International Journal of Biochemistry and Cell Biology 110 (2019) 50–58

Contents lists available at ScienceDirect

International Journal of Biochemistry and Cell Biology

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

The landscape of chimeric RNAs in bladder urothelial carcinoma

Dingjun Zhu[a,b], Sandeep Singh[b], Xu Chen[a], Zaosong Zheng[a], Jian Huang[a], Tianxin Lin[a,*], Hui Li[b,c,**,1]

[a] Department of Urology, Sun Yat-sen Memorial Hospital, 107th Yanjiangxi Road, Yuexiu District, 510120, Guangzhou, Guangdong Province, China
[b] Department of Pathology, School of Medicine, University of Virginia, Charlottesville, VA 22908, United States
[c] School of Basic Medical Sciences, Zhengzhou University, Zhengzhou, 450001, China

ARTICLE INFO

Keywords:
Bladder cancer
Gene fusion
Chimeric RNAs
Intergenic splicing
cis-SAGe

ABSTRACT

Background: Gene fusions and products have been identified as oncogenic drivers in many cancers, making them attractive diagnostic markers and therapeutic targets. However, the landscape of fusion transcripts in bladder cancer has not been fully characterized.

Methods: To identify fusion transcripts with potential therapeutic or diagnostic values, TCGA bladder urothelial carcinoma RNA-sequencing dataset was used. In order to avoid false positives, we applied multiple criteria including filtering out fusions detected in normal samples from GTEx dataset. We validated a subset of candidate fusions with a collection of bladder cancer and adjacent normal samples.

Result: We identified 19,547 high confidence fusion genes from 414 bladder cancer samples. After filtering off M/M fusions, fusions in GTEx normal samples, and occurrence frequency < 5, we obtained a list of 271 gene fusions, 13 of which were novel and specific to cancer samples. Six of those fusions were validated using cell lines and clinical samples. We discovered that two chimeric RNAs, *BCL2L2-PABPN1* and *CHFR-GOLGA3*, were detected to be expressed significantly higher in bladder cancer samples compared to adjacent normal samples. Impressively, the wild-type of the parental genes were not differentially expressed. Mechanistically, we demonstrated that these two fusions are generated by cis-splicing between adjacent genes. These two fusions were detected mainly in the fraction of cell nucleus, suggesting a potential long noncoding RNA role.

Conclusion: Our findings provide a panoramic view of the landscape of chimeric RNAs in bladder cancer. Some frequent chimeric RNAs are generated by intergenic splicing, and represent a new repertoire for potential biomarkers.

1. Introduction

Bladder cancer represents the 7th most common malignancy in men and 11th in both genders worldwide with approximately 549,393 new cases, and causes 199,922 deaths per year (Bray et al., 2018). The length of bladder cancer survival is strongly influenced by cancer stage at diagnosis. During the early stages, 5-year survival is more than 95%, but when the cancer metastasize, this rate dramatically drops to 4.8%. Even with the improvement of multiple therapeutic approaches, such as surgery, chemotherapy and immunotherapy, the mortality rate of bladder cancer has not decreased significantly over the years (Racioppi et al., 2012). Therefore, exploitation of new diagnostic and therapeutic strategy is needed.

Gene fusions in cancers have been known for over half a century. The oncogenic fusion *BCR-ABL1* resulted from t(9,22) translocation in chronic myeloid leukemia provides a paradigm for cancer diagnosis and treatment (Rowley, 1973). With the application of high-throughput sequencing, numerous gene fusions have been identified in hematological malignancies, sarcomas and solid cancers. Fusion RNAs and protein products often play causal roles in tumorigenesis and therefore represent ideal diagnostic and therapeutic targets. Although a few gene fusions have been found in bladder cancer, such as *HNRNPA1-ALK* and *FGFR3-TACC3*, and characterized as oncogenic drivers and potential therapeutic targets (Inamura et al., 2017; Kumar-Sinha et al., 2015; Lombardi et al., 2017), highly frequent fusions are rare. Recently, we and others have shown that chimeric fusion RNAs can be found in various cells and tissues, and some are shown to be the products of intergenic splicing instead of chromosomal rearrangement (Li et al.,

* Corresponding author.
** Corresponding author at: Department of Pathology, School of Medicine, University of Virginia, Charlottesville, VA 22908, United States.
E-mail addresses: lintx@mail.sysu.edu.cn (T. Lin), hl9r@virginia.edu (H. Li).
[1] Web: http://lilab.medicine.virginia.edu.

https://doi.org/10.1016/j.biocel.2019.02.007
Received 1 November 2018; Received in revised form 20 February 2019; Accepted 23 February 2019
Available online 25 February 2019

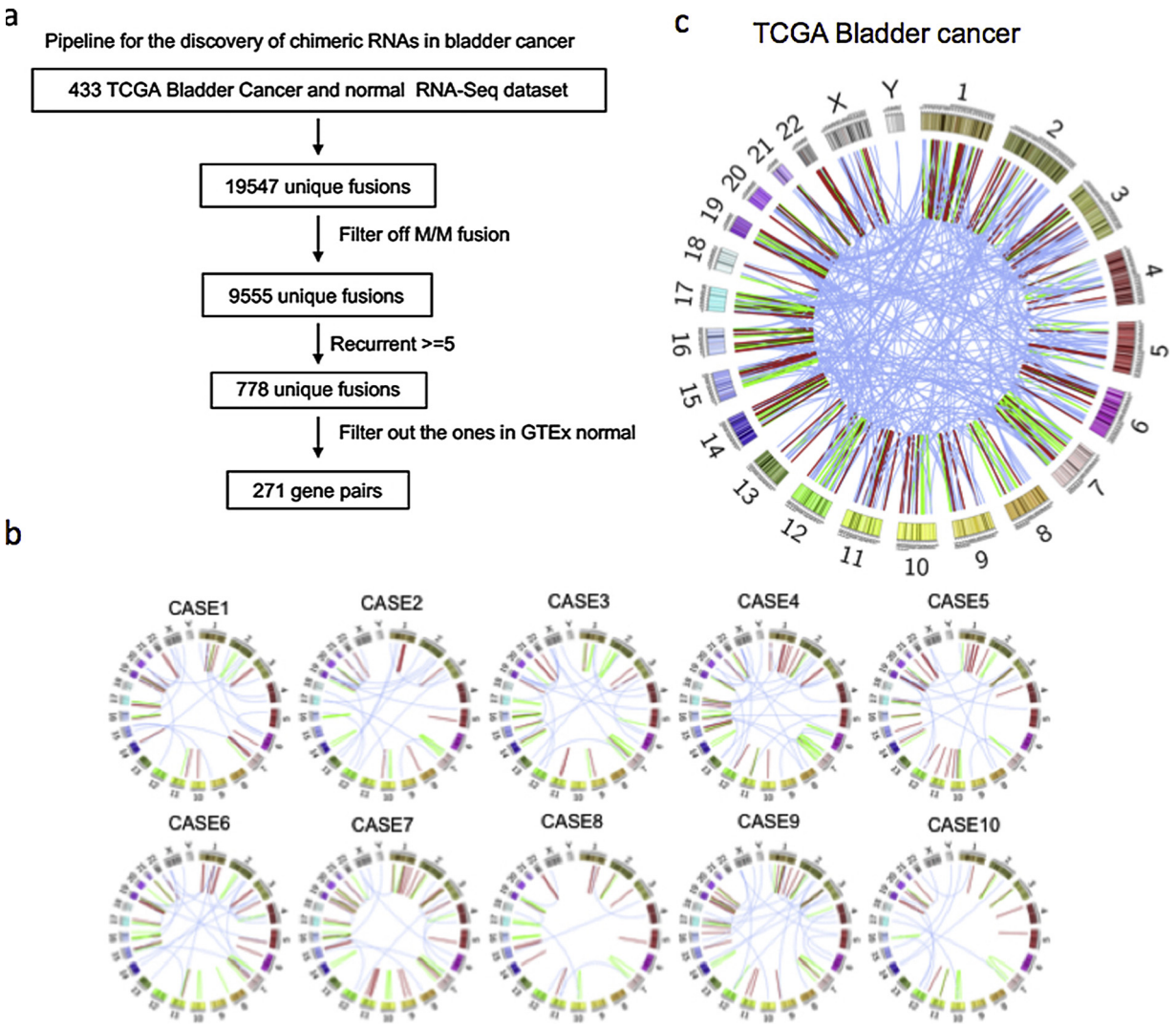

Fig. 1. Discovery of chimeric RNAs in bladder cancer.
(A) The pipeline for the discovery of chimeric RNAs in bladder cancer. TCGA RNA-Sequencing data of 414 bladder urothelial carcinoma samples and 19 relative adjacent normal samples was analyzed. After filtering out "M/M" fusions, the fusions with a recurrent frequency lower than 5, and the fusions found in the GTEx normal RNA-seq dataset, we obtained 271 unique chimeric RNAs. (B) Circos plot depicting chimeric RNAs from ten samples as examples. Lines denote the chimeric RNAs connecting two parental genes. (C) Circos plot depicting all the recurrent non M/M chimeric RNAs identified from the TCGA bladder cancer RNA-seq dataset.

2008a; Chase et al., 2010; Yuan et al., 2013; Wu et al., 2014; Zaphiropoulos, 2012). Here we deep-mined chimeric RNAs from TCGA muscle-invasive bladder cancer study RNA-seq dataset, and identified several chimeric RNAs with potential diagnostic value. Importantly, they are products of cis-splicing between adjacent genes, instead of canonical chromosomal rearrangement, supporting that intergenically spliced chimeric RNAs represent a new source of potential biomarkers.

2. Materials and methods

2.1. Cell culture

Bladder cancer cell lines T24, EJ, SV-HUC-1 were purchased from ATCC (American Type Culture Collection). EJ Cells were cultured in DMEM (Dulbecco's modified Eagle's medium) (Gibco) plus 10%FBS (Fetal Bovine Serum) (Invitrogen), 1% pen/strep, and 1% L-glutamine. SV-HUC-1 cells were maintained in F12 K medium (Gibco) containing 10%FBS, and 1% pen/strep. T24 cells were maintained in RPMI1640 medium (Gibco) containing 10%FBS, and 1% pen/strep. The cells were cultured at 37 °C and 5% CO_2.

2.2. Clinical samples

14 pairs of bladder urothelial carcinoma tissues and adjacent normal tissues were obtained from the Department of Pathology at the University Hospital of the University of Virginia. The use of the human clinical samples was approved by the IRB committee of University of Virginia. 16 additional pairs of bladder urothelial carcinoma tissues and adjacent normal tissues were obtained from the Sun Yat-sen Memorial Hospital, Sun Yat-sen University, with patient consent. All samples were immediately snap-frozen in liquid nitrogen and stored at −80 °C until required.

2.3. Bioinformatics

Raw sequencing data (RNA-Seq) of 408 bladder cancer patients (433 samples) was downloaded from TCGA portal (https://portal.gdc.cancer.gov/). The data consists of 414 cancer samples and 19 matched normal samples. NGS QC Toolkit (Patel and Jain, 2012) was used with default parameters to filter the raw reads and the good quality reads were given to EricScript software (version 0.5.5) (Benelli et al., 2012) for predicting fusion RNAs for all the 433 samples. EricScript was used

Table 1
Candidate chimeric RNAs predicted to be only in TCGA bladder cancer samples with an occurrence frequency above 10.

	GeneName1	GeneName2	chr1	chr2	Fusion type	Cancer (414)	Normal (19)
1	EEF1D	NAPRT1	8	8	Read-Through	32	0
2	BCL2L2	PABPN1	14	14	Read-Through	19	0
3	CHFR	GOLGA3	12	12	Read-Through	12	0
4	CHCHD10	VPREB3	22	22	intra-others	16	0
5	SLC2 A11	MIF	22	22	Read-Through	43	0
6	ACKR2	KRBOX1	3	3	Read-Through	10	0
7	SIRPB2	NSFL1C	20	20	Read-Through	11	0
8	ACKR2	ZNF662	3	3	intra-others	10	0
9	HOXB7	HOXB9	17	17	intra-others	16	0
10	NREP	C11orf74	5	11	interchr	15	0
11	SCAMP1	SYT1	5	12	interchr	11	0
12	TTLL11-IT1	TTLL11	9	9	Read-Through	25	0
13	KIAA1984	RABL6	9	9	Read-Through	42	0

with default parameters and hg19 genome was used as the reference genome. Fusion RNAs with EricScore < 0.6 were discarded. Based on the breakpoint position of the upstream and downstream gene partners, each fusion RNA was classified into "E/E", "E/M", "M/E" or "M/M" categories where "E" and "M" represents that the breakpoint occurred at the ends and middle of the exon respectively. End positions of the exons were considered from within a range of ± 2 bp of the actual end position. The in-frame/frame-shift potential of the fusion RNA was predicted using python script "predict_frame.py" of the FusionCatcher software [doi: https://doi.org/10.1101/011650]. For simplicity, the fusion type categories "Cis" and "intra-chromosomal" from the EricScript software were merged into "INTRA-Others" while "Read-Through" and "inter-chromosomal" (represented as INTERCHR) were kept as is.

2.4. RNA extraction, nuclear and cytoplasm extraction, PCR and Real-Time PCR

Bladder samples were grinded in liquid nitrogen. RNAs were extracted using TRIzol reagent (Invitrogen) and cDNAs were generated by cDNA synthesis Kit (Bioline, United States), following the manufacturer's instruction. Nuclear and cytoplasm RNAs extraction was performed as previously described (Qin et al., 2017). Real-time PCR was carried out on ABI StepOne Plus real time PCR system (Applied Biosystems, USA) using SYBR mix kit (Thermo). Primers used in this study are listed in Supplementary Table S1. Following PCR and gel electrophoresis, all purified DNA bands were sent for Sanger sequencing by Eton Bioscience INC.

2.5. Statistical analyses

The quantitative results were presented as the mean ± standard error of the mean (SEM). Wilcoxon matched-pairs signed rank test was used to identify statistically significant data, reported as the mean ± standard error of the mean (SEM). GraphPad Prism 7.0 (GraphPad Software, Inc., San Diego, CA, USA) was used for statistical analyses. For all analyses, $p < 0.05$ was considered statistically significant.

3. Results

3.1. Identification of chimeric RNAs in bladder cancer

An overview of the discovery pipeline of chimeric RNAs is shown in Fig. 1A. We analyzed a TCGA (The Cancer Genome Atlas) mRNA sequencing dataset consisting of 414 bladder urothelial carcinoma samples and 19 normal adjacent samples (Robertson et al., 2017). A total of 19,547 candidate chimeric transcripts were identified by the bioinformatics software tool, Ericscript (Benelli et al., 2012). To reduce the false-positive prediction rate, we filtered fusion RNAs according to the pattern of how two parental genes are connected. The chimeras were categorized according to the junction position of two parental genes: both sides being the end of known exon/intron boundaries (E/E), both sides falling into the middle of exons (M/M), one side being exon/intron boundary and the other not (E/M or M/E). Since we have proved previously that the M/M type of fusion transcripts have much lower validation rate compared to the other two types of fusions (Qin et al., 2015), we filtered off 9992 fusions of the M/M category. We then selected those fusions with recurrent frequency ≥5 and ended up with 778 unique fusion RNAs. Circos plots were used to depict these chimeric RNAs in the TCGA study (Fig. 1B and C). Furthermore, we removed fusions overlapping between bladder samples and The Genotype-Tissue Expression (GTEx) normal samples (manuscript in preparation). After applying these filters, 271 unique fusion gene pairs were found to be only present in the TCGA bladder samples (Fig. 1A). We went further to compare the occurrence of these fusions in bladder cancer versus adjacent normal tissues, and discovered 13 fusions only detectable in bladder cancer samples with recurrent frequency ≥10 (Table 1).

3.2. The landscape of chimeric RNAs and parental genes in bladder cancer

We examined the landscape of the fusion RNAs from three angles, and at three different levels (Fig. 2A). First, we characterized the fusions according to the chromosomal locations of their parental genes: parental genes located on different chromosomes (INTERCHR), neighboring genes transcribing the same strand (Read-Through), and other fusions with parental genes on the same chromosome (INTRA-Others). For all of the fusions, INTERCHR is the most prominent group (67%), and Read-Through is the least common group (2%). However, as the M/M fusions were filtered out, the inter-chromosomal group shrunk (54%), and Read-Through and INTRA-Others became more abundant (14% and 32% respectively). The percentage of the Read-Through group became even larger, when both "non M/M" and "recurrent fusion" filters were applied: it became similar to the size of the INTERCHR group (37% vs. 43%).

Second, as described above, based on the junction position relative to the exon of the parental genes, we categorized the chimeric RNAs into E/E, E/M or M/E, and M/M groups. Among all the fusions, the biggest category was M/M fusions (51%), E/M and M/E being about 25 and 10% each, and E/E 14% (Fig. 2A). After we filtered out M/M, and less frequent fusions (< 5), E/E fusions were significantly enriched (48%).

Lastly, we categorized the fusions according to their 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) (this category includes fusion RNAs whose junction sequence fall into untranslated region or one or

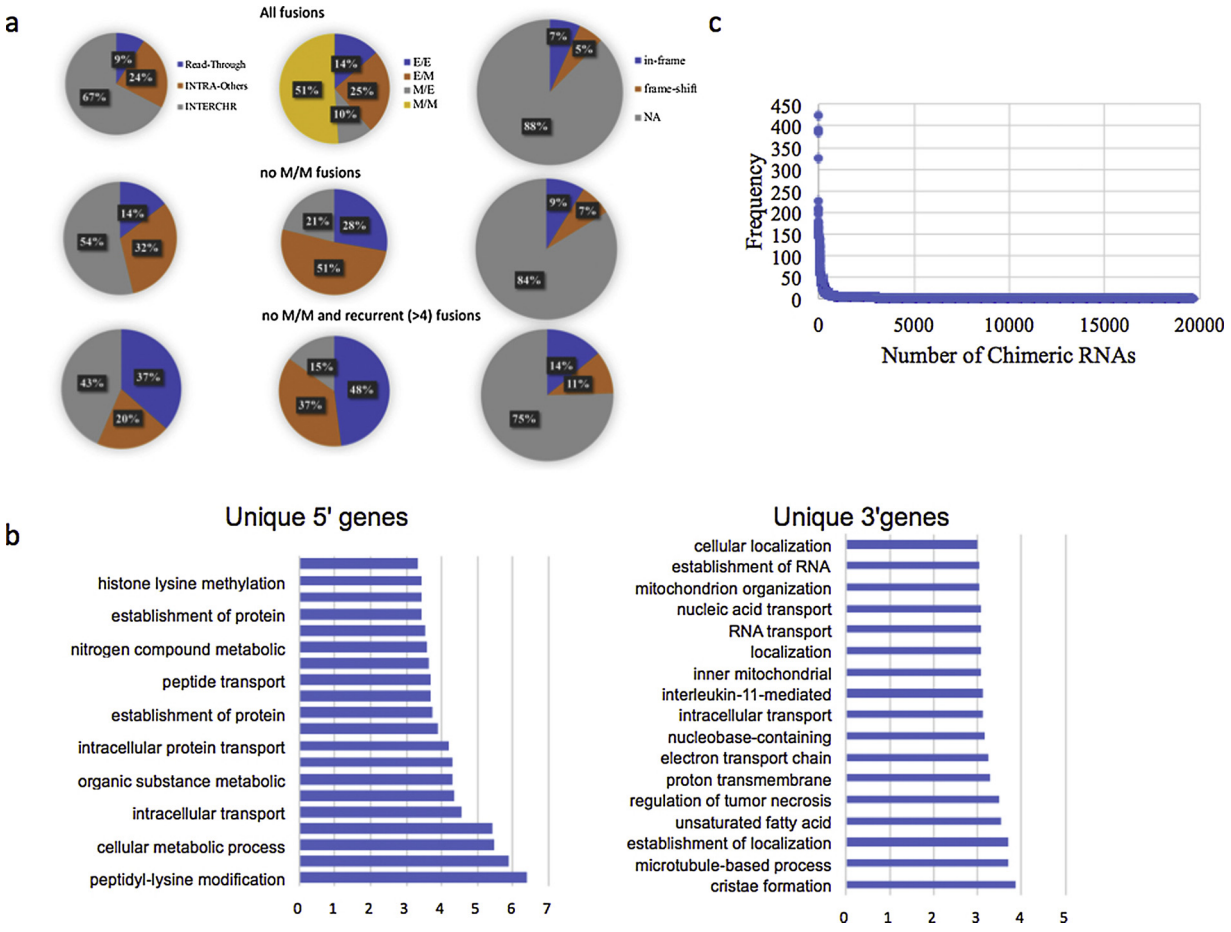

Fig. 2. The landscape of chimeric RNAs and their parental genes in bladder cancer.
(A) Distributions of chimeric RNAs from the TCGA data set. Chimeric RNAs are categorized based on their fusion junction position, fusion type, and fusion protein coding potential. When the criteria of "non-M/M", and "recurrence" applied, more E/E, Read-Through, and in-frame fusions were enriched. (B) Gene ontology analyses of the 5′ and 3′ parental genes involved in non M/M fusion RNAs in bladder cancer. Plotted are statistical significance (-Log10(p-value)) of the top 20 terms. (C) The frequency of chimeric fusion RNAs detected in bladder cancer samples.

both parental genes is lncRNA). When all the fusions were examined, the number of NA is the largest (88%). Frame-shift fusions are about the same frequency as in-frame fusions (5% vs. 7%). After filtering out M/M fusions, the NA portion became smaller (84%). When both "non-M/M" and "recurrent" filters were used, the NA fusions were further reduced (75%), and the in-frame and frame-shift became more common (14% and 11%).

We searched gene ontology terms using Gorilla (Eden et al., 2009) for the parental genes involved in the non-M/M, recurrent fusion RNAs. The top 20 GO terms for the 5′ gene include histone methylation, and transportation, whereas the 3′ genes are enriched for localization and transportation (Fig. 2C). For comparison, we analyzed RNA-Seq from 424 TCGA hepatocellular carcinoma (HCC) and 201 TCGA cervical cancer (CESC) RNA-Seq datasets. More metabolic-related GO terms were enriched in HCC, together with viral processing, symbiosis, and interspecies interaction terms (Fig. S1), whereas more viral processing terms enriched in CESC datasets (Fig. S2).

3.3. Validation of the highly recurrent chimeric RNAs in bladder cancer

The majority of the chimeric RNAs were identified in one or two samples (Fig. 2C). We then investigated the correlation between the fusion RNAs' occurrence and different clinical parameters, such as cancer stage, pathological type, and lymphatic metastasis. No obvious correlation was observed. Among the 13 recurrent chimeric RNAs that were detected only in bladder cancer samples, but not in the 19 adjacent normal samples, suitable PCR primers covering the fusion junction site could be designed for nine of them. We then performed RT-PCR and subsequently Sanger sequencing using mixed cDNA samples of cancer, normal, and cell lines. Six out of the nine chimeric RNAs were validated, and confirmed by Sanger sequencing (Fig. 3A and 3B).

3.4. BCL2L2-PABPN1 and CHFR-GOLGA3 were preferentially expressed in bladder cancer samples

Using 14 pairs of bladder cancer and adjacent normal samples obtained from the University of Virginia Hospital, we compared the expression of the chimeric RNAs in clinical samples. Among the chimeras, *BCL2L2-PABPN1* (Fig. 4A) was detected in the majority of the tumor samples, but only in two normal adjacent samples at lower levels (Fig. 4B). Consistently, Wilcoxon matched-pairs signed rank test revealed statistical significance ($P = 0.001$). Interestingly, the expression of the two wild-type parental genes showed no statistical significance between the tumor and normal samples (Fig. 4C and D).

Another chimera, *CHFR-GOLGA3* (Fig. 5A), was also found to be preferentially expressed in the bladder cancer samples, and only once in a tumor-matched sample, albeit at a lower level compared to the tumor (Fig. 5B). Significant difference of the expression was found between the groups of tumor and normal ($P = 0.0039$). Similar to *BCL2L2-PABPN1*, the expression of the two wild-type parental genes of *CHFR-GOLGA3* also had no statistical significance between tumor and normal groups (Fig. 5C and D).

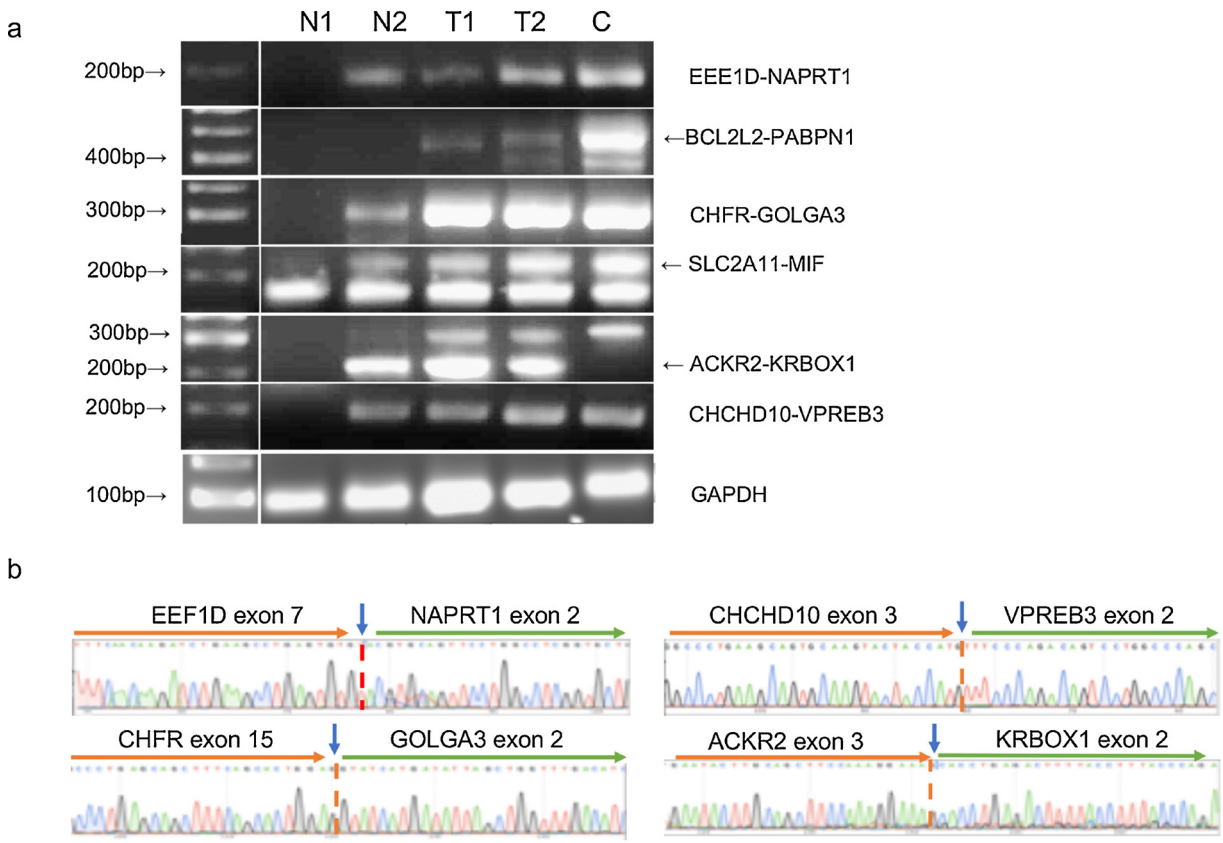

Fig. 3. Validation of the highly recurrent chimeric RNAs in bladder cancer.
(A) Electrophoresis gel images of RT-PCR product of the six candidate chimeric RNAs together with internal control, *GAPDH* (N: mixture of cDNA products from normal bladder samples; T: cDNA mixture from bladder tumor samples; C: cDNA mixture of bladder cancer cell lines). (B) Sanger sequencing graphs showing the verification of some chimeric RNAs. The junction sites were marked by dotted lines.

All the other four chimeras were found to be not differentially expressed, thus not pursued further (Fig. S3).

We then obtained another 16 pairs of samples from Sun Yat-sen Memorial Hospital, and repeated the experiments at Sun Yat-sen University to validate the above findings. We found similar trend for *BCL2L2-PABPN1* in that it was expressed at significantly higher levels in tumors than the normal matched samples ($p = 0.0175$). However, both chimeras were detected in a similar number of normal and tumor samples, and *CHFR-GOLGA3* was not found to be differentially expressed between the two groups (Fig. S4).

3.5. Both BCL2L2-PABPN1 and CHFR-GOLGA3 are products of cis-splicing between adjacent genes (cis-SAGe)

Both parental genes of *BCL2L2-PABPN1* and *CHFR-GOLGA3* are adjacent to each other on the same chromosome region, so they were classified as Read-Through. To validate that these two fusions are indeed generated by cis-SAGe, where a premature RNA is transcribed through the gene boundary and exons belonging to two parental genes are spliced together. We used the following PCR-based assay. DNase-I was first used to remove potential DNA contamination in the RNA extracted from two bladder cancer cell lines (EJ and SV-HUC1). We then used a reverse primer annealing to the downward exon neighboring to the junction site to perform Reverse Transcription (RT). Primers covering fragment of 5′ gene was then used for RT-PCR. In the absence of avian myeloblastosis virus (AMV) reverse transcriptase, no signal was observed, indicating that DNase-I digestion was complete. We can only detected signals in the presence of the AMV-RT enzyme with primer pairs to amplify a fragment of cDNA covering exon15 and intron15 of *BCL2L2,* or exon3 and intron3 of *CHFR* (Fig. 6B and C). This confirms the presence of a precursor RNA transcribing from exon3 of *BCL2L2* to exon2 of *PABPN1*, and a precursor RNA transcribing from exon15 of *CHFR* to exon2 of *GOLGA3*. We, therefore, conclude that both chimeric RNAs are products of cis-SAGe.

3.6. Chimeric RNA BCL2L2-PABPN1 and CHFR-GOLGA3 may function as non-coding RNAs

In order to gain some insights on the biological function of these two chimeric RNAs, we examined the expression of these chimeric RNAs in cell nucleus and cytoplasm fractions. Using a protein coding gene, *GAPDH*, and a long non-coding RNA, *MALAT1* as reference controls, we found that both *BCL2L2-PABPN1* and *CHFR-GOLGA3* were preferentially detected in the cell nuclear in both T24 and SV-HUC1 cells, which indicates that these chimeric RNAs may act as non-coding RNAs (Fig. 6D). In contrast, most of the wild-type parental genes were found in the cytoplasmic faction.

4. Discussion

Gene fusions generated by chromosomal rearrangements are considered hallmarks of cancer, thus representing ideal biomarkers and/or therapeutic targets (Rabbitts, 1994; Heim and Mitelman, 2008). However, recurrent gene fusions are rare in bladder cancers (Cancer Genome Atlas Research, N. et al., 2017). On the other hand, recent work on RNA trans-splicing (Yuan et al., 2013; Li et al., 2008b; Finta and Zaphiropoulos, 2002) and intergenic cis-splicing (Qin et al., 2015; Zhang et al., 2012; Kumar-Sinha et al., 2012) supports a new paradigm for intergenic splicing processes, which can also generate fusion products in the absence of chromosomal rearrangement at the DNA level.

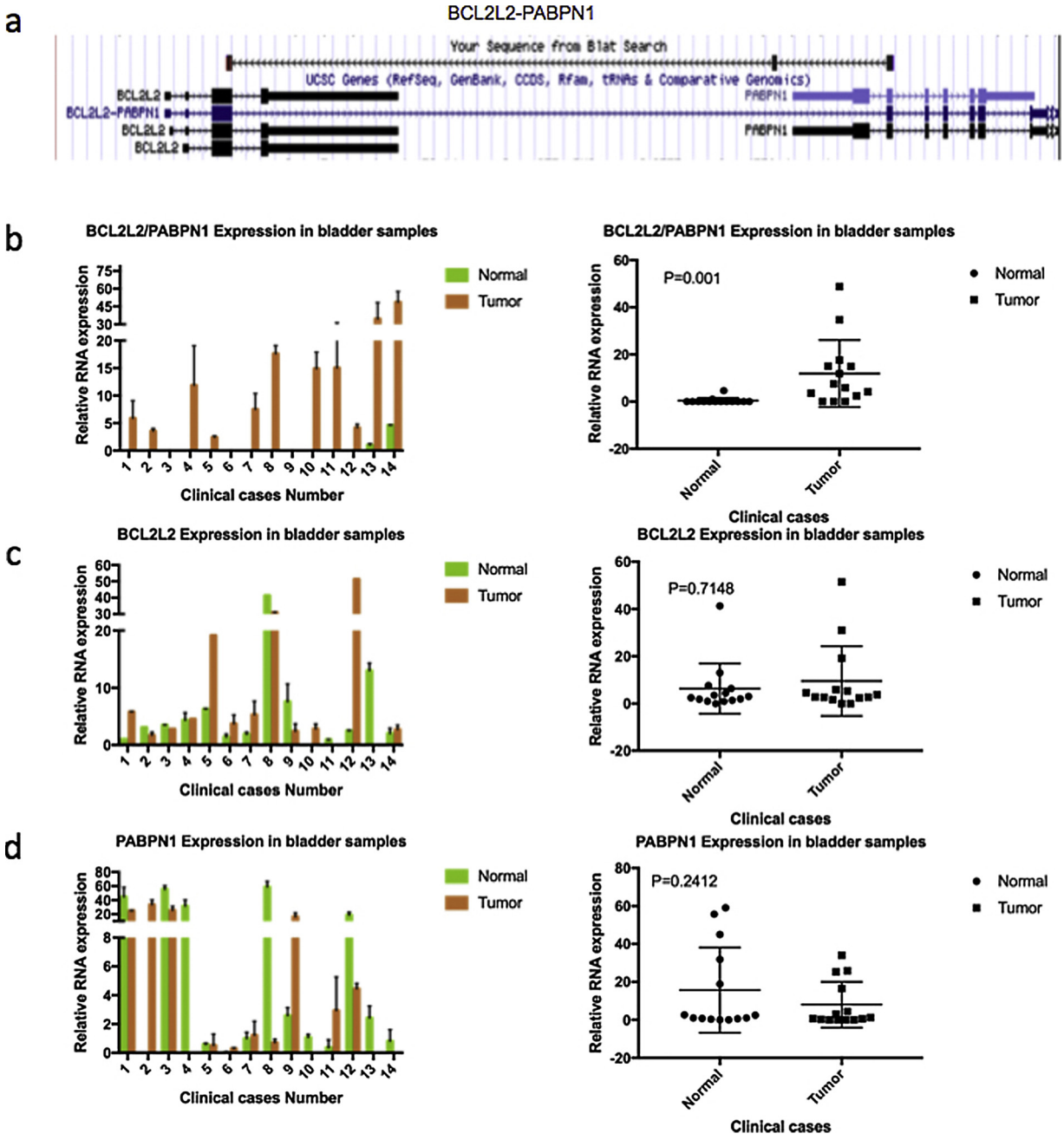

Fig. 4. *BCL2L2-PABPN1*, but not the wild-type parental gene transcripts was preferentially detected in bladder cancers.
(A) Structure of the *BCL2L2-PABPN1* fusion. The reverse primer was used for Sanger sequencing. (B) The fusion was frequently detected by RT-PCR in bladder tumor tissues (T), but not in the adjacent normal tissues (N) (left). The quantification of the fusion levels between normal and tumor groups (right). *P* value was calculated by Wilcoxon matched-pairs signed rank test. (C) and (D) Such differential expression was not observed with the wild-type parental gene *BCL2L2* (C), or *PABPN1* (D).

We believe that such intergenically spliced chimeric RNAs represent a new repertoire of cancer biomarkers and/or therapeutic targets. In this study, we identified six recurrent chimeric RNAs (> 10). Among them, *BCL2L2-PABPN1* and *CHFR-GOLGA3* are preferentially detected in bladder cancers than the normal margin samples. Instead of being a product of chromosomal rearrangement, both fusion RNAs are products of cis-SAGe.

We performed analyses on the landscape of bladder cancer chimeric RNAs on three levels, and from three angles. A large number of the chimeric RNAs are individualized, or only occurring in a small number of samples (< 5). These less frequent chimeras also tend to be M/M fusions, and belong to the category of INTERCHR. Based on our previous experience, this category of chimeric RNAs tend to have a lower validation rate (Qin et al., 2015). It is thus possible that a subset of them are false positives. On the other hand, it is known that many alternative splicings are individualized (Mele et al., 2015). Because we do not have access to the original TCGA samples, we cannot formally test them.

Both *BCL2L2-PABPN1* and *CHFR-GOLGA3* chimeric RNAs were found to be more preferentially detected in tumor samples collected and tested at the University of Virginia than in those collected and tested at Sun Yat-sen University. Several factors may contribute to the discrepancy: 1) different patient population (American vs. Chinese); 2) different researchers who extracted samples; and 3) different real-time PCR equipment. In the future, a larger sample size and more uniformed procedures are needed for direct comparison.

With the development of next generation sequencing and

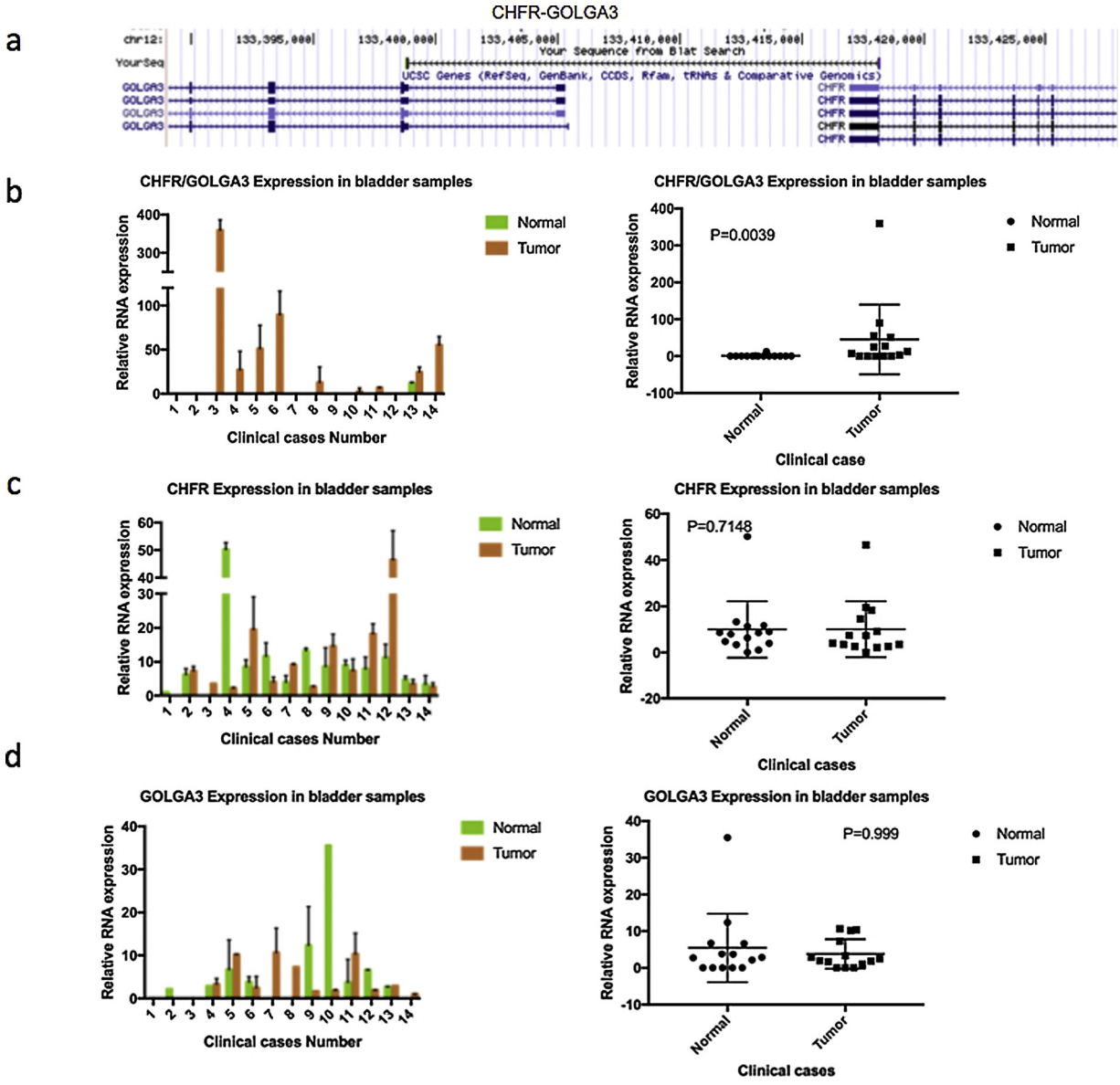

Fig. 5. CHFR-GOLGA3, but not the wild-type parental gene transcripts was preferentially detected in bladder cancers.
(A) Structure of the *CHFR-GOLGA3* fusion. The forward primer was used for Sanger sequencing. (B) The fusion was frequently detected by RT-PCR in bladder tumor tissues (T), but not in the adjacent normal tissues (N) (left). The quantification of the fusion levels between normal and tumor groups (right). *P* value was calculated by Wilcoxon matched-pairs signed rank test. (C) and (D) Such differential expression was not observed with the wild-type parental gene *CHFR* (C), or *GOLGA3* (D).

advancements in bioinformatic tools, the identification of diagnostic and prognostic biomarkers has been accelerated. These include proteomic, genomic, epigenetic and transcriptomic biomarkers (Cheng and Iyer, 2018). In terms of gene fusion biomarkers, it has been shown that urothelial cancer cell lines expressing the *FGFR3–TACC3* or *FGFR3-BAIAP2L1* fusion genes are very sensitive to FGFR-selective agents. Also the detection of the fusions facilitates the selection of patients for FGFR-targeted therapy (Williams et al., 2013). Another report validated and characterized four novel gene fusions *SEPT90-CYHR*, *IGF1R-TTC23*, *SYT8-TNNI2* and *CASZ1-DFFA* in bladder cancer (Kekeeva et al., 2016). Here we showed both *BCL2L2-PABPN1* and *CHFR-GOLGA3* are products of cis-SAGe, and they were preferentially expressed in bladder cancer samples. In contrast, the wild-type parental genes of both fusions were not differentially expressed in bladder cancer versus normal controls. This phenomenon is reminiscent of another chimeric RNA, *SLC45 A3-ELK4* (e1e2 form) in prostate cancer, whose expression correlates with the clinical Gleason Score. Interestingly, neither wild-type *SLC45 A3* or *ELK4* had such a correlation (Zhang et al., 2012). These findings support the notion that chimeric RNAs represent a new repertoire of biomarkers, which would be missed if only canonical transcripts are considered.

Funding sources

This work was supported by NIH NCI grant CA190713 (HL), and the National Natural Science Foundation of China (Grant No. 81825016, 81702523, 81772719) (TL).

Declarations of interests

None of the authors have any conflict to disclose.

Acknowledgments

High-performance computing systems and services were provided by the Data Science Institute and the other Computation and Data

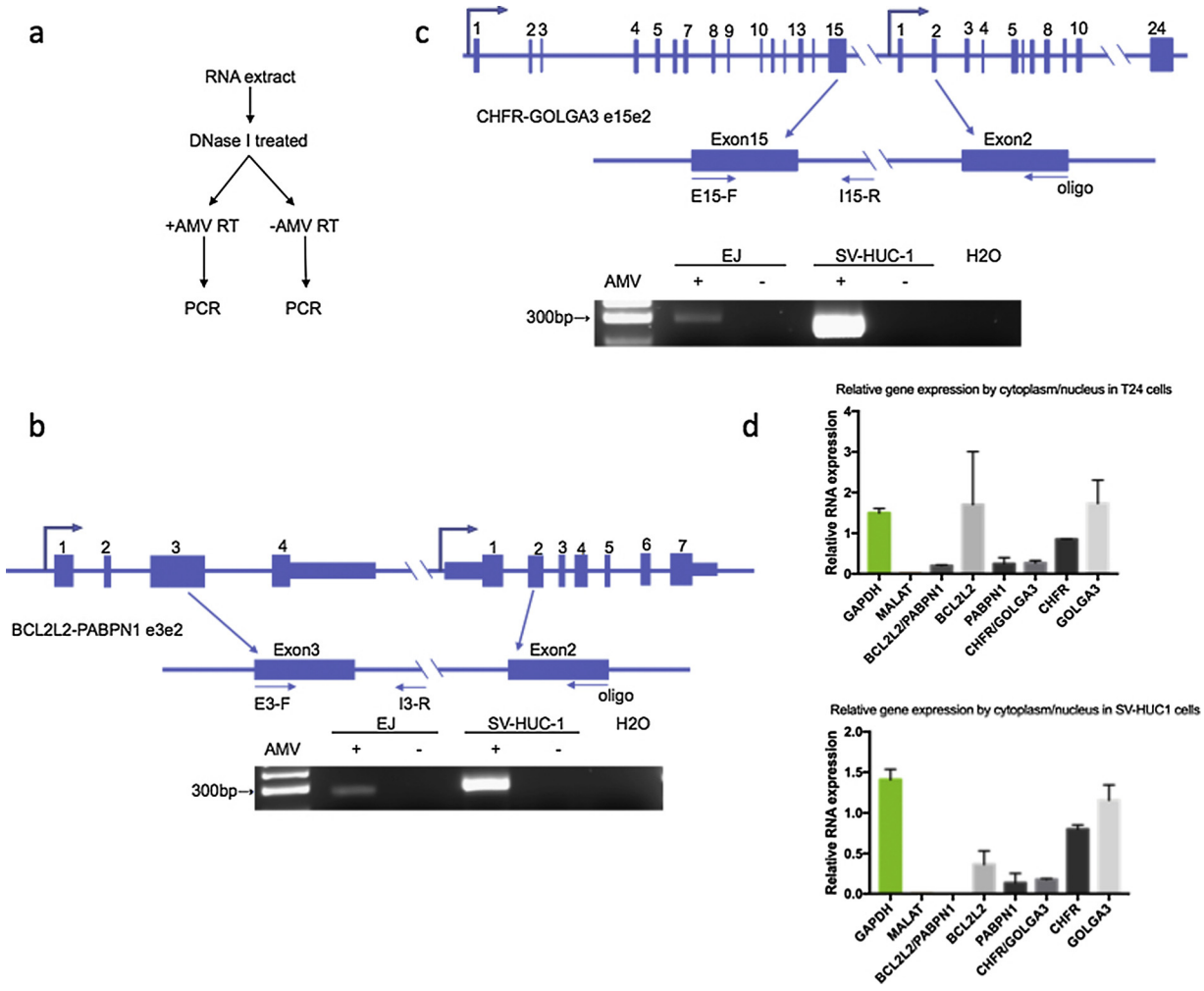

Fig. 6. *BCL2L2-PABPN1*, and *CHFR-GOLGA3* fusions are the products of cis-splicing between adjacent genes, and likely function as long non-coding RNAs. (A) Flow for the mechanistic study of the two chimeras. RNAs were extracted, followed by DNaseI treatment. Samples were then separated into two groups: with or without AMV RT enzyme. (B) The *BCL2L2-PABPN1* fusion involves the joining of the exon3 of *BCL2L2* and the exon2 of *PABPN1*. Blocks represent exons. Lines represent introns or intergenic region. The arrowhead indicates the oligo used for reverse transcription. F and R primers anneal to exon3 and intron3 of *BCL2L2* respectively. RNAs from two bladder cell lines were first treated with DNaseI. They were then separated into two groups: with or without AMV RT enzyme. The correct product was only seen in the samples with AMV-RT enzyme. (C) The *CHFR-GOLGA3* fusion involves the joining of the exon15 of *CHFR* and the exon2 of *GOLGA3*. F and R primers anneal to exon15 and intron15 of *CHFR* respectively. The correct product was only seen in the samples with AMV-RT enzyme. (D) Both fusions were enriched in the nuclear fraction. T24 and SV-HUC1 cells were fractionated into nuclear and cytoplasmic parts. RNA from each part was extracted, and reverse transcribed. *BCL2L2-PABPN1*, *BCL2L2*, *PABPN1*, *CHFR-GOLGA3*, *CHFR*, and *GOLGA3* transcripts were measure by qRT-PCR. Known protein coding RNA, *GAPDH*, and long non-coding RNA, *MALAT1* were used as controls. Ratios of RNAs in cytoplasm and nucleus were plotted.

Resource Exchange (CADRE) partner organizations at the University of Virginia. We thank the Biorepository and Tissue Research Facility (BTRF) at the University of Virginia for providing the clinical samples. We thank Drs. Fujun Qin, Zhongqiu Xie, and Hao Wu for their technical assistance and helpful suggestions.

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.2019.02.007.

References

Benelli, M., Pescucci, C., Marseglia, G., et al., 2012. Discovering chimeric transcripts in paired-end RNA-seq data by using EricScript. Bioinformatics 28, 3232–3239. https://doi.org/10.1093/bioinformatics/bts617.

Bray, F., Ferlay, J., Soerjomataram, I., et al., 2018. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J. Clin. https://doi.org/10.3322/caac.21492.

Cancer Genome Atlas Research, N, Albert Einstein College of, M, Analytical Biological, S, et al., 2017. Integrated genomic and molecular characterization of cervical cancer. Nature 543, 378–384. https://doi.org/10.1038/nature21386.

Chase, A., Ernst, T., Fiebig, A., et al., 2010. TFG, a target of chromosome translocations in lymphoma and soft tissue tumors, fuses to GPR128 in healthy individuals. Haematologica 95, 20–26 doi:haematol.2009.011536 [pii]10.3324/haematol.2009.011536.

Cheng, M.L., Iyer, G., 2018. Novel biomarkers in bladder cancer. Urol. Oncol. 36, 115–119. https://doi.org/10.1016/j.urolonc.2018.01.009.

Eden, E., Navon, R., Steinfeld, I., Lipson, D., Yakhini, Z., 2009. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinf. 10, 48 doi:1471-2105-10-48 [pii]10.1186/1471-2105-10-48.

Finta, C., Zaphiropoulos, P.G., 2002. Intergenic mRNA molecules resulting from trans-splicing. J. Biol. Chem. 277, 5882–5890.

Heim, S., Mitelman, F., 2008. Molecular screening for new fusion genes in cancer. Nat. Genet. 40, 685–686.

Inamura, K., Kobayashi, M., Nagano, H., et al., 2017. A novel fusion of HNRNPA1-ALK in inflammatory myofibroblastic tumor of urinary bladder. Hum. Pathol. 69, 96–100. https://doi.org/10.1016/j.humpath.2017.04.022.

Kekeeva, T., Tanas, A., Kanygina, A., et al., 2016. Novel fusion transcripts in bladder cancer identified by RNA-seq. Cancer Lett. 374, 224–228. https://doi.org/10.1016/j.canlet.2016.02.010.

Kumar-Sinha, C., Kalyana-Sundaram, S., Chinnaiyan, A.M., 2012. SLC45A3-ELK4 chimera in prostate cancer: spotlight on cis-splicing. Cancer Discov. 2, 582–585. https://doi.org/10.1158/2159-8290.cd-12-0212.

Kumar-Sinha, C., Kalyana-Sundaram, S., Chinnaiyan, A.M., 2015. Landscape of gene fusions in epithelial cancers: seq and ye shall find. Genome Med. 7, 129. https://doi.org/10.1186/s13073-015-0252-1.

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

Li, H., Wang, J., Mor, G., Sklar, J., 2008b. A neoplastic gene fusion mimics trans-splicing of RNAs in normal human cells. Science 321, 1357–1361. https://doi.org/10.1126/science.1156725.

Lombardi, B., Ashford, P., Moya-Garcia, A.A., et al., 2017. Unique signalling connectivity of FGFR3-TACC3 oncoprotein revealed by quantitative phosphoproteomics and differential network analysis. Oncotarget 8, 102898–102911. https://doi.org/10.18632/oncotarget.22048.

Mele, M., Ferreira, P.G., Reverter, F., et al., 2015. Human genomics. The human transcriptome across tissues and individuals. Science 348, 660–665. https://doi.org/10.1126/science.aaa0355.

Patel, R.K., Jain, M., 2012. NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. PLoS One 7, e30619. https://doi.org/10.1371/journal.pone.0030619.

Qin, F., Song, Z., Babiceanu, M., et al., 2015. Discovery of CTCF-Sensitive cis-spliced fusion RNAs between adjacent genes in human prostate cells. PLoS Genet. 11, e1005001. https://doi.org/10.1371/journal.pgen.1005001PGENETICS-D-14-03081. [pii].

Qin, F., Zhang, Y., Liu, J., Li, H., 2017. SLC45A3-ELK4 functions as a long non-coding chimeric RNA. Cancer Lett. 404, 53–61. https://doi.org/10.1016/j.canlet.2017.07.007.

Rabbitts, T.H., 1994. Chromosomal translocations in human cancer. Nature 372, 143–149.

Racioppi, M., D'Agostino, D., Totaro, A., et al., 2012. Value of current chemotherapy and surgery in advanced and metastatic bladder cancer. Urol. Int. 88, 249–258. https://doi.org/10.1159/000335556.

Robertson, A.G., Kim, J., Al-Ahmadie, H., et al., 2017. Comprehensive molecular characterization of muscle-invasive bladder cancer. Cell 171, 540–556. https://doi.org/10.1016/j.cell.2017.09.007. e525.

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

Williams, S.V., Hurst, C.D., Knowles, M.A., 2013. Oncogenic FGFR3 gene fusions in bladder cancer. Hum. Mol. Genet. 22, 795–803. https://doi.org/10.1093/hmg/dds486.

Wu, C.S., Yu, C.Y., Chuang, C.Y., et al., 2014. Integrative transcriptome sequencing identifies trans-splicing events with important roles in human embryonic stem cell pluripotency. Genome Res. 24, 25–36 doi:gr.159483.113 [pii]10.1101/gr.159483.113.

Yuan, H., Qin, F., Movassagh, M., et al., 2013. A chimeric RNA characteristic of rhabdomyosarcoma in normal myogenesis process. Cancer Discov. 3, 1394–1403 doi:2159-8290.CD-13-0186 [pii]10.1158/2159-8290.CD-13-0186.

Zaphiropoulos, P.G., 2012. Trans-splicing in higher eukaryotes: implications for cancer development? Front. Genet. 2, 92. https://doi.org/10.3389/fgene.2011.00092.

Zhang, Y., Gong, M., Yuan, H., et al., 2012. Chimeric transcript generated by cis-splicing of adjacent genes regulates prostate cancer cell proliferation. Cancer Discov. 2, 598–607. https://doi.org/10.1158/2159-8290.cd-12-0042.