Genome-wide identification of novel long non-coding RNAs and their possible roles in hypoxic zebrafish brain

Bodhisattwa Banerjee [1*], Debaprasad Koner [2], David Karasik [1], Nirmalendu Saha [2]

[1]Musculoskeletal Genetics Laboratory, The Azrieli Faculty of Medicine, Bar-Ilan University, Safed 1311502, Israel

[2]Department of Zoology, North-Eastern Hill University, Shillong 793022, India

*Corresponding author:

Dr. Bodhisattwa Banerjee

Musculoskeletal Genetics Laboratory,

The Azrieli Faculty of Medicine,

Bar-Ilan University,

Safed 1311502, Israel

Abstract:

Long non-coding RNAs (lncRNAs) are the master regulators of numerous biological processes. Hypoxia causes oxidative stress with severe and detrimental effects on brain function and acts as a critical initiating factor in the pathogenesis of Alzheimer's disease (AD). However, no data are available on the roles of lncRNAs and their target genes under hypoxia in the zebrafish brain. From the RNA-Seq in the forebrain (Fb), midbrain (Mb), and hindbrain (Hb) regions of hypoxic and normoxic zebrafish, we identified 8114, 7775, and 7816 novel lncRNA transcripts in the Fb, Mb, and Hb regions, respectively. In comparison to the normoxic group, hypoxic fish showed differential expression of 357, 553, and 786 novel lncRNA genes in the Fb, Mb, and Hb regions, respectively. Co-expression network analysis identified a series of lncRNA-mRNA pairs in all the regions, which were enriched for AD. In conclusion, functional annotation of lncRNAs induced by hypoxia provided knowledge on potential regulators for genes participating in AD pathogenesis.

Keywords: long non-coding RNAs; hypoxia; Alzheimer's disease; brain regions; zebrafish

1. Introduction

Non-coding RNAs (ncRNAs) were previously considered as "transcriptional noise" and deemed as "non-functional". Recent technical advances in the high-throughput genomic platforms have revealed that despite active transcription of about 85% of the human genome [1,2], only 1–2% of it is protein-coding. Thus the vast majority of transcripts are represented by ncRNAs. Two major classes of ncRNAs have been classified according to their length: small ncRNAs (under 200 nt) and long non-coding RNAs (lncRNAs) (200 nt to as long as several kilobase pairs). Similar to the mRNAs, most of the lncRNAs undergo splicing, 5′ capping, 3′ polyadenylation and are usually transcribed by RNA polymerase II. Like coding transcripts, lncRNAs, too, have epigenetic markers [3] and may contain polymorphisms [4]. The lncRNAs are emerging as critical regulators of major biological processes impacting development, differentiation, and disease [5,6].

Hypoxia is the condition where there is an inadequate oxygen supply in tissues. Approximately 20% of the human body's energy budget is consumed by the brain [7], and thus adequate supply of oxygen is critical for the normal functioning of the neurons. It has been suggested that hypoxia plays an essential role in the initiation of Alzheimer's' disease (AD) pathogenesis as hypoxic conditions increase free radical production in the electron transport chain in the mitochondria leading to increased oxidative stress [8,9].

The zebrafish is a small, hardy freshwater fish native to South Asian countries like India, Nepal, Pakistan, and Bangladesh. Originally zebrafish was used as a model organism for the study of vertebrate development. However, in recent decades large-scale genetic screens were able to identify hundreds of mutant phenotypes similar to human diseases, including the AD [8,10,11]. Several reported studies support the viability of the zebrafish model as an alternative way for a better understanding of neurodegenerative diseases, such as AD [12]. Nevertheless, for studies on all these diseases and AD, the focal spot is primarily on the protein-coding genes. However, similar studies using the zebrafish to probe lncRNA function are still at infancy, mostly because of a lack of proper annotation and systematic survey.

To date, only a few studies have systematically determined the lncRNA landscape in the zebrafish. Ulitsky et al. [13] used chromatin marks and poly(A) site mapping as well as RNA-seq data from three developmental stages and reported more than 550 distinct lncRNAs in zebrafish. Among which 29 lncRNAs showed sequence similarity with their mammalian orthologs.

Similarly, two more studies by Pauli et al. [14] and Kaushik et al. [15] identified 1133 and 442 long non-coding transcripts from eight developmental stages of zebrafish embryogenesis and different adult tissues, respectively.

However, there are several lacunae in these previous studies. Moreover, no data are available on the lncRNAs expressed in sub-regions of the fish brain and their expression in these brain regions. Therefore, in this study, we used a strand-specific RNA-sequencing (ssRNA-Seq) method to identify and characterize novel lncRNAs (previously unreported) and their potential target protein-coding genes in three different brain regions of the zebrafish. Furthermore, to elucidate the roles of these lncRNAs and their potential target genes under hypoxia, we constructed a co-expression network of lncRNAs and mRNAs in these brain sub-regions.

2. Results

2.1 Genome-wide identification of novel lncRNAs

From the adult zebrafish brain regions (Fb, Mb, and Hb) RNA-Seq library, a total of 307 million reads were obtained. GC content and Q30 content in these data averaged at 43.8% and 93.94%, respectively (Table 1). An average of 84.63%, 85.10%, and 86.78% of reads mapped uniquely to the zebrafish genome (GRCz11) in the forebrain, midbrain, and hindbrain regions, respectively (Table S3). We have used stringent criteria to develop a pipeline for identifying novel lncRNAs (Fig. 1). StringTie with --merge option was used to merge all the GTF files from the three brain sub-regions and produce a non-redundant set of transcripts of the zebrafish brain. Using the filtering pipeline (Fig. 1), a total of 35809 lncRNA transcripts were identified. Furthermore, using the BLASTN search (E value $1e^{-5}$) against the lncRNA sequences of the ZFLNC database resulted in the detection of 6739 novel lncRNA transcripts (5620 novel lncRNA genes) (Supplementary file S1). Among these novel lncRNAs majority (61%) were intronic, 24% were intergenic, 10% sense, and 5% were antisense (Fig. 2A). The majority (58%) of the novel lncRNAs identified in this study were 200 – 600 bp in length, whereas only 4% of these were longer than 2000 bp in size. Furthermore, we identified the common lncRNAs in the Fb, Mb, and Hb regions and showed that 208 lncRNAs were mutual among all three regions (Fig. 2C).

2.2 Functional annotation of the novel lncRNAs

The protein-coding genes which were present 10 kb up or downstream of the novel lncRNAs were considered as the *cis* target genes. We could identify 3602 potential target protein-coding

genes (Supplementary file S2). To obtain further insight into the functional role of these lncRNAs, we performed GO (Biological process) and KEGG pathway enrichment analysis (Fig. 3 and Supplementary file S3 – S4). 30 terms were enriched in the GO biological process that includes 'Nervous system development', 'Cell adhesion', 'Neuron development', 'Synaptic signaling'. Similarly, 662 genes comprising 18 functional categories were enriched in the KEGG pathway analysis, which includes the 'Calcium signaling pathway, 'GnRH signaling pathway', 'ErbB signaling pathway', and 'MAPK signaling pathway'.

2.3 Differential expression of novel lncRNAs and mRNAs under hypoxia

Differential expression of the lncRNA and mRNA genes in three brain sub-regions were measured using edgeR tool. The lncRNAs and mRNAs in the hypoxic group (1.15±0.6 mg/l of dissolved oxygen) of fish were compared to the normoxic (7.6±0.4 mg/l of dissolved oxygen) group and the genes with fold change >2 and FDR <0.05 were considered as differentially expressed (DE). A total of 357, 553, and 786 novel lncRNA (Supplementary file S5A – S5C) and 2186, 1957, and 5441 mRNA DE genes were detected in the Fb, Mb and Hb regions, respectively (Supplementary file S6A – S6C). Furthermore, we compared the DE lncRNAs and mRNAs between the three brain sub-regions. Venn diagram (Fig. 4A) showed that 136 DE lncRNAs were shared between all the three brain regions, whereas 1292 DE mRNAs were common in all the brain regions (Fig. 4B). We used hierarchical cluster analysis to illustrate the expression patterns of DE lncRNAs and DE mRNAs. The heat map of the top 500 DE lncRNA and mRNA genes showed a clear separation between the normoxic and hypoxic groups in the three brain sub-regions (Fig. 5). Many DE protein-coding genes were implicated in neurodegenerative diseases like AD (i.e., *bace2*, *mapta*, *maptb*, *apoeb*, *foxo3b*, *slc2a3b*, etc.) (Supplementary file S5 – S7).

Moreover, enrichment analysis of the DE mRNAs in each brain sub-region was performed using GO biological process, KEGG, and Disease.ZFIN databases. The GO enrichment analysis of the Fb and Mb region showed that the highest number of upregulated mRNAs were significantly associated with 'developmental process', 'anatomical structure development', 'multicellular organism development' (Supplementary Table S3 and S6). In contrast, downregulated mRNAs in Fb were associated with 'DNA-templated transcription, initiation', 'chromatin silencing'. Six KEGG pathways that included the 'MAPK signaling pathway', 'cellular senescence', 'FOXO signaling pathway' were enriched in the Fb region (Supplementary

Table S4). Only four KEGG pathways, including 'ECM-receptor interaction', 'P53 signaling pathway', 'FOXO signaling pathway', were enriched in the Mb region (Supplementary Table S7). In the Hb region, 15 GO terms were enriched. Among these eight were significantly upregulated, including 'anion transport', 'Circadian regulation of gene expression', 'Nervous system development' (Supplementary Table S9). Ten KEGG pathways were also enriched in the Hb region (Supplementary Table S10). Most interestingly, the highest number of DE mRNAs of Fb and Hb regions when using Disease.ZFIN was significantly enriched in type 2 diabetes mellitus and AD, respectively (Supplementary Table S5 and S11). Furthermore, the highest number of DE mRNAs in Hb were enriched clearly in several neurodegenerative diseases, such as AD, Huntington's disease, and Parkinson's disease (Supplementary Table S11).

2.4 Construction of co-expression network of lncRNAs and mRNAs

Using weighted correlation network analysis in WGCNA, we created clusters (modules) of highly correlated lncRNA and protein-coding genes. A uniquely colored label was assigned to each of the modules (Fig. 6A). The lncRNAs and the correlated mRNAs were grouped into 8 clusters, with a range of 45 (pink) to 341 (turquoise) genes in each cluster module (Fig. 6B). All 8 modules produced a co-expression network between the DE novel lncRNAs and mRNAs. However, the blue module, followed by the turquoise module had the two highest number of such interactions. Based on the blue module, 17 lncRNAs produced a co-expression network with 161 mRNAs (Fig. 6C). Enrichment analysis of the blue module against the GO biological process and Disease.ZFIN databases were performed. The genes were enriched for 13 GO terms, including 'Multicellular organism development', 'Generation of neurons', 'Nervous system development' (Supplementary Table S12). Remarkably, the highest number of genes were enriched for AD, followed by 'Brain ischemia' when we used Disease.ZFIN database (Supplementary Table S13). The co-expression network in each of the brain regions was also identified (Supplementary figs. S1A – S1C). A total of 142, 116, 664 DE novel lncRNA and mRNA genes were correlated in Fb, Mb, and Hb regions, respectively. The co-expression networks in Fb and Hb regions were enriched for AD (Supplementary Tables S14 and S16). Whereas, in the Mb region, Myocardial infarction was the top enriched disease when using Disease.ZFIN database (Supplementary Table S15).

2.5 Validation of RNA-Seq data by RT-qPCR

An equal number (four) of DE lncRNAs and DE mRNAs were randomly selected from the Fb, Mb, and Hb regions, and their expression was validated by qPCR analysis. The expression patterns of the selected lncRNAs and mRNAs showed a similar trend in these two methods. These results validate the reliability of the RNA-Seq data.

3. Discussion

Among the small animal models, zebrafish has gained popularity as one of the valuable ones, which are relevant for a study of vertebrate gene function in general. Thus, a high number (~70%) of human genes have at least a single zebrafish orthologue [16]. Moreover, zebrafish and the human brain have a high conservation in the neuroanatomical structures [17,18] and neurochemical pathways [19]. The zebrafish forebrain medial, dorsal, and lateral pallium has strong similarities with the amygdala, isocortex, and hippocampus, in humans, respectively [20,21]. Moreover, despite the size difference of particular zebrafish brain structures, major brain components, such as the diencephalon, the cerebellum is highly conserved with the human brain [22]. Despite lacking the complex advanced cognitive behaviors evident in rodent models, zebrafish has proven to be very informative for the study of human diseases, including AD [8]. Multiple studies have established the zebrafish as an important model for human disease-related protein-coding genes. In contrast, knowledge of biology, function, and potential of lncRNAs as biomarkers and treatment targets in genetic diseases is still in its infancy. Very few studies have determined the lncRNA landscape in the zebrafish, and there are no reports on the sub-region specific expression of lncRNAs in the zebrafish brain. The only study elucidating brain lncRNAs of zebrafish was conducted by Kaushik et al. [15]. However, there were two significant problems in their study: (1) the method used for RNA isolation. They used the poly-A enrichment method. At times (but not always), lncRNAs are processed similarly to the mRNAs (e.g., capped, spliced, and polyadenylated) [23]. Hence, poly-A enrichment of the RNA sample is not the ideal method to capture all the lncRNAs expressed. (2) They have used an older zebrafish genome version (Zv9) for alignment. The latest genome version and annotations should be used for accurate identification of lncRNAs [24]. So to avoid these potential inaccuracies, in this study, we used ssRNA-Seq, the latest zebrafish genome (GRCz11), and a stringent filtering pipeline to systematically identify novel lncRNAs, and determined their differential expression pattern in

three sub-regions of the brain under hypoxia. The fish brain model of hypoxia was obtained by the method described by Nik et al., 2014 [25].

A total of 3602 protein-coding genes were identified as the potential *cis* targets of the novel lncRNAs in the total brain. Enrichment analysis of these targets showed that a majority of genes were implicated in neuronal development and neuron differentiation. Nevertheless, there was another set of genes that were involved in synaptic and trans-synaptic signaling. KEGG pathway enrichment showed that most genes were enriched for calcium signaling and MAPK signaling pathways. Numerous studies had revealed that multiple lncRNAs were involved in mammalian neurogenesis [26,27] as well as in neural development [28]. Consequently, from this study, it could be inferred that most of the novel lncRNAs detected in this study were involved in the regulatory role of neurodevelopmental procedures. We identified a largely region-specific expression of the lncRNAs. Altogether, the expression profile of Fb, Mb, and Hb region was distinct. Although a relatively small number of novel lncRNAs (208) were shared among all three different regions indicating that these lncRNAs might have critical regulatory roles for maintenance of all these regions, this was consistent with a significant number of studies showing that lncRNAs have brain region-specific expression pattern [15,29–31]. The knowledge that the functional roles of many lncRNAs in the brain are spatially specific is relevant to understand the roles of their target protein-coding genes.

Several studies had implicated brain hypoxia for the progression of AD. Many cardiovascular risk factors, as well as vascular and traumatic brain injury, have a strong correlation with AD pathogenesis [28]. Hypoxia in the human brain due to the failure of vascular function leads to a regionalized effect, which extends progressively with a decrease in vascular functions [32]. The DE mRNA genes, mostly in the Fb and Mb regions in this study, were also enriched for vasculature development, blood vessel morphogenesis, and diseases like pulmonary hypertension and coronary artery disease, indicating a role in cardiovascular diseases. Most lncRNA expressions in the brain of AD patients show a decreasing pattern [33,34]. Our data follow a similar trend, where the majority of the novel lncRNAs were downregulated in hypoxic fish brains.

The lncRNAs can regulate, and their expression is highly associated with the expression of their neighboring protein-coding genes [35]. In the current study, the Hb region has a higher number of correlated DE lncRNA and mRNA genes compared to the other two regions owing to

its much higher number of DE genes. However, we could not accurately pinpoint the reason. In the Fb regions, the correlated DE mRNA targets showed a regulatory role in AD. Two lncRNAs MSTRG.9523 and MSTRG.11854 showed a positive correlation with the zebrafish orthologues (*mapta* and *maptb*) of the human microtubule-associated protein tau (*MAPT*) and both the mRNAs were upregulated. Nik et al. [25] observed that hypoxia caused the higher expression of these two genes and concordant decrease of zebrafish orthologue of human *TRA2B*, *tra2b*. In our current study, *tra2b* was predicted as a target gene of another lncRNA, MSTRG.18722, in the Fb region. Interestingly, both the lncRNA and the potential target gene *tra2b* were downregulated. The dysregulated protein MAPT aggregates and forms lesions in neurons or glial cells and are associated with multiple diseases like frontotemporal dementia, Pick disease, progressive supranuclear palsy [36], and AD [37]. Moreover, the *bace2* gene was also identified as the target of lncRNA MSTRG.18722. The gene *bace2* is the zebrafish homolog of human *BACE2* gene, which is involved in AD pathogenesis [38]. Bace2 efficiently cleaves the amyloid beta A4 precursor protein (APP) in the middle of the Aβ domain [38]. Enrichment analysis of the co-network module using Disease.ZFIN database predicted AD as the top enriched disease. Furthermore, in the Hb region, the lncRNA MSTRG.11575 was identified as the potential regulator of the *apoeb* gene. The gene *apoeb* is a zebrafish orthologue of human APOE (apolipoprotein E), one of the most substantial genetic risk factors for AD [39,40]. Although minimal research has yet been conducted on APOE functions in zebrafish, some studies elucidating the *apoea* [41] and *apoeb* [42] reported their expression in retina, yolk syncytial layer and microglial cell [43]. Additionally, MSTRG.704 and MSTRG.2761 lncRNAs were in a co-expression network with two DE genes, namely *psen1* and *foxo3b*. Nik et al. [44] have shown an upregulation of *psen1*, *psen2*, *appa*, *appb*, and *bace2* in the zebrafish adult brain under hypoxia. Moreover, human brains showed a similar pattern of expression [45–47]. Of note, the genes in the co-expression module were enriched for AD, Parkinson's disease, and cerebral ischemia, among other neurodegenerative disorders.

Curiously, in the Mb region, three lncRNAs, namely MSTRG.28309, MSTRG.13434, and MSTRG.21390 were identified as the probable regulator of the circadian clock genes, *per1a, per1b*, and *per3* genes. The Period (*PER1*, *PER2*, and *PER3*) and Cryptochrome (*CRY1*, *CRY2*) genes coordinate biological rhythms by repressing the expression of *CLOCK* and *BMAL1* genes in a periodic manner [48]. Moreover, multiple studies have demonstrated that acute and chronic

hypoxia changes the expression of *per1a* and *per1b* in zebrafish, whereas the hypoxia response is, in turn, regulated by the clock genes [48–50]. Similarly, another circadian rhythmic gene *exorh* was regulated by MSTRG.7006. The pineal opsin gene *exorh* is expressed rhythmically, and is consecutively, negatively regulated by the circadian regulatory gene *per3* [51]. Neurodegenerative diseases, such as AD, affects the proper functioning of circadian clock genes and may lead to disruption of the timely organization amid different areas of the brain [52].

4. Conclusion

In summary, we performed ssRNA-Seq in the three brain regions, namely Fb, Mb, and Hb of normoxic and hypoxic zebrafish and identified novel lncRNAs in the total brain as well as specific sub-regions. The functional enrichment studies of the potential *cis* targets of the novel lncRNAs were involved in neuronal development and differentiation pathways. These genes were also identified as enriched for calcium and MAPK signaling pathways. Under hypoxia, several lncRNAs and mRNAs were differentially expressed. The lncRNA expression pattern showed a difference between the brain regions. It could be inferred from the current study that in the Fb and Hb regions, the potential lncRNA targets were the genes implicated in the AD pathogenesis. Interestingly, in the Mb region, the potential target genes were involved in the circadian rhythm and regulation of circadian gene expression. From the present data, we could identify lncRNAs MSTRG.9523, MSTRG.11854, and MSTRG.18722 in the Fb region and lncRNAs MSTRG.11575, MSTRG.704, and MSTRG.9890 in the Hb region as the potential regulators of several AD genes (*mapta*, *maptb*, *tra2b*, *bace2*, *apoeb*, *psen1*, *foxo3b*). However, further studies are necessary to elaborate on the specific functions of these lncRNAs in the pathogenesis of hypoxia-induced AD.

5. Materials and methods

5.1 Zebrafish maintenance and experimental procedures

The adult wild-type zebrafish were maintained in our laboratory as per the approved guidelines of the Institutional Animal Ethics Committee of North-Eastern Hill University. Zebrafish were bred and maintained at 28°C on a 14 h light/10 h dark cycle. 1-year-old adult zebrafish (AB strain) were used for all experiments (n =24). No sex differentiation of the fish was performed. The experiments conducted under low oxygen conditions were performed following the method described by Nik et al., 2014 [25]. Briefly, oxygen was depleted by bubbling nitrogen gas through the water, and dissolved oxygen level (DOL) was measured by a

dissolved oxygen meter. The DOL in the hypoxia group was 1.15 ± 0.6 mg/l. Another group of fish (n = 12) was kept under normal oxygen conditions (6.6 ± 0.45 mg/l). The hypoxic conditions were maintained for 3 hours, and immediately after the treatment, three biological replicates, each consisting of 3 zebrafish, were used for isolation of RNA. A total of 18 fish (9 hypoxic and 9 normoxic) were sacrificed after anesthetizing with 0.03% tricaine (Sigma-Aldrich, Germany), and brains were separated and placed in ice-cold HBSS (1X Hank's basal salt solution, 2.5 mM HEPES-KOH, pH 7.4, 35 mM glucose). Three sub-regions of the brain were dissected, and RNA was isolated using the NucleoSpin RNA XS kit (Macherey-Nagel, Germany) following the manufacture's protocol. The RNA quality and concentration were measured in the QIAxcel Advanced system (Qiagen, Germany) using QIAxcel RNA QC Kit v2.0 (Qiagen) and Qubit 3.0 fluorometer (Thermo Fisher Scientific, USA). The RNA samples with an RNA integrity score (RIS) value of more than eight were used in further processing.

5.2 Library construction and sequencing

Ribosomal RNA was depleted from the total RNA using the Ribo-Zero Gold rRNA removal kit (Human/Mouse/Rat) (Illumina, USA). The ribo-depleted RNA was used to create an ssRNA-Seq library, according to the TruSeq® Stranded Total RNA sample preparation guide. The sequencing was performed on the NovaSeq 6000 system (Illumina) using 2×100 bp chemistry by Macrogen Inc., South Korea. The sequencing data have been deposited in the NCBI's Gene Expression Omnibus with the GEO accession number **GSE153687**.

5.3 Transcriptome assembly and novel lncRNA detection

Reads with a quality score of ≥30 were selected for the downstream analysis. The raw reads were aligned to the latest zebrafish genome (GRCz11) from ENSEMBL using STAR v2.7.3 [53], and transcripts were assembled by StringTie [54] using the reference annotation from Ensembl. The expression level of each transcript was calculated and normalized as transcripts per million (TPM). The transcripts of < 200 bp length were excluded from further analysis. Several stringent filtration steps were applied to obtain the lncRNAs. At first, the combined data from gffcompare were used in the FEELnc tool [55] to obtain a potential list of lncRNAs. Next, the coding potential of the assigned lncRNAs was again evaluated using CPC2 [56] and CNIT [57] tool. Any transcript assigned as coding by any of the tools were discarded. Furthermore, these lncRNAs were evaluated by Pfam search, and transcripts with E value ≥ $1e^{-3}$ were omitted from further analysis. Finally, the RepeatMasker tool was utilized to identify the sequences with

repetitive regions, and those transcripts were excluded. The total lncRNAs obtained were matched against the lncRNAs from the ZFLNC database [58] using BLASTN (coverage ≥ 55% and E value ≤ $1e^{-6}$), and those without a match were deemed as candidate novel lncRNAs in zebrafish brain regions. The novel lncRNAs were classified according to their localization and the direction of transcription of proximal RNA transcripts using the FEELnc_classifier module of the FEELnc pipeline [55].

5.4 Prediction of lncRNA targets

Potential target protein-coding genes of the novel lncRNAs were predicted on the basis of *cis* function. The closest protein-coding gene 10 kb upstream and downstream of lncRNAs were screened using the FEELnc pipeline [55]. Functional enrichment of these target genes using GO biological process and KEGG pathways were done using the ShinyGO v0.61 web-server [59].

5.5 Differential expression of lncRNAs

The raw read counts of the lncRNAs and mRNAs were determined by the featureCounts program [60], and differential expression analysis was performed using edgeR in Blast2GO [61]. Benjamini-Hochberg method for controlling the false discovery rate (FDR) was used for adjusting the *p*-value. Only the genes with FDR < 0.05 and fold change > 2 were considered as differentially expressed (DE). Enrichment analysis of the DE lncRNA-targeted protein-coding genes among GO biological process, KEGG pathways, and Disease.ZFIN databases were conducted using iDEP (integrated Differential Expression and Pathway analysis) web-server [62].

5.6 Co-expression network construction and functional enrichment analysis

The co-expression network of DE lncRNAs and mRNAs were constructed using Weighted Gene Co-expression Network Analysis (WGCNA) package [63] on the iDEP web-server. All parameters were set as the default. Different modules were detected using the dynamic tree cutting method, and the network of each module was visualized using Cytoscape v3.8 [64].

5.7 Validation of expression of lncRNAs

The differential expression profiles of lncRNAs and mRNAs between three brain regions were validated using a total of 12 lncRNAs and mRNAs (4 from each region) by qPCR analysis. The lncRNA and mRNA-specific primers were designed by Primer-BLAST (Supplementary Table S1). Total RNA isolated earlier using the NucleoSpin RNA XS kit (Macherey-Nagel) were converted to the first-strand cDNA following the method described earlier [65]. qPCR reactions

were performed in a StepOne plus Real time PCR system (Thermo Fisher Scientific) using SYBR® Premix Ex Taq™ II (Takara Clontech, Japan). Melting curve analysis was used to re-confirm the amplification of only a single PCR product. Fold changes of genes in hypoxic fish compared to normoxic controls were calculated using the modified $\Delta\Delta C_T$ method [66].

Acknowledgments

This study was supported by a National postdoctoral fellowship project sanctioned to BB (PDF/2016/002013) by the SERB, New Delhi, India. We sincerely acknowledge the Computer Centre, North-Eastern Hill University, India, for providing the SULEKOR High-Performance Computing Facility for bioinformatics works. We thank Dr. Sumit Mukherjee, Bar-Ilan University, Israel, for his suggestions on bioinformatics works.

References

[1] S. Djebali, C.A. Davis, A. Merkel, A. Dobin, T. Lassmann, A. Mortazavi, Landscape of transcription in human cells, Nature. 489 (2012). doi:10.1038/nature11233.

[2] M.J. Hangauer, I.W. Vaughn, M.T. McManus, Pervasive Transcription of the Human Genome Produces Thousands of Previously Unidentified Long Intergenic Noncoding RNAs, PLoS Genet. 9 (2013) e1003569. doi:10.1371/journal.pgen.1003569.

[3] S. Sati, S. Ghosh, V. Jain, V. Scaria, S. Sengupta, Genome-wide analysis reveals distinct patterns of epigenetic features in long non-coding RNA loci, Nucleic Acids Res. 40 (2012). doi:10.1093/nar/gks776.

[4] H. Ling, R. Spizzo, Y. Atlasi, M. Nicoloso, M. Shimizu, R.S. Redis, N. Nishida, R. Gafa, J. Song, Z. Guo, C. Ivan, E. Barbarotto, I. De Vries, X. Zhang, M. Ferracin, M. Churchman, J.F. van Galen, B.H. Beverloo, M. Shariati, F. Haderk, M.R. Estecio, G. Garcia-Manero, G.A. Patijn, D.C. Gotley, V. Bhardwaj, I. Shureiqi, S. Sen, A.S. Multani, J. Welsh, K. Yamamoto, I. Taniguchi, M.-A. Song, S. Gallinger, G. Casey, S.N. Thibodeau, L. Le Marchand, M. Tiirikainen, S.A. Mani, W. Zhang, R. V. Davuluri, K. Mimori, M. Mori, A.M. Sieuwerts, J.W.M. Martens, I. Tomlinson, M. Negrini, I. Berindan-Neagoe, J.A. Foekens, S.R. Hamilton, G. Lanza, S. Kopetz, R. Fodde, G.A. Calin, CCAT2, a novel noncoding RNA mapping to 8q24, underlies metastatic progression and chromosomal instability in colon cancer, Genome Res. 23 (2013) 1446–1461. doi:10.1101/gr.152942.112.

[5] R.A. Flynn, H.Y. Chang, Long noncoding RNAs in cell-fate programming and reprogramming, Cell Stem Cell. (2014). doi:10.1016/j.stem.2014.05.014.

[6] J.J. Quinn, H.Y. Chang, Unique features of long non-coding RNA biogenesis and function, Nat. Rev. Genet. (2016). doi:10.1038/nrg.2015.10.

[7] S. Herculano-Houzel, Scaling of brain metabolism with a fixed energy budget per neuron: Implications for neuronal activity, plasticity and evolution, PLoS One. 6 (2011) e17514. doi:10.1371/journal.pone.0017514.

[8] M. Newman, E. Ebrahimie, M. Lardelli, Using the zebrafish model for Alzheimer's disease research, Front. Genet. 5 (2014) 189. doi:10.3389/fgene.2014.00189.

[9] E.L. Bell, T.A. Klimova, J. Eisenbart, C.T. Moraes, M.P. Murphy, G.R.S. Budinger, N.S. Chandel, The Qo site of the mitochondrial complex III is required for the transduction of hypoxic signaling via reactive oxygen species production , J. Cell Biol. 177 (2007) 1029–1036. doi:10.1083/jcb.200609074.

[10] G.J. Lieschke, P.D. Currie, Animal models of human disease: Zebrafish swim into view, Nat. Rev. Genet. 8 (2007) 353–367. doi:10.1038/nrg2091.

[11] K. Dooley, L.I. Zon, Zebrafish: a model system for the study of human disease, Curr. Opin. Genet. Dev. 10 (2000) 252–256. doi:10.1016/S0959-437X(00)00074-5.

[12] E.M. Caramillo, D.J. Echevarria, Alzheimer's disease in the zebrafish: Where can we take it?, Behav. Pharmacol. 28 (2017) 179–186. doi:10.1097/FBP.0000000000000284.

[13] I. Ulitsky, A. Shkumatava, C.H. Jan, H. Sive, D.P. Bartel, Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution, Cell. 147 (2011) 1537–1550. doi:10.1016/j.cell.2011.11.055.

[14] A. Pauli, E. Valen, M.F. Lin, M. Garber, N.L. Vastenhouw, J.Z. Levin, L. Fan, A. Sandelin, J.L. Rinn, A. Regev, A.F. Schier, Systematic identification of long noncoding RNAs expressed during zebrafish embryogenesis., Genome Res. 22 (2012) 577–91. doi:10.1101/gr.133009.111.

[15] K. Kaushik, V.E. Leonard, S. Kv, M.K. Lalwani, S. Jalali, A. Patowary, A. Joshi, V. Scaria, S. Sivasubbu, Dynamic expression of long non-coding RNAs (lncRNAs) in adult zebrafish., PLoS One. 8 (2013) e83616. doi:10.1371/journal.pone.0083616.

[16] K. Howe, M.D. Clark, C.F. Torroja, J. Torrance, C. Berthelot, M. Muffato, J.E. Collins, S. Humphray, K. McLaren, L. Matthews, S. McLaren, I. Sealy, M. Caccamo, C. Churcher, C. Scott, J.C. Barrett, R. Koch, G.-J. Rauch, S. White, W. Chow, B. Kilian, L.T. Quintais, J.A. Guerra-Assunção, Y. Zhou, Y. Gu, J. Yen, J.-H. Vogel, T. Eyre, S. Redmond, R. Banerjee, J. Chi, B. Fu, E. Langley, S.F. Maguire, G.K. Laird, D. Lloyd, E. Kenyon, S. Donaldson, H. Sehra, J. Almeida-King, J. Loveland, S. Trevanion, M. Jones, M. Quail, D. Willey, A. Hunt, J. Burton, S. Sims, K. McLay, B. Plumb, J. Davis, C. Clee, K. Oliver, R. Clark, C. Riddle, D. Elliott, G. Threadgold, G. Harden, D. Ware, S. Begum, B. Mortimore, G. Kerry, P. Heath, B. Phillimore, A. Tracey, N. Corby, M. Dunn, C. Johnson, J. Wood, S. Clark, S. Pelan, G. Griffiths, M. Smith, R. Glithero, P. Howden, N. Barker, C. Lloyd, C. Stevens, J. Harley, K. Holt, G. Panagiotidis, J. Lovell, H. Beasley, C. Henderson, D. Gordon, K. Auger, D. Wright, J. Collins, C. Raisen, L. Dyer, K. Leung, L. Robertson, K. Ambridge, D. Leongamornlert, S. McGuire, R. Gilderthorp, C. Griffiths, D. Manthravadi, S. Nichol, G. Barker, S. Whitehead, M. Kay, J. Brown, C. Murnane, E. Gray, M. Humphries, N. Sycamore, D. Barker, D. Saunders, J. Wallis, A. Babbage, S. Hammond, M. Mashreghi-Mohammadi, L. Barr, S. Martin, P. Wray, A. Ellington, N. Matthews, M. Ellwood, R. Woodmansey, G. Clark, J.D. Cooper, A. Tromans, D. Grafham, C. Skuce, R. Pandian, R. Andrews, E. Harrison, A. Kimberley, J. Garnett, N. Fosker, R. Hall, P. Garner, D. Kelly, C. Bird, S. Palmer, I. Gehring, A. Berger, C.M. Dooley, Z. Ersan-Ürün, C. Eser, H. Geiger, M. Geisler, L. Karotki, A. Kirn, J. Konantz, M. Konantz, M. Oberländer, S. Rudolph-Geiger, M. Teucke, C. Lanz, G. Raddatz, K. Osoegawa, B. Zhu, A. Rapp, S. Widaa, C. Langford, F. Yang, S.C. Schuster, N.P. Carter, J. Harrow, Z. Ning, J. Herrero, S.M.J. Searle, A. Enright, R. Geisler, R.H.A. Plasterk, C. Lee, M. Westerfield, P.J. de Jong, L.I. Zon, J.H. Postlethwait, C. Nüsslein-Volhard, T.J.P. Hubbard, H.R. Crollius, J. Rogers, D.L. Stemple, The zebrafish reference genome sequence and its relationship to the human genome, Nature. 496 (2013) 498–503.

doi:10.1038/nature12111.
[17] T. Mueller, M.F. Wullimann, An Evolutionary Interpretation of Teleostean Forebrain Anatomy, Brain. Behav. Evol. 74 (2009) 30–42. doi:10.1159/000229011.
[18] M.F. Wullimann, B. Rupp, H. Reichert, M.F. Wullimann, B. Rupp, H. Reichert, The brain of the zebrafish Danio rerio: an overview, in: Neuroanat. Zebrafish Brain, Birkhäuser Basel, 1996: pp. 7–17. doi:10.1007/978-3-0348-8979-7_4.
[19] M.F. Wullimann, B. Rupp, H. Reichert, M.F. Wullimann, B. Rupp, H. Reichert, Introduction: neuroanatomy for a neurogenetic model system, in: Neuroanat. Zebrafish Brain, Birkhäuser Basel, 1996: pp. 1–5. doi:10.1007/978-3-0348-8979-7_1.
[20] M.R. Braford, Jr., Comparative Aspects of Forebrain Organization in the Ray-Finned Fishes: Touchstones or Not?, Brain. Behav. Evol. 46 (1995) 259–274. doi:10.1159/000113278.
[21] S. Saleem, R.R. Kannan, Zebrafish: an emerging real-time model system to study Alzheimer's disease and neurospecific drug discovery, Cell Death Discov. 4 (2018) 45. doi:10.1038/s41420-018-0109-7.
[22] V. Tropepe, H.L. Sive, Can zebrafish be used as a model to study the neurodevelopmental causes of autism?, Genes, Brain Behav. 2 (2003) 268–281. doi:10.1034/j.1601-183X.2003.00038.x.
[23] M. Rutenberg-Schoenberg, A.N. Sexton, M.D. Simon, The Properties of Long Noncoding RNAs That Regulate Chromatin, Annu. Rev. Genomics Hum. Genet. 17 (2016) 69–94. doi:10.1146/annurev-genom-090314-024939.
[24] T. Weirick, G. Militello, R. Müller, D. John, S. Dimmeler, S. Uchida, The identification and characterization of novel transcripts from RNA-seq data., Brief. Bioinform. 17 (2015) 1–8. doi:10.1093/bib/bbv067.
[25] S.H.M. Nik, M. Newman, S. Ganesan, M. Chen, R. Martins, G. Verdile, M. Lardelli, Hypoxia alters expression of Zebrafish Microtubule-associated protein Tau (mapta, maptb) gene transcripts, BMC Res. Notes. 7 (2014) 1–9. doi:10.1186/1756-0500-7-767.
[26] D. Antoniou, A. Stergiopoulos, P.K. Politis, Recent advances in the involvement of long non-coding RNAs in neural stem cell biology and brain pathophysiology, Front. Physiol. 5 (2014) 155. doi:10.3389/fphys.2014.00155.
[27] A.D. Ramos, F.J. Attenello, D.A. Lim, Uncovering the roles of long noncoding RNAs in neural development and glioma progression, Neurosci. Lett. 625 (2016) 70–79. doi:10.1016/j.neulet.2015.12.025.
[28] T.C. Roberts, K. V. Morris, M.J.A. Wood, The role of long non-coding RNAs in neurodevelopment, brain function and neurological disease, Philos. Trans. R. Soc. B Biol. Sci. 369 (2014). doi:10.1098/rstb.2013.0507.
[29] S.J. Liu, T.J. Nowakowski, A.A. Pollen, J.H. Lui, M.A. Horlbeck, F.J. Attenello, D. He, J.S. Weissman, A.R. Kriegstein, A.A. Diaz, D.A. Lim, Single-cell analysis of long non-coding RNAs in the developing human neocortex, Genome Biol. 17 (2016) 67. doi:10.1186/s13059-016-0932-1.
[30] M.N. Ziats, O.M. Rennert, Aberrant Expression of Long Noncoding RNAs in Autistic Brain, J. Mol. Neurosci. 49 (2013) 589–593. doi:10.1007/s12031-012-9880-8.
[31] B.M. Kadakkuzha, X.-A. Liu, J. McCrate, G. Shankar, V. Rizzo, A. Afinogenova, B. Young, M. Fallahi, A.C. Carvalloza, B. Raveendra, S. V Puthanveettil, Transcriptome analyses of adult mouse brain reveal enrichment of lncRNAs in specific brain regions and neuronal populations., Front. Cell. Neurosci. 9 (2015) 63. doi:10.3389/fncel.2015.00063.

[32] J.L. Eberling, W.J. Jagust, B.R. Reed, M.G. Baker, Reduced temporal lobe blood flow in alzheimer's disease, Neurobiol. Aging. 13 (1992) 483–491. doi:10.1016/0197-4580(92)90076-A.
[33] X. Zhou, J. Xu, Identification of Alzheimer's disease–associated long noncoding RNAs, Neurobiol. Aging. 36 (2015) 2925–2931. doi:10.1016/j.neurobiolaging.2015.07.015.
[34] C. Shi, L. Zhang, C. Qin, Long non-coding RNAs in brain development, synaptic biology, and Alzheimer's disease, Brain Res. Bull. 132 (2017) 160–169. doi:10.1016/j.brainresbull.2017.03.010.
[35] C.P. Ponting, P.L. Oliver, W. Reik, Evolution and Functions of Long Noncoding RNAs, Cell. 136 (2009) 629–641. doi:10.1016/j.cell.2009.02.006.
[36] R. Brandt, M. Hundelt, N. Shahani, Tau alteration and neuronal degeneration in tauopathies: mechanisms and models, Biochim. Biophys. Acta - Mol. Basis Dis. 1739 (2005) 331–354. doi:10.1016/j.bbadis.2004.06.018.
[37] K. Yasojima, E.G. McGeer, P.L. McGeer, Tangled areas of Alzheimer brain have upregulated levels of exon 10 containing tau mRNA, Brain Res. 831 (1999) 301–305. doi:10.1016/S0006-8993(99)01486-9.
[38] J. Walter, Fishing for function - distinct roles of Bace1 and Bace2 in Zebrafish development, J. Neurochem. 127 (2013) 435–437. doi:10.1111/jnc.12200.
[39] T.-P. V. Huynh, A.A. Davis, J.D. Ulrich, D.M. Holtzman, Apolipoprotein E and Alzheimer's disease: the influence of apolipoprotein E on amyloid-β and other amyloidogenic proteins, J. Lipid Res. 58 (2017) 824–836. doi:10.1194/jlr.R075481.
[40] M.E. Belloy, V. Napolioni, M.D. Greicius, A Quarter Century of APOE and Alzheimer's Disease: Progress to Date and the Path Forward, Neuron. 101 (2019) 820–838. doi:10.1016/j.neuron.2019.01.056.
[41] P.A. Raymond, L.K. Barthel, R.L. Bernardos, J.J. Perkowski, Molecular characterization of retinal stem cells and their niches in adult zebrafish, BMC Dev. Biol. 6 (2006) 36. doi:10.1186/1471-213X-6-36.
[42] Z. Pujic, Y. Omori, M. Tsujikawa, B. Thisse, C. Thisse, J. Malicki, Reverse genetic analysis of neurogenesis in the zebrafish retina, Dev. Biol. 293 (2006) 330–347. doi:10.1016/j.ydbio.2005.12.056.
[43] K.N. Veth, J.R. Willer, R.F. Collery, M.P. Gray, G.B. Willer, D.S. Wagner, M.C. Mullins, A.J. Udvadia, R.S. Smith, S.W.M. John, R.G. Gregg, B.A. Link, Mutations in Zebrafish lrp2 Result in Adult-Onset Ocular Pathogenesis That Models Myopia and Other Risk Factors for Glaucoma, PLoS Genet. 7 (2011) e1001310. doi:10.1371/journal.pgen.1001310.
[44] S.H. Moussavi Nik, L. Wilson, M. Newman, K. Croft, T.A. Mori, I. Musgrave, M. Lardelli, The BACE1-PSEN-AβPP regulatory axis has an ancient role in response to low oxygen/oxidative stress, J. Alzheimer's Dis. 28 (2012) 515–530. doi:10.3233/JAD-2011-110533.
[45] R. Pluta, Pathological Opening of the Blood-Brain Barrier to Horseradish Peroxidase and Amyloid Precursor Protein following Ischemia-Reperfusion Brain Injury, Chemotherapy. 51 (2005) 223–226. doi:10.1159/000086924.
[46] R. De Gasperi, M.A.G. Sosa, S. Dracheva, G.A. Elder, Presenilin-1 regulates induction of hypoxia inducible factor-1α: Altered activation by a mutation associated with familial Alzheimer's disease, Mol. Neurodegener. 5 (2010) 38. doi:10.1186/1750-1326-5-38.
[47] X. Zhang, W. Le, Pathological role of hypoxia in Alzheimer's disease, Exp. Neurol. 223

(2010) 299–303. doi:10.1016/j.expneurol.2009.07.033.
[48] A.M. Sandbichler, B. Jansen, B.A. Peer, M. Paulitsch, B. Pelster, M. Egg, Metabolic Plasticity Enables Circadian Adaptation to Acute Hypoxia in Zebrafish Cells, Cell. Physiol. Biochem. 46 (2018) 1159–1174. doi:10.1159/000489058.
[49] Y. Wu, D. Tang, N. Liu, W. Xiong, H. Huang, Y. Li, Z. Ma, H. Zhao, P. Chen, X. Qi, E.E. Zhang, Reciprocal Regulation between the Circadian Clock and Hypoxia Signaling at the Genome Level in Mammals, Cell Metab. 25 (2017) 73–85. doi:10.1016/j.cmet.2016.09.009.
[50] M. Egg, L. Köblitz, J. Hirayama, T. Schwerte, C. Folterbauer, A. Kurz, B. Fiechtner, M. Möst, W. Salvenmoser, P. Sassone-Corsi, B. Pelster, Linking Oxygen to Time: The Bidirectional Interaction Between the Hypoxic Signaling Pathway and the Circadian Clock, Chronobiol. Int. 30 (2013) 510–529. doi:10.3109/07420528.2012.754447.
[51] L.X. Pierce, R.R. Noche, O. Ponomareva, C. Chang, J.O. Liang, Novel functions for Period 3 and Exo-rhodopsin in rhythmic transcription and melatonin biosynthesis within the zebrafish pineal organ, Brain Res. 1223 (2008) 11–24. doi:https://doi.org/10.1016/j.brainres.2008.05.020.
[52] F. Bellanti, G. Iannelli, M. Blonda, R. Tamborra, R. Villani, A. Romano, S. Calcagnini, G. Mazzoccoli, M. Vinciguerra, S. Gaetani, A.M. Giudetti, G. Vendemiale, T. Cassano, G. Serviddio, Alterations of Clock Gene RNA Expression in Brain Regions of a Triple Transgenic Model of Alzheimer's Disease, J. Alzheimer's Dis. 59 (2017) 615–631. doi:10.3233/JAD-160942.
[53] A. Dobin, C.A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, T.R. Gingeras, STAR: ultrafast universal RNA-seq aligner, Bioinformatics. 29 (2013) 15–21. doi:10.1093/bioinformatics/bts635.
[54] M. Pertea, D. Kim, G.M. Pertea, J.T. Leek, S.L. Salzberg, Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown, Nat. Protoc. 11 (2016) 1650–1667. doi:10.1038/nprot.2016.095.
[55] V. Wucher, F. Legeai, B. Hédan, G. Rizk, L. Lagoutte, T. Leeb, V. Jagannathan, E. Cadieu, A. David, H. Lohi, S. Cirera, M. Fredholm, N. Botherel, P.A.J. Leegwater, C. Le Béguec, H. Fieten, J. Johnson, J. Alföldi, C. André, K. Lindblad-Toh, C. Hitte, T. Derrien, FEELnc: A tool for long non-coding RNA annotation and its application to the dog transcriptome, Nucleic Acids Res. 45 (2017) gkw1306. doi:10.1093/nar/gkw1306.
[56] Y.J. Kang, D.C. Yang, L. Kong, M. Hou, Y.Q. Meng, L. Wei, G. Gao, CPC2: A fast and accurate coding potential calculator based on sequence intrinsic features, Nucleic Acids Res. 45 (2017) W12–W16. doi:10.1093/nar/gkx428.
[57] J.-C. Guo, S.-S. Fang, Y. Wu, J.-H. Zhang, Y. Chen, J. Liu, B. Wu, J.-R. Wu, E.-M. Li, L.-Y. Xu, L. Sun, Y. Zhao, CNIT: a fast and accurate web tool for identifying protein-coding and long non-coding transcripts based on intrinsic sequence composition, Nucleic Acids Res. 47 (2019) W516–W522. doi:10.1093/nar/gkz400.
[58] X. Hu, W. Chen, J. Li, S. Huang, X. Xu, X. Zhang, S. Xiang, C. Liu, ZFLNC: a comprehensive and well-annotated database for zebrafish lncRNA, Database. 2018 (2018). doi:10.1093/database/bay114.
[59] S.X. Ge, D. Jung, R. Yao, ShinyGO: a graphical gene-set enrichment tool for animals and plants, Bioinformatics. 36 (2019) 2628–2629. doi:10.1093/bioinformatics/btz931.
[60] Y. Liao, G.K. Smyth, W. Shi, featureCounts: an efficient general purpose program for assigning sequence reads to genomic features, Bioinformatics. 30 (2013) 923–930.

doi:10.1093/bioinformatics/btt656.
[61] A. Conesa, S. Götz, J.M. García-Gómez, J. Terol, M. Talón, D. Genómica, I. Valenciano, D.I. Agrarias, U.P. De Valencia, M. Robles, Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research, Bioinformatics. 21 (2005) 3674–3676. doi:10.1093/bioinformatics/bti610.
[62] S.X. Ge, E.W. Son, R. Yao, iDEP: an integrated web application for differential expression and pathway analysis of RNA-Seq data, BMC Bioinformatics. 19 (2018) 534. doi:10.1186/s12859-018-2486-6.
[63] P. Langfelder, S. Horvath, WGCNA: an R package for weighted correlation network analysis, BMC Bioinformatics. 9 (2008) 559. doi:10.1186/1471-2105-9-559.
[64] P. Shannon, A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, Cytoscape: a software environment for integrated models of biomolecular interaction networks., Genome Res. 13 (2003) 2498–504. doi:10.1101/gr.1239303.
[65] B. Banerjee, D. Koner, R. Hasan, N. Saha, Molecular characterization and ornithine-urea cycle genes expression in air-breathing magur catfish (Clarias magur) during exposure to high external ammonia, Genomics. 112 (2020) 2247–2260. doi:10.1016/j.ygeno.2019.12.021.
[66] K.J. Livak, T.D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the 2−ΔΔCT method, Methods. 25 (2001) 402–408. doi:http://dx.doi.org/10.1006/meth.2001.1262.

Figure legends:

Fig. 1. Pipeline for the identification of novel lncRNAs from the normoxic and hypoxic zebrafish brain regions.

Fig. 2. Classification of lncRNAs. (A) Numbers of lncRNAs in each of the four main classes, as defined by their genomic location relative to neighboring or overlapping genes. (B) Length distribution of lncRNAs. X-axis: length of lncRNAs in bp; Y-axis: count. (C) Venn diagram showing novel lncRNAs in three different brain regions. Fb: forebrain, Mb: midbrain, Hb: hindbrain.

Fig. 3. Hierarchical clustering tree summarizing the correlation among significantly enriched pathways of candidate lncRNAs protein-coding targets. Pathways with many shared genes are clustered together. Bigger dots indicate more significant P-values. (A) Enrichment using GO Biological process. (B) Enrichment using KEGG pathway.

Fig. 4. (A) Venn diagram showing the DE lncRNAs in three different brain regions between hypoxic and normoxic fish (B) Venn diagram showing the DE mRNAs in three different brain regions between hypoxic and normoxic fish. Fbc: forebrain normoxic, Fbt: forebrain hypoxic, Mbc: midbrain normoxic, Mbt: midbrain hypoxic, Hbc: hindbrain normoxic, Hbt: hindbrain hypoxic

Fig. 5. Hierarchical clustering with heatmap. (A) DE lncRNAs in all three brain tissues between hypoxic and normoxic fish. (B) DE mRNAs in all three brain tissues between hypoxic and normoxic fish. The red represents upregulated expression, and the green code represents downregulated expression.

Fig. 6. (A) Hierarchical clustering dendrogram of co-expressed lncRNAs and mRNAs in modules. This dendrogram was generated by unsupervised hierarchical clustering of genes using topological overlap. The colored strips below each dendrogram demonstrate the module designation that was identified through the clusters of co-expressed genes. (B) Heatmap of identified modules. (C) The co-expression network of lncRNAs and protein-coding genes from Blue module. The yellow triangles denote the lncRNAs.

Fig. 7. The RT-qPCR validation results of DE lncRNAs for (A) Fb (C) Mb and (E) Hb and DE mRNAs for (B) Fb, (D) Mb, and (F) Hb regions of the zebrafish brain, respectively.

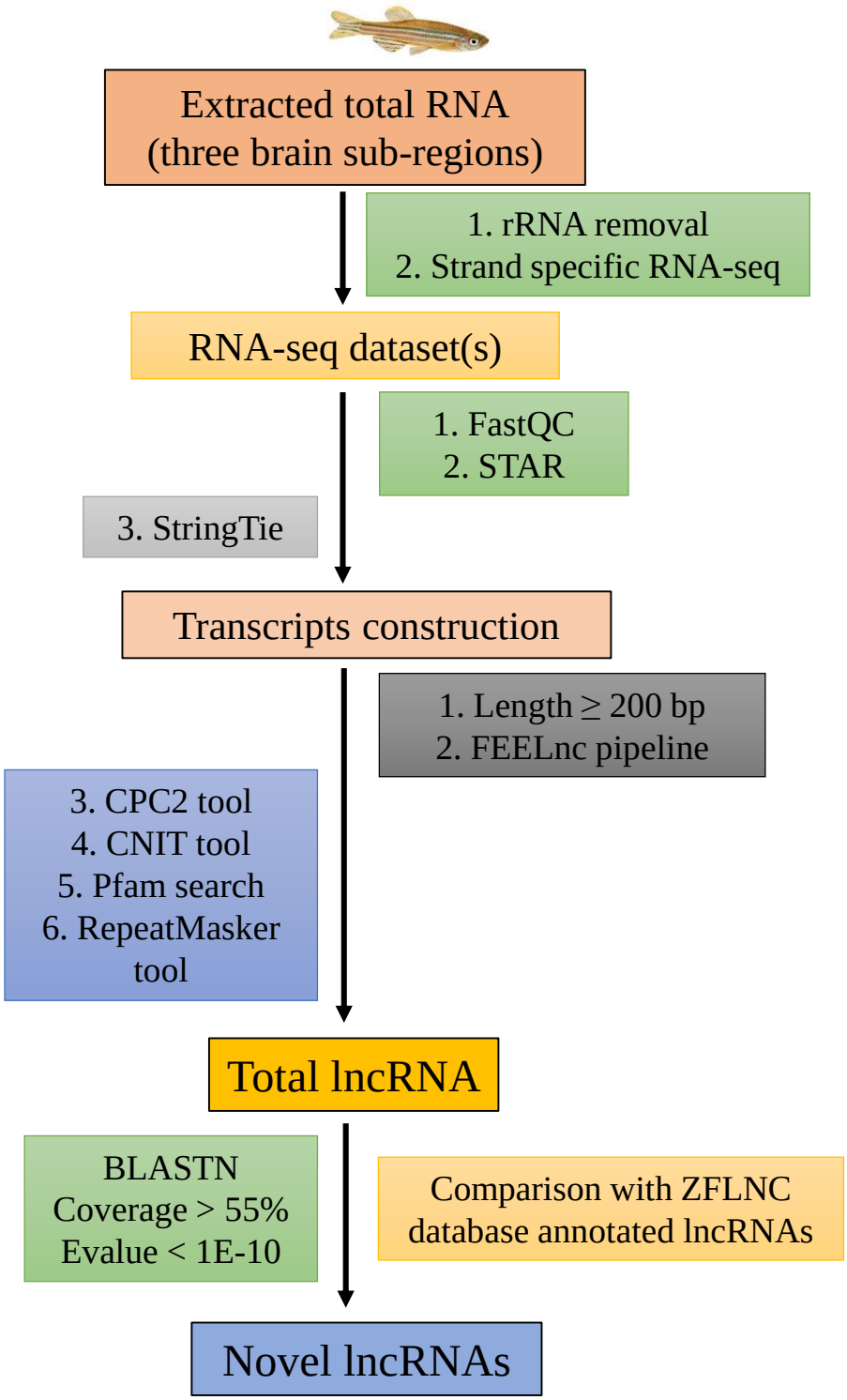
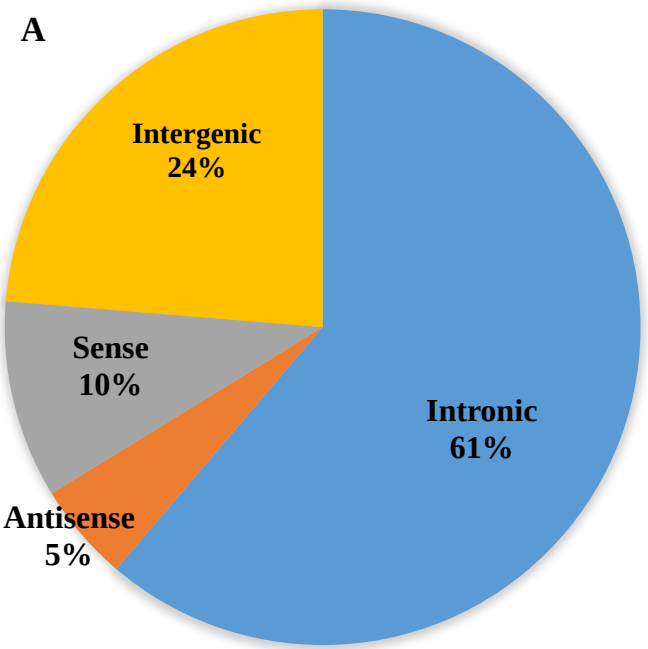
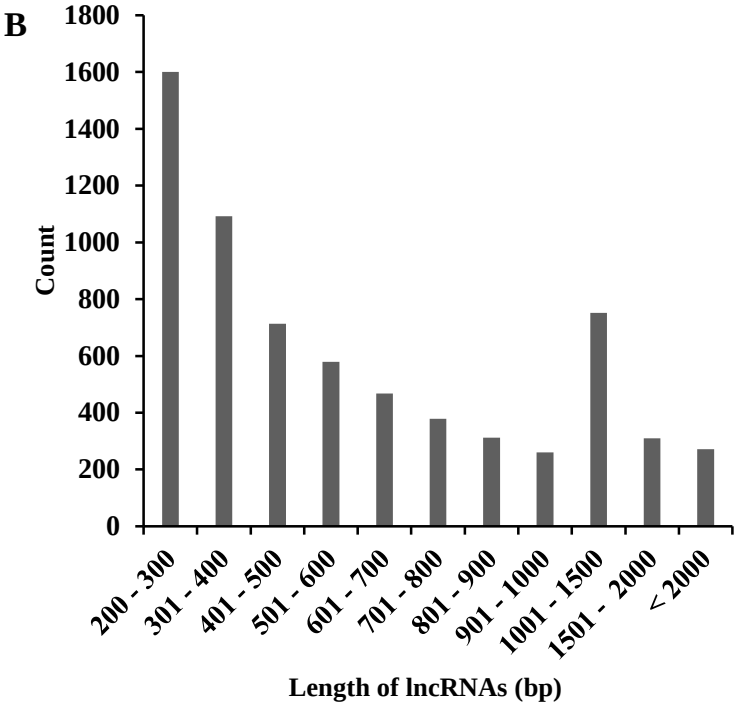

Extracted total RNA
(three brain sub-regions)
1. rRNA removal
2. Strand specific RNA-seq
RNA-seq dataset(s)
1. FastQC
2. STAR
3. StringTie
Transcripts construction
1. Length ≥ 200 bp
2. FEELnc pipeline
3. CPC2 tool
4. CNIT tool
5. Pfam search
6. RepeatMasker tool
Total lncRNA
BLASTN
Coverage > 55%
Evalue < 1E-10
Comparison with ZFLNC database annotated lncRNAs
Novel lncRNAs

Fig. 1


A
Intergenic
24%
Sense
10%
Antisense
5%
Intronic
61%


B
1800
1600
1400
1200
1000
800
600
400
200
0
Count
200 - 300
301 - 400
401 - 500
501 - 600
601 - 700
701 - 800
801 - 900
901 - 1000
1001 - 1500
1501 - 2000
< 2000
Length of lncRNAs (bp)

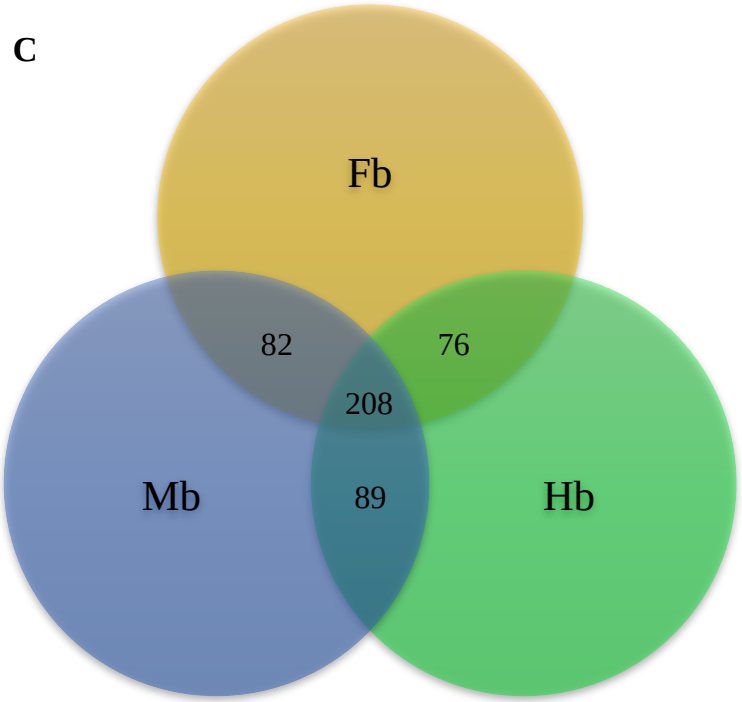

C
Fb
Mb
Hb
82
76
208
89

Fig. 2

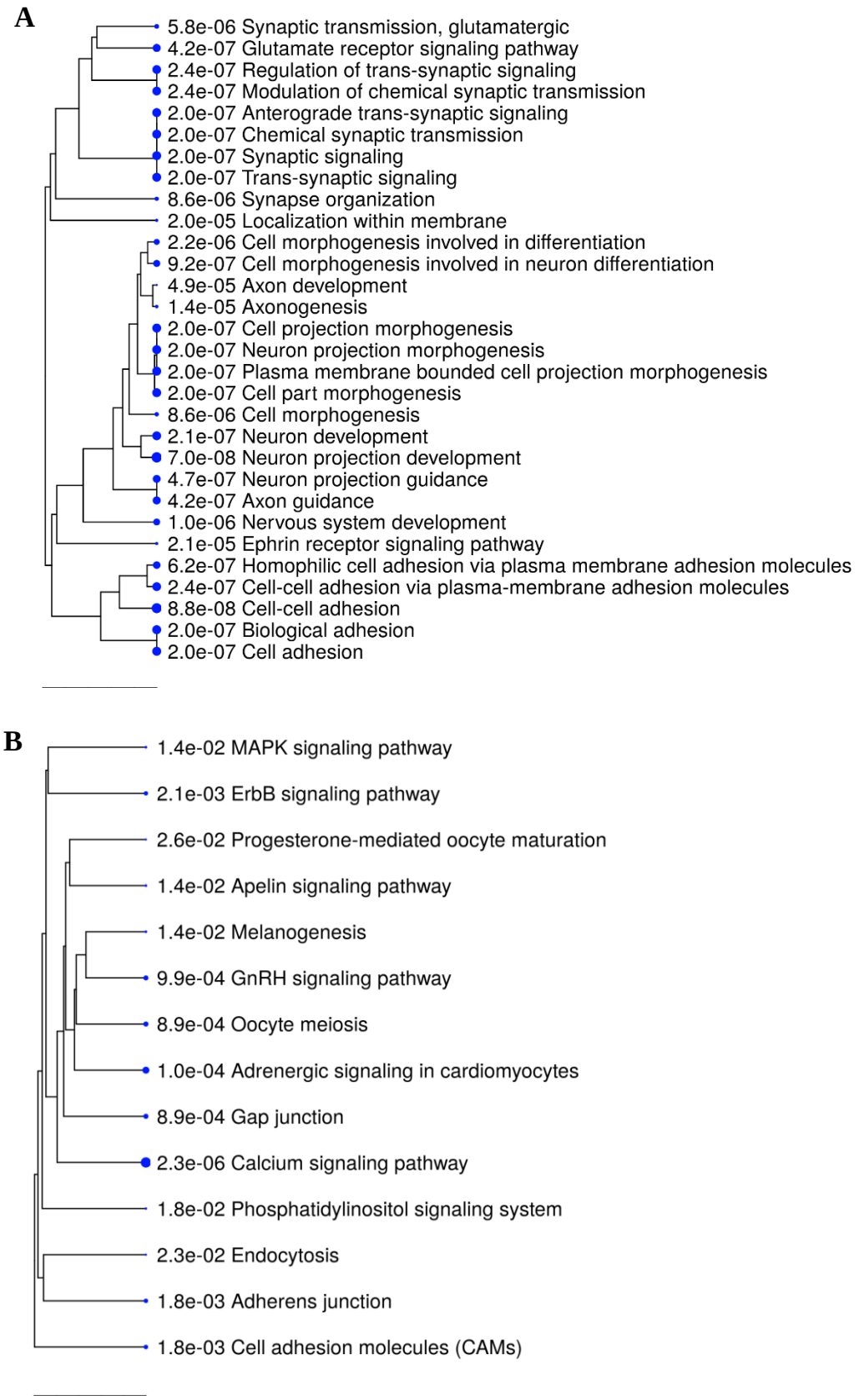

A
5.8e-06 Synaptic transmission, glutamatergic
4.2e-07 Glutamate receptor signaling pathway
2.4e-07 Regulation of trans-synaptic signaling
2.4e-07 Modulation of chemical synaptic transmission
2.0e-07 Anterograde trans-synaptic signaling
2.0e-07 Chemical synaptic transmission
2.0e-07 Synaptic signaling
2.0e-07 Trans-synaptic signaling
8.6e-06 Synapse organization
2.0e-05 Localization within membrane
2.2e-06 Cell morphogenesis involved in differentiation
9.2e-07 Cell morphogenesis involved in neuron differentiation
4.9e-05 Axon development
1.4e-05 Axonogenesis
2.0e-07 Cell projection morphogenesis
2.0e-07 Neuron projection morphogenesis
2.0e-07 Plasma membrane bounded cell projection morphogenesis
2.0e-07 Cell part morphogenesis
8.6e-06 Cell morphogenesis
2.1e-07 Neuron development
7.0e-08 Neuron projection development
4.7e-07 Neuron projection guidance
4.2e-07 Axon guidance
1.0e-06 Nervous system development
2.1e-05 Ephrin receptor signaling pathway
6.2e-07 Homophilic cell adhesion via plasma membrane adhesion molecules
2.4e-07 Cell-cell adhesion via plasma-membrane adhesion molecules
8.8e-08 Cell-cell adhesion
2.0e-07 Biological adhesion
2.0e-07 Cell adhesion
B
1.4e-02 MAPK signaling pathway
2.1e-03 ErbB signaling pathway
2.6e-02 Progesterone-mediated oocyte maturation
1.4e-02 Apelin signaling pathway
1.4e-02 Melanogenesis
9.9e-04 GnRH signaling pathway
8.9e-04 Oocyte meiosis
1.0e-04 Adrenergic signaling in cardiomyocytes
8.9e-04 Gap junction
2.3e-06 Calcium signaling pathway
1.8e-02 Phosphatidylinositol signaling system
2.3e-02 Endocytosis
1.8e-03 Adherens junction
1.8e-03 Cell adhesion molecules (CAMs)

Fig. 3

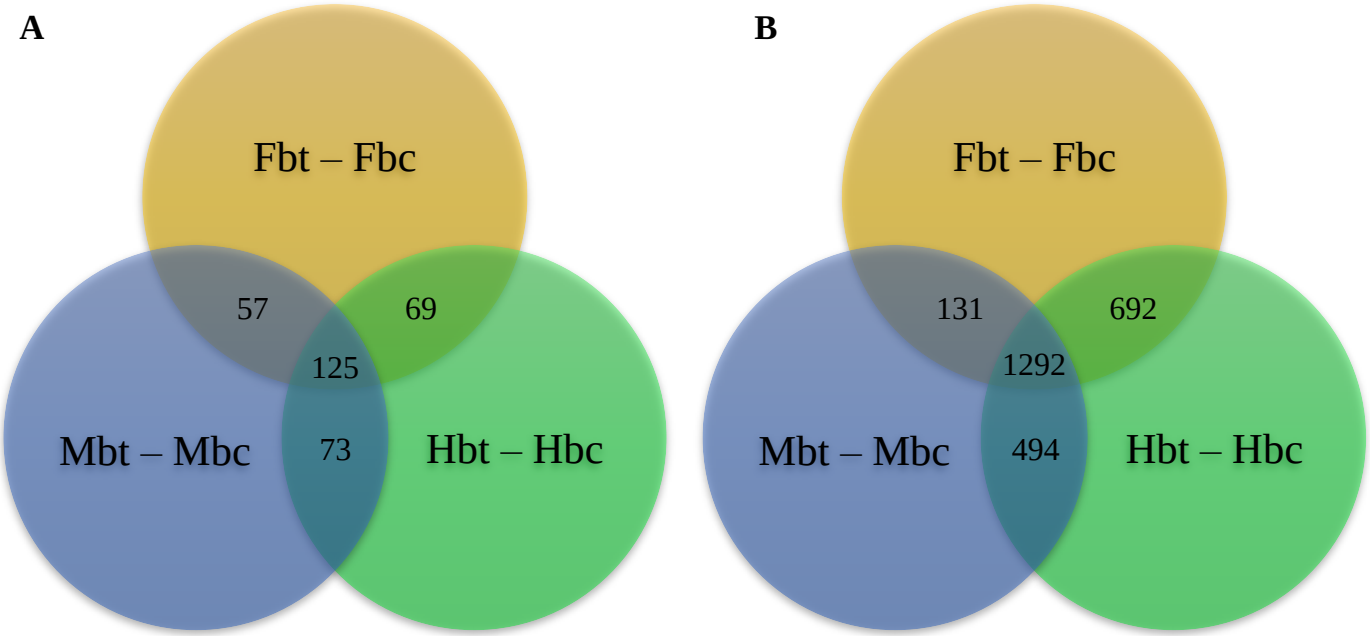

A
Fbt – Fbc
57
69
125
Mbt – Mbc
73
Hbt – Hbc
B
Fbt – Fbc
131
692
1292
Mbt – Mbc
494
Hbt – Hbc

Fig. 5

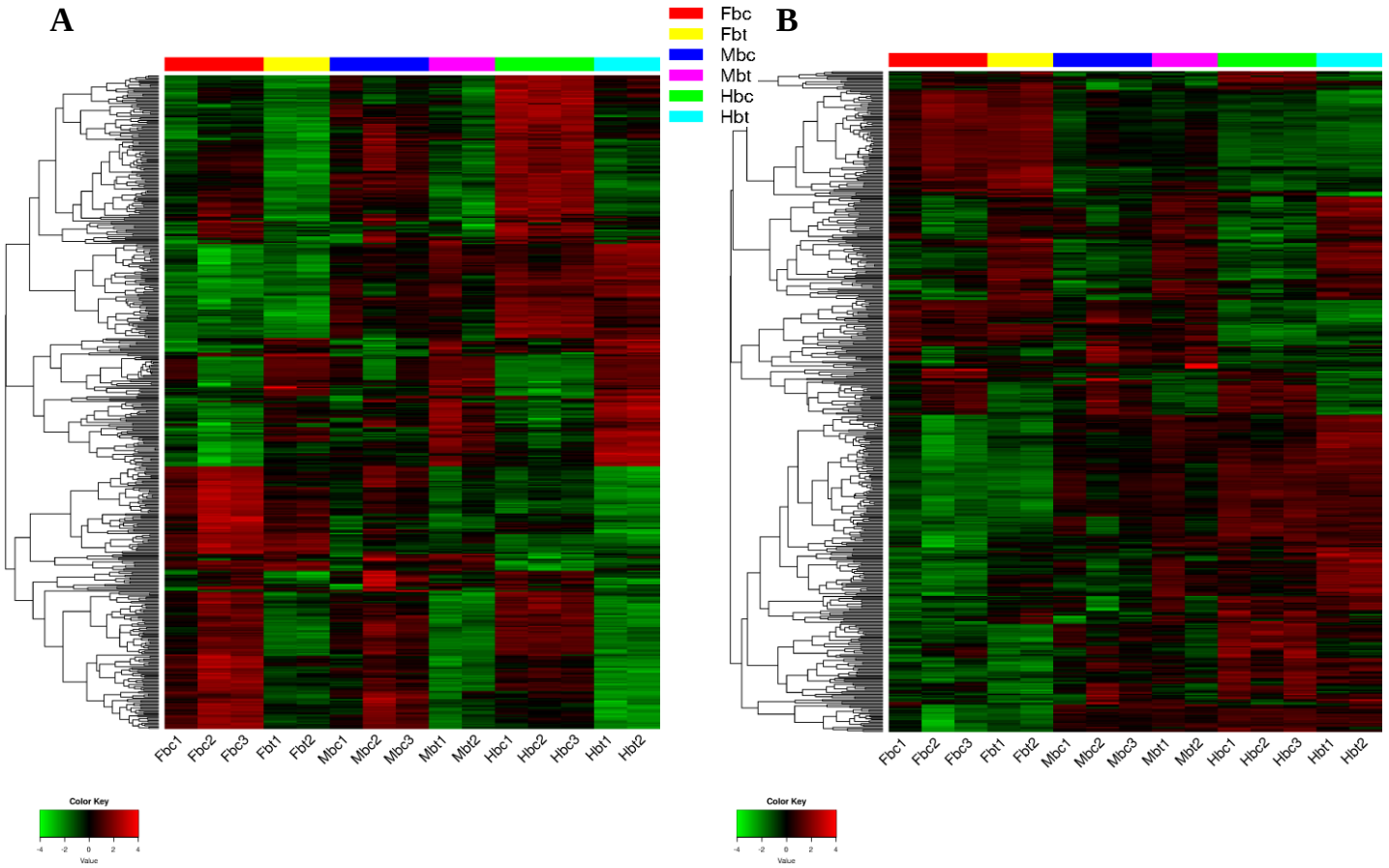

A
B
Fbc
Fbt
Mbc
Mbt
Hbc
Hbt
Fbc1
Fbc2
Fbc3
Fbt1
Fbt2
Mbc1
Mbc2
Mbc3
Mbt1
Mbt2
Hbc1
Hbc2
Hbc3
Hbt1
Hbt2
Color Key
-4
-2
0
2
4
Value

Fig. 5

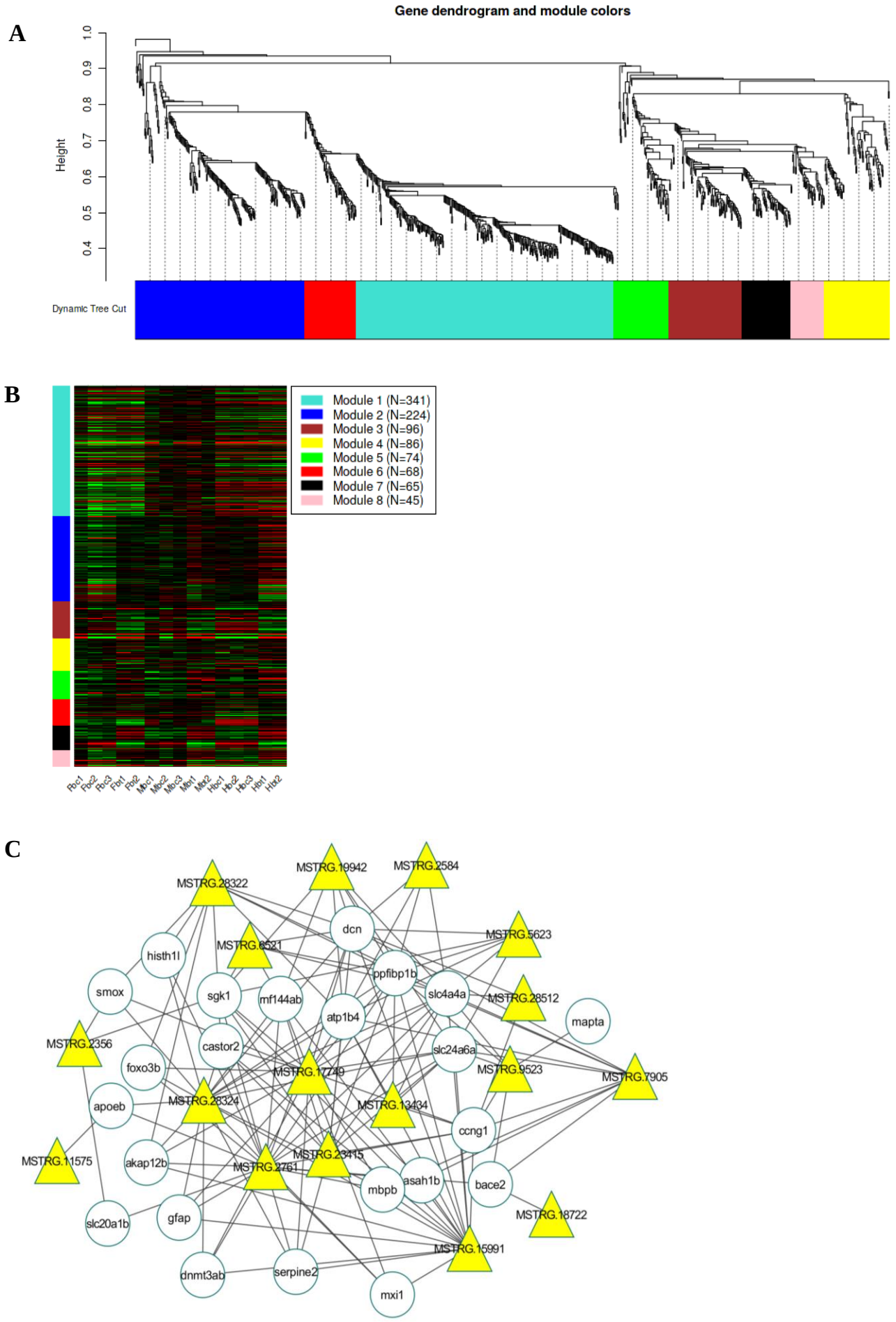

A
Gene dendrogram and module colors
Height
1.0
0.9
0.8
0.7
0.6
0.5
0.4
Dynamic Tree Cut
B
Module 1 (N=341)
Module 2 (N=224)
Module 3 (N=96)
Module 4 (N=86)
Module 5 (N=74)
Module 6 (N=68)
Module 7 (N=65)
Module 8 (N=45)
Fbc1
Fbc2
Fbc3
Fbt1
Fbt2
Mbc1
Mbc2
Mbc3
Mbt1
Mbt2
Hbc1
Hbc2
Hbc3
Hbt1
Hbt2
C
MSTRG.19942
MSTRG.2584
MSTRG.28322
dcn
MSTRG.5623
histh1l
MSTRG.6521
ppfibp1b
smox
sgk1
mf144ab
slc4a4a
MSTRG.28512
atp1b4
mapta
MSTRG.2356
castor2
slc24a6a
foxo3b
MSTRG.17749
MSTRG.9523
MSTRG.7905
apoeb
MSTRG.28324
MSTRG.13434
ccng1
MSTRG.11575
akap12b
MSTRG.23415
MSTRG.2761
mbpb
asah1b
bace2
MSTRG.18722
slc20a1b
gfap
MSTRG.15991
dnmt3ab
serpine2
mxi1

Fig. 6

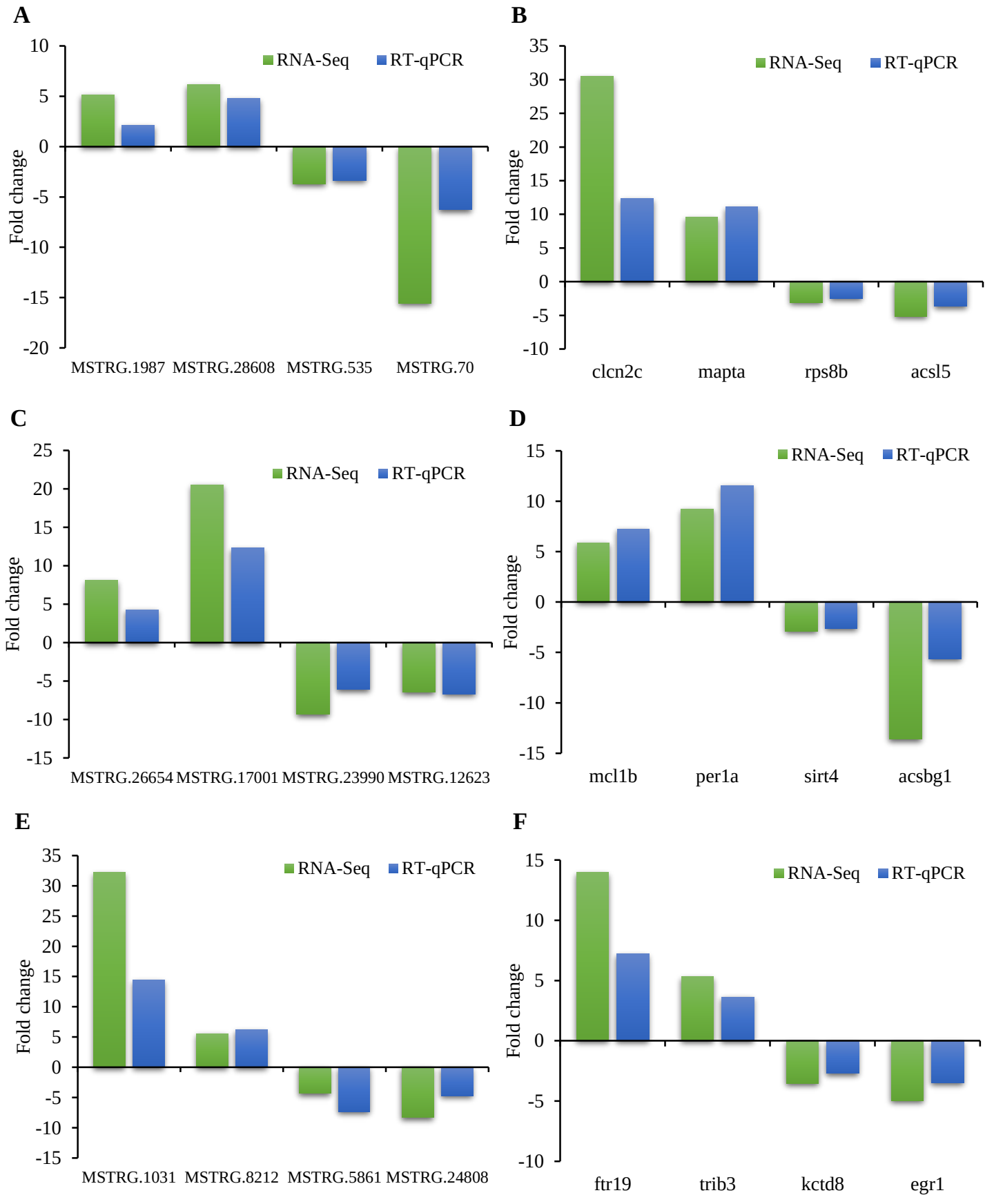
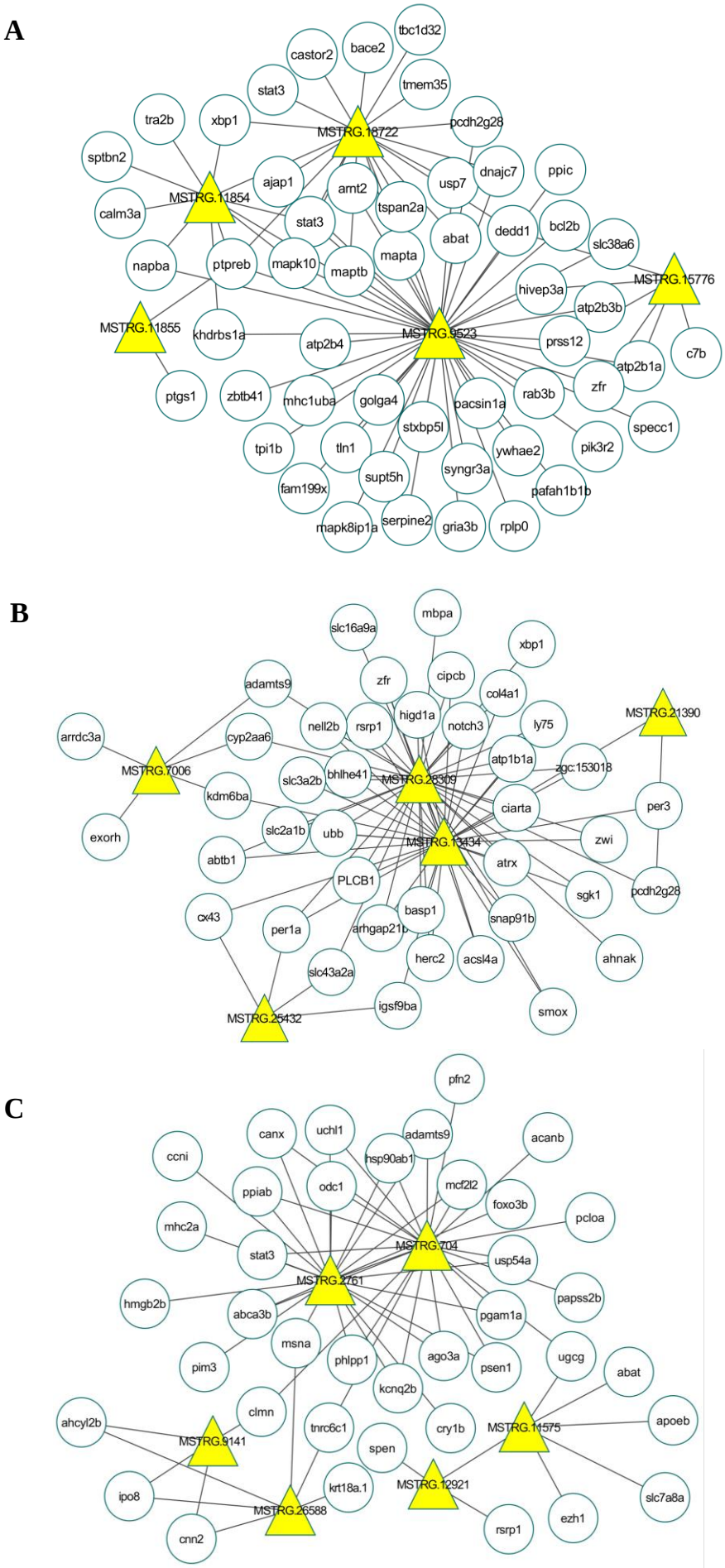

A
RNA-Seq
RT-qPCR
Fold change
10
5
0
-5
-10
-15
-20
MSTRG.1987
MSTRG.28608
MSTRG.535
MSTRG.70
B
RNA-Seq
RT-qPCR
Fold change
35
30
25
20
15
10
5
0
-5
-10
clcn2c
mapta
rps8b
acsl5
C
RNA-Seq
RT-qPCR
Fold change
25
20
15
10
5
0
-5
-10
-15
MSTRG.26654
MSTRG.17001
MSTRG.23990
MSTRG.12623
D
RNA-Seq
RT-qPCR
Fold change
15
10
5
0
-5
-10
-15
mcl1b
per1a
sirt4
acsbg1
E
RNA-Seq
RT-qPCR
Fold change
35
30
25
20
15
10
5
0
-5
-10
-15
MSTRG.1031
MSTRG.8212
MSTRG.5861
MSTRG.24808
F
RNA-Seq
RT-qPCR
Fold change
15
10
5
0
-5
-10
ftr19
trib3
kctd8
egr1

Fig. 7



Supplementary fig. S1. The co-expression network of DE novel lncRNAs and mRNAs. The yellow triangles denote lncRNAs and circles denote mRNAs. (A) Fb region (B) Mb region (C) Hb region. Yellow Triangle: lncRNAs; Circle: mRNAs