

REVIEW

Open Access



Transcriptomics and Epigenomics in head and neck cancer: available repositories and molecular signatures

Mara S. Serafini^{1†}, Laura Lopez-Perez^{2†}, Giuseppe Fico², Lisa Licitra^{3,4}, Loris De Cecco^{1†} and Carlo Resteghini^{3*†} 

Abstract

For many years, head and neck squamous cell carcinoma (HNSCC) has been considered as a single entity. However, in the last decades HNSCC complexity and heterogeneity have been recognized. In parallel, high-throughput *omics* techniques had allowed picturing a larger spectrum of the behavior and characteristics of molecules in cancer and a large set of omics web-based tools and informative repository databases have been developed. The objective of the present review is to provide an overview on biological, prognostic and predictive molecular signatures in HNSCC. To contextualize the selected data, our literature survey includes a short summary of the main characteristics of omics data repositories and web-tools for data analyses. The timeframe of our analysis was fixed, encompassing papers published between January 2015 and January 2019. From more than 1000 papers evaluated, 61 omics studies were selected: 33 investigating mRNA signatures, 11 and 13 related to miRNA and other non-coding-RNA signatures and 4 analyzing DNA methylation signatures. More than half of identified signatures (36) had a prognostic value but only in 10 studies selection of a specific anatomical sub-site (8 oral cavity, 1 oropharynx and 1 both oral cavity and oropharynx) was performed. Noteworthy, although the sample size included in many studies was limited, about one-half of the retrieved studies reported an external validation on independent dataset(s), strengthening the relevance of the obtained data. Finally, we highlighted the development and exploitation of three gene-expression signatures, whose clinical impact on prognosis/prediction of treatment response could be high. Based on this overview on *omics-related* literature in HNSCC, we identified some limits and strengths. The major limits are represented by the low number of signatures associated to DNA methylation and to non-coding RNA (miRNA, lncRNA and piRNAs) and the availability of a single dataset with multiple omics on more than 500 HNSCC (i.e. TCGA). The major strengths rely on the integration of multiple datasets through meta-analysis approaches and on the growing integration among *omics* data obtained on the same cohort of patients. Moreover, new approaches based on artificial intelligence and informatic analyses are expected to be available in the next future.

Keywords: Head and neck squamous cell carcinoma, HNSCC, Transcriptomics, Epigenomics, Oral cavity, Oropharynx, Larynx, Hypopharynx, Prognosis, HPV

Background

Head and neck squamous cell carcinoma (HNSCC) is the seventh most frequent cancer, with a worldwide incidence of 0.7 million new cases per year, and a low 5-year survival rate for both localized and advanced disease (69 and 34%, respectively) [1]. For several years,

HNSCC has been considered as a single entity, since all sub-sites (i.e. oral cavity, oropharynx, larynx, hypopharynx) share a common epithelial precursor. Based on this assumption, treatment and biological analyses were mostly applied with no distinction for each of the sub-sites. However, clinical-pathological features and molecular changes, driving carcinogenesis [2], have helped in recognizing HNSCC complexity and heterogeneity. In addition, The Human Genome Project in 2003 [3] and following developments of next-generation sequencing (NGS) technologies have generated a cascade of high-

* Correspondence: carlo.resteghini@istitutotumori.mi.it

[†]Mara S. Serafini, Laura Lopez-Perez, Loris De Cecco and Carlo Resteghini contributed equally to this work.

³Head and Neck Medical Oncology Department, Fondazione IRCCS Istituto Nazionale dei Tumori di Milano, Milan, Italy

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



throughput methodologies, altogether named *omics*. *Omics* have substantially led biology understanding to a deeper level for several cancer types, including HNSCC. In the present paper, we reviewed the main *omics* methodologies and the available resources for retrieving and analyzing *omics* data. Additionally, we updated our previous work [4] with the most recent published data in the context of HNSCC *Transcriptomics* and *Epigenomics*, considering these reviews as a continuum. The objective of the present work is to comprehensively review available information on transcriptomics and epigenomics in HNSCC to provide an overview on biological, prognostic and predictive molecular signatures.

Main Omics methodologies

Biology is the result of the presence, expression, interaction, and regulation of different types of molecules. For their ability to account such a complexity, *omics* technologies have grown over the last two decades and they are now highly intertwined with other biological functional analysis [5]. Considering the classical cellular workflow of transcription (from DNA to mRNA) and translation (from mRNA to protein), *omics* can be presented as follows: i) *Genomics* has been introduced as the first high-throughput *omics* technique that impacted several aspects of clinical activity. It analyses the whole sequences of coding and non-coding portions of the genome, and targeted sequences (such as exome or clinical exome sequences). *Genomics* allows the identification of possibly relevant variants, such as single nucleotide polymorphisms (SNPs), copy number variation (CNV), mutations and translocations; ii) *Transcriptomics* involves all the RNA transcripts (with a particular attention in the last decade to mRNA, and more recently to long non-coding RNA [lncRNA]), monitor their differences in expression and infer the impacts of their alteration; iii) *Epigenomics* essentially studies DNA methylation variations and the functional consequences of the spatial behavior of the DNA (see also Table 1). Moreover, other cellular molecules have been analyzed by high-throughput methodologies and entered in the *omics* sciences, such as proteins, metabolites in general and lipids in particular (*Proteomics*, *Metabolomics*, *Lipidomics*). Recently, the *omics* suffix was also applied to emerging non-molecular fields: 'radiomics', the high-throughput mining of quantitative image features from clinically used medical imaging [6] and 'metagenomics', the assessment of the microbial communities inhabitant of the human body. More details about the characteristics of these other *omics* areas and methods are available elsewhere [7].

Available resources for retrieving and analyzing Omics data

The application of high-throughput techniques requires high computational capacity and expertise in handling large amounts of data. Consequently, repositories for

omics have been created worldwide (Tables 2, 3). Most of these repositories are publicly accessible and useful for data consulting. The ArrayExpress archive is one of the ELIXIR Core Data Resources and stores data derived from array- and sequence-based experiments. Researchers can upload data if the provided content is compliant with the Minimum Information About a Microarray Experiment (MIAME) and the Minimum Information About a Next-generation Sequencing Experiment (MINSEQE) standards [8]. ArrayExpress experiment results are available as: i) metadata information with the experiment description, protocol procedures, sample annotations and author information; ii) raw experiment data; iii) processed data. ArrayExpress enables access to BioSamples [9], another ELIXIR repository, providing a store to collect metadata, about biological samples. Gene Expression Omnibus (GEO) is a public repository supported by the National Cancer Center for Biotechnology Information (NCBI) and it archives MIAME- and MINSEQE-compliant functional genomics data of all organisms. Data derived from array- and sequence-based analyses are available, comprising dataset information, experiment variable subsets, expression value measurements, gene symbols and, comprehensive gene annotation. Additionally, GEO offers several functionalities for data analysis through GEO DataSet database, such as gene search, comparison of samples sets, inspection of cluster heat-maps, execution of experimental design and value distribution with box plot visualization support. Another available repository is The Cancer Genome Atlas (TCGA), which contains only human cancer data and, for this reason, differs from the previously described repository. TCGA was born as collaboration between the National Cancer Institute (NCI) and the National Human Genome Research Institute (NHGRI) and was upgraded and merged with the Pan-Cancer Atlas [10]. Both TCGA and Pan-Cancer Atlas offer a reclassification of human tumor types based on molecular similarity, a molecular landscape of the oncogenic processes and a comprehensive analysis of tumor signaling pathways. Only TCGA and Pan-Cancer consortium members have the access to submit *omics* data and data upload is continuously in progress. TCGA dataset system contains 25 human cancer types and it is provided free of charge. Its exploration is supported by descriptive charts. A controlled access is required for data downloading. Another important repository is the Functional Annotation of the Mammalian Genome (FANTOM), an international research consortium that encompasses the field of transcriptome analyses. The project delivered the FANTOM5 collection, a data series supporting structure of mammalian transcriptome atlases in diverse cell types. FANTOM5 data contains: Cap Analysis of Gene Expression (CAGE) and annotation tables; pathway enrichment and co-expression cluster analysis; enhancers; results of *de-novo*

Table 1 The main *omics* techniques and their characteristics: the biological material analyzed, the major methodologies applied and the type of information achievable with them

Omics name	Material analyzed	Methodologies	Type of information
Genomics	DNA	Whole-genome sequencing	Complete genome sequence for identification of coding and non-coding sequence variants
		Exome sequencing or clinical exome-sequencing	Sequencing of protein-coding regions of the genome or genes known to be associated to a clinical phenotype
		Comparative genomic hybridization (aCGH)	Analysis of copy number alterations in the genome (comparison with a reference genome/sample)
Epigenomics/Transcriptomics	DNA/RNA	Whole or targeted DNA methylation	Identification of disfunctions in regulatory elements
		Whole or targeted miRNA	Sequencing or detection of differential expression of miRNA
		Whole or targeted long noncoding RNA	Sequencing or detection of differential expression of long noncoding RNA
		Whole or targeted mRNA	Sequencing or detection of differential expression of coding genes

and motif activity analysis; sample ontology and ontology term enrichment; CAGE peaks identified by specific classifier and visualization tools.

In parallel, a large set of omics web-based tools and an increasing amount of informative repository databases have been developed (Table 3). ELIXIR [11] is an inter-governmental organization, composed by 23 members and over 180 research organizations among Europe. It is a Core Data resource with several web-based bioinformatics tools such as: i) Ensembl, a browser for DNA sequences and assemblies, provided by international projects on vertebrate genomes that accommodates annotated genes, computes multiple alignments, predicts regulatory function and collects disease data; ii) European Genome-phenome Archive (EGA), a web-tool, providing information from genetic and phenotypic data

coming from biomedical research projects; iii) Rfam, a database collecting multiple sequence alignments, consensus secondary structures and covariance models (CMs) for non-coding RNA families; and iv) RNACentral, provided by collaborating groups (ENA, Ensembl, GENCODE, miRBase), bringing integrated access to a comprehensive and up-to-date set of non-coding RNA sequences. Furthermore, a number of web-based tools or software querying TCGA are available: i) The Cancer Omics Atlas (TCOA), providing useful functions complementary to other existing tools, such as querying of gene, miRNA and protein expression, somatic mutations (based on a single molecule or cancer type correlations of gene–gene, miRNA–miRNA, protein–protein, gene–miRNA and gene–protein), and their correlation with survival prognosis in cancer patients; ii) Broad Institute,

Table 2 Main public repositories and their features

SOURCE	ArrayExpress	GEO (Gene expression omnibus) - NCBI (The National Center for Biotechnology Information)	TCGA (The Cancer Genome Atlas)	FANTOM5
Link	www.ebi.ac.uk/arrayexpress	www.ncbi.nlm.nih.gov/geo/summary	portal.gdc.cancer.gov	fantom.gsc.riken.jp
Repository data	Public archive of Genomics/Functional Genomics Data all organisms, normal and disease associated	Public archive of Functional Genomics Data of all organisms, normal and disease associated	Public dataset of Omics Data of Human Cancer	International consortium of functional annotations of mammalian coding and non coding portion of genome
Submitting authors	All the researchers, following the MIAME and MINSEQE rules	All the researchers, following the MIAME and MINSEQE rules	Only TCGA/Pancancer consortium members	Only FANTOM5 consortium members
Access for downloading	Free	Free	Required for downloading controlled-access data	Free
Type of available data	Array- and sequence-based data. Meta-analysis not accepted. Updated daily	Array- and sequence-based data. Meta-analysis not accepted. Updated daily	Omics data generated by TCGA Consortium and Pancancer Atlas. Continuously in progress	Transcriptomics data. Updated monthly
Query types	See Elixir Core Data Resource in Table 3	Download analysis (find genes, compare 2 sets of samples, cluster heatmaps, experiment design and value distribution)	See tools for querying TCGA in Table 3	Visualization tools such as ZENBU, a system for data integration, analysis and visualization of NGS based data

Table 3 Details of the largest and most utilized web-tools for omics analyses and a list of free access repositories to retrieve useful information related to mRNA, microRNA and other non coding-RNA

SOURCE	link/access	Type of tool/repository	Type of available data
Elixir Core Data Resource	www.ensembl.org/index free access	Genome browser that: annotates genes, computes multiple alignments; predicts regulatory function and collects disease data.	Vertebrates organisms data
European Genome-phenome Archive (EGA)	www.ebi.ac.uk/ega/ login requested	Database containing all types of sequence and genotype experiments, including case-control, population, and family studies.	Human, normal and disease associated
Rfam	rfam.xfam.org/ free access	Database of non-coding RNA families	Vertebrates organisms
RNAcentral	rnacentral.org/ free access	Web tool with access to a set of non-coding RNA sequences, allowing: sequence search, public Postgres database and genome browser. It enables integrated text search, sequence similarity search, bulk downloads, and programmatic data access.	Vertebrates and other organisms
Tools for querying TCGA	www.tcoacpu.edu.cn/ free access	This tool provides useful functions complementary to other existing tools, for fast and straightforward querying of TCGA	Human cancer data
Broad Institute	http://gdac.broadinstitute.org/ Required TCGA controlled-access data	Web tool to systematize analysis from TCGA, with the following functionalities: gene expression Viewer, data analysis (overview, results, methods and data), and data download.	Human cancer data
Oncolnc	http://www.oncolnc.org/ free access to survival data (raw data not accessible)	Web tool analyzes TCGA in terms of survival by selecting the target genes.	Human cancer data
TCGA Batch Effects Viewer	http://bioinformatics.mdanderson.org/tcgambatch/	Webtool designed to help assess, diagnose and correct for any batch effects in TCGA data.	Human cancer data
cBioPortal	http://www.cbioportal.org/ ; Open source license via GitHub	Software allowing integrative analysis of cancer genomics	Human cancer data

allowing systematic analysis on TCGA data and comparison with other diseases; iii) OncoLnc, analyzing patients survival (Kaplan-Meier curves) according to mRNA, miRNA, lncRNA expression levels; and iv) TCGA Batch Effects Viewer [12], a tool specifically designed to diagnose and correct for any batch effects in TCGA data; v) cBioPortal [13], a software allowing genomic analysis both from population or a single patient of multiple cancer types. In addition to these two main types of resources (i.e. ELIXIR and TCGA querying tools), it should be mentioned: i) MiTranscriptome [14], a catalog of human long poly-adenylated RNA transcripts, from samples encompassing different cancer and tissue types; ii) KM plotter [15], a tool assessing the effect of genes/miRNA on overall survival data for biomarker discovery; iii) Bioconductor, an open-source tool based on R programming language for the analysis and comprehension of high-throughput data and enabling generation of workflows for multiple data types, data preprocessing, statistical inference, regression, network analysis, machine learning, multiomics integration and visualization. For further information about other tools, databases and websites also see the following reviews [16–18].

Strategy of search and selection of studies

Literature surveys of HNSCC genomics [2] and proteomics [19] have already been conducted and published in the past years. For this reason, we decided to focus on HNSCC transcriptomics and epigenomics studies, characterizing signatures related to biology, prognosis and prediction of treatment response. The timeframe of our analysis was fixed, encompassing papers published between January 2015 and January 2019. The purpose of this choice was to partially overlap with a previous review on transcriptomics data [4] and was dictated by the evidence that both epigenomics (DNA methylation) and transcriptomics based on non-coding RNA (miRNA, lncRNA and piRNAs) are advancing and growing only in recent years. A web-based search has been performed in the following databases: Pubmed, ArrayExpress and GEO. The combination of the following keywords has been used: “gene expression” or “methylation” or “miRNA” or “transcriptomics” or “sequencing” or “microarray” AND “head and neck cancer” or “HNSCC”. The title and the abstract of all potentially relevant studies were assessed for their contents before the retrieval of full articles. The full text of each selected study was carefully evaluated. Eligible studies were required to meet the following inclusion criteria: publication which data has been obtained using HNSCC tumor tissue; the number of cases per each analysis had to be ≥ 40 . Moreover, the following exclusion criteria were applied: non-English publication; case reports, letters and reviews;

expression studies of individual preselected candidate gene/miRNA/DNA methylation site; and data obtained on not human samples/cell lines/different patient materials (serum, plasma, saliva). Papers that fulfilled the inclusion criteria were processed for data.

Analysis of recent Epigenomics and Transcriptomics data

Over more than 1000 papers have been analyzed from which we retrieved a total of 61 omics studies fulfilling inclusion and exclusion criteria [20–80].

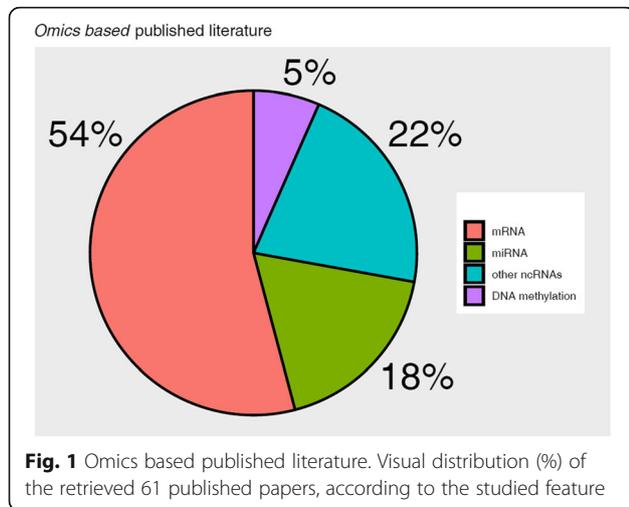
The selected studies are listed in the Additional file 1: Table S1, subdivided according to the analyzed feature, such as mRNA, miRNA, non-coding RNAs and DNA methylation, respectively. This subdivision was univocal for 54 studies. Conversely, authors of 6 studies analyzed more than one feature and it was decided to classify them on the basis of the feature with higher relevance, according to paper aim. Data in the supplementary tables are reported as: i) the name of the identified signature according to the authors of publication; ii) the type of feature analyzed; iii) the information provided by the signature (biology, prognosis, prediction of treatment response); iv) selection based on anatomical site (oral cavity, oropharynx, larynx, hypopharynx) or HPV-status for data analysis; v) the ID of the dataset; and vi) availability of independent validation.

A large portion of the studies regarded mRNA signatures (33/61) [19–51], while a minor portion of the studies considered miRNA signatures (11/61) [52–62] and other non-coding RNA signatures (13/61) [63–76]. Only 4/61 [77–80] studies identified signatures by a high-throughput omics analysis of methylated DNA. The percentage distribution of the studies according to the type of features analyzed is reported in Fig. 1. The information provided by the signature, i.e. biology, prognosis and prediction of treatment response, are depicted in Fig. 2.

Most of the identified signatures had a prognostic value (36/61); on the contrary, only a minority (7/61) was related to prediction of treatment response. No signature derived from the study of other non-coding RNAs or DNA methylation had prognostic or predictive value.

A total of 21 studies selected HNSCC samples based on specific anatomical site or HPV status, while no selection was applied in the remaining 40 studies. Another remarkable aspect, regarding datasets of the analyzed studies, is the presence or absence of internal validation in the same publication: only half of the studies included in our analyses performed a validation in independent datasets.

Noteworthy, the majority of transcriptomic and epigenomic datasets used to define or validate the signature under evaluation were generated by TCGA. TCGA datasets were the only ones included in meta-analyses or



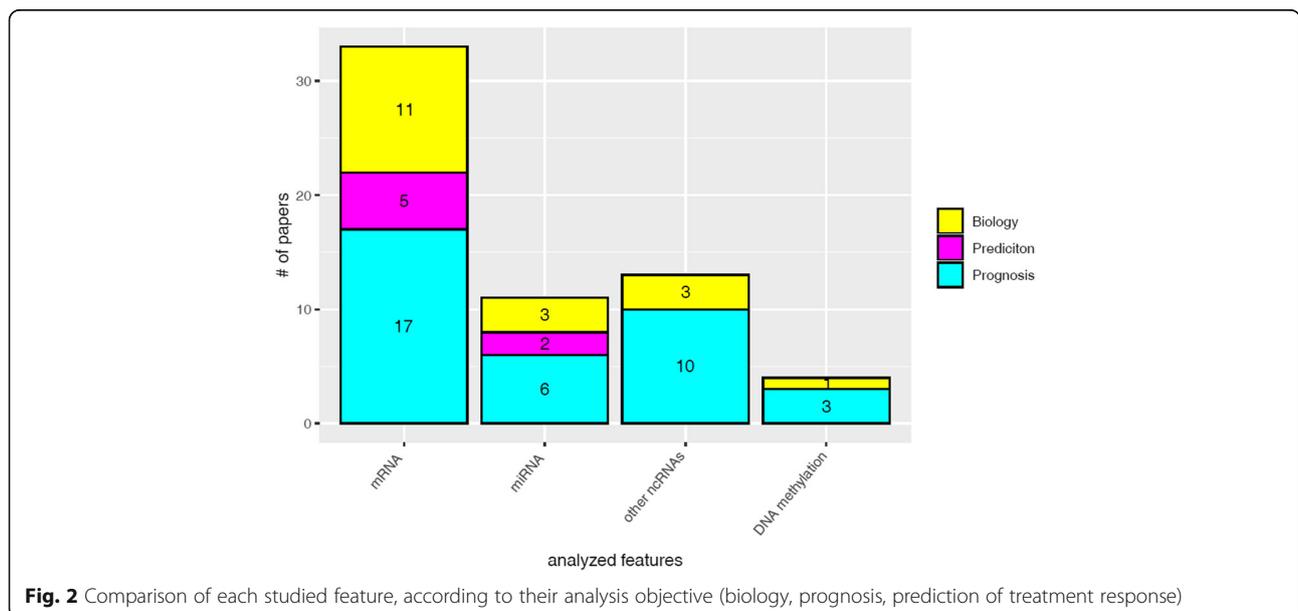
validation set in 46/61 studies: mRNA (24/33); miRNA expression (7/11); DNA methylation (3/4); and 12/13 studies on non-coding RNA.

Even though all the selected studies deserve attention, a detailed analysis of each one is out of scope of this survey. However, we decided to comment the development and exploitation of three gene-expression signatures, whose clinical impact on prognosis/prediction of treatment response could be high. Two signatures were initially identified by analysis of HNSCC cell lines. These gene classifiers/indexes were subsequently tested in historical retrospective clinical cohorts and validated in prospective clinical studies, showing promising prognostic or predictive ability. The first signature is the radio-sensitivity index (RSI), whose development and clinical validation in three data sets of rectal, esophageal and

HNSCC was originally described in 2009 [81]. The RSI was further commented in 2017 [82] and constituted the pillar for proposing a genome-based model for adjusting radiotherapy dose (GARD) as it was applied to a large retrospective, cohort-based study [83]. At present, a company (Precision Genomic Radiation Therapy platform: pGRT™) developed a mathematical approach to the integration of genomics into radiation treatment and planning; this application is central to the Cvergenx patented RSI and GARD (<https://www.cvergenx.com/>). At the present, this nomogram is in development for clinical purposes in other tumor types, but not in HNSCC.

The second signature is the 15-gene hypoxia classifier, first described between 2010 and 2011 [84, 85] and then validated as prognostic signature in retrospective series [86]. A patent application is currently pending on this method for determination of clinically relevant hypoxia in cancer specimen (WO/2012/146259). The clinical relevance of this signature is highlighted by the decision to conduct a double blind randomized multicenter phase III study, the Intergroup EORTC-1219-ROG-HNCG/DAHANCA-29 trial (NCT01880359). This study is designed to prospectively evaluate if nimorazole, a radio-sensitizer, can improve the effect of accelerated concomitant chemo-radiotherapy with cisplatin on the locoregional control rate in patients with newly diagnosed HPV 16 negative stage III-IV carcinoma of the larynx, oropharynx or hypopharynx. The study is designed to stratify patients according to the 15-gene signature in order to determine if the treatment benefit is larger in patients who carry a hypoxia- profile.

The third signature [32], published in 2016 by some of the Authors of the present paper, analysed HNSCC gene



expression of patients with short and prolonged responses to cetuximab- and platinum-based chemotherapy. Basal subtype traits, including signatures of EGFR signalling and hypoxic differentiation, characterized patients with long response whereas short-response patients showed RAS activation. These results were commented upon [87] as an important step in the identification of candidate predictive biomarkers of response to cetuximab–platinum therapy in recurrent/metastatic HNSCC patients. Furthermore, the predictive power of the signature were refined by the creation of a common network with results from miRNA analyses of the same sample set [58]. At last, a validation was successfully completed in different sample sets of recurrent/metastatic HNSCC treated with different anti-EGFR agent, specifically the phase II PANI01 trial employing panitumumab [88] and a window of opportunity trial evaluating pre-operative afatinib [89]. Despite differences in clinical settings and anti-EGFR inhibitor treatments, prediction of response by the previously identified Cluster 3 signature and selected miRNAs was comparable. Cluster 3 signature is characterized, beside hypoxia, by others functional pathways including altered metabolism.

Conclusions and future directions

In the last decade, several progresses have been achieved not only in the methodology for “omics” analyses but also in availability of data repositories and web-based tools for the storage and the analysis of the enormous amount of data generated. Despite these progresses, the present literature revision highlighted that most of published works on HNSCC are not omics-based. In fact, we were able to retrieve only 61 such studies out of more than 1000 that were initially identified in our research.

Based on this overview on omics-related literature in HNSCC, we identified three major limits: i) the classical epigenomics area (DNA methylation) and the omics based on non-coding RNA (miRNA, lncRNA and piRNAs) have been analyzed quite recently; ii) the limited sample size included in most “omics” studies; iii) the largest dataset for HNSCC at present available is TCGA; and iv) even if there was an evident expansion of omics-related HNSCC publications starting from 2017, this increase should be partially attributed to the reiterated bioinformatic analysis of the TCGA dataset. To overcome the issues, the integration of multiple datasets through a meta-analysis approach has been reported to offer advantages, improving the reliability of results [21, 52]. However, some important aspects included in the more recent analyses should be highlighted. The presence of HPV infection in HNSCC, especially those arising from the oropharynx, has a well-known and profound impact on prognosis. The recently released 8th edition of the

American Joint Committee on Cancer (AJCC) staging system has introduced major differences in oropharynx squamous cell carcinoma, now staged according to p16 status [90]. Few recent omics-based analyses dissected the biologic aspect underlying this phenomenon [91] and very recent data indicate a clear association between subtypes and different prognosis [52]. A deeper knowledge of molecular biology and mechanisms of carcinogenesis in HPV-related HNSCC will be critical in order to further differentiate patient’s prognosis and therefore improve disease management. Of note, in the context of growing epidemics [92] the identification of ideal candidate for safe de-escalated therapy should be focused on genomic and molecular factors in order to achieve a successful application of the precision medicine ideal [93]. Another point of emphasis derived from the present overview is the growing integration among omics data obtained on the same cohort of patients. Even if these experiences are still limited for HNSCC and in terms of types of omics employed and amount of published studies [23, 27, 43, 58, 66, 76], these initiatives enabling to better dissect cancer complexity deserve further investigations. Moreover, we can expect that new approaches based on artificial intelligence will be available in the next future dealing with more complex data even integrating multi-omics layers [94]. At present, a particular method, self-organizing maps (SOM)-machine learning offers a practical solution when hundreds of samples are profiled for thousands of genes as microarray/RNAseq and a number of studies on different cancer types proved its robustness [18, 19]. As an example SOM enabled to separate oropharynx p16 positive tumors in three clusters with different prognosis [52]. Future informatic analyses are expected to: i) identify and implement services to retrieve omics data from public repositories; ii) harmonize omics data in order to merge different data sources in one integrated, HNSCC-specific dataset; and iii) explore the resulting dataset with dedicated techniques. Finally, we have to acknowledge not only the important anatomical site specific contribute of TCGA on HNSCC [95], but also the recent contribution of the Cancer Genome Atlas Pan-Cancer analysis project, which, by a multiplatform analysis of different cancer types [96, 97], revealed a molecular classification within and across tissues of origin. In particular, the analyses of 12 and 42 different cancer types by Campbell et al. [97] and Chen et al. [28], respectively, enabled to reveal that: i) squamous cell cancers from different tissue sites may be distinguished from other cancers and may be subclassified molecularly by squamous cell pathways and programs providing candidates for therapy; and ii) a small subset of HNSCC expresses evident traits of neuro-endocrinicity. In addition to the Cancer Genome Atlas upcoming data, in the next years we hope to witness a surge of new omics-based

analyses in HNSCC, and based on new, large and rigorously clinically annotated datasets. An example is represented by the European Commission funded project named “Big Data and Models for Personalized Head and Neck Cancer Decision Support (BD2Decide)” (Clinical-Trial.gov Identifier NCT02832102, <http://www.bd2decide.eu/>). The project, started on 2016 and expected to be concluded at the end of 2019, aims at the definition of a prognostic tool based on the integration of multi-omics analyses of a large dataset of locoregionally advanced HNSCC.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s41199-020-0047-y>.

Additional file 1. Supplementary Table

Abbreviations

AJCC: American Joint Committee on Cancer; GEO: Gene Expression Omnibus; HNSCC: Head and neck squamous cell carcinoma; HPV: Human papilloma virus; GEO; TGCA; SOM; Elixir; SOM: Self-Organizing Map; TGCA: The Cancer Genome Atlas

Acknowledgments

We thank Silvana Canevari (Fondazione IRCCS Istituto Nazionale dei Tumori Milano, Italy) for fruitful discussions and critical reading of the manuscript. We would like also to thank all the cited authors for their informative papers that enabled us to organize this review and we apologize for potential missing citations.

We would also like to thank Luca Giacomelli, PhD, and Aashni Shah (Polistudium srl) for editorial assistance, funded by internal funds. We would like also to thank Andrea Careno for the creation and editing of the images.

Authors' contributions

LLP and MSS drafted the manuscript. LL, GF, LDC and CR supervised the manuscript writing. LDC and CR corrected and finalized the text. All the authors read and approved the final version.

Funding

This study is partially supported by: Associazione Italiana Ricerca Cancro (AIRC) grant number AIRC IG 18519 to LDC; European Union's Horizon 2020 research and innovation program under grant agreement No.689715 to LL.

Availability of data and materials

This is a review article and there is no raw data related to this manuscript for data sharing.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

Author details

¹Integrated Biology Platform, Department of Applied Research and Technology Development, Fondazione IRCCS Istituto Nazionale dei Tumori di Milano, Milan, Italy. ²Life Supporting Technologies, Universidad Politécnica de Madrid, Madrid, Spain. ³Head and Neck Medical Oncology Department,

Fondazione IRCCS Istituto Nazionale dei Tumori di Milano, Milan, Italy.

⁴University of Milan, Milan, Italy.

Received: 3 October 2019

Published online: 21 January 2020

References

1. Global Cancer Observatory. <https://gco.iarc.fr/>. Accessed 5 July 2019.
2. Leemans CR, Snijders PJF, Brakenhoff RH. The molecular landscape of head and neck cancer. *Nat Rev Cancer*. 2018;18(5):269–82. <https://doi.org/10.1038/nrc.2018.11>.
3. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature*. 2004;431(7011):931–45. <https://doi.org/10.1038/nature03001>.
4. Tonella L, Giannoccaro M, Alfieri S, Canevari S, De Cecco L. Gene expression signatures for head and neck Cancer patient stratification: are results ready for clinical application? *Curr Treat Options in Oncol*. 2017;18(5):32. <https://doi.org/10.1007/s11864-017-0472-2>.
5. OMICS, BIOINFORMATICS, COMPUTATIONAL BIOLOGY. <http://alttox.org/mapp/emerging-technologies/omics-bioinformatics-computational-biology/>.
6. Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol*. 2017;14(12):749–62. <https://doi.org/10.1038/nrclinonc.2017.141>.
7. Gasperskaja E, Kučinskis V. The most common technologies and tools for functional genome analysis. *Acta medica Litua*. 2017;24(1):1–11. <https://doi.org/10.6001/actamedica.v24i1.3457>.
8. GEO and MIAME - GEO - NCBI. <https://www.ncbi.nlm.nih.gov/geo/info/MIAME.html>. Accessed 5 July 2019.
9. BioSamples & EMBL-EBI. <https://www.ebi.ac.uk/biosamples/>. Accessed 5 July 2019.
10. Pan-Cancer Atlas. <https://www.cell.com/pb-assets/consortium/pancanceratlas/pancani3/index.html>. Accessed 5 July 2019.
11. Elixir Tools Platform. <https://elixir-europe.org/platforms/tools>. Accessed 15 Nov 2018.
12. Batch Effects Viewer. <https://bioinformatics.mdanderson.org/BatchEffectsViewer/>. Accessed 5 July 2019.
13. cBioPortal. <https://www.cbioportal.org/>. Accessed 15 Nov 2018.
14. MiTranscriptome. <http://mitranscriptome.org/>. Accessed 15 Nov 2018.
15. Kaplan Meier-plotter. <http://kmplot.com/analysis/>. Accessed 5 July 2019.
16. Yang Y, Dong X, Xie B, et al. Databases and web tools for Cancer genomics study. *Genomics Proteomics Bioinformatics*. 2015;13(1):46–50. <https://doi.org/10.1016/j.gpb.2015.01.005>.
17. Pavlopoulou A, Spandidos DA, Michalopoulos I. Human cancer databases (Review). *Oncol Rep*. 2015;3–18. <https://doi.org/10.3892/or.2014.3579>.
18. Kristensen VN, Lingjærde OC, Russnes HG, Vollan HKM, Frigessi A, Børresen-Dale A-L. Principles and methods of integrative genomic analyses in cancer. *Nat Rev Cancer*. 2014;14(5):299–313. <https://doi.org/10.1038/nrc3721>.
19. Rezende TMB, de Souza FM, Franco OL. Head and neck cancer: proteomic advances and biomarker achievements. *Cancer*. 2010;116(21):4914–25. <https://doi.org/10.1002/cncr.25245>.
20. Wichmann G, Rosolowski M, Krohn K, et al. The role of HPV RNA transcription, immune response-related gene expression and disruptive TP53 mutations in diagnostic and prognostic profiling of head and neck cancer. *Int J Cancer*. 2015;137(12):2846–57. <https://doi.org/10.1002/ijc.29649>.
21. De Cecco L, Nicolau M, Giannoccaro M, et al. Head and neck cancer subtypes with biological and clinical relevance: Meta-analysis of gene-expression data. *Oncotarget*. 2015;6(11):9627–42. <https://doi.org/10.18632/oncotarget.3301>.
22. Keck MK, Zuo Z, Khattry A, et al. Integrative analysis of head and neck cancer identifies two biologically distinct HPV and three non-HPV subtypes. *Clin Cancer Res*. 2015;21(4):870–81. <https://doi.org/10.1158/1078-0432.CCR-14-2481>.
23. Zhang W, Edwards A, Fang Z, Flemington EK, Zhang K. Integrative genomics and Transcriptomics analysis reveals potential mechanisms for favorable prognosis of patients with HPV-positive head and neck carcinomas. *Sci Rep*. 2016;6(April):1–14. <https://doi.org/10.1038/srep24927>.
24. Zhang Y, Koneva LA, Virani S, et al. Subtypes of HPV-positive head and neck cancers are associated with HPV characteristics, copy number alterations, PIK3CA mutation, and pathway signatures. *Clin Cancer Res*. 2016;22(18):4735–45. <https://doi.org/10.1158/1078-0432.CCR-16-0323>.
25. Cao B, Wang Q, Zhang H, Zhu G, Lang J. Two immune-enhanced molecular subtypes differ in inflammation, checkpoint signaling and outcome of

- advanced head and neck squamous cell carcinoma. *Oncoimmunology*. 2018;7(2). <https://doi.org/10.1080/2162402X.2017.1392427>.
26. Chen YP, Wang YQ, Lv JW, et al. Identification and validation of novel microenvironment-based immune molecular subgroups of head and neck squamous cell carcinoma: implications for immunotherapy. *Ann Oncol*. 2019;30(1):68–75. <https://doi.org/10.1093/annonc/mdy470>.
 27. Costa RL, Boroni M, Soares MA. Distinct co-expression networks using multi-omic data reveal novel interventional targets in HPV-positive and negative head-and-neck squamous cell cancer. *Sci Rep*. 2018;8(1):1–13. <https://doi.org/10.1038/s41598-018-33498-5>.
 28. Chen F, Zhang Y, Gibbons DL, et al. Pan-cancer molecular classes transcending tumor lineage across 32 cancer types, multiple data platforms, and over 10,000 cases. *Clin Cancer Res*. 2018;24(9):2182–93. <https://doi.org/10.1158/1078-0432.CCR-17-3378>.
 29. Irimie AI, Braicu C, Cojocneanu R, et al. Differential effect of smoking on gene expression in head and neck Cancer patients. *Int J Environ Res Public Health*. 2018;15(7):1558. <https://doi.org/10.3390/ijerph15071558>.
 30. Saidak Z, Clatot F, Chatelain D, Galmiche A. A gene expression profile associated with perineural invasion identifies a subset of HNSCC at risk of post-surgical recurrence. *Oral Oncol*. 2018;86:53–60. <https://doi.org/10.1016/j.oraloncology.2018.09.005>.
 31. Sanati N, Iancu OD, Wu G, Jacobs JE, McWeeney SK. Network-based predictors of progression in head and neck squamous cell carcinoma. *Front Genet*. 2018;9:183. <https://doi.org/10.3389/fgene.2018.00183>.
 32. Bossi P, Bergamini C, Siano M, et al. Functional genomics uncover the biology behind the responsiveness of head and neck squamous cell Cancer patients to Cetuximab. *Clin Cancer Res*. 2016;22(15):3961–70. <https://doi.org/10.1158/1078-0432.CCR-15-2547>.
 33. Foy J-P, Bazire L, Ortiz-Cuaran S, et al. A 13-gene expression-based radioresistance score highlights the heterogeneity in the response to radiation therapy across HPV-negative HNSCC molecular subtypes. *BMC Med*. 2017;15(1):165. <https://doi.org/10.1186/s12916-017-0929-y>.
 34. Chen X, Yan B, Lou H, et al. Immunological network analysis in HPV associated head and neck squamous cancer and implications for disease prognosis. *Mol Immunol*. 2018;96(February):28–36. <https://doi.org/10.1016/j.molimm.2018.02.005>.
 35. Chen J, Fu G, Chen Y, Zhu G, Wang Z. Gene-expression signature predicts survival benefit from postoperative chemoradiotherapy in head and neck squamous cell carcinoma. *Oncol Lett*. 2018;16(2):2565–78. <https://doi.org/10.3892/ol.2018.8964>.
 36. You GR, Cheng AJ, Lee LY, et al. Prognostic signature associated with radioresistance in head and neck cancer via transcriptomic and bioinformatic analyses. *BMC Cancer*. 2019;19(1):1–11. <https://doi.org/10.1186/s12885-018-5243-3>.
 37. Tawk B, Schwager C, Deffaa O, et al. Comparative analysis of transcriptomics based hypoxia signatures in head- and neck squamous cell carcinoma. *Radiother Oncol*. 2016;118(2):350–8. <https://doi.org/10.1016/j.radonc.2015.11.027>.
 38. Guo W, Chen X, Zhu L, Wang Q. A six-mRNA signature model for the prognosis of head and neck squamous cell carcinoma. *Oncotarget*. 2017; 8(55):94528–38. <https://doi.org/10.18632/oncotarget.21786>.
 39. Mes SW, Te BD, Poli T, et al. Prognostic modeling of oral cancer by gene profiles and clinicopathological co-variables. *Oncotarget*. 2017;8(35):59312–23. <https://doi.org/10.18632/oncotarget.19576>.
 40. Qiu Z, Sun W, Gao S, et al. A 16-gene signature predicting prognosis of patients with oral tongue squamous cell carcinoma. *PeerJ*. 2017;5:e4062. <https://doi.org/10.7717/peerj.4062>.
 41. Shen S, Bai J, Wei Y, et al. A seven-gene prognostic signature for rapid determination of head and neck squamous cell carcinoma survival. *Oncol Rep*. 2017;38(6):3403–11. <https://doi.org/10.3892/or.2017.6057>.
 42. Yang B, Chen Z, Huang Y, Han G, Li W. Identification of potential biomarkers and analysis of prognostic values in head and neck squamous cell carcinoma by bioinformatics analysis. *Onco Targets Ther*. 2017;10:2315–21. <https://doi.org/10.2147/OTT.S135514>.
 43. Zhang ZL, Zhao LJ, Chai L, et al. Seven lncRNA-mRNA based risk score predicts the survival of head and neck squamous cell carcinoma. *Sci Rep*. 2017;7(1):309. <https://doi.org/10.1038/s41598-017-00252-2>.
 44. Lee DJ, Eun Y-GG, Rho YS, et al. Three distinct genomic subtypes of head and neck squamous cell carcinoma associated with clinical outcomes. *Oral Oncol*. 2018;85(January):44–51. <https://doi.org/10.1016/j.oraloncology.2018.08.009>.
 45. Namani A, Maturi Rahaman M, Chen M, Tang X. Gene-expression signature regulated by the KEAP1-NRF2-CUL3 axis is associated with a poor prognosis in head and neck squamous cell cancer. *BMC Cancer*. 2018;18(1):46. <https://doi.org/10.1186/s12885-017-3907-z>.
 46. Schmidt S, Linge A, Zwanenburg A, et al. Development and validation of a gene signature for patients with head and neck carcinomas treated by postoperative radio (chemo)therapy. *Clin Cancer Res*. 2018;24(6):1364–74. <https://doi.org/10.1158/1078-0432.CCR-17-2345>.
 47. Tian S, Meng G, Zhang W. A six-mRNA prognostic model to predict survival in head and neck squamous cell carcinoma. *Cancer Manag Res*. 2019;11: 131–42. <https://doi.org/10.2147/CMAR.S185875>.
 48. Wintergerst L, Selmsberger M, Maihofer C, et al. A prognostic mRNA expression signature of four 16q24.3 genes in radio (chemo)therapy-treated head and neck squamous cell carcinoma (HNSCC). *Mol Oncol*. 2018;12(12): 2085–101. <https://doi.org/10.1002/1878-0261.12388>.
 49. Xu X, Li M, Hu J, et al. Expression profile analysis identifies a two-gene signature for prediction of head and neck squamous cell carcinoma patient survival. *J Cancer Res Ther*. 2018;14(7):1525. https://doi.org/10.4103/jcrt.jcrt_557_18.
 50. Wang W, Lim WK, Leong HS, et al. An eleven gene molecular signature for extra-capsular spread in oral squamous cell carcinoma serves as a prognosticator of outcome in patients without nodal metastases. *Oral Oncol*. 2015;51(4):355–62. <https://doi.org/10.1016/j.oraloncology.2014.12.012>.
 51. Zhao X, Sun S, Zeng X, Cui L. Expression profiles analysis identifies a novel three-mRNA signature to predict overall survival in oral squamous cell carcinoma. *Am J Cancer Res*. 2018;8(3):450–61. <http://www.ncbi.nlm.nih.gov/pubmed/29637000>. Accessed 24 July 2019.
 52. Locati LD, Serafini MS, Iannò MF, et al. Mining of self-organizing map gene-expression portraits reveals prognostic stratification of HPV-positive head and neck squamous cell carcinoma. *Cancers (Basel)*. 2019;11(8):1057. <https://doi.org/10.3390/cancers11081057>.
 53. An F, Zhang Z, Xia M, Xing L. Subpath analysis of each subtype of head and neck cancer based on the regulatory relationship between miRNAs and biological pathways. *Oncol Rep*. 2015;34(4):1745–54. <https://doi.org/10.3892/or.2015.4150>.
 54. Krishnan AR, Zheng H, Kwok JG, et al. A comprehensive study of smoking-specific microRNA alterations in head and neck squamous cell carcinoma. *Oral Oncol*. 2017;72:56–64. <https://doi.org/10.1016/j.oraloncology.2017.07.009>.
 55. Wong N, Khwaja SS, Baker CM, et al. Prognostic microRNA signatures derived from the Cancer genome atlas for head and neck squamous cell carcinomas. *Cancer Med*. 2016;5(7):1619–28. <https://doi.org/10.1002/cam4.718>.
 56. Sass S, Pitea A, Unger K, Hess J, Mueller NS, Theis FJ. MicroRNA-target network inference and local network enrichment analysis identify two microRNA clusters with distinct functions in head and neck squamous cell carcinoma. *Int J Mol Sci*. 2015;16(12):30204–22. <https://doi.org/10.3390/ijms161226230>.
 57. Chen L, Wen Y, Zhang J, et al. Prediction of radiotherapy response with a 5-microRNA signature-based nomogram in head and neck squamous cell carcinoma. *Cancer Med*. 2018;7(3):726–35. <https://doi.org/10.1002/cam4.1369>.
 58. De Cecco L, Giannoccaro M, Marchesi E, et al. Integrative miRNA-Gene Expression Analysis Enables Refinement of Associated Biology and Prediction of Response to Cetuximab in Head and Neck Squamous Cell Cancer. *Genes (Basel)*. 2017;8(1). <https://doi.org/10.3390/genes8010035>.
 59. Citron F, Armenia J, Franchin G, et al. An integrated approach identifies mediators of local recurrence in head and neck squamous carcinoma. *Clin Cancer Res*. 2017;23(14):3769–80. <https://doi.org/10.1158/1078-0432.CCR-16-2814>.
 60. Hess J, Unger K, Maihofer C, et al. A five-microRNA signature predicts survival and disease control of patients with head and neck cancer negative for HPV infection. *Clin Cancer Res*. 2019;25(5):1505–16. <https://doi.org/10.1158/1078-0432.CCR-18-0776>.
 61. Hui L, Wu H, Yang N, Guo X, Jang X. Identification of prognostic microRNA candidates for head and neck squamous cell carcinoma. *Oncol Rep*. 2016; 35(6):3321–30. <https://doi.org/10.3892/or.2016.4698>.
 62. Nunez Lopez YO, Victoria B, Golusinski P, Golusinski W, Masternak MM. Characteristic miRNA expression signature and random forest survival analysis identify potential cancer-driving miRNAs in a broad range of head and neck squamous cell carcinoma subtypes. *Rep Pract Oncol Radiother*. 2018;23(1):6–20. <https://doi.org/10.1016/j.rpor.2017.10.003>.
 63. Shi H, Chen J, Li Y, et al. Identification of a six microRNA signature as a novel potential prognostic biomarker in patients with head and neck squamous cell carcinoma. *Oncotarget*. 2016;7(16):21579–90. <https://doi.org/10.18632/oncotarget.7781>.
 64. Nohata N, Abba MC, Gutkind JS. Unraveling the oral cancer lncRNAome: identification of novel lncRNAs associated with malignant progression

- and HPV infection. *Oral Oncol.* 2016;59:58–66. <https://doi.org/10.1016/j.oraloncology.2016.05.014>.
65. Zou AE, Zheng H, Saad MA, et al. The non-coding landscape of head and neck squamous cell carcinoma. *Oncotarget.* 2016;7(32). <https://doi.org/10.18632/oncotarget.9979>.
66. Zhao G, Fu Y, Su Z, Wu R. How long non-coding RNAs and MicroRNAs mediate the endogenous RNA Network of head and neck squamous cell carcinoma: a comprehensive analysis. *Cell Physiol Biochem.* 2018;50(1):342–52. <https://doi.org/10.1159/000494009>.
67. Cao W, Liu JN, Liu Z, et al. A three-lncRNA signature derived from the Atlas of ncRNA in cancer (TANRIC) database predicts the survival of patients with head and neck squamous cell carcinoma. *Oral Oncol.* 2017;65:94–101. <https://doi.org/10.1016/j.oraloncology.2016.12.017>.
68. Feng L, Houck JR, Lohavanichbutr P, Chen C. Transcriptome analysis reveals differentially expressed lncRNAs between oral squamous cell carcinoma and healthy oral mucosa. *Oncotarget.* 2017;8(19):31521–31. <https://doi.org/10.18632/oncotarget.16358>.
69. Firmino N, Martinez VD, Rowbotham DA, Enfield KSS, Bennewith KL, Lam WL. HPV status is associated with altered PIWI-interacting RNA expression pattern in head and neck cancer. *Oral Oncol.* 2016;55:43–8. <https://doi.org/10.1016/j.oraloncology.2016.01.012>.
70. de Lena PG, Paz-Gallardo A, Paramio JM, García-Escudero R. Clusterization in head and neck squamous carcinomas based on lncRNA expression: molecular and clinical correlates. *Clin Epigenetics.* 2017;9(1):1–11. <https://doi.org/10.1186/s13148-017-0334-6>.
71. Krishnan AR, Korrapati A, Zou AE, et al. HHS Public Access; 2018. p. 68–75. <https://doi.org/10.1016/j.oraloncology.2016.12.022.Smoking>.
72. Diao P, Song Y, Ge H, et al. Identification of 4-lncRNA prognostic signature in head and neck squamous cell carcinoma. *J Cell Biochem.* 2019;120(6):10010–20. <https://doi.org/10.1002/jcb.28284>.
73. Liu G, Zheng J, Zhuang L, et al. A prognostic 5-lncRNA expression signature for head and neck squamous cell carcinoma. *Sci Rep.* 2018;8(1):15250. <https://doi.org/10.1038/s41598-018-33642-1>.
74. Wang P, Jin M, Sun C, et al. A three-lncRNA expression signature predicts survival in head and neck squamous cell carcinoma (HNSCC). *Biosci Rep.* 2018;38(6):BSR20181528. <https://doi.org/10.1042/bsr20181528>.
75. Zhao C, Zou H, Wang J, Shen J, Liu H. A three long noncoding RNA-based signature for Oral squamous cell carcinoma prognosis prediction. *DNA Cell Biol.* 2018;37(11):888–95. <https://doi.org/10.1089/dna.2018.4317>.
76. Fang XN, Yin M, Li H, et al. Comprehensive analysis of competitive endogenous RNAs network associated with head and neck squamous cell carcinoma. *Sci Rep.* 2018;8(1):1–13. <https://doi.org/10.1038/s41598-018-28957-y>.
77. Degli Esposti D, Sklias A, Lima SC, et al. Unique DNA methylation signature in HPV-positive head and neck squamous cell carcinomas. *Genome Med.* 2017;9(1):33. <https://doi.org/10.1186/s13073-017-0419-z>.
78. Shen S, Wang G, Shi Q, et al. Seven-CpG-based prognostic signature coupled with gene expression predicts survival of oral squamous cell carcinoma. *Clin Epigenetics.* 2017;9(1):88. <https://doi.org/10.1186/s13148-017-0392-9>.
79. Chen Y, Hei N, Zhao J, et al. A two-CpG-based prognostic signature for oral squamous cell carcinoma overall survival. *J Cell Biochem.* 2019;120(6):9082–90. <https://doi.org/10.1002/jcb.28182>.
80. Brennan K, Koenig JL, Gentles AJ, Sunwoo JB, Gevaert O. Identification of an atypical etiological head and neck squamous carcinoma subtype featuring the CpG island methylator phenotype. *EBioMedicine.* 2017;17:223–36. <https://doi.org/10.1016/j.ebiom.2017.02.025>.
81. Eschrich SA, Pramana J, Zhang H, et al. A gene expression model of intrinsic tumor Radiosensitivity: prediction of response and prognosis after Chemoradiation. *Int J Radiat Oncol.* 2009;75(2):489–96. <https://doi.org/10.1016/j.ijrobp.2009.06.014>.
82. Caudell JJ, Torres-Roca JF, Gillies RJ, et al. The future of personalised radiotherapy for head and neck cancer. *Lancet Oncol.* 2017;18(5):e266–73. [https://doi.org/10.1016/S1470-2045\(17\)30252-8](https://doi.org/10.1016/S1470-2045(17)30252-8).
83. Scott JG, Berglund A, Schell MJ, et al. A genome-based model for adjusting radiotherapy dose (GARD): a retrospective, cohort-based study. *Lancet Oncol.* 2017;18(2):202–11. [https://doi.org/10.1016/S1470-2045\(16\)30648-9](https://doi.org/10.1016/S1470-2045(16)30648-9).
84. Sørensen BS, Toustrup K, Horsman MR, Overgaard J, Alsnér J. Identifying pH independent hypoxia induced genes in human squamous cell carcinomas in vitro. *Acta Oncol (Madr).* 2010;49(7):895–905. <https://doi.org/10.3109/02841861003614343>.
85. Toustrup K, Sorensen BS, Nordsmark M, et al. Development of a hypoxia gene expression classifier with predictive impact for hypoxic modification of radiotherapy in head and neck Cancer. *Cancer Res.* 2011;71(17):5923–31. <https://doi.org/10.1158/0008-5472.CAN-11-1182>.
86. Toustrup K, Sørensen BS, Metwally MAHH, et al. Validation of a 15-gene hypoxia classifier in head and neck cancer for prospective use in clinical trials. *Acta Oncol (Madr).* 2016;55(9–10):1091–8. <https://doi.org/10.3109/0284186X.2016.1167959>.
87. Chau NG, Hammerman PS. Heads up! Predictive gene signatures in head and neck Cancer may be coming soon. *Clin Cancer Res.* 2016;22(15):3710–2. <https://doi.org/10.1158/1078-0432.CCR-16-0582>.
88. Siano M, Espeli V, Mach N, et al. Gene signatures and expression of miRNAs associated with efficacy of panitumumab in a head and neck cancer phase II trial. *Oral Oncol.* 2018;82:144–51. <https://doi.org/10.1016/j.oraloncology.2018.05.013>.
89. Machiels J-P, Bossi P, Menis J, et al. Activity and safety of afatinib in a window preoperative EORTC study in patients with squamous cell carcinoma of the head and neck (SCCHN). *Ann Oncol.* 2018;29(4):985–91. <https://doi.org/10.1093/annonc/mdy013>.
90. AJCC Cancer Staging Manual, 8th Edition. <http://cancerstaging.org/references-tools/deskreferences/Pages/default.aspx>.
91. Lawrence MSMSMS, Sougnez C, Lichtenstein L, et al. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature.* 2015;517(7536):576–82. <https://doi.org/10.1038/nature14129>.
92. Marur S, D'Souza G, Westra WH, Forastiere AA. HPV-associated head and neck cancer: a virus-related cancer epidemic. *Lancet Oncol.* 2010;11(8):781–9. [https://doi.org/10.1016/S1470-2045\(10\)70017-6](https://doi.org/10.1016/S1470-2045(10)70017-6).
93. Orlandi E, Licitra L. Personalized medicine and the contradictions and limits of first-generation deescalation trials in patients with human papillomavirus-positive oropharyngeal cancer. *JAMA Otolaryngol - Head Neck Surg.* 2018; 144(2):99–100. <https://doi.org/10.1001/jamaoto.2017.2308>.
94. Resteghini C, Trama A, Borgonovi E, et al. Big data in head and neck Cancer. *Curr Treat Options in Oncol.* 2018;19(12):62. <https://doi.org/10.1007/s11864-018-0585-2>.
95. Cancer Genome Atlas Research Network, Weinstein JN, Collisson EA, et al. The Cancer genome atlas pan-Cancer analysis project. *Nat Genet.* 2013; 45(10):1113–20. <https://doi.org/10.1038/ng.2764>.
96. Hoadley KA, Yau C, Wolf DM, et al. Multiplatform analysis of 12 Cancer types reveals molecular classification within and across tissues of origin. *Cell.* 2014; 158(4):929–44. <https://doi.org/10.1016/j.cell.2014.06.049>.
97. Campbell JD, Yau C, Bowlby R, et al. Genomic, Pathway Network, and Immunologic Features Distinguishing Squamous Carcinomas. *Cell Rep.* 2018; 23(1):194–212.e6. <https://doi.org/10.1016/j.celrep.2018.03.063>.

Publisher's Note

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

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

