International Journal of Biochemistry and Cell Biology 111 (2019) 19–26

Contents lists available at ScienceDirect

International Journal of Biochemistry and Cell Biology

journal homepage: www.elsevier.com/locate/biocel

Identification of chimeric RNAs in human infant brains and their implications in neural differentiation

Yue Tang[a,b,c], Shanshan Ma[a], Xinxin Wang[d], Qu Xing[a], Tuanjie Huang[a], Hongtao Liu[a], Qinghua Li[a], Yanting Zhang[a], Kun Zhang[a], Minghao Yao[a], Greta Luyuan Yang[e], Hui Li[b,f], Xingxing Zang[g], Bo Yang[h], Fangxia Guan[a,c,i,*]

[a] School of Life Sciences, Zhengzhou University, Zhengzhou, Henan, China
[b] School of Basic Medical Sciences, Zhengzhou University, Zhengzhou, Henan, China
[c] Henan Provincial People's Hospital, Zhengzhou, 450003, Henan, China
[d] Department of Gynecology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, China
[e] Stuyvesant High School, NY, USA
[f] Department of Pathology, School of Medicine, University of Virginia, Charlottesville, VA, USA
[g] Einstein College of Medicine, Department of Microbiology and Immunology, 1300 Morris Park Ave, Bronx, NY, 10461, USA
[h] Department of Neurosurgery, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, China
[i] The Translational Medical Center of First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, China

ARTICLE INFO

Keywords:
Chimeric RNAs
DUS4L-BCAP29
Brain
Mesenchymal stem cells
Neural differentiation

ABSTRACT

Chimeric RNAs are transcripts composed of RNA fragments from different genes and are traditionally well-known cancer-causing genetic events. Recent studies show chimeric RNAs being present in multiple non-neoplastic tissues and cells, suggesting that at least some may have roles in normal physiology. However, chimeric RNAs and their implications in brain development and neural differentiation have not been formally studied. Here, we firstly characterized the landscape of chimeric RNAs in human infant brain tissues and identified 599 chimeric RNAs. Through a series of filtering, 22 were selected and tested in a neural differentiation process starting from stem cells. Ten were validated experimentally. One of these ten chimeric RNAs, *DUS4L-BCAP29*, dramatically increased when human umbilical mesenchymal stem cells were induced for neural differentiation. Consistently, we found that overexpressed *DUS4L-BCAP29* effectively promoted neural differentiation. Our results support the important role(s) chimeric RNAs play in neural differentiation, and are consistent with the new notion that chimeric RNAs also exist in normal physiology, and likely serve biological purposes.

1. Introduction

Chimeric RNAs are transcripts composed of RNA fragments from different genes (Li et al., 2018). Bioinformatics analyses and second generation sequencing technologies have facilitate the discovery of chimeric RNAs, and have led to the rapid expansion of the repertoire of chimeric RNAs (ADAMS et al., 1991; BENELLI et al., 2012; Carrara et al., 2013; Chen et al., 2012; Chuang et al., 2016; Davidson et al., 2015; Francis et al., 2012; Ge et al., 2011; Gorohovski et al., 2017; Hoffmann et al., 2009; Kim et al., 2007; Kim and Salzberb, 2011; Kong et al., 2011; Kumar et al., 2016a, b; Langmead et al., 2009; Lee et al., 2017; Li et al., 2011; Liu et al., 2016; Lu and Matera, 2014; Mcpherson et al., 2011, 2012; Novo et al., 2007; Prakash et al., 2010; Suhre and Claverie, 2004; Wang et al., 2015; wu et al., 2014). The widespread detection of chimeric RNAs has far exceeded previous expectations, making them a newly popular focus in transcriptomics. Some chimeric RNAs have been known to contribute to cancer diagnosis and treatment. Prominent examples of this include *BCR-ABL1*, which is targeted by Gleevec in chronic myelogenous leukemia (Druker, 2001; Rowley, 1973); *EML4-ALK,* which has improved patient prognosis when treated with ALK inhibitors in non-small cell lung cancer patients (Shaw et al., 2011; Soda et al., 2007); and *SLC45A3-ELK4,* a potential biomarker in prostate cancer (Zhang et al., 2012; Rickman et al., 2009). As more and more chimeric RNAs are being identified in non-cancerous tissues and cells (Babiceanu et al., 2016; Li et al., 2008; Wu et al., 2014; Yuan et al., 2013), it is postulated that at least some may play important roles in normal physiology. Supporting this notion, are two chimeric RNAs, *CTBS-GNG5* and *CTNNBIP1-CLSTN1,* which have already been

* Corresponding author at: School of Life Sciences, Zhengzhou University, No.100 Science Avenue, Zhengzhou, 450001, Henan, China.
E-mail addresses: guanfangxia@126.com, fxguan@126.com (F. Guan).

https://doi.org/10.1016/j.biocel.2019.03.012
Received 31 October 2018; Received in revised form 6 March 2019; Accepted 30 March 2019
Available online 05 April 2019
1357-2725/

demonstrated to be involved in regulating normal cell survival, proliferation, and apoptosis (Babiceanu et al., 2016). Studying these chimeric RNAs may not only enhance our understanding of the genome and transcriptome, but also provide novel insights into physiological processes.

A large number of chimeric RNAs have been detected in 30 different normal human tissues and cells (Babiceanu et al., 2016). The profiles of these chimeric RNAs were found to be cell and tissue specific. To achieve a panoramic view of chimeric RNAs in normal brain development, we analysed the RNA-seq data of human infant brain tissues. 599 chimeric RNAs were identified, and the landscape of chimeric RNAs was characterized. A subset of them were selected and validated in a neural differentiation system using human umbilical cord-derived mesenchymal stem cells (hUC-MSCs). We then focused on one chimeric RNA, *DUS4L-BCAP29,* and found that its expression dramatically increased during differentiation. Lastly, we provided evidence that *DUS4L-BCAP29* overexpression is sufficient to promote neural differentiation. These results provide a genome level overview of the chimeric RNAs in brain development and neural differentiation, and lend support to the notion that chimeric RNAs are involved in normal physiology.

2. Results

2.1. Identification of chimeric RNAs in human infant brain tissues

Here we used nine paired-end RNA-seq of dorsolateral prefrontal cortex (DLPFC) from infant brain tissues (Supplementary Table S1) to uncover normal chimeric RNAs during brain development. The bioinformatic software, SOAPfuse (Babiceanu et al., 2016; Kumar et al., 2016a,b; Qin et al., 2015) was used to uncover a total of 599 chimeric RNAs. The landscape of the chimeras was illustrated using Circos plots (Fig. 1A). Chromosome 1 contained the most chimeric RNAs (72 chimeric RNAs), while chromosome 18 had the fewest, with only 3 chimeric RNAs. The number of chimeric RNAs correlates to the number of genes on the same chromosome ($R = 0.8357$, $p < 0.0001$) (Fig. 1B).

We examined gene ontology (GO) terms for the parental genes involved in chimeric RNAs. The predominant functions of upstream parental genes are protein polymerization and post-translational protein folding (Fig. 1C). Downstream parental genes are mainly related to small molecule or protein metabolism, and glutamate secretion (Fig. 1D).

2.2. Classification of chimeric RNAs in DLPFCs of infant brain tissues

We then used several classifications to categorize these chimeras in order to explore their potential functions and mechanisms. Chimeric RNAs were first classified according to the chromosomal location of their parental genes: parental genes located on different chromosomes (INTERCHR), neighbouring genes transcribing the same strand (INTRACHR-SS-0GAP) and other fusions with parental genes on the same chromosome (INTRACHR-OTHER). The proportions of the three types of chimeric RNAs in this study were 46%, 51% and 3%, respectively (Fig. 2A). It should be noted that over half of the chimeric RNAs identified in this study were INTRACHR-SS-0GAP (51%). Such result is in conformity with a forefather study (Babiceanu et al., 2016). INTRACHR-SS-0GAP chimeric RNAs received more attention due to their high validation rate (Huang et al., 2017; Kannan et al., 2011; Nacu et al., 2011; Xie et al., 2016).

The junction position relative to the exon of the parental genes may be associated with the generating mechanism of the fusion transcripts. Therefore we categorized the fusions accordingly: both sides being known exon boundaries (E/E); one side being exon boundary, the other not (E/M or M/E); both sides falling into the middle of exons (M/M). The proportions for these three categories in this study were 36%, 21% and 43%, respectively (Fig. 2B).

We also subdivided the fusions according to the reading frames: the known protein coding sequence of the 3′ gene uses a different reading frame than the 5′ gene (frame-shift); the known reading frame of the 3′ gene is the same as the 5′gene (in-frame); no effect on the reading frame of the parental genes (NA). NA could occur when the fusion junction falls into the 3′ UTR of the 3′ parental gene, or no known protein coding sequence for the 3′gene, or the 5′ portion of the fusion will not affect the CDS of the 3′ gene; a very small population of fusions fell into the 'both' category, which could be in-frame or frame-shift depending on the alternative splicing isoforms of the parental genes. The proportions of these four subcategories in this study were 21%, 15%, 63%, and 1%, respectively (Fig. 2C). The in-frame fusions may be the candidates that fulfil functional roles as fusion proteins.

2.3. Validations of selected chimeric RNAs

In order to validate a subset of chimeric RNAs potentially found in infant brain tissue, we isolated and cultured the most promising mesenchymal stem cells derived from human umbilical cord (hUC-MSCs) (Supplementary Fig. S1) and their neural differentiated cells (dhUC-MSCs) (Supplementary Fig. S2) that most closely match the developmental stage of the infant brain we work with. From the above-mentioned analyses, we selected 22 chimeric RNAs that are in both the INTRACHR-SS-0GAP category and the in-frame category for validation (Supplementary Table S2). We focused on these neighbouring gene transcripts because of their potential high validation rate (Kowalski et al., 1999; Qin et al., 2016a,b), as well as fusion proteins. Ten chimeric RNAs were correctly amplified by RT-PCR (Fig. 3) and their sequences confirmed via Sanger sequencing (Supplementary Fig. S3). *PBXIP1-PMVK* and *MLLT6-CISD3* were only found in hUC-MSCs. The other eight chimeric RNAs were detected in both hUC-MSCs and differentiated hUC-MSCs. These findings support the notion that chimeric RNAs do exist in normal cells.

2.4. Further characterizations of the validated chimeric RNAs

Among the ten validated chimeric RNAs, nine of them involved the second-to-last exon of their 5′ parental gene fusing to the second exon of the 3′ parental gene (Supplementary Table S3). This configuration is consistent with previous reports (Babushok et al., 2007; Akiva et al., 2006; Kim et al., 2014; Qin et al., 2015; Varley et al., 2014), and further supports the idea that certain exons are favoured for the formation of this type of fusion RNA. The distance between the neighbouring genes involved in these chimeric RNAs were less than the average genomic distance between adjacent parent genes (Fig. 4A), which is again consistent with the previous findings in other tissues and cells (Qin et al., 2015; Prakash et al., 2010; Varley et al., 2014). This supports the idea that more closely arranged adjacent parental genes are more likely to form this type of fusions.

We compared the relative expression level of the ten validated chimeric RNAs to their parental genes' expression by their FKPM (Fragments Per Kilobase Of Exon Per Million Fragments Mapped). Six fusions are expressed at relatively high levels (above 10% relative to their 5′ parental gene expression) (Fig. 4B). Similarly, one fusion is expressed at a level greater than 10% of its corresponding 3′ parental genes (Fig. 4C).

2.5. DUS4L-BCAP29 promotes neural differentiation

Among the ten validated chimeric RNAs, *DUS4L-BCAP29* was the only one whose expression was dramatically increased during neural differentiation. Our previous study proved that this chimeric RNA is a product of cis-splicing of adjacent genes (cis-SAGe) (Tang et al., 2017). The results of qRT-PCR further revealed that *DUS4L-BCAP29* was highly expressed in neural differentiated hUC-MSCs when compared to hUC-MSCs ($p < 0.05$). In contrast, the expression of parental gene DUS4L

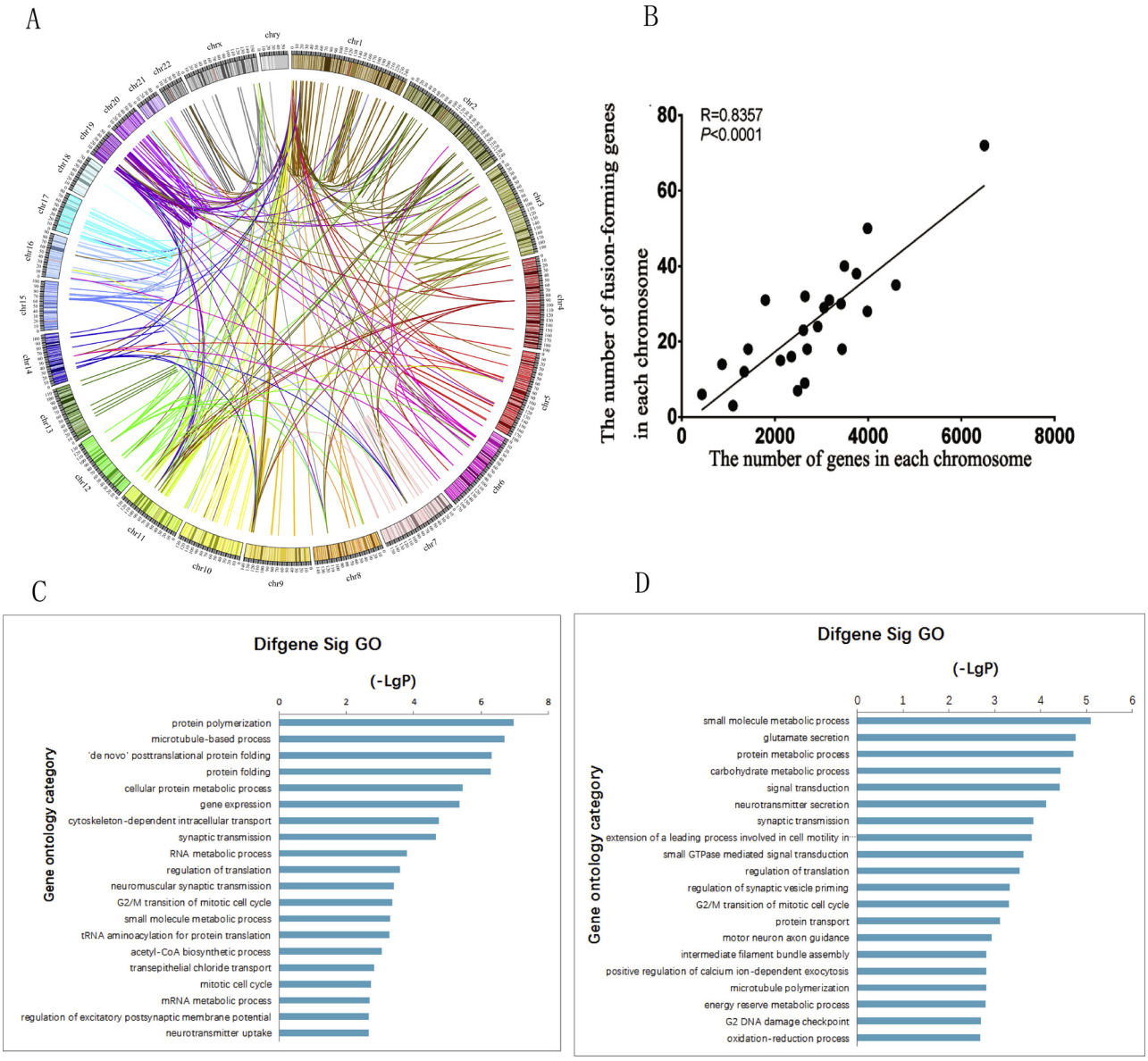


Fig. 1. Identification of chimeric RNAs in infant brain tissue. A: Chimeric RNAs were plotted on Circos plots. The fused transcripts are illustrated here as a line that connects two parental genes. B: The density of genes participating in fusion formation correlates to the overall gene density on individual chromosomes. C:Gene ontology terms enriched in 5′ parental genes involved in forming chimeric RNAs. D: Gene ontology terms enriched in 3′ parental genes involved in forming chimeric RNAs.

was hardly changed (BCAP29 was not detected) (Fig. 5).The differentiation brought about a proportional increase in the ratio of *DUS4L-BCAP29* to DUS4L(Supplementary Fig. S4).

To further explore whether this chimeric RNA is involved in neural differentiation, we constructed a *DUS4L-BCAP29* expression vector and transfected it into hUC-MSCs (Supplementary Fig. S5). Under the neuronal induction conditions (Wang et al., 2016, 2018), the neural differentiation efficiency of the *DUS4L-BCAP29* overexpression group was significantly enhanced when compared to the empty vector transfected control group. Immunofluorescence results demonstrated that the expression of neural markers, doublecortin (DCX) and neuron-specific enolase (NSE) were significantly increased ($p < 0.05$) (Fig. 6A and B). qRT-PCR results showed that Neurogenin-2 (Ngn2) and mammalian achaete-scute homologue-1 (Mash1) were significantly increased, while the expression of Neurogenin1 (Ngn1) was decreased ($p < 0.05$) (Fig. 6C). We repeated the experiment using a commercial mesenchymal stem cell neurogenic differentiation medium (C-28015, PromoCell), and observed very similar results with these markers by qRT-PCR (Supplementary Fig. S6).

In a previous study, we performed microarray analyses on gastric cell line sample transfected with the control siRNA, or si*DUS4L-BCAP29*, or siDUS4L (Tang et al., 2017). There were a few unique changes in signaling pathways in the samples transfected with the si-*DUS4L-BCAP29* (Supplementary Fig. S7). Among them, two signaling pathways (MAPK and JAK-STAT signaling pathways)are related to neural differentiation (Zang et al., 2008; Todd et al., 2016). The previous research of our group has also shown that the JAK2-STAT3 signaling pathway is involved in transformation between neural subtypes (Chi et al., 2017). Using Western blot analyses, we found that there was no significant change in phosphorylated p38MAPK expression, while phosphorylated STAT3 was decreased ($p < 0.05$) in the *DUS4L-BCAP29* overexpressed group (Fig. 6D and E).

3. Discussion

With the increasing numbers of chimeric RNAs identified in normal

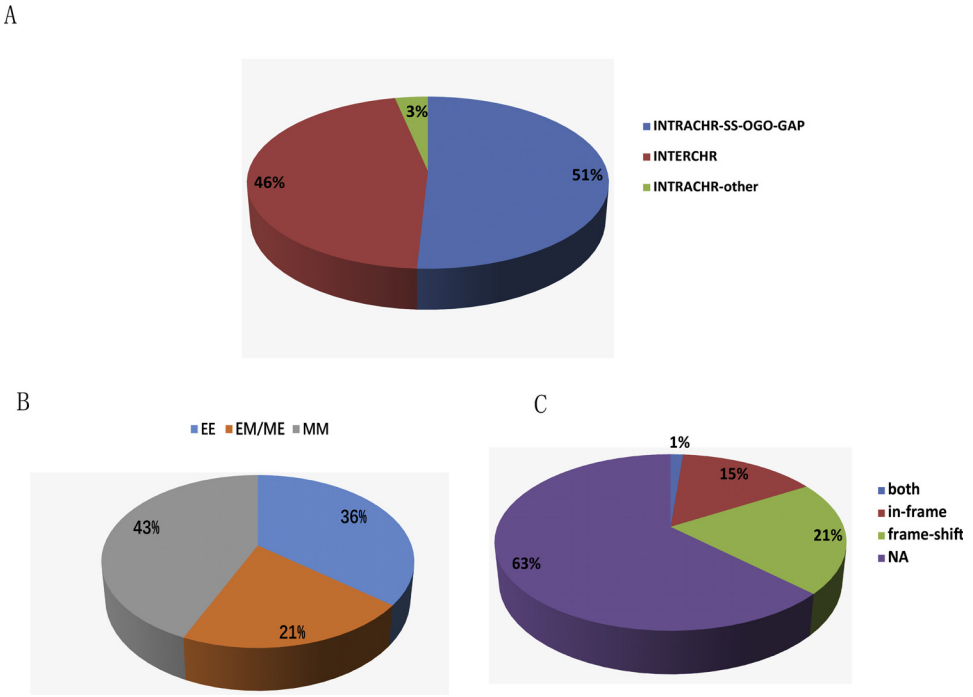
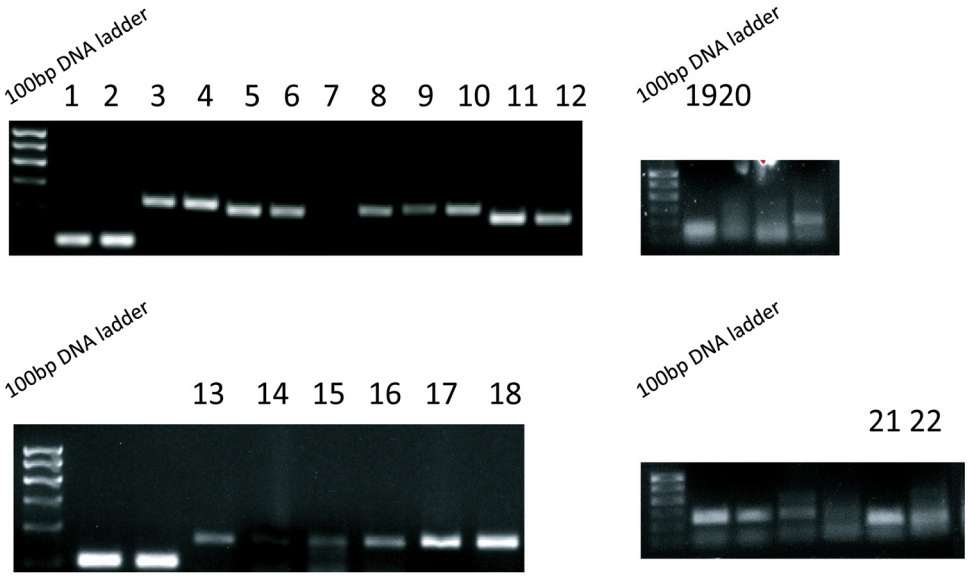


Fig. 2. Three categories of chimeric RNAs in infant brain tissue. A: Distribution of chimeric RNAs according to the chromosomal location of the parental genes. INTERCHR:fusions involving parental genes located on different chromosomes; INTRACHR-SS-0GAP: fusions involving neighboring genes transcribing the same strand; and INTRACHR-OTHER: other fusions with parental genes on the same chromosome. B: Distribution of chimeric RNAs according to the junction position relative to the parental exons. E/E: both 5′and 3′using known exon boundaries; E/M or M/E: one side using known exon boundary, the other not; M/M: both sides fall in the middle of known exons. C: Distribution of the chimeric RNAs according to their protein-coding potential: the known protein coding sequence of the 3′gene uses a different reading frame than the 5′ gene (frame-shift); the known reading frame of the 3′gene is the same as the 5′gene (in-frame); no effect on the reading frame of the parental genes (NA).



1 GAPDH in hUC-MSCs
2 GAPDH in dhUC-MSCs
3 ADSL-SGSM3 in hUC-MSCs
4 ADSL-SGSM3 in dhUC-MSCs
5 CTNNBIP1-CLSTN1 in hUC-MSCs
6 CTNNBIP1-CLSTN1 in dhUC-MSCs
7 DUS4L-BCAP29 in hUC-MSCs
8 DUS4L-BCAP29 in dhUC-MSCs
9 KIF1B-PGD in hUC-MSCs
10 KIF1B-PGDin dhUC-MSCs
11 SYNJ2BP-COX16 in hUC-MSCs
12 SYNJ2BP-COX16 in dhUC-MSCs
13 ZNF606-C19orf18 in hUC-MSCs
14 ZNF606-C19orf18 in dhUC-MSCs
15 ANAPC15-LAMTOR1 in hUC-MSCs
16 ANAPC15-LAMTOR1 in dhUC-MSCs
17 CTBS-GNG5 in hUC-MSCs
18 CTBS-GNG5 in dhUC-MSCs
19 PBXIP1-PMVK in hUC-MSCs
20 PBXIP1-PMVK in dhUC-MSCs
21 MLLT6-CISD3 in hUC-MSCs
22 MLLT6-CISD3 in dhUC-MSCs

Fig. 3. Validation of the candidate chimeric RNAs. Detection of the candidate chimeric RNAs in hUC-MSCs and neural differentiated cells by RT-PCR and followed by agarose electrophoresis. GAPDH was used as an internal control.

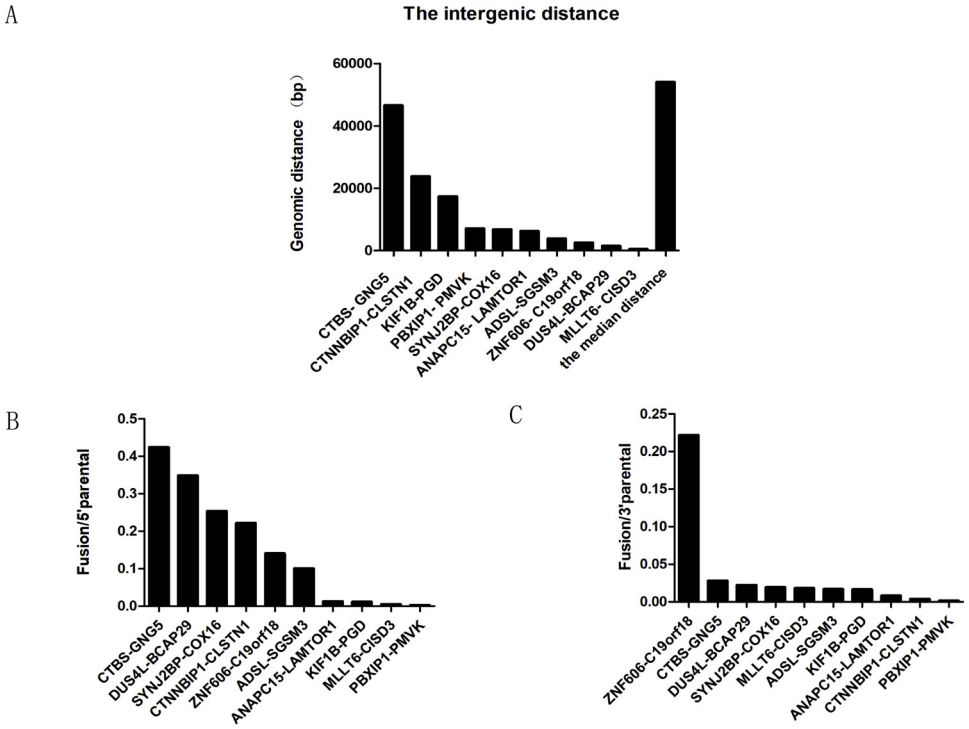


Fig. 4. Characteristics of the ten chimeric RNAs. A: Intergenic distance of the ten chimeric RNAs (distance between the end of the 5′ parental gene to the beginning of the 3′ gene). B: The expression ratios of fusion FKPM *versus* the FKPM of corresponding 5′parental gene were plotted following the order of highest to lowest. C: The expression ratios of fusion FKPM *versus* the FKPM of corresponding 3′parental gene were plotted following the order of highest to lowest.

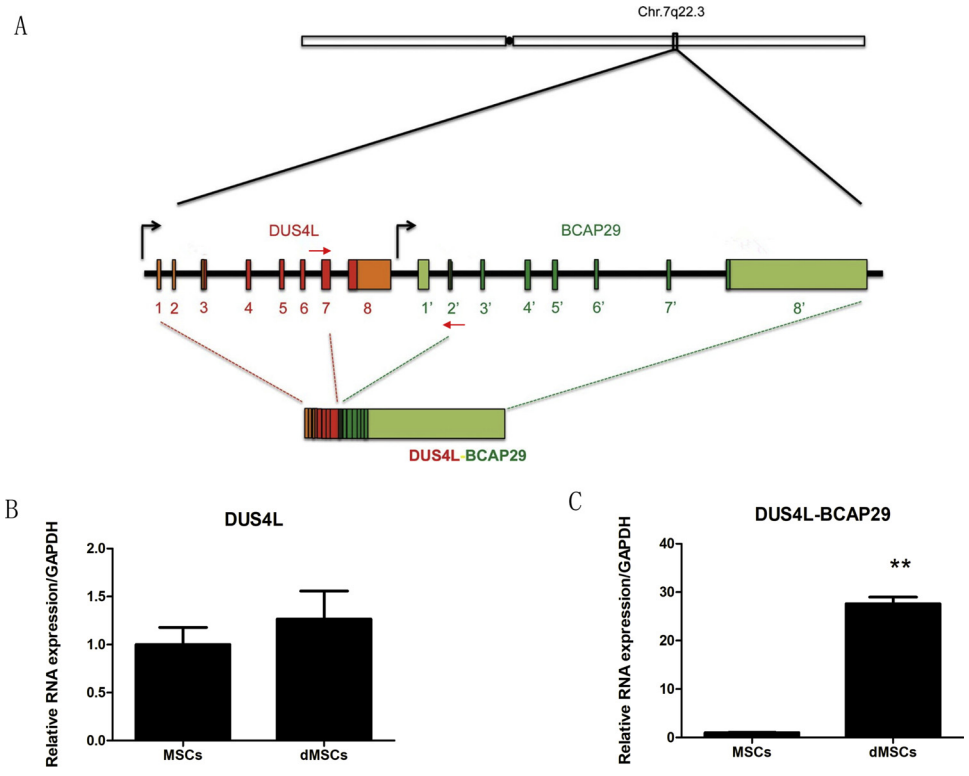


Fig. 5. Detection of DUS4L and *DUS4L-BCAP29* in hUC-MSCs and their neural differentiated counterparts by qRT-PCR. A: Gene structure diagram of the fusion and parental genes. Colored blocks represent exons, whereas lines represent introns and the intergenic region. The black arrows represent the primers used for wild type RNAs. The red arrows represent the primers are used for fusion RNAs. B: Relative expression level of DUS4L. C: Relative expression level of *DUS4L-BCAP29*. **: *vs.* MSCs $p < 0.001$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

tissues and cells, chimeric RNAs are now appreciated to be not unique to cancer (Yuan et al., 2013; Li et al., 2008; Tang et al., 2017). Chimeric RNA profiles differ from tissue to tissue. Thus, it is necessary to define the profiles of chimeric RNAs in specific tissues and cell types. Because there is currently no panoramic information reported in regards to chimeric RNAs in human brain development and neural differentiation, our study focused on infant brain tissues, and validated ten chimeric RNA candidates in an hUC-MSCs differentiation system. Among them, *DUS4L-BCAP29* was once thought to be prostate and gastric cancer specific (Kim et al., 2014). We recently showed that *DUS4L-BCAP29* expression levels are similar in both normal prostate and gastric tissues, as well as their cancer counterparts (Tang et al., 2017). In addition, this chimeric RNA plays a critical role in non-cancer prostate and gastric epithelial cell lines (Tang et al., 2017). Here we first demonstrated that its expression is dramatically increased during neural differentiation of hUC-MSC. We then showed that its overexpression promoted neural differentiation. These findings suggest the multifaceted function of the chimera.

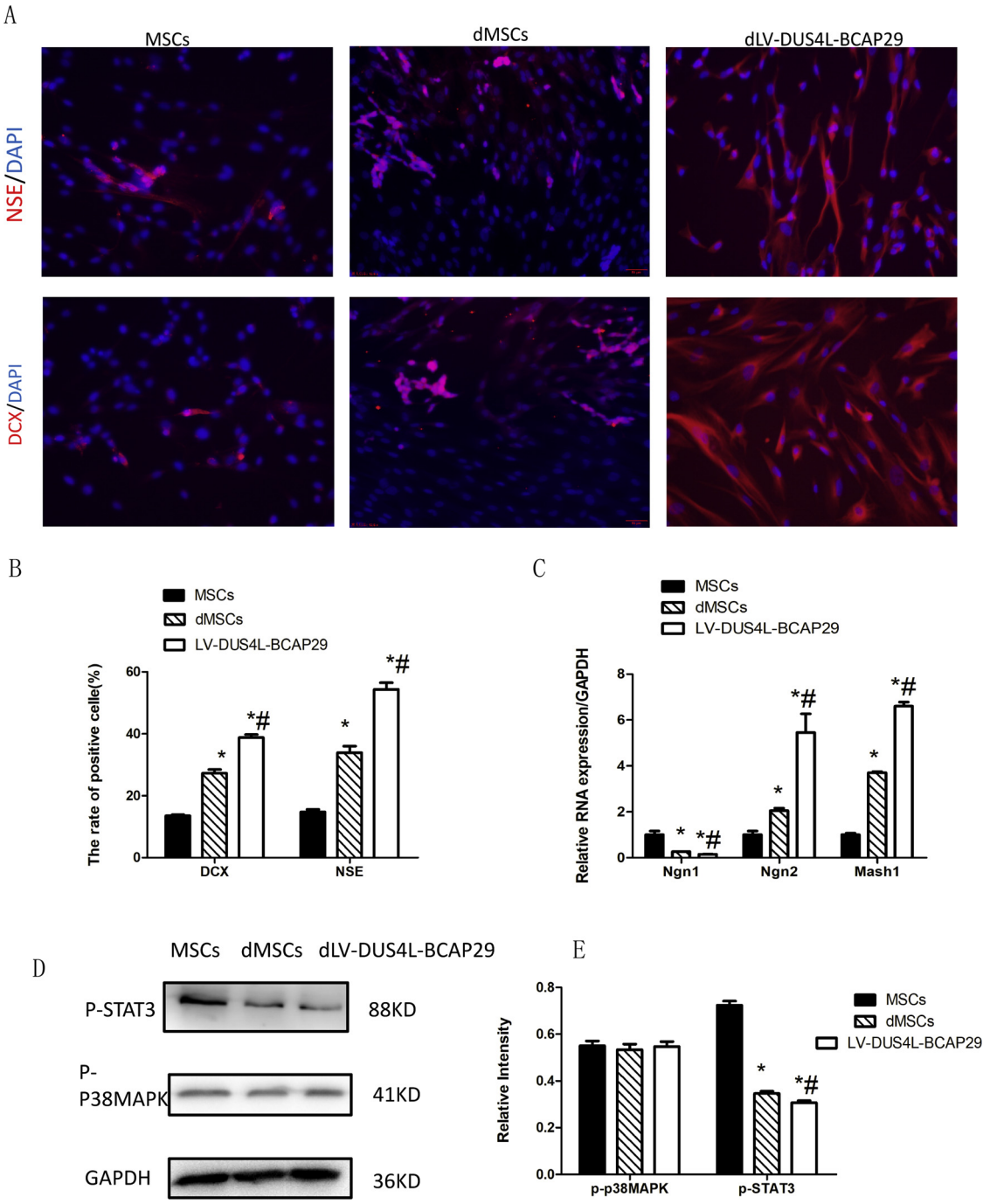


Fig. 6. DUS4L-BCAP29 stimulates neural lineage differentiation of hUC-MSCs. A: Representative immunofluorescence pictures of the expression of DCX and NSE (Red) in the hUC-MSCs. Nuclei were counterstained with DAPI (blue). Scale bar = 50 μm. B: Percentages of cells positive for specific neural markers (DCX/NSE) *vs.* total cells detected by immunofluorescence. C: The relative gene expression of Ngn1, Ngn2 and Mash1 in the different groups, which is normalized to the house-keeping gene GAPDH. D: Western blot results of the expression of p-STAT3, p-P38MAPK and internal control GAPDH. E: Bar graph of the gray intensity of Western blot. * : *vs.* MSCs, $p < 0.05$. # : *vs.* dMSCs, $p < 0.05$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

We noticed that there is at least one possible CTCF binding site in the intergenic regions of each of the chimeric RNAs (data not shown). CTCF and its binding to the intergenic region have been shown to regulate cis-splicing between adjacent genes (cis-SAGe) (Qin et al., 2015). Whether all the INTRACHR-SS-0GAP chimeras are cis-SAGe fusions, and whether they are regulated by CTCF require further investigation.

Even though a large number of chimeric RNAs are being identified in non-cancer tissues and cells, only a few have been shown to be functionally relevant. One bottleneck for this new field of research is determining the role that these chimeras play in normal physiological conditions. Here, we showed that *DUS4L-BCAP29* is expressed in infant brain tissues and that it is also involved in neural differentiation. These findings and other evidence support the notion that at least a subset of chimeric RNAs may have roles in crucial cellular processes and may expand the functional genome. Future studies involving more in depth experiments pertaining to the functional mechanisms of chimeric RNAs in the developing human brain are necessary.

4. Materials and methods

4.1. Bioinformatics analysis

Deep RNA sequencing data from SRA were mapped to Human genome version hg19 and analyzed using the SOAPfuse software to identify chimeric RNA. The putative fusions derived from the union of two parental genes were generated from multiple SOAPfuse runs (Jia et al., 2013). These fusions were presented using Circos plots. The fused genes are illustrated as a line that connects the two parental genes.

4.2. Isolation, and culture of hUC-MSCs

The hUC-MSCs were isolated as previously described (Ma et al., 2014). The hUC-MSCs were labelled with the following antibodies: CD29-PE (BD Bioscience, USA)、CD45-PE (Bio legend, USA)、HLA-DR-PE (Bio legend, USA)、CD34-APC (BD Bioscience, USA)、CD73-APC (Bio legend, USA)、CD105-APC (Bio legend, USA)、CD44-FITC (BD Bioscience, USA)、CD90-FITC (BD Bioscience, USA) and HLA-ABC-FITC (BD Bioscience, USA) before being analysed using a BD Accuri™ C6 Flow Cytometer System (Becton-Dickinson, USA).

The cells were cultured in DMEM/F12 media containing 10% fetal bovine serum (Gibco®, Invitrogen, Carlsbad, CA), 100 units/mL penicillin, and 100 μg/mL streptomycin, at 37 °C in a 5% CO_2 humidified incubator. The hUC-MSCs at passage 4(P4) were used for a typical experiment. The study was approved by the Ethics Committees of the First affiliated hospital of Zhengzhou University and consent was obtained from hUC-MSC donors.

4.3. Neural differentiation

Neural differentiation of hUC-MSCs was performed as previously described (Ma et al., 2014).

4.4. RT-PCR and Sanger sequencing

Total RNA was extracted with RNA Extraction Kit (Beibei Biotech, China), according to the manufacturer's instructions. We designed primers flanking the fusion junction to amplify fusion transcripts (Supplementary Table S2). Following RT-PCR and gel electrophoresis, purified DNA bands were sent for Sanger sequencing by Sangon Biotech.

4.5. Real-time PCR

The qPCR experiments were conducted with an ABI 7500 real time PCR system (Applied Biosystems, USA) according to the previous protocols (Yang et al., 2014). The primers used for *DUS4L-BCAP29* and DUS4L are the same as previously described (Tang et al., 2017).

4.6. Lentiviral vectors and mesenchymal stem cell infection

The fragment of human *DUS4L–BCAP29* from previous experimentation (Tang et al., 2017) was inserted into EcoRI/XhoI cloning sites of pEZ-Lv203. The purified lentiviral particles and the empty control vector were produced by Genecopoeia (Rockville, MD, USA). The hUC-MSCs were infected with the lentiviral particles or the negative controls in the presence of polybrene. Cells stably expressing the transgenes were selected with puromycin. The expression of *DUS4L–BCAP29* in hUC-MSCs was examined by immunofluorescence and qPCR.

4.7. Immunofluorescence staining

The hUC-MSCs on a 24-well plate were fixed at room temperature with 4% paraformaldehyde for 30 min. Non-specific staining was blocked by incubation in 10% normal goat serum (Sigma) and 0.2% Triton X-100 for 10 min at room temperature. The cells were then stained with primary antibodies against Doublecortin (DCX,1:100, SantaCruz, USA), or Neuron Specific Enolase (NSE, 1:100, Proteintech Group, China) overnight at 4 °C. After washing with PBS, the cells were incubated with Cy3-conjugated anti-rabbit anti IgG (1:1,000, Molecular Probes, USA) for 1 h, followed by DAPI staining (Biotech, China).

4.8. Statistical analysis

The level of statistical significance was set at $p < 0.05$. Independent paired t-test and Pearson's correlation coefficient were used for different data sets in order to test for group differences.

Conflicts of interest

We declare that there is no conflict of interest in relation to the work described.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (81601078, 81471306). This work is supported Open Research Fund Program of the State Key Laboratory of Virology of China (2018KF004) .We thank Xingxing Zang and Loryn Facemire for the help in editing the manuscript.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.biocel.2019.03.012.

References

Adams, M.D., Kelley, J.M., Gocayne, J.D., Dubnick, M., Polymeropoulos, M.H., Xiao, H., Merril, C.R., Wu, A., Olde, B., Moreno, R.F., Et, A., 1991. Complementary DNA sequencing: expressed sequence tags and human genome project. J. Science 252 (5013), 1651–1656.

Akiva, P., Toporik, A., Edelheit, S., Peretz, Y., Diber, A., Shemesh, R., Novik, A., Sorek, R., 2006. Transcription-mediated gene fusion in the human genome. J. Genome Res. 16 (1), 30–36.

Babiceanu, M., Qin, F., Xie, Z., Jia, Y., Lopez, K., Janus, N., Facemire, L., Kumar, S., Pang, Y., Qi, Y., Lazar, I.M., Li, H., 2016. Recurrent chimeric fusion RNAs in non-cancer tissues and cells. J. Nucleic Acids Res. 44 (6), 2859–2872.

Babushok, D.V., Ohshima, K., Ostertag, E.M., Chen, X., Wang, Y., Mandal, P.K., Okada, N., Abrams, C.S., Kazazian, H.J., 2007. A novel testis ubiquitin-binding protein gene arose by exon shuffling in hominoids. J. Genome Res. 17 (8), 1129–1138.

Benelli, M., Pescucci, C., Marseglia, G., Severgnini, M., Torricelli, F., Magi, A., 2012. Discovering chimeric transcripts in paired-end RNA-seq data by using EricScript. J. Bioinform. 28 (24), 3232–3239.

Carrara, M., Beccuti, M., Lazzarato, F., Cavallo, F., Cordero, F., Donatelli, S., Calogero, R.A., 2013. State-of-the-art fusion-finder algorithms sensitivity and specificity. J. Biomed Res. Int., 340620.

Chen, K., Wallis, J.W., Kandoth, C., Kalicki-Veizer, J.M., Mungall, K.L., Mungall, A.J., Jones, S.J., Marra, M.A., Ley, T.J., Mardis, E.R., Wilson, R.K., Weinstein, J.N., Ding, L., 2012. BreakFusion: targeted assembly-based identification of gene fusions in whole transcriptome paired-end sequencing data. J. Bioinformatics 28 (14), 1923–1924.

Chi, L., Fan, B., Feng, D., Chen, Z., Liu, Z., Hui, Y., Xu, X., Ma, L., Fang, Y., Zhang, Q., Jin, G., Liu, L., Guan, F., Zhang, X., 2017. The dorsoventral patterning of human forebrain follows an activation/transformation model. J. Cereb Cortex 27 (5), 2941–2954.

Chuang, T.J., Wu, C.S., Chen, C.Y., Hung, L.Y., Chiang, T.W., Yang, M.Y., 2016. NCLscan: accurate identification of non-co-linear transcripts (fusion, trans-splicing and circular RNA) with a good balance between sensitivity and precision. J. Nucleic Acids Res. 44 (3), e29.

Davidson, N.M., Majewski, I.J., Oshlack, A., 2015. JAFFA: High sensitivity transcriptome-focused fusion gene detection. J. Genome Med. 7 (1), 43.

Druker, B.J., 2001. Current treatment approaches for chronic myelogenous leukemia. J. Cancer J 7 (Suppl. 1), S14–S18.

Francis, R.W., Thompson-Wicking, K., Carter, K.W., Anderson, D., Kees, U.R., Beesley, A.H., 2012. FusionFinder: a software tool to identify expressed gene fusion candidates from RNA-seq data. J. PLoS One 7 (6), e39987.

Ge, H., Liu, K., Juan, T., Fang, F., Newman, M., Hoeck, W., 2011. FusionMap: detecting fusion genes from next-generation sequencing data at base-pair resolution. J. Bioinform. 27 (14), 1922–1928.

Gorohovski, A., Tagore, S., Palande, V., Malka, A., Raviv-Shay, D., Frenkel-Morgenstern, M., 2017. ChiTaRS-3.1-the enhanced chimeric transcripts and RNA-seq database matched with protein-protein interactions. J. Nucleic Acids Res. 45 (D1), D790–D795.

Hoffmann, S., Otto, C., Kurtz, S., Sharma, C.M., Khaitovich, P., Vogel, J., Stadler, P.F., Hackermuller, J., 2009. Fast mapping of short sequences with mismatches, insertions and deletions using index structures. J. PLoS Comput Biol. 5 (9), e1000502.

Huang, R., Kumar, S., Li, H., 2017. Absence of correlation between chimeric RNA and aging. J. Genes (Basel) 8 (12).

Jia, W., Qiu, K., He, M., Song, P., Zhou, Q., Zhou, F., Yu, Y., Zhu, D., Nickerson, M.L., Wan, S., Liao, X., Zhu, X., Peng, S., Li, Y., Wang, J., Guo, G., 2013. SOAPfuse: an algorithm for identifying fusion transcripts from paired-end RNA-seq data. J. Genome Biol. 14 (2), R12.

Kannan, K., Wang, L., Wang, J., Ittmann, M.M., Li, W., Yen, L., 2011. Recurrent chimeric RNAs enriched in human prostate cancer identified by deep sequencing. J. Proc. Natl. Acad. Sci. U. S. A. 108 (22), 9172–9177.

Kim, D., Salzberg, S.L., 2011. TopHat-fusion: an algorithm for discovery of novel fusion transcripts. J. Genome Biol. 12 (8), R72.

Kim, D.S., Huh, J.W., Kim, H.S., 2007. HYBRIDdb: a database of hybrid genes in the human genome. J. BMC Genom. 8, 128.

Kim, H.P., Cho, G.A., Han, S.W., Shin, J.Y., Jeong, E.G., Song, S.H., Lee, W.C., Lee, K.H., Bang, D., Seo, J.S., Kim, J.I., Kim, T.Y., 2014. Novel fusion transcripts in human gastric cancer revealed by transcriptome analysis. J. Oncogene 33 (47), 5434–5441.

Kong, F., Zhu, J., Wu, J., Peng, J., Wang, Y., Wang, Q., Fu, S., Yuan, L.L., Li, T., 2011. dbCRID: a database of chromosomal rearrangements in human diseases. J. Nucleic Acids Res. 39 (Database issue), D895–D900.

Kowalski, P.E., Freeman, J.D., Mager, D.L., 1999. Intergenic splicing between a HERV-H endogenous retrovirus and two adjacent human genes. J. Genom. 57 (3), 371–379.

Kumar, S., Razzaq, S.K., Vo, A.D., Gautam, M., Li, H., 2016a. Identifying fusion transcripts using next generation sequencing. J. Wiley Interdiscip Rev RNA 7 (6), 811–823.

Kumar, S., Vo, A.D., Qin, F., Li, H., 2016b. Comparative assessment of methods for the fusion transcripts detection from RNA-Seq data. Yan Jiu Bao Gao 6, 21597.

Langmead, B., Trapnell, C., Pop, M., Salzberg, S.L., 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. J. Genome Biol. 10 (3), R25.

Lee, M., Lee, K., Yu, N., Jang, I., Choi, I., Kim, P., Jang, Y.E., Kim, B., Kim, S., Lee, B., Kang, J., Lee, S., 2017. ChimerDB 3.0: an enhanced database for fusion genes from cancer transcriptome and literature data mining. J. Nuceic Acids Res. 45 (D1), D784–D789.

Li, H., Wang, J., Mor, G., Sklar, J., 2008. A neoplastic gene fusion mimics trans-splicing of RNAs in normal human cells. J. Science 321 (5894), 1357–1361.

Li, Y., Chien, J., Smith, D.I., Ma, J., 2011. FusionHunter: identifying fusion transcripts in cancer using paired-end RNA-seq. J. Bioinform. 27 (12), 1708–1710.

Li, Z., Qin, F., Li, H., 2018. Chimeric RNAs and their implications in cancer. J. Curr Opin Genet Dev. 48, 36–43.

Liu, S., Tsai, W.H., Ding, Y., Chen, R., Fang, Z., Huo, Z., Kim, S., Ma, T., Chang, T.Y., Priedigkeit, N.M., Lee, A.V., Luo, J., Wang, H.W., Chung, I.F., Tseng, G.C., 2016. Comprehensive evaluation of fusion transcript detection algorithms and a meta-caller to combine top performing methods in paired-end RNA-seq data. J. Nucleic Acids Res. 44 (5), e47.

Lu, Z., Matera, A.G., 2014. Vicinal: a method for the determination of ncRNA ends using chimeric reads from RNA-seq experiments. J. Nucleic Acids Res. 42 (9), e79.

Ma, S., Liang, S., Jiao, H., Chi, L., Shi, X., Tian, Y., Yang, B., Guan, F., 2014. Human umbilical cord mesenchymal stem cells inhibit C6 glioma growth via secretion of dickkopf-1 (DKK1). J. Mol. Cell. Biochem. 385 (1–2), 277–286.

Mcpherson, A., Hormozdiari, F., Zayed, A., Giuliany, R., Ha, G., Sun, M.G., Griffith, M., Heravi, M.A., Senz, J., Melnyk, N., Pacheco, M., Marra, M.A., Hirst, M., Nielsen, T.O., Sahinalp, S.C., Huntsman, D., Shah, S.P., 2011. deFuse: an algorithm for gene fusion discovery in tumor RNA-seq data. J. PLoS Comput Biol. 7 (5), e1001138.

Mcpherson, A., Wu, C., Wyatt, A.W., Shah, S., Collins, C., Sahinalp, S.C., 2012. nFuse: discovery of complex genomic rearrangements in cancer using high-throughput sequencing. J. Genome Res. 22 (11), 2250–2261.

Nacu, S., Yuan, W., Kan, Z., Bhatt, D., Rivers, C.S., Stinson, J., Peters, B.A., Modrusan, Z., Jung, K., Seshagiri, S., Wu, T.D., 2011. Deep RNA sequencing analysis of readthrough gene fusions in human prostate adenocarcinoma and reference samples. J. BMC Med Genom. 4, 11.

Novo, F.J., de Mendibil, I.O., Vizmanos, J.L., 2007. TICdb: a collection of gene-mapped translocation breakpoints in cancer. J. BMC Genom. 8, 33.

Prakash, T., Sharma, V.K., Adati, N., Ozawa, R., Kumar, N., Nishida, Y., Fujikake, T., Takeda, T., Taylor, T.D., 2010. Expression of conjoined genes: another mechanism for gene regulation in eukaryotes. J. PLoS One 5 (10), e13284.

Qin, F., Song, Z., Babiceanu, M., Song, Y., Facemire, L., Singh, R., Adli, M., Li, H., 2015. Discovery of CTCF-sensitive cis-spliced fusion RNAs between adjacent genes in human prostate cells. J. PLoS Genet. 11 (2), e1005001.

Qin, F., Song, Z., Chang, M., Song, Y., Frierson, H., Li, H., 2016a. Recurrent cis-SAGe chimeric RNA, D2HGDH-GAL3ST2, in prostate cancer. J. Cancer Lett. 380 (1), 39–46.

Qin, F., Song, Y., Zhang, Y., Facemire, L., Frierson, H., Li, H., 2016b. Role of CTCF in regulating SLC45A3-ELK4 chimeric RNA. J. PLoS One 11 (3), e150382.

Rickman, D.S., Pflueger, D., Moss, B., Vandoren, V.E., Chen, C.X., de la Taille, A., Kuefer, R., Tewari, A.K., Setlur, S.R., Demichelis, F., Rubin, M.A., 2009. SLC45A3-ELK4 is a novel and frequent erythroblast transformation-specific fusion transcript in prostate cancer. J. Cancer Res. 69 (7), 2734–2738.

Rowley, J.D., 1973. Letter: a new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and giemsa staining. J. Nature 243 (5405), 290–293.

Shaw, A.T., Yeap, B.Y., Solomon, B.J., Riely, G.J., Gainor, J., Engelman, J.A., Shapiro, G.I., Costa, D.B., Ou, S.H., Butaney, M., Salgia, R., Maki, R.G., Varella-Garcia, M., Doebele, R.C., Bang, Y.J., Kulig, K., Selaru, P., Tang, Y., Wilner, K.D., Kwak, E.L., Clark, J.W., Iafrate, A.J., Camidge, D.R., 2011. Effect of crizotinib on overall survival in patients with advanced non-small-cell lung cancer harbouring ALK gene rearrangement: a retrospective analysis. J. Lancet Oncol. 12 (11), 1004–1012.

Soda, M., Choi, Y.L., Enomoto, M., Takada, S., Yamashita, Y., Ishikawa, S., Fujiwara, S., Watanabe, H., Kurashina, K., Hatanaka, H., Bando, M., Ohno, S., Ishikawa, Y., Aburatani, H., Niki, T., Sohara, Y., Sugiyama, Y., Mano, H., 2007. Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. J. Nat. 448 (7153), 561–566.

Suhre, K., Claverie, J.M., 2004. FusionDB: a database for in-depth analysis of prokaryotic gene fusion events. J. Nucleic Acids Res. 32 (Database issue), D273–D276.

Tang, Y., Qin, F., Liu, A., Li, H., 2017. Recurrent fusion RNA DUS4L-BCAP29 in non-cancer human tissues and cells. J. Oncotarget 8 (19), 31415–31423.

Todd, L., Squires, N., Suarez, L., Fischer, A.J., 2016. Jak/Stat signaling regulates the proliferation and neurogenic potential of muller glia-derived progenitor cells in the avian retina. J. Sci Rep. 6, 35703.

Varley, K.E., Gertz, J., Roberts, B.S., Davis, N.S., Bowling, K.M., Kirby, M.K., Nesmith, A.S., Oliver, P.G., Grizzle, W.E., Forero, A., Buchsbaum, D.J., Lobuglio, A.F., Myers, R.M., 2014. Recurrent read-through fusion transcripts in breast cancer. J. Breast Cancer Res. Treat 146 (2), 287–297.

Wang, Y., Wu, N., Liu, J., Wu, Z., Dong, D., 2015. FusionCancer: a database of cancer fusion genes derived from RNA-seq data. J. Diagn Pathol. 10, 131.

Wang, X., Ma, S., Meng, N., Yao, N., Zhang, K., Li, Q., Zhang, Y., Xing, Q., Han, K., Song, J., Yang, B., Guan, F., 2016. Resveratrol exerts dosage-dependent effects on the self-renewal and neural differentiation of hUC-MSCs. J. Mol. Cells 39 (5), 418–425.

Wang, X., Ma, S., Yang, B., Huang, T., Meng, N., Xu, L., Xing, Q., Zhang, Y., Zhang, K., Li, Q., Zhang, T., Wu, J., Yang, G.L., Guan, F., Wang, J., 2018. Resveratrol promotes hUC-MSCs engraftment and neural repair in a mouse model of Alzheimer's disease. J. Behav Brain Res. 339, 297–304.

Wu, C.S., Yu, C.Y., Chuang, C.Y., Hsiao, M., Kao, C.F., Kuo, H.C., Chuang, T.J., 2014. Integrative transcriptome sequencing identifies trans-splicing events with important roles in human embryonic stem cell pluripotency. J. Genome Res. 24 (1), 25–36.

Xie, Z., Babiceanu, M., Kumar, S., Jia, Y., Qin, F., Barr, F.G., Li, H., 2016. Fusion transcriptome profiling provides insights into alveolar rhabdomyosarcoma. J. Proc. Natl. Acad. Sci. U. S. A. 113 (46), 13126–13131.

Yang, H., Tang, Y., Guo, W., Du, Y., Wang, Y., Li, P., Zang, W., Yin, X., Wang, H., Chu, H., Zhang, G., Zhao, G., 2014. Up-regulation of microRNA-138 induce radiosensitization in lung cancer cells. J. Tumour Biol. 35 (7), 6557–6565.

Yuan, H., Qin, F., Movassagh, M., Park, H., Golden, W., Xie, Z., Zhang, P., Sklar, J., Li, H., 2013. A chimeric RNA characteristic of rhabdomyosarcoma in normal myogenesis process. J. Cancer Discov. 3 (12), 1394–1403.

Zang, Y., Yu, L.F., Pang, T., Fang, L.P., Feng, X., Wen, T.Q., Nan, F.J., Feng, L.Y., Li, J., 2008. AICAR induces astroglial differentiation of neural stem cells via activating the JAK/STAT3 pathway independently of AMP-activated protein kinase. J. Biol Chem. 283 (10), 6201–6208.

Zhang, Y., Gong, M., Yuan, H., Park, H.G., Frierson, H.F., Li, H., 2012. Chimeric transcript generated by cis-splicing of adjacent genes regulates prostate cancer cell proliferation. J. Cancer Discov. 2 (7), 598–607.