Molecular Neurobiology (2019) 56:1035–1043
https://doi.org/10.1007/s12035-018-1147-0

Deep Sequencing Reveals Uncharted Isoform Heterogeneity of the Protein-Coding Transcriptome in Cerebral Ischemia

Sunil Bhattarai[1] · Ahmed Aly[1] · Kristy Garcia[1] · Diandra Ruiz[1] · Fabrizio Pontarelli[1] · Ashutosh Dharap[1]

Received: 12 March 2018 / Accepted: 22 May 2018 / Published online: 3 June 2018

Abstract

Gene expression in cerebral ischemia has been a subject of intense investigations for several years. Studies utilizing probe-based high-throughput methodologies such as microarrays have contributed significantly to our existing knowledge but lacked the capacity to dissect the transcriptome in detail. Genome-wide RNA-sequencing (RNA-seq) enables comprehensive examinations of transcriptomes for attributes such as strandedness, alternative splicing, alternative transcription start/stop sites, and sequence composition, thus providing a very detailed account of gene expression. Leveraging this capability, we conducted an in-depth, genome-wide evaluation of the protein-coding transcriptome of the adult mouse cortex after transient focal ischemia at 6, 12, or 24 h of reperfusion using RNA-seq. We identified a total of 1007 transcripts at 6 h, 1878 transcripts at 12 h, and 1618 transcripts at 24 h of reperfusion that were significantly altered as compared to sham controls. With isoform-level resolution, we identified 23 splice variants arising from 23 genes that were novel mRNA isoforms. For a subset of genes, we detected reperfusion time-point-dependent splice isoform switching, indicating an expression and/or functional switch for these genes. Finally, for 286 genes across all three reperfusion time-points, we discovered multiple, distinct, simultaneously expressed and differentially altered isoforms per gene that were generated via alternative transcription start/stop sites. Of these, 165 isoforms derived from 109 genes were novel mRNAs. Together, our data unravel the protein-coding transcriptome of the cerebral cortex at an unprecedented depth to provide several new insights into the flexibility and complexity of stroke-related gene transcription and transcript organization.

Keywords Mouse · Ischemia · Cerebral cortex · RNA-sequencing · Gene expression

Introduction

Stroke is one of the leading causes of death and long-term disability in the USA. Loss of blood supply to the brain due to arterial blockage results in a cascade of damaging events such as excitotoxicity and oxidative stress that eventually result in brain damage and neurological deficit [1–3]. The underlying gene-expression changes that occur during an ischemic attack have been under intense focus over the last few decades [4–12]. A better understanding of the brain's innate molecular response to ischemia will improve our understanding of the pathophysiology of stroke and help identify novel targets for effective neuroprotective strategies against the stroke injury. To date, nearly all of the high-throughput studies evaluating transcriptomic profiles in experimental stroke have utilized probe-based techniques such as microarray [4, 5, 7, 12–15]. Although microarrays have contributed significantly to our understanding of gene expression changes in the post-stroke brain, this methodology is limited in that one can only explore the changes in the expression of previously known transcripts for which probes have been incorporated onto the microarray. This precludes the assessment of a vast portion of the unannotated transcriptome. Recent advances in genome-wide sequencing have enabled us to overcome this limitation and probe deeper into the transcriptome. Taking advantage of this technological advance, we recently published the first study on the genome-wide evaluation of the noncoding transcriptome in the post-stroke cerebral cortex in which we

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s12035-018-1147-0) contains supplementary material, which is available to authorized users.

✉ Ashutosh Dharap
ashutosh.dharap@hackensackmeridian.org

[1] Laboratory for Stroke Research and Noncoding RNA Biology, JFK Neuroscience Institute, HackensackMeridian Health JFK Medical Center, 65 James Street, Edison, NJ 08820, USA

discovered a large number of novel stroke-responsive long noncoding RNAs that had escaped prior detection [6]. No study to date has undertaken a similar, comprehensive genome-wide approach to evaluate the protein-coding transcriptome in the post-stroke cortex. Therefore, our goal here was to conduct an in-depth evaluation of the protein-coding transcriptome in the mouse cortex following transient focal ischemia utilizing RNA-sequencing (RNA-seq) to examine gene expression at the isoform-level at multiple time-points of reperfusion (6, 12, and 24 h). Overall, this study provides an updated and detailed catalog of the post-stroke protein-coding transcriptome in the adult mouse cortex using the latest annotations from mouse genome build GRCm38/mm10 and can therefore serve as a valuable resource for future studies incorporating gene expression changes in the post-stroke brain.

Methods

Animals

Adult male C57BL/6N mice (20–25 g; Taconic Biosciences, USA) were cared for in accordance with the Guide for the Care and Use of Laboratory Animals, US Department of Health and Human Services Publication number 86-23 (revised 1986). Protocols were approved by the Institutional Animal care and Use Committee of Seton Hall University. All animals were housed in light-controlled and temperature-controlled rooms with ad libitum access to food and water.

Transient Focal Ischemia

Animals were randomly assigned to control or treatment groups ($n = 12$/group). In the treatment groups, transient focal ischemia was induced as described previously [6]. Briefly, animals were anesthetized under 5% isoflurane on a pre-warmed heating pad, and a longitudinal cervical midline incision was made followed by careful dissection of the perivascular structures to expose the common carotid artery and the external carotid artery (ECA). A nick was made in the ECA, and a sterile 6-0 silicon-coated monofilament (Ethicon Inc., Germany) was inserted and advanced through the internal carotid artery up to the middle cerebral artery (MCA) to occlude blood flow into the MCA. Regional cerebral blood flow, mean arterial pressure, and arterial gases were monitored, and rectal temperature was maintained at 37.0 ± 0.5 °C during surgery. After 1 h of occlusion, the filament was removed and reperfusion was allowed for 6, 12, or 24 h. The incision was sutured, and animals were returned to their cages. For the sham cohort (control group), all the surgical steps were followed with the exception of MCA occlusion. In the ischemic groups, only those animals that displayed symptoms of stroke such as failure to extend the right paw and circling to the right were included for further study. Each brain was rapidly isolated (those showing a hemorrhage were excluded) and coronally sectioned on a mouse brain matrix. Cortical tissue from the mid-coronal sections was harvested and snap-frozen in liquid nitrogen for downstream experiments, and the adjacent sections were used for TTC staining to confirm the ischemic injury.

RNA Isolation

Total RNA was extracted from the tissues using the mirVana total RNA isolation kit (Thermo Fisher, USA) as per the manufacturer's instructions, followed by DNase I treatment for 20 min at 37 °C to remove genomic contamination. The DNase-treated RNA was purified using the RNA Clean and Concentrator kit (Zymo Research, USA) and eluted in RNAse-free water. The quantity, purity, and integrity of the RNA were evaluated using Qubit 2.0 (Thermo Fisher, USA) and Agilent Bioanalyzer (Agilent Technologies, USA).

RNA-seq and Data Analysis

RNA-seq was performed as described previously [6]. Briefly, RNA was extracted from three independent biological replicates per group, converted into cDNA libraries, and sequenced on a NextSeq500 instrument to generate 50 bp paired-end reads to yield on average 40.6 million reads per sample. TopHat and Bowtie2 were used to align the reads to the mouse reference genome GRCm38/mm10. Cufflinks was used to assemble the reads into transcript isoforms using mouse transcript annotations from Ensemble_81 as a guide. Cuffdiff was used to calculate the FPKM values and to calculate the differential expression of transcripts between the samples. The differentially expressed genes are expressed as absolute fold-change ≥ 1.5 with Benjamini-Hochberg FDR corrected threshold of $q < 0.05$ as a measure of significance. A small number of transcripts were not detected in the sham samples (FPKM = 0) but were highly expressed in the MCAO samples. In order to be able to calculate fold-change values for these transcripts, we assigned a heuristically determined non-zero FPKM value of 0.009 to each of these transcripts in the sham samples, as follows: we averaged FPKM values falling in the range of 0–1 from all the samples across all of the groups (sham, 6, 12, and 24 h), yielding an average FPKM value of 0.39 per sample. This number was divided by the average number of reads per group (40 million) to yield the final FPKM value of 0.009. The raw datasets from the RNA-seq experiments are deposited in the NCBI Gene Expression Omnibus (GEO) repository under the accession number GSE112348.

RT-PCR

RNA was isolated from the samples and converted to cDNA using the iScript cDNA synthesis kit (Biorad, USA). Fifty nanograms of cDNA was loaded into each 50 μl PCR reaction and amplified on a SimpliAmp Thermocycler (Thermo Fisher Scientific, USA) for 35 cycles using the Dream Taq Green PCR Mastermix (Thermo Fisher Scientific, USA) following the manufacturer's protocol. The PCR products were resolved on a 1.5% agarose gel and visualized on a Syngene PXi imaging system (Syngene, USA). Primer sequences were as follows: Pkp4 Fwd: 5-CGTGAAGGAGCAGG TGTTTA-3; Pkp4 Rev: 5- GATTCTGCTCCAAG CCTACAT-3; Khdrbs3 Fwd: 5-GGCAAGAGGAATGG ACTAACT-3 and Khdrbs3 Rev: 5-AGAAGCAG GTCCTTTGTCTG-3; Xkr4 Fwd: 5-TCCTCCAC TCTTCTTTCATAAACAT-3; Xkr4 Rev: 5-ACTTGAAG CAGGTCCCTTTAC-3; Cep170b Fwd: 5-AGCGTCCC AAGGACCTG-3; Cep170b Rev: 5-GGGCAGTC CCACCAACA-3; Mtch1 Fwd: 5-GCGATGAGATGTGG GAGATAAG-3; Mtch1 Rev: 5-ATGACCCACCTTCT AAACACG-3; Tnc Fwd: 5-CCACGGCCATGGGTTCT-3; Tnc Rev: 5-GAAACTCTCCACCTGAGCAGTA-3; Fbxo34 Fwd: 5-CGTGCAGAACTGCTGAGT-3; Fbxo34 Rev: 5-AGGCCTAGGAAGCCTGT-3.

Quantitative Real-Time PCR

Total RNA was converted into cDNA using iScript cDNA synthesis kit (Bio-Rad Laboratories, USA) and evaluated on a StepOne Real-Time System (Applied Biosystems, USA) using iTaq Universal SYBR Green Supermix (Bio-Rad Laboratories, USA) with standard cycling conditions. Per sample, 25 ng of starting cDNA material was mixed with gene-specific forward and reverse primers and SYBR Green Supermix in a 20 μl reaction. β-actin mRNA was quantified as an endogenous control. Fold-change differences across two groups were determined using the ΔΔCt method. Primer sequences were as follows: Hmox1 Fwd: 5- AGCAGAAC CAGCCTGAACTA -3; Hmox1 Rev: 5- GGAAGCCA TCACCAGCTTAAA-3; Map3k6 Fwd: 5- AACGCGGA TGTAGTGGTG -3 and Map3k6 Rev: 5- CACAATCC GAGTTCTTCTGGA -3; Il1a Fwd: 5- TGACCTGC AACAGGAAGTAAA -3; Il1a Rev: 5- TCGGTCTC ACTACCTGTGAT-3; Cep170b Fwd: 5-AGCGTCCC AAGGACCTG-3; Cep170b Rev: 5-GGGCAGTCCCACCA ACA-3; Mtch1 Fwd: 5-GCGATGAGATGTGGGAGATA AG-3; Mtch1 Rev: 5-ATGACCCACCTTCTAAACACG-3; Fbxo34 Fwd: 5-CGTGCAGAACTGCTGAGT-3; Fbxo34 Rev: 5-AGGCCTAGGAAGCCTGT-3; Xkr Fwd: 5-TCCTCCACTCTTCTTTCATAAACAT-3; Xkr Rev: 5-ACTTGAAGCAGGTCCCTTTAC-3. Pkp4 Fwd: 5-CGTGAAGGAGCAGGTGTTTA-3; Pkp4 Rev: 5-GATTCTGCTCCAAGCCTACAT-3. β-actin Fwd: 5-CGGTTCCGATGCCCTGAGGCTCTT -3; β-actin Rev: 5-CGTCACACTTCATGATGGAATTGA-3.

Statistics

The RNA-seq data analysis pipeline is based on the Tuxedo software package which implements in-built peer-reviewed statistical methods. Cuffdiff was used to test for differential transcript expression between the groups (three independent biological replicates per group) with Benjamini-Hochberg FDR-corrected threshold of $q < 0.05$ applied as a measure of significance. For quantitative real-time PCR experiments, at least two independent experiments, each with duplicate samples, were performed. Significant differences between two groups were determined using unpaired two-tailed t test with $p < 0.05$ as a measure of significance. The data are presented as mean ± SD.

Results

Differential Expression of mRNA Isoforms during Stroke

Using our RNA-seq pipeline, we detected 20,748 genes and 56,586 isoforms in the sham group; 22,192 genes and 60,023 isoforms in the 6 h group; 21,771 genes and 59,539 isoforms in the 12 h group; and 21,576 genes and 59,020 isoforms in the 24 h group. To generate a list of protein-coding transcripts, all noncoding RNAs were filtered out of the datasets. The resulting datasets were used to evaluate differential expression across the groups. This resulted in total of 1007 isoforms at 6 h, 1878 isoforms at 12 h, and 1618 isoforms at 24 h of reperfusion that showed significant differential expression (absolute fold-change ≥ 1.5; $q < 0.05$) as compared to sham controls. For the full lists of differentially altered isoforms at each time-point, see Supplementary Tables 1, 2, and 3.

Temporal Characteristics of Post-Stroke Isoform Expression

Comparisons of isoform expression profiles between the different reperfusion time-points showed that the overall number of altered isoforms increased from 1007 at the 6-h time-point to 1878 isoforms at 12 h, and then slightly decreased to 1618 isoforms at the 24-h time-point (Fig. 1a). When analyzed based on direction of expression (up/down), the number of upregulated isoforms increased through the first 12 h of reperfusion (from 549 isoforms at 6 h to 790 isoforms at 12 h) and then plateaued at 24 h time-point to 789 isoforms. Similarly, the downregulated isoforms showed increasing numbers through 6 h (458 isoforms) and 12 h (1088 isoforms) of

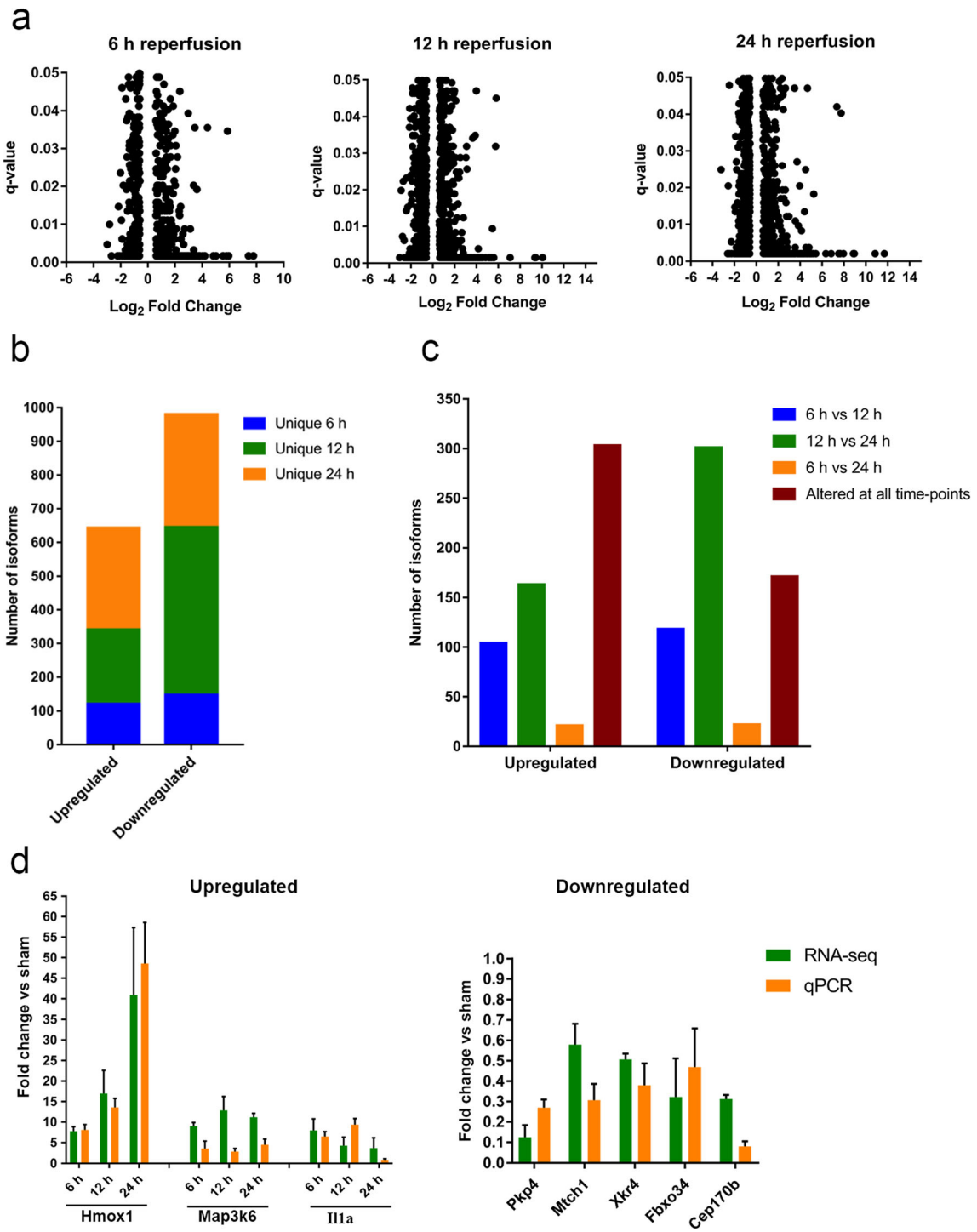

Fig. 1 Distribution of transcript isoforms differentially altered as a function of reperfusion time after MCAO. **a** Scatter plot showing the distribution of all of the significantly altered isoforms at each reperfusion time-point. **b** The number of isoforms that are uniquely expressed at each reperfusion time-point as compared to sham controls. **c** Temporal distribution of the isoforms that are significantly altered across two or more time-points. **d** Real-time PCR validation of the expression of key stroke-responsive genes. The data were derived from at least two independent experiments conducted in duplicate and are presented as mean ± SD

reperfusion, but dipped significantly at 24 h (829 isoforms) of reperfusion. We observed that a large number of isoforms were expressed exclusively at each of the time-points. A total of 121, 220, and 302 isoforms were exclusively upregulated and 147, 498, and 335 isoforms were exclusively downregulated at 6, 12, or 24 h of reperfusion, respectively (Fig. 1b). The remaining isoforms were commonly expressed at two or more time-points as follows: 222 transcripts (104 upregulated

and 118 downregulated) commonly expressed at 6 and 12 h; 464 transcripts (163 upregulated and 301 downregulated) commonly expressed at 12 and 24 h; 43 transcripts (21 upregulated and 22 downregulated) commonly expressed at 6 and 24 h; and 474 transcripts (303 upregulated and 171 downregulated) commonly expressed across all three reperfusion time-points (Fig. 1c). The expression changes of selected stroke-relevant genes were verified using real-time PCR (Fig. 1d). The complete list of altered transcripts and their temporal distributions are available in Supplementary Table 4.

Isoforms Derived via Alternative Splicing

Of all the isoforms altered at each of the reperfusion time-points, 32 isoforms at 6 h, 61 isoforms at 12 h, and 25 isoforms at 24 h were derived via alternative splicing (Supplementary Table 5). Interestingly, we observed that some genes yielded not one but two significantly altered splice isoforms at a given time-point (Table 1). At 6 h of reperfusion, the genes Prdm2 and Gria4 gave rise to two significantly altered isoforms each. For Prdm2, one isoform was upregulated, whereas the other was downregulated. For Gria4, both the isoforms were downregulated. At 12 h of reperfusion, the gene Tpm3 yielded two significantly altered isoforms of which one was upregulated and the other was downregulated. At 24 h of reperfusion, the gene Khdrbs3 yielded two significantly altered isoforms of which one was upregulated and the other was downregulated (Table 1).

Novel Splice Variants in Stroke

Further analysis of the differentially altered splice variants revealed that a number of them had not been previously reported. At 6 h of reperfusion, eight splice variants (three upregulated and five downregulated) were novel; at 12 h of reperfusion, 13 splice variants (three upregulated and ten downregulated) were novel; and at 24 h of reperfusion, four splice variants (two upregulated and two downregulated) were novel (Table 2). Barring the few instances detailed below, all of the novel isoforms were exclusively expressed at their respective time-points.

As noted earlier, at the 6-h time-point, the genes Prdm2 and Gria4 yielded two splice isoforms each. Of these, one isoform for each of the genes was novel. For Prdm2, the novel isoform was upregulated, whereas the annotated isoform was downregulated. For Gria4, both the novel and annotated isoforms were downregulated. At the 12-h time-point, of the two splice isoforms generated by the Tpm3 gene, one was novel. This novel isoform was downregulated, whereas the annotated isoform was upregulated. At the 24-h time-point, of the two splice isoforms for the Khdrbs3 gene, one was novel. Interestingly, the novel isoform was upregulated at both 12 and 24 h of reperfusion, whereas the annotated isoform, while detectable at 12 h, was not significantly altered in its expression, and by 24 h was significantly downregulated.

PCR Verification of the Novel Splice Isoforms

In order to verify that the novel splice isoforms identified by our RNA-seq pipeline are in fact bona fide splice variants, we examined the expression of these transcripts using PCR with primers designed to overlap the novel splice junctions (Fig. 2). We randomly selected six novel splice isoforms (from Table 2) and probed their expression at the corresponding time-points. We found that all of the novel splice variants tested were readily detected confirming the presence of the novel splice junctions as observed in our RNA-seq data.

Isoforms Generated via Alternative Transcription Start Sites (TSS) and Termination Sites (TTS)

A recent study evaluating 23 human cell-types concluded that for most protein-coding genes in human, transcript isoform differences are derived from alternative start and termination sites of transcription, rather than alternative splicing [16]. When we examined the genes yielding multiple differentially altered isoforms in our datasets, we found that this held true in the mouse cortex as well. Across the three reperfusion time-points, a total of 595 isoforms (103 at 6 h, 299 at 12 h, and 193

Table 1 Genes yielding two significantly altered splice variants at a given reperfusion time-point

Transcript ID	Gene	Time-point	Fold change
MC_00122827	Prdm2	6 h	1.9
MC_00122826	Prdm2	6 h	− 1.5
MC_00174667	Gria4	6 h	− 1.8
MC_00174665	Gria4	6 h	− 4.3
MC_00102475	Tpm3	12 h	3.2
MC_00102476	Tpm3	12 h	− 1.8
MC_00054928	Khdrbs3	24 h	2.0
MC_00054929	Khdrbs3	24 h	− 6.0

Table 2 Novel splice variants identified at each of the reperfusion time-points

Transcript ID	Gene	Fold change
6 h		
MC_00036829	Asap2	2.5
MC_00122827	Prdm2	1.9
MC_00169726	Ppan	2.7
MC_00085277	Pkp4	− 8.0
MC_00174665	Gria4	− 4.3
MC_00137595	Cux1	− 4.1
MC_00143763	Sgce	− 3.7
MC_00006921	Xkr4	− 1.9
12 h		
MC_00058296	Foxred2	5.1
MC_00054928	Khdrbs3	2.2
MC_00066558	Fam120b	2.2
MC_00166346	Ankrd10	− 8.3
MC_00060567	Prkdc	− 3.9
MC_00054119	Drosha	− 3.7
MC_00039182	Cep170b	− 3.2
MC_00048758	Fbxo34	− 3.1
MC_00109840	Fnbp1l	− 2.3
MC_00006921	Xkr4	− 2.0
MC_00102476	Tpm3	− 1.8
MC_00071365	Mtch1	− 1.7
MC_00053381	Pcdh9	− 1.6
24 h		
MC_00119272	Tnc	6.6
MC_00054928	Khdrbs3	2.0
MC_00096226	Meis2	− 6.2
MC_00160906	Plekha7	− 2.5

at 24 h) were derived via alternative transcription start or termination sites (Fig. 3a; full list available in Supplementary Table 6), whereas just 118 isoforms were derived via alternative splicing as noted earlier. Of the former, at 6 h of reperfusion, 61 isoforms were derived due to alternative TSS and 42 isoforms were derived due to alternative TTS; at 12 h of reperfusion, 211 isoforms were derived due to alternative TSS and 88 isoforms were derived due to alternative TTS; and at 24 h of reperfusion, 126 isoforms were derived due to alternative TSS and 67 isoforms were derived due to alternative TTS (Fig. 3a). Overall, the number of isoforms derived via alternative TSSs was higher than that derived by alternative TTSs at each time-point. This indicates that the predominant mechanism by which genes simultaneously transcribe multiple isoforms is via different start sites. Next, we examined how many of these were novel transcripts. At 6 h, 30 out of 103 isoforms were novel (21 derived via alternative TSS and 9 derived via alternative TTS); at 12 h, 84 out of 299 isoforms were novel (67 derived via alternative TSS and 17 derived via alternative TTS); and at 24 h, 51 out of 193 isoforms were novel (40 originated from alternative TSS and 11 originated from alternative TTS) (Fig. 3b). Consistent with the overall trend, the number of novel isoforms derived via alternative TSSs was substantially higher than that derived by alternative TTSs at each time-point. The distribution of isoforms based on origin (splicing or alternative TSS/TTS) for all of the significantly altered transcripts at each reperfusion time-point is shown in Fig. 3c. The regulatory mechanisms such as alternative promoters or enhancers or specific transcription factors driving these gene expression patterns need to be elucidated in future mechanistic studies.

Discussion

In the current study, we used RNA-seq to comprehensively investigate for the first time the post-stroke protein-coding transcriptome in a genome-wide manner in the cerebral cortex.

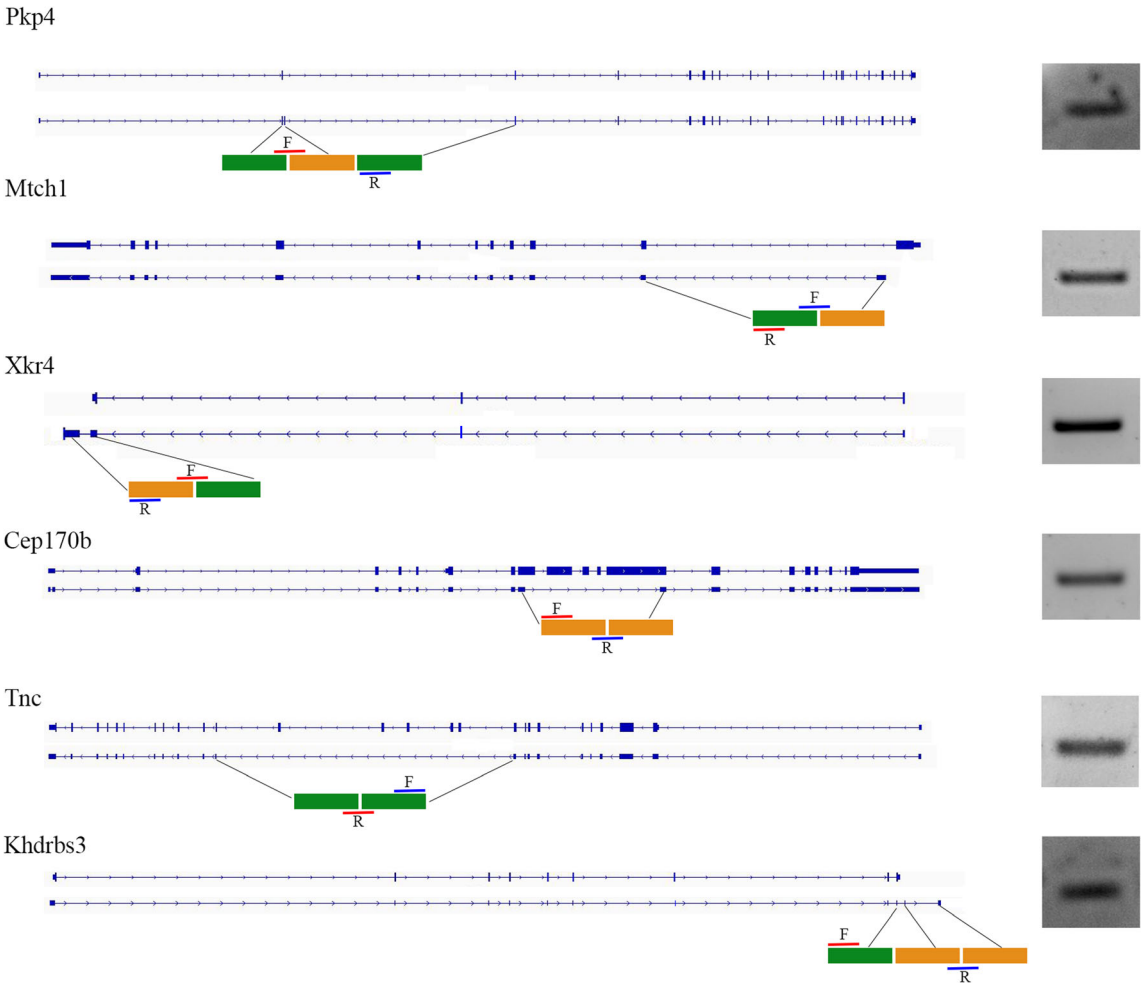

Fig. 2 RT-PCR verification of novel splice isoforms. The structures of the annotated isoforms (top) and novel isoforms (bottom) for each gene as visualized in IGV are presented. Colored rectangular boxes illustrate the exons against which the primers were designed. Novel exons are represented by orange boxes and annotated exons are represented by green boxes. For the Tnc gene, although both of the exons are annotated, the arrangement of those exons is unique such that it results in a novel splice junction. The PCR products for the transcripts are shown adjacent to their respective transcript structures. F, forward primer; R, reverse primer

Overall, our data was in agreement with previous studies that utilized other methodologies such as microarrays and PCR arrays. For example, chemokine genes such as Cxcl2 and Cxcl16, and genes from the tumor necrosis factor family such as Tnfrsf1a and Tnfrsf12a that were upregulated during stroke [4, 5, 17–19] were also upregulated in our datasets. Other genes such as Hspb1, Hmox-1, Mmp3, Mmp14, S100a11, Anax3, and Anax2 that are known to be significantly altered in the post-ischemic brain [4, 5, 20, 21] were similarly altered in our datasets. A key strength of our approach (which is also a key differentiator from previous studies) was that out of the various possible splice isoforms that may be generated by each gene via alternative splicing, we were able to converge on the specific splice isoform that is preferentially altered during stroke. As a result of this level of resolution, we were also able to identify new splice isoforms for some genes.

Alternative splicing enables the cell to generate a diverse complement of transcripts from a limited number of genes, which in turn enables diversification of function. Previous work has shown that alternative splicing is most prevalent in the mammalian brain as compared to any other organ [22]. The molecular diversity created by these widespread splicing events has been shown to have a significant functional impact on the brain via differential protein activity. For example, alternative splicing of the neuroligin gene (which encodes a postsynaptic adhesion protein) in neurons was shown to result in distinct molecular interactions at the synapse that, depending on the splice variant, altered post-synaptic differentiation fate [23]. Thus, the neurons were able to selectively modulate their synaptic interactions by simply deploying alternative splicing of the neuroligin gene.

The impact of alternative splicing during stroke is not yet fully understood. We found that while numerous genes generated multiple splice isoforms, for the vast majority of them, only one of the isoforms was significantly altered after stroke. For a small fraction of genes, we observed a deviation in this

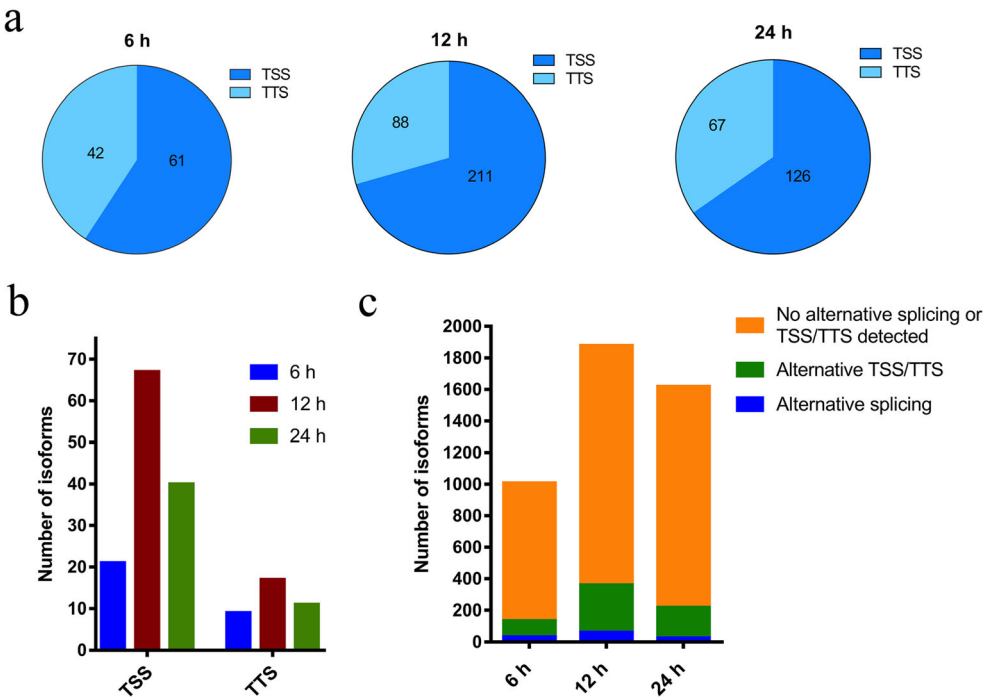

Fig. 3 Isoforms generated by the use of alternative TSS or TTS. **a** Distribution of mRNA isoforms derived via alternative TSS or TTS at each reperfusion time-point. The dark blue regions of the pie charts represent isoforms derived from alternative TSS and the light blue regions represent isoforms derived from alternative TTS. **b** Graph showing the time-point-based distribution of the subset of alternative TSS/TTS-derived isoforms that are novel. **c** Graph showing the isoform distribution across the total number of significantly altered transcripts at each reperfusion time-point. TSS, transcription start site; TTS, transcription termination site

trend wherein two isoforms were significantly altered during stroke. A closer examination of this subset of genes revealed that the two isoforms showed opposite direction of expression at a given time-point, i.e., one of the isoforms was upregulated whereas the other was downregulated, indicating isoform switching. The only exclusion to this was for the Gria4 gene at the 6-h time-point wherein both the isoforms were significantly downregulated. Of particular significance to isoform switching is our discovery of novel splice variants for a number of genes, which has the potential to alter the interpretation of the expression patterns for these genes. For example, those genes that expressed significantly altered novel isoforms but did not express the annotated isoforms would normally be interpreted as unaltered using traditional probe-based methods that are designed to detect only the annotated isoforms. Or in the instances where the novel isoform was significantly upregulated but the annotated isoform was significantly downregulated, without the ability to detect the novel isoform, such a gene would be interpreted as being downregulated when in fact it is significantly upregulated. An example of this in our study is seen for the Khdrbs3 gene which showed a significantly upregulated *novel* isoform at both the 12- and 24-h time-points, but also showed a significantly downregulated *annotated* isoform at 24 h. Without the knowledge of the novel isoform and the ability to detect it, the interpretation for this gene would be that it is downregulated at 24 h of reperfusion. However, our RNA-seq datasets revealed that this gene was in fact upregulated at both the 12- and 24-h time-points via the novel isoform. Another such example is that of the Prdm2 gene, which showed a significantly upregulated *novel* isoform as well as a significantly downregulated *annotated* isoform at 6 h of reperfusion. By detecting only the annotated isoform, this gene would be interpreted as being downregulated. However, because our RNA-seq data revealed a novel isoform that is significantly upregulated, an alternative interpretation of this gene is that it is upregulated via isoform switching at 6 h of reperfusion. Such examples highlight the significance of our study and underscore the importance of comprehensively evaluating gene expression using unbiased deep sequencing strategies. Further work needs to be done to determine how such isoform switching influences the resulting proteins and their functions and what significance it holds for the progression of the post-stroke pathophysiology.

Finally, the detection of multiple, simultaneously expressed mRNA isoforms generated via alternative transcription start or stop sites in a gene offers new insights for the interpretation of gene expression and function. It is possible that the protein products of the multiple isoforms act synergistically to carry out the gene's functions. Such a mechanism may be used by the cell to modulate gene dosage, foster transcript redundancy, or fine-tune its activities. Alternatively, the proteins derived from the various isoforms may have disparate functions. These possibilities may not necessarily be mutually exclusive and/or may be manifested in a gene-specific or condition-specific manner. Furthermore, the differences in the 5′ and 3′ sequence composition of the various transcript isoforms may result in differential post-transcriptional regulation with functional consequences.

Overall, our in-depth analysis of the protein-coding transcriptome after cerebral ischemia adds new information to the repertoire of gene expression data in stroke. It also raises new questions. For example, how do the novel mRNA isoforms affect the structures and functions of the resulting protein products? It would be expected that they would generate different protein isoforms due to shifts in the open reading frames. If so, what are the differences and how do they affect the stroke pathophysiology? These are other questions can be investigated in future studies.

Sources of Funding This study was funded by the JFK Neuroscience Institute, Edison, NJ.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interests.

References

1. Hossmann K-A (2006) Pathophysiology and therapy of experimental stroke. Cell Mol Neurobiol 26:1055–1081. https://doi.org/10.1007/s10571-006-9008-1
2. Dirnagl U, Iadecola C, Moskowitz MA (1999) Pathobiology of ischaemic stroke: an integrated view. Trends Neurosci 22:391–397. https://doi.org/10.1016/S0166-2236(99)01401-0
3. Lipton P (1999) Ischemic cell death in brain neurons. Physiol Rev 79:1431–1568. https://doi.org/10.1016/j.shpsa.2008.02.001
4. Hori M, Nakamachi T, Rakwal R, Shibato J, Nakamura K, Wada Y, Tsuchikawa D, Yoshikawa A et al (2012) Unraveling the ischemic brain transcriptome in a permanent middle cerebral artery occlusion mouse model by DNA microarray analysis. Dis Model Mech 5:270–283. https://doi.org/10.1242/dmm.008276
5. Chen MJ, Wong CHY, Peng ZF, Manikandan J, Melendez AJ, Tan TM, Crack PJ, Cheung NS (2011) A global transcriptomic view of the multifaceted role of glutathione peroxidase-1 in cerebral ischemic-reperfusion injury. Free Radic Biol Med 50:736–748. https://doi.org/10.1016/j.freeradbiomed.2010.12.025
6. Bhattarai S, Pontarelli F, Prendergast E, Dharap A (2017) Discovery of novel stroke-responsive lncRNAs in the mouse cortex using genome-wide RNA-seq. Neurobiol Dis 108:204–212. https://doi.org/10.1016/j.nbd.2017.08.016
7. Dharap A, Nakka VP, Vemuganti R (2012) Effect of focal ischemia on long noncoding RNAs. Stroke 43:2800–2802. https://doi.org/10.1161/STROKEAHA.112.669465
8. Zhang J, Yuan L, Zhang X, Hamblin MH, Zhu T, Meng F, Li Y, Chen YE et al (2016) Altered long non-coding RNA transcriptomic profiles in brain microvascular endothelium after cerebral ischemia. Exp Neurol 277:162–170. https://doi.org/10.1016/j.expneurol.2015.12.014
9. Carmichael ST, Archibeque I, Luke L, Nolan T, Momiy J, Li S (2005) Growth-associated gene expression after stroke: evidence for a growth-promoting region in peri-infarct cortex. Exp Neurol 193:291–311. https://doi.org/10.1016/j.expneurol.2005.01.004
10. Moore DF, Li H, Jeffries N, Wright V, Cooper RA Jr, Elkahloun A, Gelderman MP, Zudaire E et al (2005) Using peripheral blood mononuclear cells to determine a gene expression profile of acute ischemic stroke: a pilot investigation. Circulation 111:212–221. https://doi.org/10.1161/01.CIR.0000152105.79665.C6
11. Szaflarski J, Burtrum D, Silverstein FS (1995) Cerebral hypoxia-ischemia stimulates cytokine gene expression in perinatal rats. Stroke 26:1093–1100. https://doi.org/10.1161/01.STR.26.6.1093
12. Tang Y, Xu H, Du XL et al (2006) Gene expression in blood changes rapidly in neutrophils and monocytes after ischemic stroke in humans: a microarray study. J Cereb Blood Flow Metab 26:1089–1102. https://doi.org/10.1038/sj.jcbfm.9600264
13. Xu Z, Ford GD, Croslan DR et al (2005) Neuroprotection by neuregulin-1 following focal stroke is associated with the attenuation of ischemia-induced pro-inflammatory and stress gene expression. Neurobiol Dis 19:461–470. https://doi.org/10.1016/j.nbd.2005.01.027
14. Tang Y, Pacary E, Fréret T, Divoux D, Petit E, Schumann-Bard P, Bernaudin M (2006) Effect of hypoxic preconditioning on brain genomic response before and following ischemia in the adult mouse: identification of potential neuroprotective candidates for stroke. Neurobiol Dis 21:18–28. https://doi.org/10.1016/J.NBD.2005.06.002
15. Tang Y, Lu A, Aronow BJ, Wagner KR, Sharp FR (2002) Genomic responses of the brain to ischemic stroke, intracerebral haemorrhage, kainate seizures, hypoglycemia, and hypoxia. Eur J Neurosci 15:1937–1952. https://doi.org/10.1046/j.1460-9568.2002.02030.x
16. Reyes A, Huber W (2017) Alternative start and termination sites of transcription drive most transcript isoform differences across human tissues. Nucleic Acids Res 46:582–592. https://doi.org/10.1093/nar/gkx1165
17. Wolinski P, Glabinski A (2013) Chemokines and neurodegeneration in the early stage of experimental ischemic stroke. Mediat Inflamm 2013:727189. https://doi.org/10.1155/2013/727189
18. Rosito M, Deflorio C, Limatola C, Trettel F (2012) CXCL16 orchestrates adenosine A3 receptor and MCP-1/CCL2 activity to protect neurons from excitotoxic cell death in the CNS. J Neurosci 32:3154–3163. https://doi.org/10.1523/JNEUROSCI.4046-11.2012
19. Ueland T, Smedbakken LM, Hallén J, Atar D, Januzzi JL, Halvorsen B, Jensen JK, Aukrust P (2012) Soluble CXCL16 and long-term outcome in acute ischemic stroke. Atherosclerosis 220:244–249. https://doi.org/10.1016/j.atherosclerosis.2011.10.004
20. Stetler RA, Gao Y, Zhang L, Weng Z, Zhang F, Hu X, Wang S, Vosler P et al (2012) Phosphorylation of HSP27 by protein kinase D is essential for mediating neuroprotection against ischemic neuronal injury. J Neurosci 32:2667–2682. https://doi.org/10.1523/JNEUROSCI.5169-11.2012
21. Nimura T, Weinstein PR, Massa SM, Panter S, Sharp FR (1996) Heme oxygenase-1 (HO-1) protein induction in rat brain following focal ischemia. Mol Brain Res 37:201–208
22. Yeo G, Holste D, Kreiman G, Burge CB (2004) Variation in alternative splicing across human tissues. Genome Biol 5:R74. https://doi.org/10.1186/gb-2004-5-10-r74
23. Chih B, Gollan L, Scheiffele P (2006) Alternative splicing controls selective trans-synaptic interactions of the neuroligin-neurexin complex. Neuron 51:171–178. https://doi.org/10.1016/j.neuron.2006.06.005