

RESEARCH ARTICLE

Open Access

Events of alternative splicing in head and neck cancer via RNA sequencing – an update



Vishwas Sharma^{1†}, Amrita Nandan^{2†}, Harpreet Singh^{3,4}, Suyash Agarwal^{3,4}, Richa Tripathi⁵, Dhirendra Narain Sinha⁶ and Ravi Mehrotra^{1*}

Abstract

Background: *Alternative splicing* (AS) is a regulatory mechanism used to create many forms of mature messengers RNAs (mRNAs) from the same gene. Sequencing of RNA (RNA-Seq) is an advanced technology, which has been utilized by different studies to find AS mechanisms in head and neck cancer (HNC). Hitherto, there is no available review that could inform us of the major findings from these studies. Hence, we aim to perform a systematic literature search following PRISMA guidelines to study AS events in HNC identified through RNA-Seq studies.

Results: A total of five records were identified that utilized RNA-Seq data for identifying AS events in HNC. Five software was used in these records to identify AS events. Two genes influenced by AS i.e. *MLL3* and *RPS9* were found to be common in 4 out of 5 records. Likewise, 38 genes were identified to be similar in at least 3 records.

Conclusions: Alternative splicing in HNC is a multifaceted regulatory mechanism of gene expression. It can be studied via RNA-Seq using different bioinformatics tools. Genes *MLL3*, as well as *RPS9*, were repeatedly found to be associated with HNC, however needs further functional validation.

Keywords: Alternative splicing, Exon skipping, Head and neck cancer, Mutually exclusive exon, Retained introns, RNA-Seq, Transcriptome

Background

Alternative splicing (AS) is a key molecular mechanism through which single gene gives rise to multiple differently spliced forms of a protein [1], hence provides greater diversity at transcriptomics and proteomics level [2]. The four major subtypes of AS include a) cassette exon skipping (ES): results in the loss of an exon together with its flanking introns in the alternatively spliced mRNA [3], b) intron retention (IR): part of an exon is either spliced (similar to intron) or retained in the mature mRNA [4–7], c) mutually exclusive exon (MXE): exactly one out of two exons (or one group out of two exon groups) is retained, while the other is spliced out [6] and d) alternative 5' and 3' splice site (ASS): two or more alternative 5' splice site competes for joining with two or greater than two

alternate 3' splice site [8] (details on Additional files 1 and 2 Figure S1). Alternative splicing may result in the addition or elimination of a stop codon in the alternative coding sequence or use of a different translation termination/ initiation site due to a frameshift. Alternative splicing may also result in a change in the internal region of a gene due to an in-frame insertion or deletion [9].

Classically real-time RT-PCR or array-based methods were used to detect AS events, but had their own limitations as they required the prior knowledge of the genome sequence [10, 11]. This shortcoming has been overcome by modern sequencing of RNA (RNA-Seq) technology, which does not require the existing knowledge of genome sequence, and hence could be crucial in identifying previously unidentified genes and splicing events [5]. This technology has expedited the process of identifying all four AS subtypes in cancers [12] and has also been implemented to study the mechanism/s of AS in head and neck cancer (HNC) [13], the sixth most common cancer in men,

* Correspondence: directomicpr@gmail.com

[†]Vishwas Sharma and Amrita Nandan contributed equally to this work.

¹Department of Health Research, National Institute of Cancer Prevention and Research, Noida, Uttar Pradesh, India

Full list of author information is available at the end of the article



comprising around 6% of all cancer incidents [14]. Till date, there is no available review of the literature (Additional file 3) that could update us with the recent developments in finding AS events in HNC via RNA-Seq data. Absence of any relevant literature inspired us to perform a comprehensive review with the aim to highlight the importance of studying AS events in HNC captured through RNA-Seq data for identifying molecular clues for future research and therapy in HNC.

Methods

A comprehensive search was performed to know the AS events in HNC identified through RNA-Seq (Fig. 1 and Additional files 3 and 4Table S1). Briefly, records were screened on PubMed and Web of Science databases on the 19th of September, 2017. In addition, Google (1st November 2017) was also consulted for including records not captured through PubMed and Web of Science search. A total of 323 records were obtained from PubMed, 887 records from Web of Science and 3 additional from Google search. None of the records duplicated from databases, hence a total of 1213 records were

obtained during a comprehensive search. Similar phrases were used for searching articles in PubMed and Web of Science. The only difference is with Web of Science where records published between the years 2010 to 2017 were considered. The details of phrases used for searching records are mentioned in Additional file 3.

The records were screened using inclusion and exclusion criteria mentioned below-.

Inclusion criteria

Original studies that utilized RNA-Seq data for identifying AS events in human oral cancer (OC), oral squamous cell carcinoma (OSCC), squamous cell carcinoma of the oral cavity, squamous cell carcinoma of lip, squamous cell carcinoma of buccal mucosa, squamous cell carcinoma of tongue, carcinoma of floor of the mouth, carcinoma of palate, carcinoma of gingival, carcinoma of maxillary sinus, verrucous carcinoma, Lane tumor, adenoid cystic carcinoma, salivary duct carcinoma, salivary duct adenocarcinoma, ghost cell odontogenic carcinoma, clear cell odontogenic carcinoma, cancers of the nasal cavity, nasopharyngeal carcinoma, African jaw lymphoma, oropharyngeal cancer, laryngeal cancer, laryngeal

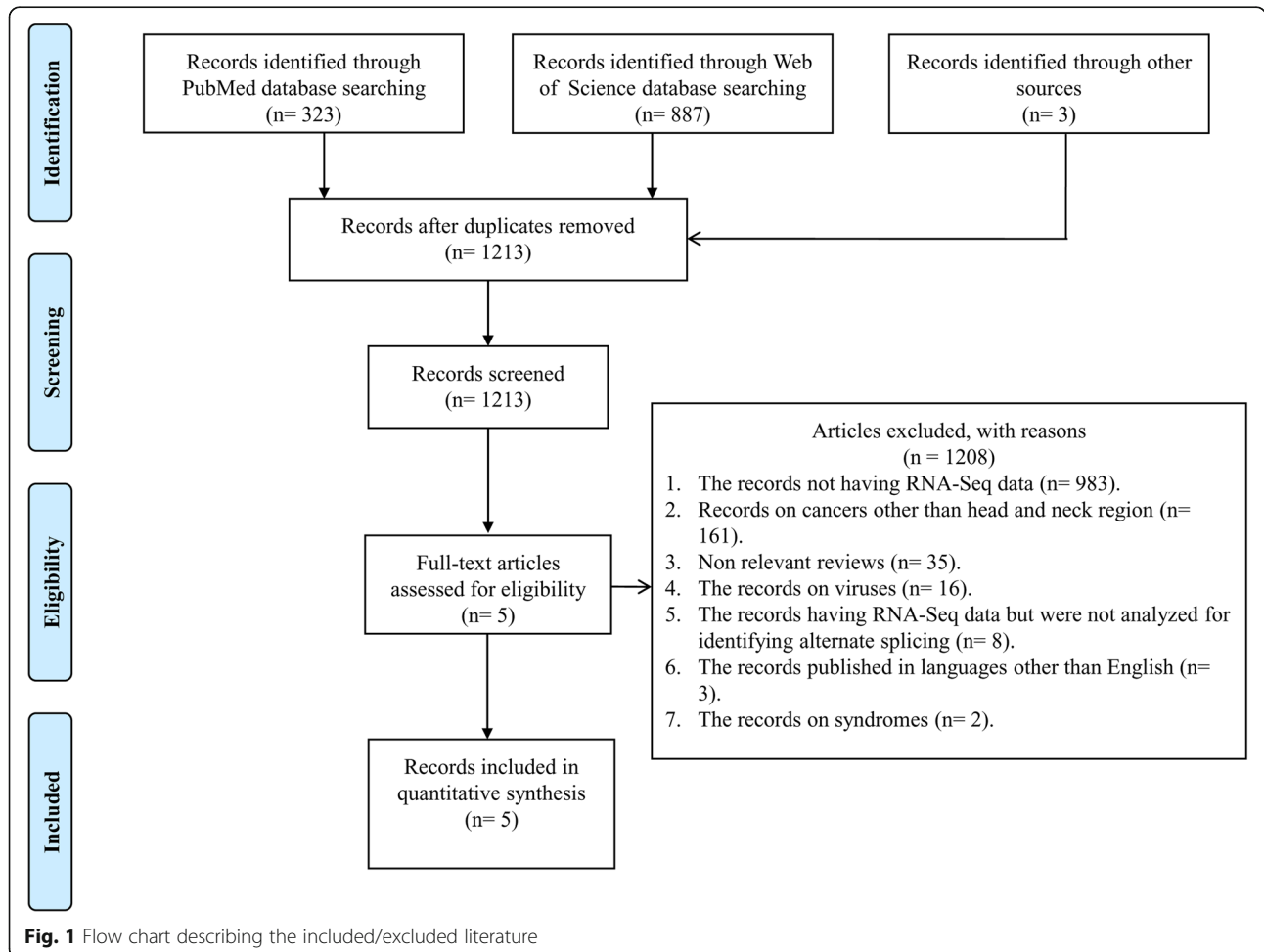


Fig. 1 Flow chart describing the included/excluded literature

squamous cell carcinoma, hypopharyngeal cancer, head and neck cancer, head and neck squamous cell carcinoma (HNSCC) were considered ($n = 5$; identified records comprises oropharyngeal squamous cell carcinoma (OPSCC), OC, buccal mucosa squamous cell carcinoma (BMSCC) and other subtypes of HNC) (Additional file 4 Table S1).

Exclusion criteria

- a) The records not having RNA-Seq data ($n = 983$).
- b) Records on cancers other than head and neck region ($n = 161$).
- c) Non relevant reviews ($n = 35$).
- d) The records on viruses ($n = 16$).
- e) The records having RNA-Seq data but were not analyzed for identifying alternative splicing ($n = 8$).
- f) The records published in languages other than English ($n = 3$).
- g) The records on syndromes ($n = 2$).

Filtering common genes having AS events

In order to find out common genes showing AS events in at least 3 out of 5 independent records, the gene list of all the 5 records was compared (Fig. 2 and Table 1).

Network and functional enrichment analysis

The network and functional enrichment analysis of all the genes, which were common in either 3 or 4 records was done using Cytoscape (version 3.7.1) which uses the

STRING database as a data source. Gene ontology and pathway enrichment analysis was carried out at FDR value ≤ 0.05 . The Cytoscape plugin, cytoHubba was used in order to explore the hub genes from the interactome network, based on eleven topological algorithms including Degree, Edge Percolated Component (EPC), Maximum Neighborhood Component (MNC), Density of Maximum Neighborhood Component (DMNC), Maximal Clique Centrality (MCC) and centralities based on shortest paths as Bottleneck (BN), EcCentricity, Closeness, Radiality, Betweenness, and Stress.

Results

Five records were identified that implemented RNA-Seq data for finding AS events in HNC (Fig. 1). (i) A recent study by Guo et al. [16] conducted on 46 HPV-positive OPSCC and 25 normal control subjects, characterized 109 splicing events on the basis of junctions location and configuration. The junctions were further characterized as ASS (39%), insertion (21%), canonical exon skipping (9%), intron retention (9%), deletion (8%), alternate end site (7%), or noncoding (3%). Of the total 109 validated splicing events, 70% were identified as noncanonical splicing events. Since the noncanonical splicing event contributes to 1% of the total splicing events in normal cells, the abundance of this phenomenon in cancerous cells indicates its importance for further research. Data analysis through MapSplice software [23] resulted in the identification of functionally active splice variant of *AKT3* in HPV

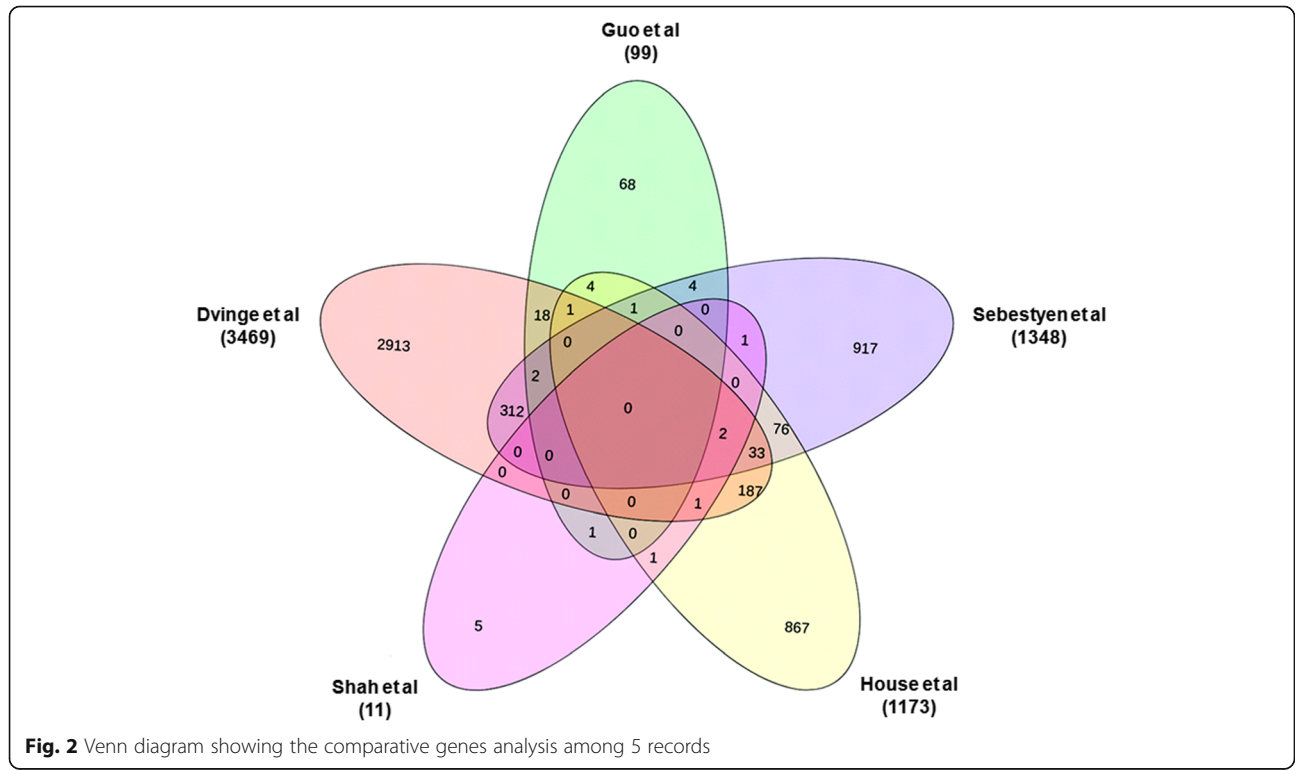


Fig. 2 Venn diagram showing the comparative genes analysis among 5 records

Table 1 Details of genes common in three or four records

	Total genes	Studies under consideration	Total number of common genes	Genes
Genes common in four studies	2	Dvinge et al. [15], Sebestyén et al. [17], House et al. [13] and Shah et al. [18]	2	MLL3, RPS9
Genes common in three studies	38	Dvinge et al. [15], Guo et al. [16] and Sebestyén et al. [17] Dvinge et al. [15], Guo et al. [16] and House et al. [13] Guo et al. [16], Sebestyén et al. [17] and House et al. [13] Dvinge et al. [15], Sebestyén et al. [17] and House et al. [13]	2 1 1 33	CLK1, CLK4 BAI2 S100A4 ATXN2L, BCLAF1, C1orf35, CCDC12, DIAPH1, EIF4G1, FXR1, HDLBP, HMG20B, HSPB1, INTS6, LAS1L, NSUN5, PCBP2, PRPF31, PRPF40B, RBM8A, RPS11, RPS18, RPS20, RPS25, RPS26, RPS27A, RPS28, SLC16A3, SLC25A11, SPTAN1, TAGLN2, TBCA, TSEN2, U2AF1, WBSR22, YTHDF3
		Dvinge et al. [15], House et al. [13] and Shah et al. [18]	1	STAT6

positive OPSCC primary tumors. The AKT is a signal transduction pathway involved in cell growth and survival in response to extracellular signals and is well known to be linked with the pathogenesis of HNSCC [16]. The observed correlation between DNA methylation and alternative start site of AS events in this study are suggested to be important for further investigation. By utilizing a unique method of analysis the authors filtered 100 tumor-specific splice variant candidates in HPV positive HNSCC. The method described in this study could be followed for the identification of novel splicing events and to better understand the complex perturbations occurring in the tumor cells. Parallel identification of correlated methylation patterns could add in defining targets for validation and therapy. Comparison of AS events between different tumor types may give a clue towards a defined structure of module followed by either the tumor cells collectively or a given group of tumor types based on origin. This module could be further validated and targeted to define a treatment regime. (ii) Studying AS events by using a computational tool SUPPA [19], Sebestyén, et al. [17] identified RNA binding motif protein 47 (*RBM47*) tumor suppressor gene to be the main common splicing factor between kidney renal papillary carcinoma and HNSCC. The study reported 141 common AS events between HNSCC and Lung Squamous Cell Carcinoma and 63 out of the 141 splicing events to be associated with RNA binding motif protein 28 (*RBM28*). The analysis was performed on 38 HNSCC and 38 control sample data downloaded from the TCGA data portal [32]. The study reported the importance of alterations in the genes encoding RNA binding proteins and that they trigger common and specific AS events in various solid tumors. Thus, suggesting a pattern of AS events followed by several solid tumors that could be studied to understand the molecular basis and oncogenic properties of tumor cells for defining treatment modalities. (iii) An IR has been recently observed to be an important factor behind the disruption of normal splicing machinery in cancerous cells [15, 33]. An important finding that abnormal RNA splicing, characterized by widespread IR events, is common across cancers was marked by a study performed by Dvinge and Bradley [15] on 40 HNC patients data downloaded from CGHub [34]. The mechanism of IR is a multifactorial process and it is suggested that transcriptional diversity in cancer cells may be driven by at least some intron-containing mRNAs that are adequately stable [15]. Whether these introns containing mRNAs are translated into proteins or degraded still remains an interesting topic of investigation. Studying AS events identified 829 cassette exons, 153 competing 5' splice sites, 202 competing 3' splice sites, 115 retained introns, 290 constitutive introns and 418 constitutive junctions. Moreover, crosstalk between RNA splicing events and epigenetics may be associated with a

diverse spectrum of mutations to RNA processing defects in cancer. The bioinformatic tool MISO [21] which is based on statistical models that estimate expression of alternatively spliced exons and isoforms and evaluates confidence in these estimates was applied for the RNA-Seq data analysis in the study. (iv) To investigate the role of CUGBP embryonic lethal abnormal vision-like family member 1 (*CELF1*), an important regulator of AS, editing and translation, in OC, House et al. [13] used RNA-Seq data to analyze AS variants in si-control and siCELF1 transfected OC cells. The AltAnalyze software which runs on two algorithms to quantify AS events, i.e. ASPIRE (Analysis of Splicing Isoform Reciprocity) and splicing index was utilized to perform the analysis [20]. The former finds AS events by matching-up the levels of the alternatively expressed exons and exon-exon junctions whereas the latter derives AS events via normalizing single exon expression to total gene expression levels. The study captured 315 AS events comprising 282 genes. Cell cycle, microtubule-based processes and translocation initiation constituted the three most significant over-represented biological functions of 282 splicing targets. The most common AS events regulated by CELF1 were ES and inclusion. CELF1 is a critical regulator of cancer cell proliferation and apoptosis. Since the authors were interested in studying CELF1 mediated AS events in three pre-mRNAs (*COL16A1*, *TACC2* and *ITGA6*) known to play a role in cell growth, they found a modest difference in the expression for *COL16A1*, *TACC2*, and *ITGA6* CELF1-mediated splicing events in HNSCC tumors in comparison to normal samples. Therefore, the rest of the data still remains to be studied for other mechanisms influenced by deregulation of CELF1 in HNSCC. (v) While analyzing data from BMSCC through SpliceMap software [22] Shah et al. [18] found 11 novel splice junctions, encompassing *IMPA2*, *CD44*, *FAM48A*, *DHX32*, *MLL3*, *STAT6*, *FAM82B*, *SPICE1*, *MEMO1*, *RPS9* and *SPINK5* genes, mostly derived from 5' ASS. All of them were known to alter the coding region and cause an alteration in protein sequence leading to gain or loss of protein function and change in specificity. Thus a relevant biological consequence of AS events was reported through this study. All the studies under consideration showed the IR event as the major component of AS in HNC, with the only exception being the study by Shah et al where IR was not detected as an AS event. This discrepancy could be due to the difference in the sequencing platform used, as others [13, 15–17] utilized Illumina however Shah et al. [18] used Roche 454 to generate the RNA-Seq data. Moreover, most of the novel splice junctions reported in the study were supported either by two reads or a single read. The AS events supported by a single read were excluded from further analysis which could be possible IR events and may have been captured if the sequencing depth was more.

Two genes, i.e. *MLL3* and *RPS9* were found to be common in 4 out of 5 records (Table 1 and Fig. 2). Similarly, 38 genes were identified to be common in 3 records (Table 1). Hence, a total of 40 genes were used for gene enrichment analysis. In the records identified, the authors analyzed the RNA-Seq data using five softwares i.e. Suppa [19], AltAnalyze [20], MISO (Mixture of Isoforms Probabilistic Model) [21] SpliceMap [22] and MapSplice [23], for finding AS events in HNC. The top 10 gene ontology enrichment pertaining to the cellular component, molecular function, and biological process are depicted in Additional file 5: Figure S2. The major pathways being regulated by these genes include ribosome, spliceosome and RNA transport (Additional file 6). The protein-protein network of all the 40 genes was (Additional file 7Figure S3) analyzed using cytoHubba resulting in 5 genes which passed more than 10 topological algorithms, out of eleven to be called as the major HUB genes which include *RPS28*, *RBM8A*, *RPS9*, *EIF4G1* and *RPS27A* (Additional file 6).

Discussion

Alternative splicing of genes is believed to be associated with the initiation, progression, and metastasis of cancer [24–26]. Differences in RNA processing may differentiate cancer and normal cells, as spliceosomal components are differentially required in cancer versus normal cells [27, 28]. Due to these differences, recent ongoing research on antitumor drugs focuses on the inhibitors of RNA splicing [29–31]. In the last few years, clinicians/researchers have utilized the power of RNA-Seq for understanding the mechanism/s of AS at genome-wide level, that was not feasible previously with classical methods. Knowledge of these events will assist us in understanding the complex regulation of gene expression in cancer. Recently, an abundance of effort has been put into extracting information about AS in HNC via RNA-Seq data, with the aim to utilize/implement this information in finding new drug targets or searching candidate molecular markers for early and accurate diagnosis of the disease. Considering the significance of the topic, the literature was reviewed and compiled to provide information on the recent ongoing in finding AS events, through RNA-Seq data, for HNC.

The subsequent literature search led to the filtration of five records associated with RNA-Seq data analysis for discovering AS events in HNC. Four records identified *MLL3* and *RPS9* as a common gene influenced by AS in HNSCC [13, 15, 17, 18]. The *MLL3* (mixed lineage leukemia-3) encodes histone-modifying enzyme characterized by a conserved catalytic SET domain and is involved in transcriptional coactivation. *MLL3* is considered as a tumor suppressor gene because of its frequent mutations in several tumor types [35, 36]. Recurrent mutations in *MLL3* have been reported in HNC in addition to several other cancers, including pancreatic, breast and gastric

[37–41]. In a study by Mountzios et al, [37] analysis of the data from “The Cancer Genome Atlas”, *MLL3* was identified with a mutation rate of 7.3% in HNC [37]. Downregulated expression of *MLL3* is reported to contribute to tumor progression in HNSCC [42]. Additionally, reduced expression of *MLL3* is reported in breast tumors and gastric cancer [43, 44]. Ribosomal Protein S9 (*RPS9*) is one of the first proteins that directly bind to the 18S rRNA and plays a crucial role in ribosome biogenesis [45–47]. Inhibition of *RPS9* activates p53 activity, it is suggested to be an efficient method to induce apoptosis in tumor cells [48]. Several studies have reported the role of *RPS9* in tumorigenesis [45, 48]. Since the protein induces the p53 pathway and was observed to be involved in HNC, futuristic validation is required for its functional importance in HNC. We have discussed genes common in 4 records, however additional 38 genes common in at least 3 records may also be of biological relevance. The records studied used Suppa [19], AltAnalyze [20], MISO [21] SpliceMap [22] and MapSplice [23], for finding AS events in HNC. Apart from that, there are also other AS detection software as listed in Additional file 8Table S2 which are already implemented in cancer research and hence could be used for finding AS in HNC.

Conclusion

The comprehensive literature search has indicated towards the gap in utilizing RNA-Seq data to identify and study AS events in HNC. Utilization of RNA-Seq data to study the AS events in HNC is crucial to understand the complex cascade of events regulating gene expression. However, in the current review only 5 articles relevant to AS events in HNC were identified, of which 3 were from patients with HNC, oral cancer, and buccal mucosa tumors, respectively, one study deals with TCGA data and the last has been conducted using cell lines. Although the data is heterogeneous, but due to the limited number of studies available, HNC was considered as an umbrella under which other subtypes such as oral cancer, and buccal mucosa tumor falls. For a more robust interpretation of genes involved in AS events in HNC, a larger number of studies at a subtype level are required from different geographical regions. Interestingly, in next-generation sequencing (NGS) based data repository such as GEO [49], SRA [50], TCGA [32], etc. RNA-Seq data of more than a hundred fifty HNC patients are available and could be analyzed through various bioinformatics softwares. We suggest comparing the data captured on the similar sequencing platform, and analyze them through multiple AS events identification software, followed by the selection of common candidate AS events. With the reducing cost of RNA-Seq technology, it is expected that in the next few years data repositories will be enriched with HNC patient's data that will help diagnostic/pharma world in finding new therapeutic and molecular targets for detecting

the disease at a preliminary stage. This will also assist clinicians in deciding their treatment and management strategies against HNC. Although RNA-Seq data has few limitations [51] but still it is a reliable biological tool for finding novel AS events. In the near future, RNA-Seq technology is expected to be a hallmark in investigating novel AS events for HNC or any other types of carcinoma.

Additional files

- Additional file 1:** Alternative splicing (AS) subtypes. (DOC 89 kb)
- Additional file 2: Figure S1.** Alternative splicing subtypes. The figure was adopted and modified from Shen et al. [52]. (TIF 501 kb)
- Additional file 3:** Criteria of screening studies. (DOC 48 kb)
- Additional file 4: Table S1.** describing the reason for including or excluding each study. (XLS 232 kb)
- Additional file 5: Figure S2.** Gene ontology enrichment graph. (TIFF 206 kb)
- Additional file 6:** Functional enrichment and hub genes. (XLS 47 kb)
- Additional file 7: Figure S3.** Network analysis of all the 40 common genes. (TIFF 3427 kb)
- Additional file 8: Table S2.** Summary of the alternative splicing detection softwares and their implementation in cancer research. (DOC 139 kb)

Abbreviations

AS: alternative splicing; ASPIRE: Analysis of Splicing Isoform Reciprocity; ASS: alternative 5' and 3' splice site; BMSCC: buccal mucosa squamous cell carcinoma; CELF1: CUGBP embryonic lethal abnormal vision-like family member 1; CGHub: Cancer Genomics Hub; ES: exon skipping; GEO: Gene Expression Omnibus; HNC: head and neck cancer; HNSCC: head and neck squamous cell carcinoma; HPV: human papillomavirus; IR: intron retention; MISO: Mixture of Isoforms probabilistic model; MLL3: mixed lineage leukemia-3; mRNA: messenger RNA; MXE: mutually exclusive exon; OC: oral cancer; OPSCC: oropharyngeal squamous cell carcinoma; OSCC: oral squamous cell carcinoma; RBM28: RNA binding motif protein 28; RBM47: RNA binding motif protein 47; RNA-Seq: Sequencing of RNA; RPS9: Ribosomal Protein S9; RT-PCR: Real Time-Polymerase Chain Reaction; SRA: Sequence Read Archive; TCGA: The Cancer Genome Atlas

Acknowledgments

The authors are thankful to Dr. Suzanne Nathan from NICPR, Noida and Dr. Asgar from IGIB, New Delhi for giving critical suggestions on the manuscript. We also thanks to Dr. Heidi Dvinge and Dr. Robert K Bradley from Fred Hutchinson Cancer Research Center, US for sharing gene list with us. The grammar was checked by 'http://www.gingersoftware.com/.

Funding

This work was funded by the Indian Council of Medical Research by grant No.BIC/11(25)/2015 and No. 2019-3557 which supported the fellowship of Dr. Vishwas Sharma who was involved in designing the study, collection of information, analysis and interpretation of data, and in writing the first draft of the manuscript.

Availability of data and materials

The data that support the findings of this study are available from Dvinge et al, Guo et al, Sebestyén et al, House et al, and Shah et al. The data i.e. gene list is available freely in Guo et al, Sebestyén et al, House et al, and Shah et al. articles. The data from Dvinge et al, article was obtained from Dr. Heidi Dvinge and Dr. Robert K Bradley hence restrictions apply to the availability of these data.

Authors' contributions

VS wrote the first draft of the review and along with HS and AN conceived the study. SA performed the gene comparison analysis and gave bioinformatic suggestions. RT reviewed and edited the manuscript. RM and DS lead, managed and wrote the final version of the manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Department of Health Research, National Institute of Cancer Prevention and Research, Noida, Uttar Pradesh, India. ²Society for Life Science and Human Health, Allahabad, Uttar Pradesh, India. ³ICMR Computational Genomics Centre, Indian Council of Medical Research, New Delhi 110029, India. ⁴Informatics, Systems and Research Management, Indian Council of Medical Research, New Delhi 110029, India. ⁵Division of Molecular Cytology, National Institute of Cancer Prevention and Research, Noida, Uttar Pradesh, India. ⁶WHO FCTC Global Knowledge Hub on Smokeless Tobacco, National Institute of Cancer Prevention and Research, Noida, Uttar Pradesh, India.

Received: 8 June 2018 Accepted: 10 May 2019

Published online: 03 June 2019

References

- Ghigna C, Valacca C, Biamonti G. Alternative splicing and tumor progression. *Curr Genomics* [Internet]. 2008;9:556–70 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19516963>.
- Cieply B, Carstens RP. Functional roles of alternative splicing factors in human disease. *Wiley Interdiscip Rev RNA* [Internet]. 6:311–26 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25630614>.
- Wang J, Ye Z, Huang TH, Shi H, Jin VX. Computational methods and correlation of exon-skipping events with splicing, transcription, and epigenetic factors. *Methods Mol Biol* [Internet]. 2017;1513:163–170. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27807836>.
- Matlin AJ, Clark F, Smith CWJ. Understanding alternative splicing: towards a cellular code. *Nat Rev Mol Cell Biol* [Internet]. 2005;6:386–398. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/15956978>.
- Bai Y, Ji S, Wang Y. IRcall and IRclassifier: two methods for flexible detection of intron retention events from RNA-Seq data. *BMC Genomics* [Internet]. 2015;16(Suppl 2):S9 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25707295>.
- Sammeth M. Complete alternative splicing events are bubbles in splicing graphs. *J Comput Biol* [Internet]. 2009;16:1117–1140. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19689216>.
- Pillmann H, Hatje K, Odronitz F, Hammesfahr B, Kollmar M. Predicting mutually exclusive spliced exons based on exon length, splice site and reading frame conservation, and exon sequence homology. *BMC Bioinformatics* [Internet]. 2011;12:270 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21718515>.
- Koren E, Lev-Maor G, Ast G. The emergence of alternative 3' and 5' splice site exons from constitutive exons. *PLoS Comput Biol* [Internet]. 2007;3:e95 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17530917>.
- Modrek B, Lee C. A genomic view of alternative splicing. *Nat Genet* [Internet]. 2002;30:13–9 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/11753382>.
- Ner-Gaon H, Fluhr R. Whole-genome microarray in Arabidopsis facilitates global analysis of retained introns. *DNA Res* [Internet]. 2006;13:111–21 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16980712>.
- Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* [Internet]. 2009;10:57–63 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19015660>.
- Feng H, Qin Z, Zhang X. Opportunities and methods for studying alternative splicing in cancer with RNA-Seq. *Cancer Lett* [Internet]. 2013;340:179–91 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23196057>.
- House RP, Talwar S, Hazard ES, Hill EG, Palanisamy V. RNA-binding protein CELF1 promotes tumor growth and alters gene expression in oral squamous cell carcinoma. *Oncotarget* [Internet]. 2015;6:43620–34 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26498364>.
- Ng S-H, Yen T-C, Liao C-T, Chang JT-C, Chan S-C, Ko S-F, et al. 18F-FDG PET and CT/MRI in oral cavity squamous cell carcinoma: a prospective study of

- 124 patients with histologic correlation. *J Nucl Med* [Internet]. 2005;46:1136–43 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16000282>.
15. Dvinge H, Bradley RK. Widespread intron retention diversifies most cancer transcriptomes. *Genome Med* [Internet]. 2015;7:45 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26113877>.
 16. Guo T, Sakai A, Afari B, Considine M, Danilova L, Favorov A V, et al. A novel functional splice variant of AKT3 defined by analysis of alternative splice expression in HPV-positive oropharyngeal cancers. *Cancer Res* [Internet]. 2017;77:5248–5258. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28733453>
 17. Sebestyén E, Singh B, Miñana B, Pagès A, Mateo F, Pujana MA, et al. Large-scale analysis of genome and transcriptome alterations in multiple tumors unveils novel cancer-relevant splicing networks. *Genome Res* [Internet]. 2016;26:732–44 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27197215>.
 18. Shah TM, Patel AK, Bhatt VD, Tripathi AK, Shah S, Shankar V, et al. The landscape of alternative splicing in buccal mucosa squamous cell carcinoma. *Oral Oncol* [Internet]. 2013;49:604–10 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23566772>.
 19. Alamancos GP, Pagès A, Trincado JL, Bellora N, Eyras E. Leveraging transcript quantification for fast computation of alternative splicing profiles. *RNA* [Internet]. 2015;21:1521–31 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26179515>.
 20. Emig D, Salomonis N, Baumbach J, Lengauer T, Konkin BR, Albrecht M. AltAnalyze and DomainGraph: analyzing and visualizing exon expression data. *Nucleic Acids Res* [Internet]. 2010;38:W755–62 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20513647>.
 21. Katz Y, Wang ET, Airoidi EM, Burge CB. Analysis and design of RNA sequencing experiments for identifying isoform regulation. *Nat Methods* [Internet]. 2010;7:1009–15 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21057496>.
 22. Au KF, Jiang H, Lin L, Xing Y, Wong WH. Detection of splice junctions from paired-end RNA-seq data by SpliceMap. *Nucleic Acids Res* [Internet]. 2010;38:4570–8 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20371516>.
 23. Wang K, Singh D, Zeng Z, Coleman SJ, Huang Y, Savich GL, et al. MapSplice: accurate mapping of RNA-seq reads for splice junction discovery. *Nucleic Acids Res* [Internet]. 2010;38:e178 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20802226>.
 24. Shapiro IM, Cheng AW, Flytzanis NC, Balsamo M, Condeelis JS, Oktay MH, et al. An EMT-driven alternative splicing program occurs in human breast cancer and modulates cellular phenotype. *PLoS Genet* [Internet]. 2011;7:e1002218 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21876675>.
 25. Yae T, Tsuchihashi K, Ishimoto T, Motohara T, Yoshikawa M, Yoshida GJ, et al. Alternative splicing of CD44 mRNA by ESRP1 enhances lung colonization of metastatic cancer cell. *Nat Commun* [Internet]. 2012;3:883 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22673910>.
 26. Oltean S, Bates DO. Hallmarks of alternative splicing in cancer. *Oncogene* [Internet]. 2014;33:5311–8 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24336324>.
 27. Chen L, Tovar-Corona JM, Urrutia AO. Increased levels of noisy splicing in cancers, but not for oncogene-derived transcripts. *Hum Mol Genet* [Internet]. 2011;20:4422–9 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21862452>.
 28. Hubert CG, Bradley RK, Ding Y, Toledo CM, Herman J, Skutt-Kakaria K, et al. Genome-wide RNAi screens in human brain tumor isolates reveal a novel viability requirement for PHF5A. *Genes Dev* [Internet]. 2013;27:1032–45 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23651857>.
 29. Bonnal S, Vigevani L, Valcárcel J. The spliceosome as a target of novel antitumour drugs. *Nat Rev Drug Discov* [Internet]. 2012;11:847–59 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23123942>.
 30. O'Brien K, Matlin AJ, Lowell AM, Moore MJ. The biflavonoid isoginkgetin is a general inhibitor of pre-mRNA splicing. *J Biol Chem* [Internet]. 2008;283:33147–54 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/18826947>.
 31. Kaida D, Motoyoshi H, Tashiro E, Nojima T, Hagiwara M, Ishigami K, et al. Spliceostatin A targets SF3b and inhibits both splicing and nuclear retention of pre-mRNA. *Nat Chem Biol* [Internet]. 2007;3:576–83 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17643111>.
 32. The cancer genome atlas National Cancer Institute. National Human Genome Research Institute. US. Available from: <https://cancergenome.nih.gov/>
 33. Jung H, Lee D, Lee J, Park D, Kim YJ, Park W-Y, et al. Intron retention is a widespread mechanism of tumor-suppressor inactivation. *Nat Genet* [Internet]. 2015;47:1242–8 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26437032>.
 34. Wilks C, Cline MS, Weiler E, Diehkans M, Craft B, Martin C, et al. The Cancer genomics hub (CGHub): overcoming cancer through the power of torrential data. *Database* (Oxford) [Internet]. 2014;2014. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25267794>
 35. Weirich S, Kudithipudi S, Kycia I, Jeltsch A. Somatic cancer mutations in the MLL3-SET domain alter the catalytic properties of the enzyme. *Clin Epigenetics* [Internet]. 2015;7:36 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25829971>.
 36. Chen C, Liu Y, Rappaport AR, Kitzing T, Schultz N, Zhao Z, et al. MLL3 is a haploinsufficient 7q tumor suppressor in acute myeloid leukemia. *Cancer Cell* [Internet]. 2014;25:652–65 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24794707>.
 37. Mountzios G, Rampias T, Psyrri A. The mutational spectrum of squamous-cell carcinoma of the head and neck: targetable genetic events and clinical impact. *Ann Oncol Off J Eur Soc Med Oncol* [Internet]. 2014;25:1889–900 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24718888>.
 38. Li B, Liu H-Y, Guo S-H, Sun P, Gong F-M, Jia B-Q. A missense mutation (S3660L) in MLL3 gene influences risk of gastric cancer. *J BUON* [Internet]. 19:394–7 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24965397>.
 39. Wang X-X, Fu L, Li X, Wu X, Zhu Z, Fu L, et al. Somatic mutations of the mixed-lineage leukemia 3 (MLL3) gene in primary breast cancers. *Pathol Oncol Res* [Internet]. 2011;17:429–33 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21116761>.
 40. Balakrishnan A, Bleeker FE, Lamba S, Rodolfo M, Daniotti M, Scarpa A, et al. Novel somatic and germline mutations in cancer candidate genes in glioblastoma, melanoma, and pancreatic carcinoma. *Cancer Res* [Internet]. 2007;67:3545–50 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17440062>.
 41. Kandath C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, et al. Mutational landscape and significance across 12 major cancer types. *Nature* [Internet]. 2013;502:333–9 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24132290>.
 42. Figueiredo DLA, do Amaral Rabello D, Zanette DL, Saggioreto FP, Mamede RCM, Zago MA, et al. Assessment of MLL methyltransferase gene expression in larynx carcinoma. *Oncol Rep* [Internet]. 2015;33:2017–22 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25633166>.
 43. Kunizaki M, Hamamoto R, Silva FP, Yamaguchi K, Nagayasu T, Shibuya M, et al. The lysine 831 of vascular endothelial growth factor receptor 1 is a novel target of methylation by SMDY3. *Cancer Res* [Internet]. 2007;67:10759–65 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/18006819>.
 44. Rabello D do A, de Moura CA, de Andrade RV, Motoyama AB, Silva FP. Altered expression of MLL methyltransferase family genes in breast cancer. *Int J Oncol* [Internet]. 2013;43:653–60 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23754336>.
 45. Cheng D-D, Zhu B, Li S-J, Yuan T, Yang Q-C, Fan C-Y. Down-regulation of RPS9 inhibits osteosarcoma cell growth through inactivation of MAPK signaling pathway. *J Cancer* [Internet]. 2017;8:2720–2728. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28928861>
 46. Lindström MS. Elucidation of motifs in ribosomal protein S9 that mediate its nucleolar localization and binding to NPM1/nucleophosmin. *PLoS One* [Internet]. 2012;7:e52476 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23285058>.
 47. O'Donohue M-F, Choessel V, Faubladier M, Fichant G, Gleizes P-E. Functional dichotomy of ribosomal proteins during the synthesis of mammalian 40S ribosomal subunits. *J Cell Biol* [Internet]. 2010;190:853–66 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20819938>.
 48. Lindström MS, Nistér M. Silencing of ribosomal protein S9 elicits a multitude of cellular responses inhibiting the growth of cancer cells subsequent to p53 activation. *PLoS One* [Internet]. 2010;5:e9578 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20221446>.
 49. Barrett T. Gene Expression Omnibus (GEO). 2013; Available from: <https://www.ncbi.nlm.nih.gov/books/NBK159736/>.
 50. Leinonen R, Sugawara H, Shumway M, International nucleotide sequence database collaboration. The sequence read archive. *Nucleic Acids Res* [Internet]. 2011;39:D19–D21. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21062823>
 51. Conesa A, Madrigal P, Tarazona S, Gomez-Cabrero D, Cervera A, McPherson A, et al. A survey of best practices for RNA-seq data analysis. *Genome Biol*

[Internet]. 2016;17(13) Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26813401>.

52. Shen S, Park JW, Huang J, Dittmar KA, Lu Z, Zhou Q, et al. MATS: a Bayesian framework for flexible detection of differential alternative splicing from RNA-Seq data. *Nucleic Acids Res* [Internet]. 2012;40:e61 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22266656>.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

