R, Konig IR, Rosand J, Clarke R, Gretarsdottir S, et al. Shared genetic susceptibility to ischemic stroke and coronary artery disease: A genome-wide analysis of common variants. *Stroke; a journal of cerebral circulation*. 2014; 45:24–36. [PubMed: 24262325]
49. Williams FM, Carter AM, Hysi PG, Surdulescu G, Hodgkiss D, Soranzo N, et al. Ischemic stroke is associated with the abo locus: The euroclot study. *Annals of neurology*. 2013; 73:16–31. [PubMed: 23381943]
50. Nakayama T. Genetic polymorphisms of prostacyclin synthase gene and cardiovascular disease. *International angiology : a journal of the International Union of Angiology*. 2010; 29:33–42. [PubMed: 20357747]
51. van Rijn MJ, Bos MJ, Yazdanpanah M, Isaacs A, Arias-Vasquez A, Koudstaal PJ, et al. Alpha-adducin polymorphism, atherosclerosis, and cardiovascular and cerebrovascular risk. *Stroke; a journal of cerebral circulation*. 2006; 37:2930–2934. [PubMed: 17082469]
52. Zafarmand MH, van der Schouw YT, Grobbee DE, de Leeuw PW, Bots ML. Alpha-adducin gly460trp variant increases the risk of stroke in hypertensive dutch women. *Hypertension*. 2008; 51:1665–1670. [PubMed: 18458162]
53. Karvinen H, Rutanen J, Leppanen O, Lach R, Levonen AL, Eriksson U, et al. Pdgf-c and -d and their receptors pdgfr-alpha and pdgfr-beta in atherosclerotic human arteries. *European journal of clinical investigation*. 2009; 39:320–327. [PubMed: 19292888]
54. Bis JC, Kavousi M, Franceschini N, Isaacs A, Abecasis GR, Schminke U, et al. Meta-analysis of genome-wide association studies from the charge consortium identifies common variants associated with carotid intima media thickness and plaque. *Nature genetics*. 2011; 43:940–947. [PubMed: 21909108]
55. Nakayama T, Soma M, Haketa A, Aoi N, Kosuge K, Sato M, et al. Haplotype analysis of the prostacyclin synthase gene and essential hypertension. *Hypertension research : official journal of the Japanese Society of Hypertension*. 2003; 26:553–557. [PubMed: 12924623]
56. Yoon JH, Abdelmohsen K, Gorospe M. Functional interactions among micrnas and long noncoding rnas. *Seminars in cell & developmental biology*. 2014; 34:9–14. [PubMed: 24965208]
57. Identification of additional risk loci for stroke and small vessel disease: A meta-analysis of genome-wide association studies. *The Lancet. Neurology*. 2016; 15:695–707. [PubMed: 27068588]

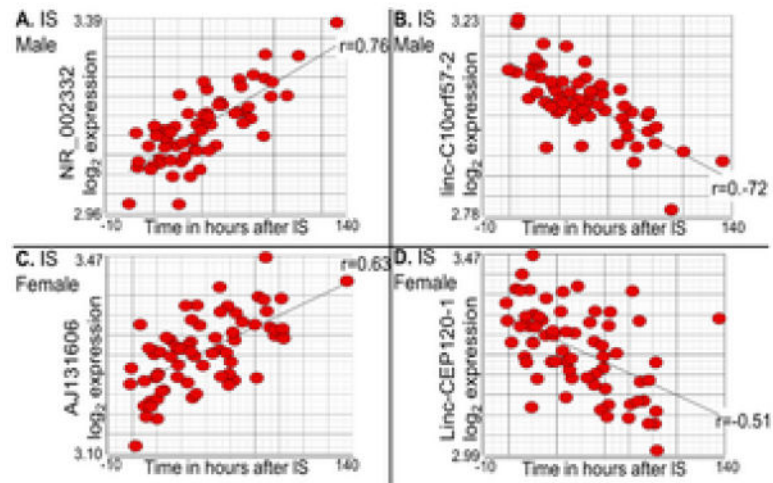


**Figure 1.**

Visualization by A. hierarchical clustering (green=decreased expression levels; red=increased expression levels) and B. PCA of male derivation and C. validation sets after removing batch effects shows separation by diagnosis based on the expression of 299 differentially expressed lncRNAs.



**Figure 2.** Visualization by A. hierarchical clustering (green=decreased expression levels; red=increased expression levels) and B. PCA of female derivation and C. validation sets after removing batch effects shows separation by diagnosis based on the expression of 97 differentially expressed lncRNAs.



**Figure 3.**

The most positive (A. lncRNA NR\_002332 and C. mRNAlike lncRNA AJ131606) and negative (B. linc-C10orf57-2 and D. linc-CEP120-1) time correlated lncRNA probesets in male (A. and B.) and female (C. and D.) whole blood samples. Scatterplots are visualized after removing effects of technical variation. Linear regression lines ( $r$ ) are shown for all correlations.

**Table 1**

Significant ( $p < 0.005$ ,  $|FC| > 1.2$ ) differentially expressed lncRNA transcript types in male and female whole blood

| lncRNA Transcript Type | Male Samples                    |                               |                               | Female Samples                  |                               |                               |
|------------------------|---------------------------------|-------------------------------|-------------------------------|---------------------------------|-------------------------------|-------------------------------|
|                        | Number, n (% total transcripts) | Increased expression level, n | Decreased expression level, n | Number, n (% total transcripts) | Increased expression level, n | Decreased expression level, n |
| Long intergenic        | 161 (53.9)                      | 101                           | 60                            | 46 (47.4)                       | 40                            | 6                             |
| mRNAlike               | 81 (27.1)                       | 24                            | 57                            | 30 (30.1)                       | 24                            | 6                             |
| Antisense              | 8 (2.7)                         | 3                             | 5                             | 2 (2.1)                         | 2                             | 0                             |
| Pseudogene             | 7 (2.3)                         | 3                             | 4                             | 6 (6.2)                         | 2                             | 4                             |
| Other                  | 49 (16.4)                       | 18                            | 31                            | 13 (13.4)                       | 8                             | 5                             |
| <b>TOTAL</b>           | <b>299</b>                      | <b>146 (48.8%)</b>            | <b>153 (51.2%)</b>            | <b>97</b>                       | <b>76 (78.4%)</b>             | <b>21 (21.6%)</b>             |

**Table 2**

LncRNA transcript types significant for changes over time ( $p < 0.005$ ) in whole blood of male and female ischemic stroke patients

| lncRNA Transcript Type | Male Samples                    |                               |                               | Female Samples                  |                               |                               |
|------------------------|---------------------------------|-------------------------------|-------------------------------|---------------------------------|-------------------------------|-------------------------------|
|                        | Number, n (% total transcripts) | Increased expression level, n | Decreased expression level, n | Number, n (% total transcripts) | Increased expression level, n | Decreased expression level, n |
| Long intergenic        | 27 (56.3)                       | 24                            | 3                             | 23 (71.9)                       | 23                            | 0                             |
| mRNAlike               | 13 (27.1)                       | 10                            | 3                             | 4 (12.5)                        | 3                             | 1                             |
| Antisense              | 1 (2.1)                         | 1                             | 0                             | 2 (6.3)                         | 2                             | 0                             |
| Pseudogene             | 1 (2.1)                         | 0                             | 1                             | 1 (3.0)                         | 1                             | 0                             |
| Other                  | 6 (12.5)                        | 5                             | 1                             | 3 (9.4)                         | 3                             | 0                             |
| <b>TOTAL</b>           | <b>48</b>                       | <b>40 (81.6%)</b>             | <b>8 (18.4%)</b>              | <b>33</b>                       | <b>32 (97.0%)</b>             | <b>1 (3.0%)</b>               |

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



**Table 3**

Significant ( $p < 0.005$ ) differentially expressed lncRNA probesets and those that changed over time in male and female whole blood samples and proximity to putative ischemic stroke or vascular risk factor-related genes and/or SNPs. (Fold Change (FC)  $> 1.2$  = increased expression level in IS vs VRFC). (Partial correlation  $r < -0.4$  = decreasing expression over time).

| Probeset, lncRNA ID              | Sex | Cyto-band | Gene/SNP in Region                    | lncRNA up-/down-stream       | lncRNA Distance (kb) from Gene / SNP | FC or r Values | Ref    |
|----------------------------------|-----|-----------|---------------------------------------|------------------------------|--------------------------------------|----------------|--------|
| PSR06037329, linc-SLC22A2        | M   | 6p25.3    | LPAL2<br>SLC22A3<br>LPA<br>rs10455872 | within<br>up<br>down<br>down | within<br>14.0<br>146.2<br>69.2      | FC=1.36        | 48     |
| PSR09027609, linc-OBP2B-1        | M   | 9q34.2    | ABO<br>rs579459                       | down<br>down                 | 44.3<br>48.1                         | FC=1.42        | 48     |
| PSR20015349, OTTHUMT-00000079682 | F   | 20q13.13  | PTGIS                                 | up                           | 132.7                                | FC=1.41        | 50     |
| PSR04025299, linc_luo_1172       | F   | 4p16.3    | ADD1                                  | down                         | 332.6                                | FC=1.22        | 51, 52 |
| PSR19028831, ENST00000507442     | M   | 19q13.32  | APOE                                  | down                         | 219.6                                | $r = -0.49$    | 54     |
|                                  |     |           | APOC1                                 | down                         | 209.7                                |                |        |
|                                  |     |           | rs445925                              | down                         | 216.4                                |                |        |
| PSR20017082, ENST0000079667      | M   | 20q13.13  | PTGIS                                 | up                           | 28.96                                | $r = -0.40$    | 55     |
| PSR04029574, NR_036641           | F   | 4q32.1    | PDGFC                                 | within                       | within                               | $r = -0.41$    | 53     |