The molecular landscape of head and neck cancer

C. René Leemans[1]*, Peter J. F. Snijders[2] and Ruud H. Brakenhoff[1]*

Abstract | Head and neck squamous cell carcinomas (HNSCCs) arise in the mucosal linings of the upper aerodigestive tract and are unexpectedly heterogeneous in nature. Classical risk factors are smoking and excessive alcohol consumption, and in recent years, the role of human papillomavirus (HPV) has emerged, particularly in oropharyngeal tumours. HPV-induced oropharyngeal tumours are considered a separate disease entity, which recently has manifested in an adapted prognostic staging system while the results of de-intensified treatment trials are awaited. Carcinogenesis caused by HPV in the mucosal linings of the upper aerodigestive tract remains an enigma, but with some recent observations, a model can be proposed. In 2015, The Cancer Genome Atlas (TCGA) consortium published a comprehensive molecular catalogue on HNSCC. Frequent mutations of novel druggable oncogenes were not demonstrated, but the existence of a subgroup of genetically distinct HPV-negative head and neck tumours with favourable prognoses was confirmed. Tumours can be further subclassified based on genomic profiling. However, the amount of molecular data is currently overwhelming and requires detailed biological interpretation. It also became apparent that HNSCC is a disease characterized by frequent mutations that create neoantigens, indicating that immunotherapies might be effective. In 2016, the first results of immunotherapy trials with immune checkpoint inhibitors were published, and these may be considered as a paradigm shift in head and neck oncology.

[1]Department of Otolaryngology/Head and Neck Surgery, VU University Medical Center.
[2]Department of Pathology, VU University Medical Center, Amsterdam, Netherlands.
*e-mail: cr.leemans@vumc.nl; rh.brakenhoff@vumc.nl

doi:10.1038/nrc.2018.11
Published online 2 Mar 2018

Head and neck squamous cell carcinomas (HNSCCs) have a yearly incidence of 600,000 cases worldwide, with 40–50% mortality[1]. The tumours originate in the epithelial cells of the mucosal linings of the upper airway and food passages (the oral cavity, oropharynx, larynx or hypopharynx), which suggests that HNSCC is a relatively homogeneous disease, as it develops from one cell type in one tissue. Rather unexpectedly, HNSCC is remarkably heterogeneous. This is in part brought about by the complex anatomical structures in which it develops but also relates to the different aetiologies and the large variety of molecular changes that drive carcinogenesis.

The disease characteristics of HNSCCs are currently undergoing a pronounced change, as our understanding of the aetiologies of this disease have progressed. The classical risk factors for HNSCC were smoking and excessive alcohol use. Currently, infection with high-risk human papillomaviruses (HPVs) causes a substantial and rising proportion of these tumours, originating mainly in the oropharynx and occurring particularly in the Western world[2]. It is now well accepted that oropharyngeal SCCs (OPSCCs) can be divided into HPV-negative (HPV–ve) and HPV-positive (HPV+ve) diseases[3,4], as the socioeconomic profiles of the patients as well as the clinical presentation, the molecular profiles[5–8] and most importantly the prognosis of the tumours differ between the two subgroups[9–11]. However, it should be noted that at present, this key distinction mainly refers to oropharyngeal tumours within HNSCCs, as this is the best-studied entity with respect to HPV involvement.

The far more favourable outcome of HPV+ve compared with HPV–ve OPSCC is so substantial that the tumour-node-metastasis (TNM) staging for HNSCC was adapted in the eighth edition to include $p16^{INK4A}$ immunostaining as a surrogate for HPV status[12] (BOX 1). Moreover, several treatment de-escalation trials of HPV+ve OPSCC have been initiated, and the results of these are now being awaited, which may lead to personalized treatment based on HPV status.

The lack of rapidly improving patient survival and personalized treatment approaches has propelled research into the molecular landscape of HNSCC. Using expression arrays and, over the past few years, RNA sequencing, subgroups of head and neck tumours characterized by gene expression patterns have been identified[8,13–16]. However, the performance of different

Box 1 | **Clinical management of head and neck cancer**

Tumours are staged according to tumour-node-metastasis (TNM) staging but can also be classified with respect to their anatomical location (in the larynx, hypopharynx, oropharynx or oral cavity), which can be further subdivided into 14 subsites according to the International Classification of Diseases (ICD)-10.

Clinical examination and modern imaging (computed tomography (CT)–magnetic resonance imaging (MRI) and ^{18}F-fluorodeoxyglucose (FDG)–positron emission tomography (PET)) are the mainstays of diagnosis and are supplemented when necessary by examination under general anaesthesia. TNM stage and anatomical site are important factors to direct multidisciplinary treatment decisions for the patient. In addition, patient factors, such as comorbidity and age, and institutional factors play a role. The newest (eighth) edition of the American Joint Committee on Cancer (AJCC) and Union for International Cancer Control (UICC) TNM classification became effective on 1 January 2018; this new version accounts for the importance of tumour depth of invasion in oral cavity tumours and has a different staging system for $p16^{INK4A}$-positive tumours of the oropharynx (see the table), where $p16^{INK4A}$ positivity is used as a surrogate marker for human papillomavirus (HPV) status, which results in a lower stage for these tumours than that assigned by the previous edition[12].

Early-stage tumours are treated with surgery or radiotherapy. The mainstays of treatment for advanced tumours are surgery combined with postoperative chemoradiotherapy or upfront chemoradiation, with surgical salvage if possible. Induction chemotherapy has not found widespread use because of the lack of benefit in previous studies[122]. Besides improving patient survival, another major aim of therapy is an improved quality of life.

Recent advances in clinical management include sentinel node biopsy (SNB), transoral robotic resections (TORS), image-based and adaptive radiotherapy and application of the epidermal growth factor receptor (EGFR)-specific antibody cetuximab in combination with radiotherapy[123,124]. The increasing array of therapeutic options necessitates optimized selection of patients to personalize treatment. At present, only the classical clinical and histopathological characteristics are used for making treatment decisions, and these classical methods have met their limitations for high-precision medicine.

Disease will recur in approximately one-half of patients. Recurrence at the locoregional site is most frequent, followed successively by distant metastases. Also, second primary cancers in the head and neck, lung or other sites occur at an annual rate of approximately 2–3%, particularly in patients with HPV-negative, non-oropharyngeal cancers[125]. Both recurrent head and neck cancer and second primary cancers are notoriously difficult to treat curatively. Palliative treatment consists of a combination of chemotherapy and cetuximab or, depending on the condition of the patient, a less toxic treatment or best supportive care.

Seventh edition TNM	Eighth edition TNM
Stage I (T1N0)	Stage I (T1–T2N0–N1)
Stage II (T2N0)	Stage II (T1–T2N2 or T3N0–N2)
Stage III (T3N0 or T1–T3N1)	Stage III (T4 or N3)
Stage IVa (T4aN0–N1 or T1–T4aN2)	Stage IV (M1)
Stage IVb (T4b or T1–T4bN3)	–
Stage IVc (M1)	–

Leukoplakia
A macroscopic white change in the mucosal linings of the upper aerodigestive tract, which is defined by the World Health Organization (WHO) as “a white plaque of questionable risk having excluded (other) known diseases or disorders that carry no increased risk for cancer”.

expression profiling platforms and analysis pipelines as well as differences in immune infiltrate and other stromal components may impact the subgroup definitions. In addition, the prognostic associations of the subgroups are variable, hampering clinical utility. Hence, in contrast to, for instance, breast cancer[17], classification based on gene expression profiles is not yet common practice for HNSCC, but these profiles are highly informative from a biological perspective.

In 2015, The Cancer Genome Atlas (TCGA) consortium published the comprehensive genomic data of 279 HNSCCs, including both HPV+ve and HPV−ve tumours[3]. The HPV−ve tumours are typically characterized by many mutations and numerous chromosomal gains and losses. Intriguingly, in this and earlier studies, a distinct subgroup of HPV−ve tumours with ‘copy number alteration (CNA)-silent’ profiles emerged, which also displayed specific mutational profiles, suggesting that these tumours form a separate genetic subgroup[3]. The number of candidate cancer driver genes in HNSCC is exploding at present and requires careful biological interpretation. In this Review, we discuss and merge existing and newly identified genomic data to obtain a novel and more integrated view on the molecular landscape of HNSCC.

Molecular landscape of HPV−ve HNSCC

Carcinogenesis

Cancer arises through the accumulation of genetic and epigenetic changes in genes whose encoded proteins act in a variety of signalling pathways, and these alterations cause the cancer-associated phenotypes[18]. These genetic changes can be broadly defined as mutations, a term mostly used in association with genes and that relates to genetic changes at the base-pair level, stretching from a single nucleotide to confined insertions and deletions (indels), and alterations encompassing gains and losses, sometimes of entire chromosomes or chromosome arms harbouring thousands of genes. How important these extensive gains and losses are in cancer evolution remains somewhat elusive, but losses are found particularly in conjunction with inactivating mutations of tumour suppressor genes, indicating that these do play a critical role[3,19]. Since the 1990s, molecular data on cancer have evolved rapidly[20].

In the clinical course of head and neck cancer, the development of precursor changes needs to be considered: a growing collection of cells with tumour-associated genetic alterations. Although most patients with HNSCC present with tumours de novo, there are precancerous lesions in the mucosal linings that are visible to the naked eye and present as white (leukoplakia) or red (erythroplakia) areas of the mucosa. These lesions may progress and develop into invasive cancers[21]. Leukoplakia is the most common precursor lesion of OPSCC, with an estimated worldwide prevalence of 2%[22]. The clinical management is to treat the lesion when possible and to analyse a tissue specimen or a biopsy sample with microscopic examination for the presence of dysplasia, graded as mild, moderate or severe, which associates with cancer risk[23,24]. Besides these macroscopically recognizable lesions, many pre-malignant mucosal abnormalities are clinically not visible and are identified only under the microscope as dysplastic mucosal epithelium. These morphological abnormalities already led in 1953 to the concept of ‘field cancerization’, both to describe these large dysplastic changes but also to explain the high frequency of tumour recurrence after HNSCC excision and the high risk of multiple independent tumours developing in the mucosal linings[25].

In the 1990s, these precancerous changes were comprehensively studied with genetic markers, and it was shown that the number of genetic changes correlates

Erythroplakia
A macroscopic red change in the mucosal linings of the upper aerodigestive tract, which is defined by the World Health Organization (WHO) as "a fiery red patch that cannot be characterized either clinically or pathologically as any other definable lesion".

with the severity of the dysplastic changes. This led to the first genetic progression model of HNSCC[26]. Loss of heterozygosity at chromosomes 3p, 9p and 17p appeared to occur in dysplasia, apparently reflecting early changes in carcinogenesis, while alterations at chromosome arms 11q, 4q and 8p were typically present in carcinomas, likely reflecting a relatively late phase in carcinogenesis[26]. Using these genetic markers, the presence of genetically altered mucosal epithelial cells could also be more comprehensively studied in the surgical margins, and it was shown that these cells form an important source of local recurrences and second primary tumours, at least in surgically treated patients[27–29]. Most recent studies have identified genetic changes at chromosome arms 3p and 9p together with mutations in *TP53* (residing at 17p) as the best predictors of malignant transformation in both leukoplakia[30] and surgical margins[31]. These lesions present an opportunity for future studies to define the timing of genetic events and completely characterize the molecular landscape of head and neck cancer[32].

But what precedes these mucosal fields of cells with tumour-associated genetic changes? Carcinogenesis, also occurring in the mucosal linings, likely starts with the accumulation of mutations in a single adult stem cell[33,34]. In 2002, small p53-immunopositive focal patches in tumour-adjacent mucosal epithelium were described[35] that contained *TP53* mutations. These mutated p53-positive patches were considered equivalent to the clones or clonal units that form the squamous epithelium of the mucosa and the skin, and consist of progeny cells originating from a single adult stem cell, which became immunodetectable by the mutation in *TP53*. Although the squamous stem cell itself cannot be identified owing to a lack of specific biomarkers, these clonal units visualized by the presence of mutated *TP53* were considered to represent the very first sign of oncogenic changes in the mucosa and led to a hypothetical patch-field-tumour-metastasis progression model for HNSCC[19]. Although this model was hypothetical at the time it was proposed, subsequent data obtained in engineered mouse models now support it. Using lineage-tracing experiments in mice, AXIN2-expressing stem cells and their progeny can be visualized, resulting in the formation of morphologically comparable patches, at least in the skin[36].

Driver genes and pathways

The genomic data of 279 HNSCCs (HPV+ve and HPV−ve) have been analysed and published[3], but data from over 500 patients are available at the TCGA network and can be queried through a variety of data portals and analysis tools. Depending on the analysis tool applied (BOX 2), between 50 and 100 genes are indicated that are substantially mutated in HNSCC and are considered as candidate cancer driver genes. However, many genes are mutated at only very low frequencies, and functional consequences of these mutations are often unclear. Many of these genes are, as yet, also not functionally linked to carcinogenesis, and without functional studies, they remain only candidates. Previously, we formulated criteria that took the form of levels of molecular evidence to indicate whether a candidate driver gene is indeed bona fide, and these might still apply today[19]. Of note, the TCGA applied a correction algorithm (BOX 2) to filter the mutation data in the published data sets. In other data sets, such correction tools might not have been employed to curate the data, and this could potentially explain the differences in the candidate cancer genes mentioned when mutation data of different studies are compared[37].

Besides mutations, genes may be activated by amplification (for example, amplification of epidermal growth factor receptor (*EGFR*)) or inactivated by heterozygous and homozygous losses (for example, loss of cyclin-dependent kinase inhibitor 2A (*CDKN2A*)) or epigenetic changes. We will discuss some of the genes (TABLE 1) that are frequently mutated, that are functionally annotated to be involved in cancer and that emerged as candidates in previous publications[3]. Many others have been discussed before by us[19] and in other publications[37,38]. The currently available molecular information can no longer be captured in a single comprehensive overview as there is simply too much data. In the supplementary information of the 2015 HNSCC genomic TCGA study, mutual exclusivity or correlations among mutated genes were analysed, which could be considered as functional associations. However, as this was not performed per genetic subgroup of CNA-high, HPV−ve; CNA-silent, HPV−ve; and HPV+ve tumours, it could be that the associations might relate in part to this genetic subclassification.

Box 2 | Driver and passenger mutations

Modern cancer genomics was initiated by large-scale Sanger sequencing of cancer genomes[126] and advanced rapidly with massively parallel sequencing (MPS)[3]. Many somatic mutations have been identified since then but often in unexpected genes such as olfactory receptor genes[127,128]. The concept of driver mutations and passenger mutations was established and was associated with the following questions: how can these be distinguished, and are there any factors that could be of use to make the distinction?

These questions were tested in a large meta-analysis of data sets from The Cancer Genome Atlas (TCGA), and algorithms to correct for the relevance of mutations were developed and presented[127]. Not unexpectedly, the frequency of mutations in cancer genomes appeared to be related to tumour type and mutagen exposure, and this phenomenon can be taken into consideration[127]. Remarkable in this study was the observation that somatic mutations accumulate in genome regions that replicate late during S phase, encompassing the regions that are not actively transcribed. This may relate to gene expression-associated DNA repair and/or nucleotide shortage during replication[127]. Furthermore, a correction algorithm was described in this study that takes into account the number of indicated biases, and when applying this tool to the sequencing data of 178 lung squamous cell carcinomas, the number of candidate cancer driver genes decreased from 450 to 11 (REF. 127). Although correction algorithms have their limitations, it is clear that mutation data should be interpreted with great caution.

Although often neglected, the interpretation of mutations may also play a role in bona fide cancer genes. Not all somatic non-synonymous mutations impact protein function, and although bioinformatic analysis of pathogenicity can be helpful, such mutations might in fact be passengers. Specifically, when MPS has been applied, another factor may come into play: the variant allele frequency (VAF), which is the ratio between the mutated and wild-type allele reads. A typical driver gene in head and neck squamous cell carcinoma (HNSCC), such as *TP53*, which mutates early in carcinogenesis, is usually mutated in one or both alleles in all cancer cells. Hence, when sequencing a tumour cell line, the VAF will be close to 50% or 100%. However, tumours contain stroma, and driver mutations may occur in subpopulations of cancer cells. For these reasons, lower VAFs, usually 5% or 10%, are often used to designate a mutation as a candidate driver. However, a low VAF may also be suggestive that a mutation is a passenger.

Cell cycle control. The early and most frequent genetic alterations in head and neck carcinogenesis are losses of chromosomes 3p, 9p and mutations of *TP53* (REFS 19,26). The tumour suppressor gene on chromosome arm 9p, more specifically 9p21, is *CDKN2A*, encoding the p16^{INK4A} protein, which binds and disrupts the cyclin D–CDK4 and cyclin D–CDK6 complexes. Loss of *CDKN2A* combined with the frequently observed amplification of cyclin D1 (*CCND1*) on 11q13 drives cells through the G1–S checkpoint of the cell cycle and contributes to unscheduled DNA replication (FIG. 1). Normally, unscheduled DNA replication leads to DNA damage and p53 activation[39], but *TP53* is also frequently inactivated in HNSCC. The p53 protein is a key tumour suppressor in many tumour types with a multitude of functions, including the induction of p21, another CDK inhibitor that arrests the cell cycle[40]. The p53 protein is also a strong inducer of apoptosis[40]. The *TP53* gene, located on chromosome 17p13, is frequently inactivated in HNSCC, mostly by missense mutations combined with allelic loss[29]. Somatic mutations in *TP53* are found in 60–80% of HNSCCs and 84% of HPV–ve HNSCCs[3,19]. Of note, the *TP53* locus typically shows allelic losses and not copy number losses in HNSCC, suggesting that it is advantageous for tumours to have a double copy of chromosome 17 or two copies of mutated *TP53* (REF. 6).

Besides these proteins acting in the cell cycle itself, the genes of several growth factor receptors have been identified as candidate cancer genes that might play a role in HNSCC, including *EGFR* and *MET*; but because of the low mutation frequency in some of these genes, and therefore their somewhat elusive role as driver genes, not all of them have been listed in TABLE 1. However, *EGFR* is frequently amplified, supporting its role as a driver gene in HNSCC[3], yet EGFR has pleiotropic functions and links to MAPK signalling, PI3K signalling[41] (see below) and direct nuclear signalling[42]. One of the targets of MAPK signalling is cyclin D1, and together with the observation that there seems to be a mutual exclusion of genetic changes in growth factor receptors and cyclin D1 amplification (see figure S7.1 in REF. 3), this suggests that the growth factor receptors are functionally linked to cell cycle control.

As cell cycle control is deregulated at the G1–S transition in HNSCC, cell cycle regulation therefore depends on the S phase, the G2–M transition checkpoint and the M phase to ensure proper cell division. Hence, these other cell cycle checkpoints may represent targetable functional mechanisms, and there are some key molecular players such as WEE1 kinase that are critical for at least G2–M checkpoint regulation and targetable by small-molecule inhibitors[43].

WNT signalling. A remarkable observation in the genomic profiling studies is the frequent number of inactivating mutations in the cadherin-related *FAT1* gene[3]. The *FAT1* gene is located on chromosome 4q35.2 and mutated in 23% of HNSCC cases and lost or deleted in 8% (cBioPortal). It encodes a large membrane protein belonging to a family of four FAT proteins and is part of the larger cadherin superfamily, as these proteins contain a multitude of E-cadherin domains[44]. E-Cadherin is the classical, epithelial prototype of the cadherins, which are calcium-dependent transmembrane adhesion molecules that can form homotypic and heterotypic adhesion structures (reviewed in REF. 44). FAT1 has previously been linked to actin dynamics and cell–cell contacts[45], and recently, a role for FAT1 in WNT signalling has been established[46]. The functional data were mainly obtained using glioblastoma cell lines rather than HNSCC cell lines, but there are no reasons at present to suggest that these findings cannot be extrapolated.

WNT signalling plays a key role in cell orientation and cell fate and thereby in stem cell maintenance in most tissues and has been predominantly linked to colorectal cancer because of the frequent mutations in the genes encoding key players of the pathway (FIG. 2a): adenomatous polyposis coli (APC), AXIN and β-catenin[47]. The β-catenin molecule is the central player in this pathway[47]. Remarkably, β-catenin has also

Table 1 | **Genes with frequent and highly significant somatic genetic changes in HPV–ve HNSCC**

Cellular process	Gene	Protein	Type of gene	Mutation frequency (%)	CNA frequency (%)
Cell cycle	*CDKN2A*	p16^{INK4A}	Tumour suppressor	22	32
	TP53	p53	Tumour suppressor	72	1.4
	CCND1	G1–S-specific cyclin D1	Oncogene	0.6	25
Growth signals	*EGFR*	EGFR	Oncogene	4	11
Survival	*PIK3CA*	Catalytic p110α subunit of class 1 PI3Ks	Oncogene	18	21
	PTEN	PTEN	Tumour suppressor	3	4
WNT signalling	*FAT1*	Protocadherin FAT1	Tumour suppressor	23	8
	AJUBA	LIM domain-containing protein AJUBA	Tumour suppressor	7*	1
	NOTCH1	NOTCH1	Tumour suppressor	18	4
Epigenetic regulation	*KMT2D*	Histone-lysine *N*-methyltransferase KMT2D	Tumour suppressor	16	0.4
	NSD1	Histone-lysine *N*-methyltransferase NSD1	Tumour suppressor	12*	0.8

Data from REF. 3. Mutation data were taken from The Cancer Genome Atlas (TCGA) ($n = 504$) using the cBioPortal. CNA, copy number alteration; EGFR, epidermal growth factor receptor. *Putative passenger mutation that requires further functional studies.

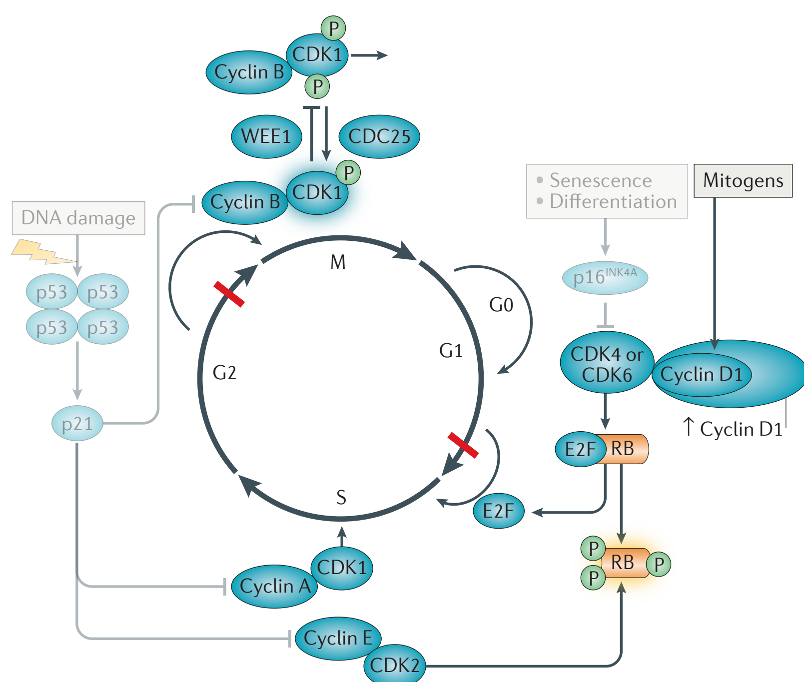

Figure 1 | **Cell cycle regulation in head and neck squamous cell carcinoma.** The cell cycle is regulated by cyclin-dependent kinase (CDK) complexes. There are two checkpoints (indicated by red bars) — one at G1–S, regulated by the transcription factor E2F and the pocket proteins, which include RB, and one at G2–M, regulated by cyclin B–CDK1. The activity of the cyclin–CDK complexes is regulated during the various phases by cyclic expression of the cyclins, inhibitors such as $p16^{INK4A}$ or p21, which is induced by p53, and phosphorylation of CDKs by, for example, WEE1 kinase. The genes CDK inhibitor 2A (*CDKN2A*), encoding $p16^{INK4A}$, and *TP53* are generally inactivated in head and neck squamous cell carcinoma (HNSCC; indicated by the faded colours), while cyclin D1 expression is increased in tumours (enlarged oval). In addition, *RB1*-inactivating mutations have been reported[3]. Hence, the G1–S checkpoint is generally bypassed in HNSCC, and the cell cycle is now coordinated at only the S and G2–M phases. This regulation is critical, as unscheduled S phase entry causes DNA replication stress and in the absence of active p53 does not lead to cell cycle arrest to enable repair. However, cancer cells still require a scheduled M phase entry with completely replicated chromosomes, otherwise cells will be unable to divide. This makes CDKs and WEE1 promising drug targets in HNSCC. Halos indicate protein activation through phosphorylation (P). CDC25, cell division cycle 25 phosphatase. Figure adapted from REF. 19, Macmillan Publishers Limited.

been identified as a binding partner of E-cadherin in adhesion plaques[44]. The relationship between β-catenin and E-cadherin is strikingly similar to the relationship between β-catenin and FAT1. Morris et al.[46] showed that FAT1 also binds β-catenin and sequesters it at the membrane, inhibiting its function in nuclear transcription. Knockdown of FAT1 expression caused β-catenin release and β-catenin-mediated signalling[46] (FIG. 2b). Hence, in parallel with the regulation of β-catenin levels by WNT, levels of β-catenin are seemingly also regulated by the cadherin family of adhesion molecules. This would imply that other cadherins and cadherin-like molecules may also function as tumour suppressor proteins, as is indeed the case for E-cadherin in specific types of cancer such as gastric cancer and breast cancer[44]. In these cancers, the *CDH1* gene encoding E-cadherin has undergone inactivating mutations like *FAT1* in HNSCC. The sequestering of β-catenin by FAT1 is less studied than that by E-cadherin, but a comparable model may be assumed[44]. Of note, E-cadherin also shows crosstalk with receptor tyrosine kinases such as EGFR, insulin-like growth factor 1 receptor (IGF1R), fibroblast growth factor receptor 2 (FGFR2) and MET (reviewed in REF. 44). Whether the same holds true for FAT1 or other cadherins remains to be determined, but it might shed a different light on these receptors and their genetic changes in HNSCC.

A second tumour suppressor protein that was identified in HNSCC and is also linked to the WNT pathway is LIM domain-containing protein AJUBA[3]. There are three AJUBA family members: AJUBA itself, LIM domain-containing protein 1 (LIMD1) and Wilms tumour protein 1-interacting protein (WTIP). They belong to the larger family of cytosolic LIM domain proteins, which share zinc-finger LIM domains as protein interaction sites and are involved in a multitude of cellular functions[48], including cell division, cell–matrix adhesion and cell–cell adhesion[49]. Furthermore, these proteins are connected to the Hippo signalling pathway[50]. In primary keratinocytes, AJUBA also interacts with cadherin-dependent cell–cell adhesive complexes through α-catenin, which shuttles to the nucleus akin to β-catenin[51] and determines cell fate during early development[52]. AJUBA has been shown to activate glycogen synthase kinase 3β (GSK3β), thereby reducing β-catenin levels through degradation and the consequent activity of the TCF family of transcription factors[53] (FIG. 2a). Thus, inactivation of AJUBA might cause reduced activity of GSK3β, which through diminished phosphorylation increases β-catenin levels and TCF family activity. *FAT1* alterations are mutually exclusive with *AJUBA* mutations, suggesting that the proteins encoded by these genes function in the same pathway[3]. However, AJUBA has a multitude of functions that might be partially redundant with other proteins, and its precise role in squamous cell carcinogenesis needs to be functionally investigated in more detail.

Nonetheless, these considerations raise the immediate question of whether the upstream part of the classical WNT–Frizzled pathway is active in squamous cells, as *APC* mutations are not found in HNSCC. Also, the role of FAT1 or, in a broader sense, the cadherin-sequestering mechanism of β-catenin and the precise role of AJUBA should be further studied specifically in squamous cells. These observations suggest an intricate inverse balance between cell–cell contacts and stemness in squamous cells, an interesting topic for further investigation.

A third tumour suppressor gene that might play a role in this context is *NOTCH1*. There are four receptors of the Notch family named NOTCH1–NOTCH4. The Notch receptors bind to membrane-bound ligands on other cells and are cleaved, and the Notch intracellular domain translocates to the nucleus and acts as a transcription factor[54,55]. *NOTCH1* was identified as a tumour suppressor gene in HNSCC[3,56,57], although its precise role remains somewhat unclear[58] and it perhaps requires a more systematic functional evaluation of the specific mutations and cellular context. NOTCH1 is a strong inducer of keratinocyte differentiation, but it was initially difficult to study by classical knockout methods

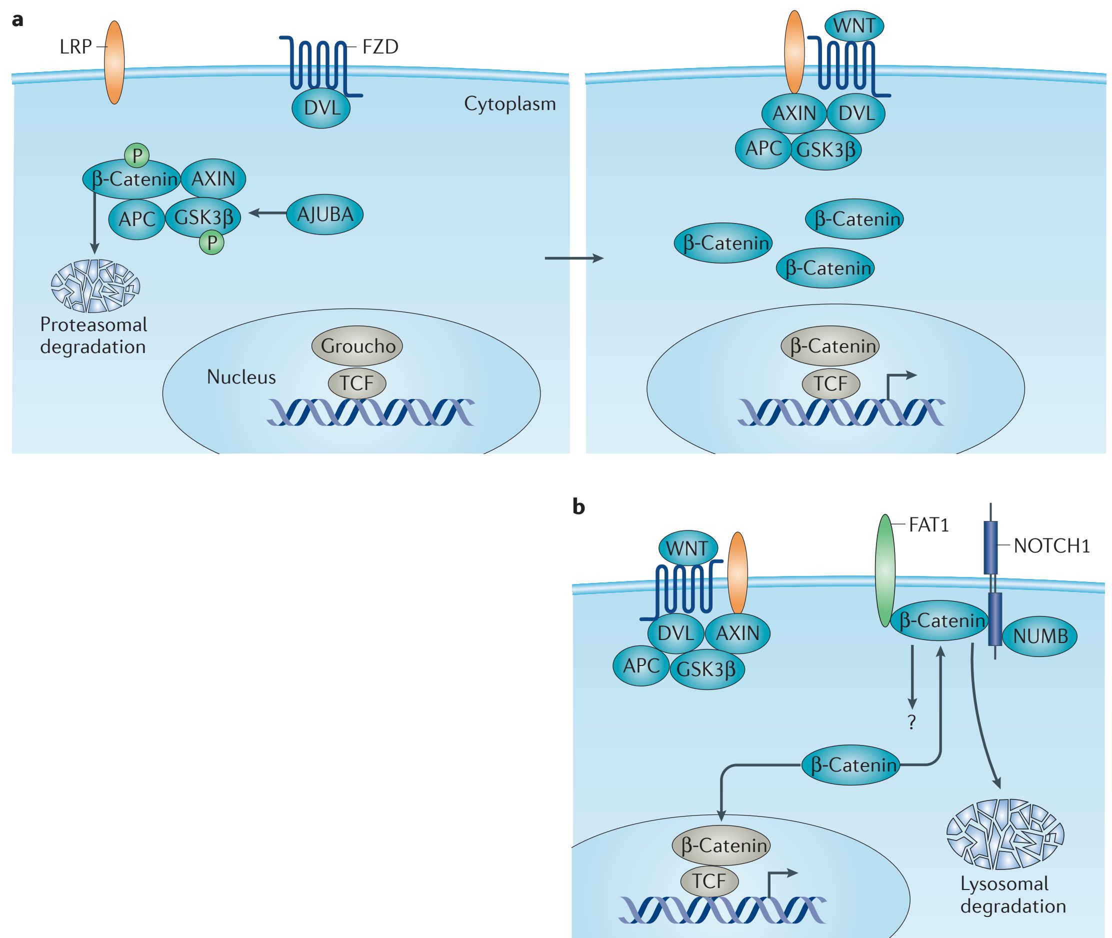

Figure 2 | **WNT signalling and the putative role of AJUBA, FAT1 and NOTCH1. a** | Canonical WNT signalling plays an important role in cell fate, cell orientation and stem cell behaviour. A key player is β-catenin. In the absence of WNT ligand, β-catenin is bound in the cytoplasm to a complex with AXIN, adenomatous polyposis coli (APC) and glycogen synthase kinase 3β (GSK3β). GSK3β phosphorylates β-catenin and induces its proteasomal degradation. LIM domain-containing protein AJUBA activates GSK3β and stimulates the phosphorylation and degradation of β-catenin. The canonical pathway is activated by WNT ligands, usually produced by neighbouring cells. Upon WNT binding to the receptor Frizzled (FZD), which is bound to Disheveled (DVL), and complexing with low-density lipoprotein receptor-related protein (LRP) co-receptors, the degradation complex is recruited to the membrane and inactivated. As a consequence, β-catenin is not degraded, translocates to the nucleus and displaces the transcriptional repressor Groucho (also known as TLE) from members of the transcription factor family TCF. The latter activate transcription of genes involved in stem cell biology and may induce proliferation, as one of the targets is the cyclin D1 (*CCND1*) gene. The WNT signalling pathway has been simplified in this schematic, as molecules such as the co-receptor leucine-rich repeat-containing G protein-coupled receptor 5 (LGR5), the ligand R-spondin and many others are also involved (see also the WNT homepage). **b** | Besides regulation by classical WNT signalling, FAT1, a cadherin-like molecule, and NOTCH1 also appear to play a role in the regulation of β-catenin, which might be important in mucosal squamous cells. Membrane-bound NOTCH1 is regulated by endosomal sorting pathways for either lysosomal degradation or recycling. The conserved endocytic adaptor proteins NUMB and NUMB-like protein (NUMBL) are involved in these processes and negatively regulate membrane-bound NOTCH1. These two proteins also play a role in the lysosomal degradation of the NOTCH1–β-catenin complex[60]. The lysosomal degradation of β-catenin by NOTCH1 binding has been demonstrated[60], but for FAT1, it is currently only hypothetical (denoted by the question mark).

in mouse models because of its lethal effect in embryogenesis[59]. However, young mice with a conditional knockout of *Notch1* show increased keratinocyte proliferation and changes in differentiation markers. Upon ageing, these mice developed corneal abnormalities and frequent skin cancer, and the loss of *Notch1* was associated with increased β-catenin signalling[59]. Whether this finding indicated a direct interaction between NOTCH1 and β-catenin was not clear from this study, and the authors also did not report the development of mucosal tumours in the mice, which may either relate to the nature of the conditional knockout using the keratin 5 promoter or to the differential roles of Notch–β-catenin signalling in skin and mucosa. Later, it was shown that Notch and β-catenin have a direct interaction: membrane-bound Notch can bind to active β-catenin[60] and regulate its expression level by degradation. Of note, in addition to the regulation of β-catenin levels,

there are more levels of molecular crosstalk between the NOTCH1 and WNT signalling pathways (reviewed in REF. 61). Furthermore, remarkable in this context is the co-occurrence of inactivating *FAT1* and *NOTCH1* alterations in HNSCC (see supplementary data in REF. 3), suggesting that the carcinogenic route through inactivation of *FAT1* is associated with inactivation of *NOTCH1* to further increase β-catenin levels. Obviously, these are all challenging hypotheses to prove that will require further investigation.

Epigenetic regulation. Gene activation is regulated in part by the opening and closing of the chromatin structure, which is governed by the modification of histone proteins. The histone proteins are enzymatically modified by acetyl and methyl groups, which influence the accessibility of DNA by transcription factors[62]. Besides changes in the histones, the DNA itself may be methylated, typically on cytosine residues in CpG dinucleotides, resulting in specific methylation profiles[62]. DNA methylation profiling and cluster analysis revealed four methylation types in HNSCC irrespective of HPV status that were assigned as hyper-methylated, hypo-methylated, normal-like and CpG island-methylated[3].

A remarkable observation was that nuclear receptor-binding SET domain-containing protein 1 (*NSD1*) and histone-lysine *N*-methyltransferase 2D (*KMT2D*; also known as *MLL2*), both encoding histone-lysine *N*-methyltransferases, were frequently mutated in HNSCC[3]. KMT2D is a histone H3K4 mono-methyltransferase and di-methyltransferase, and germline mutations in *KMT2D* cause Kabuki syndrome[63]. NSD1 is a histone H3K36 methyltransferase, and germline mutations in *NSD1* are associated with Sotos syndrome[64]. Patients with Sotos syndrome show a hypomethylation pattern in the DNA obtained from their blood, which is such a defining characteristic that it is now applied as a diagnostic assay for functional analysis of *NSD1* variants of unknown significance and suggests an association between histone H3 methylation and DNA methylation[65]. Furthermore, this finding fits entirely with the observation in the TCGA study that the DNA methylation subtype 'hypo-methylated' was highly enriched for *NSD1* mutations[3], also indicating that most of these mutations are indeed pathogenic. However, it raises questions on the functional meaning of the few *NSD1* mutations in HNSCCs that cluster to the 'hyper-methylation' type[3]. There are no mutually exclusive or co-occurring mutations that may shed light on where mutations in these two intriguing genes should be placed in HNSCC carcinogenesis.

Another part of epigenetic regulation is achieved by microRNAs (miRNAs)[62], small non-coding RNAs that are generated in a complex biogenesis pathway[66]. Multiple genes are regulated by a single miRNA, and multiple miRNAs can regulate a gene. In HPV–ve HNSCC, decreased expression of miR-let-7c-5p and miR-100-5p were significantly inversely correlated with the expression of genes encoding proteins involved in cell cycle control, which include *CDK6,* transcription factor *E2F1* and polo-like kinase 1 (*PLK1*)[3]. PLK1 had already been identified as an interesting target for chemopreventive treatment in HNSCC[67]. These data again support the critical role for abrogation of cell cycle control in head and neck carcinogenesis.

CNA-silent tumours

In 2007, a separate subgroup within the HPV–ve tumour type emerged that was characterized by very few CNAs, wild-type *TP53* and a more favourable prognosis[68]. This same subgroup was then later identified but alternatively characterized based on wild-type *TP53* and retention of chromosome 3p[69]. In the most recent molecular profiling studies of the TCGA, the existence of this subgroup of tumours was again confirmed, but some intriguing new insights were added. This subgroup of HPV–ve, wild-type *TP53*-bearing and CNA-silent tumours typically display activating *HRAS* and inactivating caspase 8 (*CASP8*) mutations and indeed showed a more favourable prognosis[3].

In our previous investigations, we further showed that this subgroup is diploid, DNA mismatch repair proficient and seems to occur more frequently in females without a history of smoking and alcohol consumption[68]. It is unclear at present whether this subgroup relates to a specific aetiological factor, what the prevalence rates are and whether they change over time, and how this subgroup impacts the data obtained in clinical trials.

Genomic profiling

Besides subgroups of tumours that can be distinguished based on HPV and copy number profile, there are several indications for genomic subgroups based on expression patterns. In 2004, subgroups of tumours were described that could be distinguished based on unsupervised clustering of RNA profiles[14], and these studies were followed by many others[8,15,16,70]. The precise clinical impact is still uncertain, but the clusters themselves appeared more or less consistent and were described as basal, mesenchymal, classical and atypical[15]. The initial studies employed gene-array data, which may be less comparable among platforms, but the most recent RNA sequencing approaches will likely improve this aspect.

These expression subgroups may become increasingly relevant. As an example, from the mutational data, it was shown that the classical subgroup, which is characterized by high expression of genes in oxidative stress response pathways (a finding that is most likely linked to smoking history), is also enriched for gene mutations in these oxidative stress response pathways. None of the mutated genes reached significance on their own, but the combination of gene mutations enriched in this subgroup is striking[3]. A critical gene is nuclear factor erythroid 2-related factor 2 (*NFE2L2*; also known as *NRF1*), which encodes a transcription factor that drives expression of genes in the oxidative stress response and which displays missense mutations in HNSCC[3]. In addition, genes encoding its inhibitors, the ubiquitin ligase components kelch-like ECH-associated protein 1 (KEAP1) and cullin 3 (CUL3), are mutated in some other HNSCCs of this classical subgroup[3]. This same phenomenon has been observed in lung cancer, which

Kabuki syndrome
A dominant autosomal disease caused by inactivating germline mutations in the histone-lysine *N*-methyltransferase 2D (*KMT2D*) gene and characterized by anatomical abnormalities and mental retardation.

Sotos syndrome
A neurological autosomal dominant disorder caused by the loss of one active copy of the nuclear receptor-binding SET domain-containing protein 1 (*NSD1*) gene and characterized by an unusual face with a large skull, acromegalic features and mental retardation.

MicroRNAs
(miRNAs). Small, 22–24 nucleotide single-stranded RNAs that bind to the 3′ untranslated region of genes and cause degradation of the transcript or a stop in translation.

HPV attributable fraction
The percentage of head and neck squamous cell carcinomas in any defined subsite that are assumingly caused by human papillomavirus (HPV).

is also a smoking-induced cancer[71]. Moreover, druggable interactions have been identified in these pathways that might develop into treatment opportunities[72,73]. Hence, with a combination of genomic data mining and translational research, unanticipated therapeutic strategies may emerge.

Molecular landscape of HPV+ve HNSCC

HPV contribution to subsites of HNSCC

A novel risk factor for HNSCC is HPV infection, particularly in the oropharynx, causing OPSCCs. Among OPSCCs, tonsillar squamous cell carcinomas (SCCs), particularly those arising from the tonsillar crypt epithelium, display the highest HPV attributable fraction[2,74–76]. In a recent comprehensive study, the HPV attributable fraction was estimated to be 22% for OPSCCs and 47% specifically for tonsillar SCCs[2]. HPV-16 is by far the most common HPV type and accounts for 83% of HPV+ve OPSCC. HPV attributable fractions vary substantially among geographic regions and have increased over time[2]. Among the regions analysed in this international study, the highest fractions were found in South America and central–eastern and northern Europe, and the lowest were found in southern Europe[2]. Worldwide, HPV attributable fractions of OPSCC increased from 7.2% in the period of 1990–1994 to 32.7% in the period of 2010–2012, likely as a result of changes in sexual behaviour[77]. As future treatment strategies of OPSCC might become dependent on HPV status, it is of the utmost importance to have diagnostic tests available that can reliably discern OPSCC tumours that are caused by and dependent on HPV (BOX 3).

For a long time, data on HPV involvement in non-oropharyngeal HNSCC was less consistent. Early studies had overestimated the role of HPV, likely through over-detection of (traces of) HPV DNA by PCR that was not causally associated with tumour development (such as viral DNA depositions from adjacent or distant sites and laboratory contamination). A systematic review in 2005 described HPV DNA prevalences of 23.5% and 24.0% for oral and laryngeal SCC, respectively[78]. A subsequent meta-analysis in 2014 revealed that substantial subsets of these HPV DNA+ve tumours were negative for mRNA expression of the viral oncogenes *E6* and *E7* (REF. 79), while the presence of *E6* and *E7* mRNA is currently considered the most definitive marker for transforming HPV infections (BOX 3) in relation to their oncogenic activity. The E7 viral oncoprotein binds and degrades the retinoblastoma pocket proteins RB, p130 (also known as RBL2) and p107 (also known as RBL1), inducing S phase-related molecular changes to support viral replication, and E6 binds and degrades p53 to prevent apoptosis by this unscheduled S phase entry[80]. Therefore, expression of these viral oncogenes in proliferating cells is causally linked to malignant transformation[81].

The most recent comprehensive analysis in 2016 yielded estimates for HPV involvement on the basis of the presence of HPV DNA plus E6 and E7 region *E6*I* mRNA of only 3.9% and 3.1% for oral and laryngeal cancers, respectively[2], with HPV-16 as the prevailing type. In terms of geographical distribution, HPV attributable fractions of both oral and laryngeal cancers were highest in South America, Central America and northern Europe[2]. Unlike in HPV-driven OPSCCs, no trends have been observed over time for HPV-driven oral and laryngeal SCCs[2].

Carcinogenesis

The natural history of HPV-induced head and neck tumours remains an enigma. In contrast to cervical carcinomas, of which development via dysplastic precursor lesions over a long time period is well known, little is known about the pathogenesis of HPV-driven HNSCCs. In fact, dysplastic lesions have rarely been found in tonsils, which represent the site from which

Box 3 | Tests for assessing HPV involvement in head and neck squamous cell carcinoma

As human papillomavirus (HPV) DNA PCR methods also detect viral DNA depositions from other sites or laboratory contamination, it alone is not sufficiently accurate[129] at identifying HPV-driven head and neck squamous cell carcinoma (HNSCC). HPV DNA in situ hybridization is more specific because the presence of HPV DNA is directly linked to the tumour cells, but it lacks sensitivity[130]. Transforming HPV infections that lead to cancer development have non-canonical regulation of expression of the viral oncogenes *E6* and *E7*, resulting in overexpression of their encoded protein products in proliferating cells[80]. Therefore, detection of *E6* and *E7* region *E6*I* mRNA by reverse transcription PCR is considered the gold standard for establishing HPV involvement[131–133], but it is technically challenging with routine formalin-fixed tissue specimens.

As increased E7 expression in proliferating cells results in epigenetic de-silencing of the cyclin-dependent kinase inhibitor 2A (*CDKN2A*) gene encoding $p16^{INK4A}$ (REF. 134), $p16^{INK4A}$ immunohistochemistry has also emerged as a valuable surrogate marker for HPV-driven oropharyngeal SCCs (OPSCCs). However, this method requires knowledge of the presence of HPV to reduce false-positivity rates[135,136]. Combined $p16^{INK4A}$ immunohistochemistry and HPV DNA PCR testing is considered a reliable strategy for diagnosis of HPV-driven OPSCC[132,136,137]. Interestingly, HPV-16 E6 seropositivity is surprisingly high in HPV-16-positive OPSCCs with a cervical cancer-like viral mRNA pattern (sensitivity of 96% at a specificity of 98%), and, therefore, measurement of serological responses to HPV-16 E6 has recently also been proposed as a way to identify HPV-driven OPSCC[138].

For the diagnosis of HPV-driven HNSCC at non-oropharyngeal sites, tests or test algorithms have not been validated so far. Many laryngeal tumours that are positive for HPV *E6*I* mRNA lack $p16^{INK4A}$ overexpression[2]. Therefore, the test algorithm involving $p16^{INK4A}$ immunohistochemistry and HPV DNA PCR double testing may not be useful for non-oropharyngeal cancers. In addition, the measurement of serological responses is not an option given the rather low seroconversion rates for these tumours, in contrast to OPSCC[138].

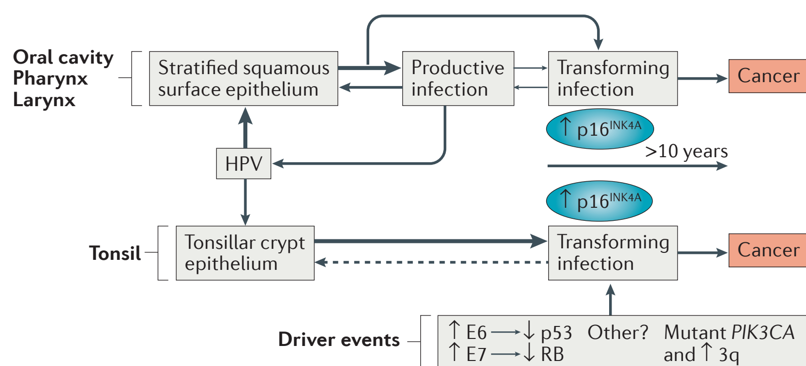

Figure 3 | **Concept of human papillomavirus-induced carcinogenesis in the head and neck region.** According to this concept, stratified squamous surface epithelium within the head and neck region is generally permissive for a productive human papillomavirus (HPV) infection. Infection rates would follow a gradient determined by exposure and immune control, the latter being higher in the oropharynx than, for example, in the oral cavity. As in the cervix, such productive infections are presumed to have a low propensity to change into a transforming infection. Productive infections may yield high viral titres (arrow back to HPV) that, following deposition, increase the likelihood of subsequent infection of tonsillar crypt epithelial cells (small arrow pointing downwards). However, these cells do not support the normal viral life cycle but are highly susceptible to a transforming HPV infection. Transforming infections are characterized by deregulated HPV E6 and E7 oncoprotein expression in proliferating cells. E7 and E6 inactivate RB and p53, respectively. Together, these features result in disruption of cell cycle regulation and inhibition of p53-mediated apoptotic responses, driving immortalization and the accumulation of epigenetic and genetic alterations necessary for progression towards cancer[80]. One of these driver events involves oncogenic activation of the PI3K pathway, often by mutation and/or amplification of *PIK3CA*, the gene encoding the catalytic p110α subunit of class 1 PI3Ks. Similar to cervical cancer development, development of HPV-driven head and neck squamous cell carcinoma (HNSCC) from a transforming infection may take more than 10 years[139]. ↑ indicates overexpression or gain; ↓ indicates loss of expression; dotted arrow indicates regression towards normalcy, which might occur in rare cases.

most HPV-driven HNSCCs develop[10]. In addition, large studies on non-malignant tonsil samples have either failed to detect HPV or found a rather low HPV prevalence[82–84]. This finding is somewhat surprising because the discontinuous organization of epithelial cells in reticular tonsillar crypt epithelium would suggest that these are more accessible to HPV than those of stratified squamous surface epithelium. Therefore, this low infection rate may reflect effective immune control in the lymphoepithelial tissue of the tonsils. Alternatively, the tonsillar reticular crypt epithelium might be non-permissive for productive HPV infections that support the viral life cycle to generate viral progeny. This seems to be different in the oral cavity, in which HPV infections are much more common[85,86]. In a study in which gargles and tonsil brushings were collected from the same cancer-free subjects ($n = 268$), HPV positivity was more than threefold higher in gargles than in tonsil brushings (13% versus 4%)[84]. On the contrary, transforming HPV infections seem apparently uncommon in the oral cavity given the rarity of HPV-driven oral SCC (HPV attributable fraction is 3.9%)[2], and although HPV infections in tonsillar crypt epithelium are rare, the proportion of infections that become transforming and consequently progressive is relatively high (HPV attributable fraction is 47%)[2].

Productive HPV infections
A type of human papillomavirus (HPV) infection that supports the normal viral life cycle and leads to viral progeny.

Transforming HPV infections
A type of human papillomavirus (HPV) infection that does not result in virus production but may instead cause malignant transformation of the infected cell.

Why tonsillar crypt epithelium would be more susceptible to transforming HPV infections is unknown, but insights may be deduced from cervical carcinogenesis. In squamous epithelium of the cervix, about 20% of HPV infections result in dysplastic lesions, most of which represent productive HPV infections[80]. Such productive lesions have histologically mild-to-moderate dysplastic features and display no signs of transformation. In such lesions, viral gene expression and resulting virus production are tightly regulated in a differentiation-dependent manner, where expression of the *E6* and *E7* genes is very low in the proliferative cell compartment and elevated in the differentiating cell compartment[80]. Only a minority of HPV infections in the cervix (estimated to be 3–5%) give rise to transforming infections[80]. These infections, in which the normal viral life cycle is aborted and E6 and E7 expression becomes strongly elevated in the proliferating cells, have a moderate-to-severe dysplastic manifestation.

The productive versus transforming fate of an HPV infection may be dependent on the infected cell type of origin. In a concept based on the findings of Herfs et al.[87], it has been proposed that specific single-layered epithelial cells of embryonic origin at the cervical squamocolumnar junction constitute highly susceptible cells for transforming infections and form the progenitor cells of most cervical carcinomas[80]. These junction cells are characterized by a specific gene expression profile and show detectable overexpression of several proteins by immunohistochemistry, including cytokeratin 7 (CK7). By contrast, productive infections are supposed to arise from infection of basal cells of the stratified squamous epithelium[80].

Of interest, a recent study using tonsillar specimens revealed strong CK7 immunostaining in reticular crypt epithelium, whereas staining was absent or patchy in the stratified squamous surface epithelium of normal tonsils[88]. In OPSCC, CK7 overexpression was significantly associated with a tonsillar subsite and HPV and p16^{INK4A} positivity. Thus, similar to cervical squamocolumnar junction cells, tonsillar crypt epithelial cells constitute a different cellular entity that may have a high transformation susceptibility. HPV-16 E6 and E7 proteins, when expressed from a heterologous promoter using lentiviral constructs, can induce an extended lifespan of both tonsillar surface and crypt epithelial cells[89]. Hence, once overexpressed, E6 and E7 can induce transformation of both cell types with comparable efficiency. Therefore, different transformation susceptibilities of crypt epithelial cells compared with surface epithelial cells may be due to the differential activity of the viral *E6* and *E7* promoters in the two cell types, and this concept is outlined in FIG. 3.

A novel element that is currently not considered in the pathogenesis model of HPV+ve tumours in general is the recent observation that HPV-16 consists of several sequence variants, and some of these sequence variants are more likely to occur in cervical cancers than others[90,91]. Similar data in HNSCC is not yet available but is expected to follow in due course.

Genetic changes and driver genes

Data from the recent comprehensive epigenetic and genomic landscape studies revealed that the genes most frequently altered in HNSCC are largely unaffected in the HPV+ve fraction[3,92]. These include *TP53* and *CDKN2A*. This finding is in line with the notion that (epi)genetic alterations leading to inactivation of p53 and p16^{INK4A}–cyclin D1–RB pathways are functionally equivalent to binding and inactivation of p53 and RB by HPV E6 and E7, respectively. Inactivation of these tumour suppressor pathways (either virally, epigenetically or genetically induced) represents the earliest hit in HNSCC development and in vitro triggered immortalization of human epithelial cells[19,93].

As in cervical carcinomas, molecular alterations in genes of the PI3K pathway constitute the most common genetic changes in HPV-driven HNSCC. These particularly involve activating mutations and amplifications of the oncogene *PIK3CA*, which encodes the catalytic p110α subunit of the class 1 PI3Ks. The *PIK3CA* gene resides at 3q26, a locus that is the most frequently gained region by far in HPV-driven HNSCC and cervical carcinomas[3]. As described in detail elsewhere[19,94], the PI3K signalling pathway affects translation and transcription of multiple targets that are involved in various cellular properties such as proliferation, survival and motility. Of note, although *PIK3CA* alterations seem to be more common in HPV+ve HNSCC[56,57,95], these are also recurrently found in HPV–ve counterparts[3,92]. Therefore, oncogenic PI3K signalling seems to constitute a driver for many HNSCCs, independent of HPV involvement.

In HPV-16-transformed and HPV-18-transformed foreskin epithelial cells, *PIK3CA* expression was found to increase with progression from an immortal to an anchorage-independent phenotype[96]. Inhibition of PI3K signalling by chemical interference or small interfering RNA (siRNA)-mediated silencing of *PIK3CA* resulted in reduced cellular viability, migration and anchorage-independent growth[96], and it can be assumed that this would also be the case in mucosal epithelial cells of the upper aerodigestive tract. Although AKT is largely regarded as the dominant mediator of oncogenic PI3K signalling in most cancers[94], it is still questionable whether this also holds true for HPV-driven HNSCC. A recent study revealed that HPV+ve OPSCC with activating *PIK3CA* mutations did not show increased phosphorylation of AKT[95]. In a mouse model of oral carcinogenesis, overexpression of *Pik3ca* was found to trigger tumour invasion and metastasis, but this was not accompanied with AKT activation[97]. Here, signalling was mediated through 3-phosphoinositide-dependent protein kinase 1 (PDK1), and PI3K–PDK1 signalling was found to interact with transforming growth factor-β (TGFβ)–SMAD3 signalling during tumour progression. Irrespective of the exact key signalling mediator, the collected data indicate that oncogenic PI3K signalling is a common additive hit during HPV-induced carcinogenesis in the head and neck region (FIG. 4). As in the cervix, PI3K alterations are already manifest in dysplastic lesions of the head and neck region in patients without evidence of invasive cancer and may mark cancer precursor lesions with a high risk of malignant transformation, irrespective of HPV status[98,99]. Additional alterations in HPV+ve HNSCC are rare, and it is still unknown to what extent these reflect crucial driver events. Noteworthy are the HPV-specific losses of chromosomal loci 14q32 and 9q, which contain the tumour necrosis factor receptor-associated factor 3 (*TRAF3*) and ataxia telangiectasia mutated (*ATM*) genes, respectively[38]. Furthermore, the apolipoprotein B mRNA-editing enzyme catalytic subunit (APOBEC)-induced mutational signature, resulting from high cytosine deaminase activity, causes a specific mutational profile in HPV+ve HNSCCs[38]. This is common in tumours with a viral origin, although other tumours may show this signature as well. APOBEC induction is a response to virus infection but may also result from other causes of increased *APOBEC* gene expression such as gene amplification[100].

Apolipoprotein B mRNA-editing enzyme catalytic subunit
(APOBEC). A class of cytidine deaminases that function in innate immunity as well as in RNA editing. Viral infections may induce high expression of these genes, causing specific mutation patterns that have also been identified in human papillomavirus (HPV)-induced head and neck cancers.

Genomic profiling

Recent data indicate that HPV+ve OPSCCs are more heterogeneous with respect to their expression patterns. Using expression profiling and cluster analysis, HPV+ve HNSCC could be subclassified into two gene profile groups[16]. In a subsequent RNA sequencing study[101], these two groups were characterized by a signature of mesenchymal and immunological response genes (named HPV-IMU) or keratinocyte differentiation and oxidative stress genes (named HPV-KRT). More detailed molecular analysis revealed that the second group had more frequently integrated HPV and had a higher ratio of spliced *E6* to full length *E6* transcripts. Moreover, this group was enriched for 3q CNAs and *PIK3CA* mutations, while HPV-IMU tumours were enriched for chromosome 16q losses[101]. A significant survival difference between these two subgroups could not be shown, but the prognostic analysis was hampered by limited outcome data and a smaller sample size than is required to evaluate HPV+ve tumours with their favourable prognosis. Hence, HPV+ve tumours can also be subclassified. Validation of these data could lead to a more refined prognostic stratification of patients who are HPV+ve, which is particularly relevant as de-intensification of treatment is currently being studied.

Genomic progression models

The initial genetic progression models[26], later adapted and simplified[19], were elementary, one-dimensional linear models. However, these simplified models no longer capture the molecular complexity of the disease. The plethora of new candidate cancer genes that are mutated in HNSCC and the apparent molecular heterogeneity require more complex and diverse models, as there are likely multiple routes leading to mucosal squamous cell transformation (FIG. 4). Whether and how this will impact clinical management remains a challenge for the future. Deregulation of the cell cycle by abrogation of the RB and p53 pathways seems to occur

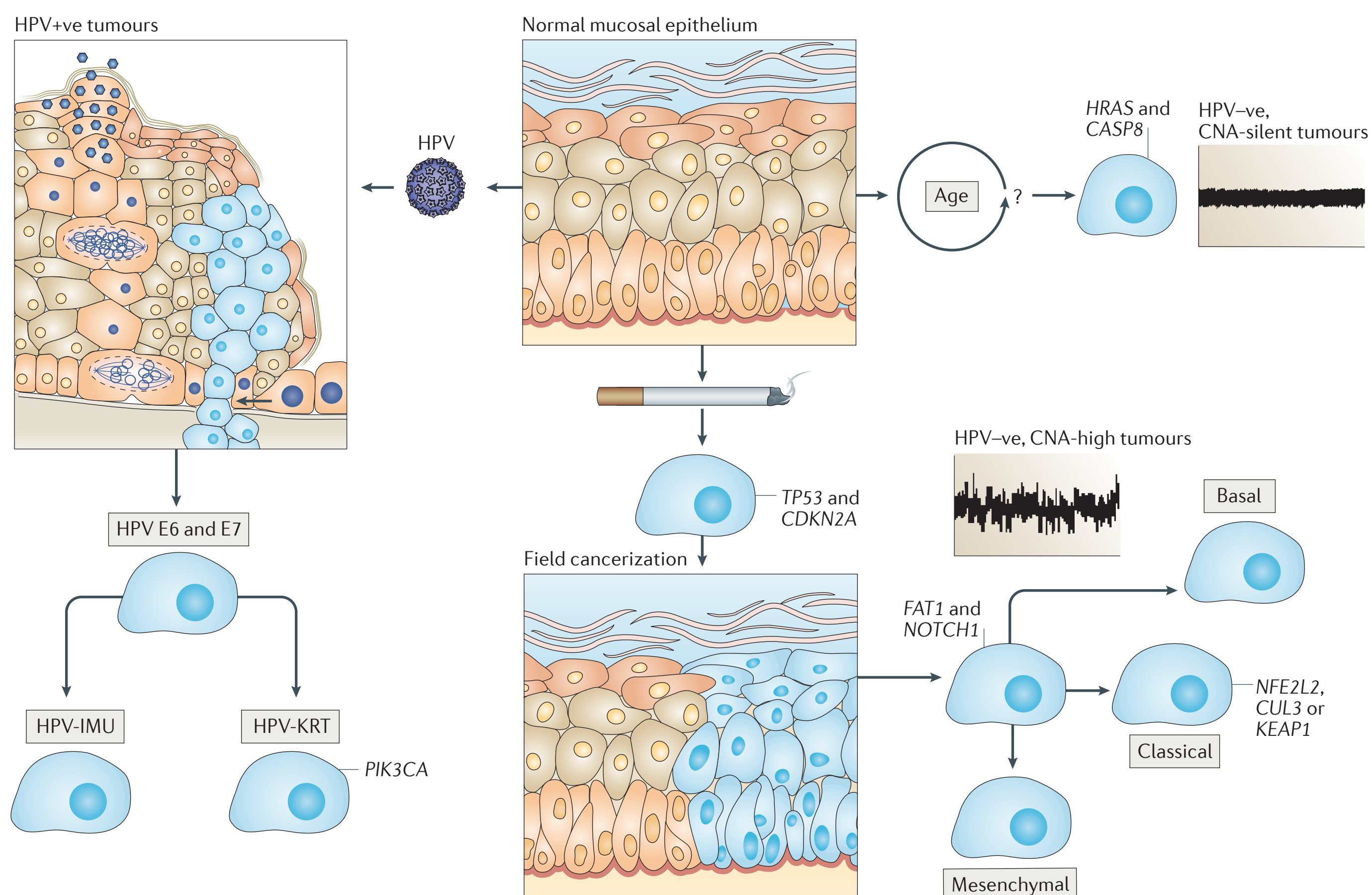

Figure 4 | **Genomic carcinogenesis models of head and neck squamous cell carcinoma.** A schematic overview of squamous cell carcinogenesis in the head and neck is shown. The main distinction among head and neck squamous cell carcinomas (HNSCCs) is the presence of three genetic subgroups: tumours that contain transcriptionally active human papillomavirus (HPV+ve), tumours that are HPV-negative (HPV–ve) and have numerous copy number alterations (CNA-high), and tumours that are HPV–ve but CNA-silent. HPV infection in oral squamous epithelium mostly leads to productive infections, while infection particularly in specific oropharyngeal crypt cells (light blue) might lead to an oncogenic event resulting in either HPV-KRT (HPV-keratinocyte differentiation and oxidative reduction process) or HPV-IMU (HPV-immune response and mesenchymal cell differentiation[101]) tumours. These subgroups have been identified by expression profiling but have not been definitively verified and are still under investigation. The p53 and RB pathways that play a key role in cell cycle control are frequently abrogated in HPV–ve tumours, except they seem to remain active in CNA-silent tumours. In addition, the aetiology of this latter subgroup remains unclear, and ageing is hypothesized to be the risk factor. Many cancer genes and pathways seem to be involved in the progression of the HPV–ve, CNA-high tumours, but *FAT1* and *NOTCH1*, which might act in the WNT–β-catenin pathway, are worth mentioning, and smoking is a known risk factor. At least three subgroups of tumours can be identified based on expression profiling, indicated as classical, basal and mesenchymal, but more may exist. The classical HPV–ve, CNA-high subgroup is characterized by nuclear factor erythroid 2-related factor 2 (NFE2L2) pathway mutations. HPV–ve tumours typically develop from mucosal precursor changes that can present as leukoplakias. Cells in these 'fields' progress to cancer by an accumulation of mutations. The current lack of data on precursor changes hampers the precise timing of events, but it is likely that the accumulation of events is the most important factor. Specific details and references are indicated in the main text. *CASP8*, caspase 8; *CDKN2A*, cyclin-dependent kinase inhibitor 2A; *CUL3*, cullin 3; *KEAP1*, kelch-like ECH-associated protein 1.

at the start of carcinogenesis in almost all HNSCCs[19]. This can then be followed by PI3K pathway activation[3]. There are no mutually exclusive mutations or concordant mutations with those in *PIK3CA*[3]. A route with FAT1–NOTCH1–AJUBA pathway alterations that impact upon β-catenin signalling might form another step in carcinogenesis. Whether the combination of all these mutations fully transforms primary mucosal keratinocytes remains to be determined. In addition, the role of EGFR and other growth factor receptors and their ligands demands further study in appropriate models by genome editing, although some data suggest that amplification or mutation of these proteins form an alternative to amplification of *CCND1* (REFS 3,19). As mentioned above, there is a separate route leading to the subgroup of HPV–ve tumours that lack CNAs and are characterized by wild-type *TP53* as well as *CASP8* and *HRAS* mutations.

Emerging developments

Intratumoural heterogeneity

A problem in treatment, in particular with respect to radiation and chemotherapeutic agents, is intratumoural heterogeneity, which may explain why cancer cells within a tumour show different therapeutic responses. In part, this may relate to different genetic changes in the tumours[102] but also to the presence of cells with stem cell properties[103–106]. Genetic heterogeneity and its potential clinical relevance to HNSCC has been shown with the development and use of a mathematical algorithm that exploits available genetic data[107–109]. However, sequencing of multiple biopsy samples and, recently, single cells from tumours and recurrences will shed more light on this phenomenon[110]. The clinical impact of the cancer stem cell concept in HNSCC remains somewhat elusive. After the publication of studies showing the existence of HNSCC stem cells in xenografted mouse models[103,104] and their associated ability to undergo epithelial to mesenchymal transition (EMT)[111], as well as initial studies showing clinical impact[105,106], this research field did not find its way to the clinic. At present, numbers of stem cells are analysed in relation to outcome[105,106] but not to their biological characteristics, which are likely more relevant. Moreover, the currently used stem cell biomarkers are not specific for stem cells only, which hampers the data interpretation and likely impacts the associations. The recent data on single cell sequencing in HNSCC might become a game changer in this respect[110].

Disease monitoring with biomarkers

Next-generation sequencing has propelled the use of 'liquid biopsies', such as saliva and blood, for the analysis of genetic markers that might predict relapse at the earliest stage. Initial studies in head and neck cancer showed the feasibility of these approaches and suggested that they will develop into reliable disease monitoring methods[112,113]. Salvage surgery after chemoradiation is most effective when recurrent disease is detected early, and the current standard for recurrence detection is a single ^{18}F-fluorodeoxyglucose (FDG)-positron emission tomography (PET) imaging at 3 months, as frequent imaging is unsuitable for disease monitoring because of costs and radiation exposure. Therefore, analysis of liquid biopsies might substantially improve follow-up management of treated patients. Circulating tumour DNA analysis, exosome profiling and analysis of tumour-educated blood platelets and nucleated blood cells are intriguing developments in this research field[112–117].

The immune landscape and immunotherapy

Head and neck cancer is an intrinsically immune-suppressing disease[118]. This observation has fostered research into the role of the immune landscape in HNSCC in relation to outcome. In addition, immune checkpoint inhibitors have emerged as novel and effective therapeutic options[119,120]. In 2016, two studies were published demonstrating long lasting responses with immune checkpoint inhibitors, but these responses occurred in only 10–20% of patients[119,120]. Differential responses are not uncommon with these treatments, as several factors play a role, including degree of tumour lymphocyte infiltration, expression of the immune checkpoint proteins, availability of neoantigens and others[121]. The challenges for the future will be the selection of patients who will benefit from such treatment, the identification of the most suitable treatment regimens and the reduction of immunosuppression in non-responding patients with head and neck cancer. The optimal application of these novel treatment options is under active investigation, but it is exciting to note that patients with such poor prognosis might still be cured.

Conclusions

Over the past decade, the role of HPV in HNSCC has changed the research field, and the impact for staging and prognosis became manifest in the new eighth edition of the TNM staging system. Personalized treatment and the de-intensification of current treatment protocols on the basis of HPV status is on the horizon. Together, this will make a consensus on HPV testing important. The lack of precursor lesions in the upper aerodigestive tract caused by HPV remains puzzling, and it is likely that there is an anatomical separation of productive infection and transforming infection. In 2015, the TCGA consortium published a comprehensive molecular catalogue of head and neck cancer, with mutations associated mostly with tumour suppressor genes in cell cycle control, cellular growth and survival, WNT–β-catenin signalling, and epigenetics. In addition, the existence of a subgroup of genetically distinct HPV–ve head and neck tumours with favourable prognosis was confirmed. The simple linear progression models need to be revised, but this revision also demands thorough functional studies. HNSCC is a disease characterized by frequent mutations that create neoantigens, and it is an exciting prospect that immunotherapies show efficacy in subgroups of patients.

1. Ferlay, J. et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int. J. Cancer* **136**, E359–E386 (2015).
2. Castellsague, X. et al. HPV involvement in head and neck cancers: comprehensive assessment of biomarkers in 3680 patients. *J. Natl. Cancer Inst.* **108**, djv403 (2016).
3. Lawrence, M. S. et al. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature* **517**, 576–582 (2015).
A report with what is currently the largest genomics data set of head and neck cancer.
4. Seiwert, T. Y. et al. Integrative and comparative genomic analysis of HPV-positive and HPV-negative head and neck squamous cell carcinomas. *Clin. Cancer Res.* **21**, 632–641 (2015).
5. Braakhuis, B. J. M. et al. Genetic patterns in head and neck cancers that contain or lack transcriptionally active human papillomavirus. *J. Natl Cancer Inst.* **96**, 998–1006 (2004).
This study presents for the first time the major molecular difference between HPV+ve and HPV–ve HNSCC.
6. Smeets, S. J. et al. Genome-wide DNA copy number alterations in head and neck squamous cell carcinomas with or without oncogene-expressing human papillomavirus. *Oncogene* **25**, 2558–2564 (2006).
7. Slebos, R. J. C. et al. Gene expression differences associated with human papillomavirus status in head and neck squamous cell carcinoma. *Clin. Cancer Res.* **12**, 701–709 (2006).
8. Wichmann, G. 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* **137**, 2846–2857 (2015).
9. Ang, K. K. et al. Human papillomavirus and survival of patients with oropharyngeal cancer. *N. Engl. J. Med.* **363**, 24–35 (2010).
10. Rietbergen, M. M. et al. Human papillomavirus detection and comorbidity: critical issues in selection of patients with oropharyngeal cancer for treatment de-escalation trials. *Ann. Oncol.* **24**, 2740–2745 (2013).

11. O'Sullivan, B. et al. Development and validation of a staging system for HPV-related oropharyngeal cancer by the International Collaboration on Oropharyngeal cancer Network for Staging (ICON-S): a multicentre cohort study. *Lancet Oncol.* **17**, 440–451 (2016). **On the basis of this paper, the TNM staging was adapted for HPV + ve oropharyngeal cancer in the eighth edition.**
12. Brierley, J. D. G., Gospodarowicz, M. K. & Wittekind, C. *TNM Classification of Malignant Tumours* 8th edn (Wiley-Blackwell, 2016).
13. Belbin, T. J. et al. Molecular classification of head and neck squamous cell carcinoma using cDNA microarrays. *Cancer Res.* **62**, 1184–1190 (2002).
14. Chung, C. H. et al. Molecular classification of head and neck squamous cell carcinomas using patterns of gene expression. *Cancer Cell* **5**, 489–500 (2004). **This is the first report to describe head and neck cancer classifications based on gene expression profiles.**
15. Walter, V. et al. Molecular subtypes in head and neck cancer exhibit distinct patterns of chromosomal gain and loss of canonical cancer genes. *PLOS ONE* **8**, e56823 (2013).
16. Keck, M. K. et al. Integrative analysis of head and neck cancer identifies two biologically distinct HPV and three non-HPV subtypes. *Clin. Cancer Res.* **21**, 870–881 (2015).
17. Schnitt, S. J. Classification and prognosis of invasive breast cancer: from morphology to molecular taxonomy. *Modern Pathol.* **23**, S60–S64 (2010).
18. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646–674 (2011).
19. Leemans, C. R., Braakhuis, B. J. M. & Brakenhoff, R. H. The molecular biology of head and neck cancer. *Nat. Rev. Cancer* **11**, 9–22 (2011).
20. Vogelstein, B. et al. Cancer genome landscapes. *Science* **339**, 1546–1558 (2013).
21. Brouns, E. R. E. A. et al. Malignant transformation of oral leukoplakia in a well-defined cohort of 144 patients. *Oral Dis.* **20**, e19–e24 (2014).
22. Petti, S. Pooled estimate of world leukoplakia prevalence: a systematic review. *Oral Oncol.* **39**, 770–780 (2003).
23. Napier, S. S. & Speight, P. M. Natural history of potentially malignant oral lesions and conditions: an overview of the literature. *J. Oral Pathol Med.* **37**, 1–10 (2008).
24. van der Waal, I. Potentially malignant disorders of the oral and oropharyngeal mucosa; terminology, classification and present concepts of management. *Oral Oncol.* **45**, 317–323 (2009).
25. Slaughter, D. P., Southwick, H. W. & Smejkal, W. Field cancerization in oral stratified squamous epithelium — clinical Implications of multicentric origin. *Cancer* **6**, 963–968 (1953).
26. Califano, J. et al. Genetic progression model for head and neck cancer: implications for field cancerization. *Cancer Res.* **56**, 2488–2492 (1996). **This study presents the first genetic progression model of head and neck cancer.**
27. Tabor, M. P. et al. Genetically altered fields as origin of locally recurrent head and neck cancer: a retrospective study. *Clin. Cancer Res.* **10**, 3607–3613 (2004).
28. Tabor, M. P. et al. Multiple head and neck tumors frequently originate from a single preneoplastic lesion. *Am. J. Pathol.* **161**, 1051–1060 (2002).
29. Tabor, M. P. et al. Persistence of genetically altered fields in head and neck cancer patients: biological and clinical implications. *Clin. Cancer Res.* **7**, 1523–1532 (2001).
30. Zhang, L. W. et al. Loss of Heterozygosity (LOH) profiles-validated risk predictors for progression to oral cancer. *Cancer Prev. Res.* **5**, 1081–1089 (2012).
31. Graveland, A. P. et al. Loss of heterozygosity at 9p and p53 immunopositivity in surgical margins predict local relapse in head and neck squamous cell carcinoma. *Int. J. Cancer* **128**, 1852–1859 (2011).
32. Campbell, J. D. et al. The case for a pre-cancer genome atlas (PCGA). *Cancer Prev. Res* **9**, 119–124 (2016).
33. Tomasetti, C., Li, L. & Vogelstein, B. Stem cell divisions, somatic mutations, cancer etiology, and cancer prevention. *Science* **355**, 1330–1334 (2017).
34. Tomasetti, C. & Vogelstein, B. Variation in cancer risk among tissues can be explained by the number of stem cell divisions. *Science* **347**, 78–81 (2015).
35. van Houten, V. M. M. et al. Mutated p53 as a molecular marker for the diagnosis of head and neck cancer. *J. Pathol.* **198**, 476–486 (2002).
36. Lim, X. H. et al. Interfollicular epidermal stem cells self-renew via autocrine Wnt signaling. *Science* **342**, 1226–1230 (2013).
37. Beck, T. N. & Golemis, E. A. Genomic insights into head and neck cancer. *Cancers Head Neck* **1**, 1 (2016).
38. Hayes, D. N., Van Waes, C. & Seiwert, T. Y. Genetic landscape of human papillomavirus-associated head and neck cancer and comparison to tobacco-related tumors. *J. Clin. Oncol.* **33**, 3227–3236 (2015).
39. Toledo, L., Neelsen, K. J. & Lukas, J. Replication catastrophe: when a checkpoint fails because of exhaustion. *Mol. Cell* **66**, 735–749 (2017).
40. Levine, A. J. & Oren, M. The first 30 years of p53: growing ever more complex. *Nat. Rev. Cancer* **9**, 749–758 (2009).
41. Normanno, N. et al. Epidermal growth factor receptor (EGFR) signaling in cancer. *Gene* **366**, 2–16 (2006).
42. Lin, S. Y. et al. Nuclear localization of EGF receptor and its potential new role as a transcription factor. *Nat. Cell Biol.* **3**, 802–808 (2001).
43. Moser, R. et al. Functional kinomics identifies candidate therapeutic targets in head and neck cancer. *Clin. Cancer Res.* **20**, 4274–4288 (2014).
44. van Roy, F. & Berx, G. The cell-cell adhesion molecule E-cadherin. *Cell. Mol. Life Sci.* **65**, 3756–3788 (2008).
45. Tanoue, T. & Takeichi, M. Mammalian Fat1 cadherin regulates actin dynamics & cell-cell contact. *J. Cell Biol.* **165**, 517–528 (2004).
46. Morris, L. G. T. et al. Recurrent somatic mutation of FAT1 in multiple human cancers leads to aberrant WNT activation. *Nat. Genet.* **45**, 253–261 (2013). **This is an interesting study on the role of the *FAT1* tumour suppressor gene in cancer, including head and neck cancer, and its functional association with WNT signalling.**
47. Clevers, H., Loh, K. M. & Nusse, R. An integral program for tissue renewal and regeneration: WNT signaling and stem cell control. *Science* **346**, 54 (2014).
48. Schimizzi, G. V. & Longmore, G. D. Ajuba proteins. *Curr. Biol.* **25**, R445–R446 (2015).
49. Hirota, T. et al. Aurora-A and an interacting activator, the LIM protein Ajuba, are required for mitotic commitment in human cells. *Cell* **114**, 585–598 (2003).
50. Sun, G. P. & Irvine, K. D. Ajuba family proteins link JNK to Hippo signaling. *Sci. Signal.* **6**, ra81 (2013).
51. Marie, H. et al. The LIM protein Ajuba is recruited to cadherin-dependent cell junctions through an association with alpha-catenin. *J. Biol. Chem.* **278**, 1220–1228 (2003).
52. Kanungo, J., Pratt, S. J., Marie, H. & Longmore, G. D. Ajuba, a cytosolic LIM protein, shuttles into the nucleus and affects embryonal cell proliferation and fate decisions. *Mol. Biol. Cell* **11**, 3299–3313 (2000).
53. Haraguchi, K. et al. Ajuba negatively regulates the Wnt signaling pathway by promoting GSK-3 beta-mediated phosphorylation of beta-catenin. *Oncogene* **27**, 274–284 (2008).
54. Kopan, R. & Ilagan, M. X. G. The canonical NOTCH signaling pathway: unfolding the activation mechanism. *Cell* **137**, 216–233 (2009).
55. Ntziachristos, P., Lim, J. S., Sage, J. & Aifantis, I. From fly wings to targeted cancer therapies: a centennial for NOTCH signaling. *Cancer Cell* **25**, 318–334 (2014).
56. Stransky, N. et al. The mutational landscape of head and neck squamous cell carcinoma. *Science* **333**, 1157–1160 (2011).
57. Agrawal, N. et al. Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in *NOTCH1*. *Science* **333**, 1154–1157 (2011).
58. Sun, W. Y. et al. Activation of the NOTCH pathway in head and neck cancer. *Cancer Res.* **74**, 1091–1104 (2014).
59. Nicolas, M. et al. Notch1 functions as a tumor suppressor in mouse skin. *Nat. Genet.* **33**, 416–421 (2003).
60. Kwon, C. et al. Notch post-translationally regulates beta-catenin protein in stem and progenitor cells. *Nat. Cell Biol.* **13**, 1244–1251 (2011). **This paper describes a very intriguing new role for NOTCH1, highlighting the interaction of NOTCH1 with WNT signalling.**
61. Borggrefe, T. et al. The Notch intracellular domain integrates signals from Wnt, Hedgehog, TGF beta/BMP and hypoxia pathways. BBA Mol. *Cell Res.* **1863**, 303–313 (2016).
62. Perri, F. et al. Epigenetic control of gene expression: potential implications for cancer treatment. *Crit. Rev. Oncol. Hemat.* **111**, 166–172 (2017).
63. Ng, S. B. et al. Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. *Nat. Genet.* **42**, 790–793 (2010).
64. Kurotaki, N. et al. Haploinsufficiency of NSD1 causes Sotos syndrome. *Nat. Genet.* **30**, 365–366 (2002).
65. Choufani, S. et al. NSD1 mutations generate a genome-wide DNA methylation signature. *Nat. Commun.* **6**, 10207 (2015).
66. Lin, S. B. & Gregory, R. I. MicroRNA biogenesis pathways in cancer. *Nat. Rev. Cancer* **15**, 321–333 (2015).
67. de Boer, D. V. et al. Targeting PLK1 as a novel chemopreventive approach to eradicate prenoplastic mucosal changes in the head and neck. *Oncotarget* **8**, 97928–97940 (2017).
68. Smeets, S. J., Braakhuis, B. J. M., Ylstra, B., Leemans, C. R. & Brakenhoff, R. H. *TP53* mutations are associated with a particular pattern of genomic imbalances in head and neck squamous cell carcinoma. *Cell. Oncol.* **29**, 160–160 (2007). **This study reports the existence of a novel, genetically defined subgroup of head and neck cancers.**
69. Gross, A. M. et al. Multi-tiered genomic analysis of head and neck cancer ties TP53 mutation to 3p loss. *Nat. Genet.* **46**, 939–943 (2014).
70. De Cecco, L. et al. Head and neck cancer subtypes with biological and clinical relevance: meta-analysis of gene-expression data. *Oncotarget* **6**, 9627–9642 (2015).
71. Wilkerson, M. D. et al. Lung squamous cell carcinoma mRNA expression subtypes are reproducible, clinically Important, and correspond to normal cell types. *Clin. Cancer Res.* **16**, 4864–4875 (2010).
72. Romero, R. et al. Keap1 loss promotes Kras-driven lung cancer and results in dependence on glutaminolysis. *Nat. Med.* **23**, 1362–1368 (2017).
73. Bar-Peled, L. et al. Chemical proteomics identifies druggable vulnerabilities in a genetically defined cancer. *Cell* **171**, 696–709 (2017).
74. Gillison, M. L. et al. Evidence for a causal association between human papillomavirus and a subset of head and neck cancers. *J. Natl Cancer Inst.* **92**, 709–720 (2000). **This is the very first report on the favourable prognosis of HPV-induced head and neck cancers.**
75. Begum, S., Cao, D. F., Gillison, M., Zahurak, M. & Westra, W. H. Tissue distribution of human papillomavirus 16 DNA integration in patients with tonsillar carcinoma. *Clin. Cancer Res.* **11**, 5694–5699 (2005).
76. Kim, S. H. et al. HPV integration begins in the tonsillar crypt and leads to the alteration of p16, EGFR and c-myc during tumor formation. *Int. J. Cancer* **120**, 1418–1425 (2007).
77. D'Souza, G. et al. Case-control study of human papillomavirus and oropharyngeal cancer. *N. Engl. J. Med.* **356**, 1944–1956 (2007).
78. Kreimer, A. R., Clifford, G. M., Boyle, P. & Franceschi, S. Human papillomavirus types in head and neck squamous cell carcinomas worldwide: a systematic review. *Cancer Epidem. Biomar.* **14**, 467–475 (2005).
79. Ndiaye, C. et al. HPV DNA, E6/E7 mRNA, and p16(INK4a) detection in head and neck cancers: a systematic review and meta-analysis. *Lancet Oncol.* **15**, 1319–1331 (2014).
80. Steenbergen, R. D. M., Snijders, P. J. F., Heideman, D. A. M. & Meijer, C. J. L. M. Clinical implications of (epi)genetic changes in HPV-induced cervical precancerous lesions. *Nat. Rev. Cancer* **14**, 395–405 (2014).
81. Rampias, T., Sasaki, C., Weinberger, P. & Psyrri, A. E6 and E7 gene silencing and transformed phenotype of human papillomavirus 16-positive oropharyngeal cancer cells. *J. Natl Cancer Inst.* **101**, 412–423 (2009).
82. Chen, R. W. et al. Presence of DNA of human papillomavirus 16 but no other types in tumor-free tonsillar tissue. *J. Clin. Microbiol.* **43**, 1408–1410 (2005).
83. Palmer, E. et al. Human papillomavirus infection is rare in nonmalignant tonsil tissue in the UK: Implications for tonsil cancer precursor lesions. *Int. J. Cancer* **135**, 2437–2443 (2014).
84. Combes, J. D. et al. Prevalence of human papillomavirus in tonsil brushings and gargles in cancer-free patients: The SPLIT study. *Oral Oncol.* **66**, 52–57 (2017).

85. Kreimer, A. R. et al. Oral human papillomavirus in healthy individuals: a systematic review of the literature. *Sex. Transm. Dis.* **37**, 386–391 (2010).
86. Gillison, M. L. et al. Prevalence of oral HPV infection in the United States, 2009–2010. *JAMA* **307**, 693–703 (2012).
87. Herfs, M. et al. A discrete population of squamocolumnar junction cells implicated in the pathogenesis of cervical cancer. *Proc. Natl Acad. Sci. USA* **109**, 10516–10521 (2012).
This is an intriguing study that describes a subgroup of epithelial cells in the cervix characterized by specific molecular markers and that seem susceptible to HPV-mediated carcinogenesis.
88. Woods, R. S. R. et al. Cytokeratin 7 in oropharyngeal squamous cell carcinoma: a junctional biomarker for human papillomavirus-related tumors. *Cancer Epidem. Biomar.* **26**, 702–710 (2017).
89. Kang, S. Y. et al. Characterization of epithelial progenitors in normal human palatine tonsils and their HPV16 E6/E7-induced perturbation. *Stem Cell Rep.* **5**, 1210–1225 (2015).
90. Mirabello, L. et al. HPV16 sublineage associations with histology-specific cancer risk using HPV whole-genome sequences in 3200 women. *J. Natl. Cancer Inst.* **108**, djw100 (2016).
91. Mirabello, L. et al. HPV16 E7 Genetic conservation is critical to carcinogenesis. *Cell* **170**, 1164–1174 (2017).
92. Chung, C. H. et al. Genomic alterations in head and neck squamous cell carcinoma determined by cancer gene-targeted sequencing. *Ann. Oncol.* **26**, 1216–1223 (2015).
93. Smeets, S. J. et al. Immortalization of oral keratinocytes by functional inactivation of the p53 and pRb pathways. *Int. J. Cancer* **128**, 1596–1605 (2011).
94. Bader, A. G., Kang, S. Y., Zhao, L. & Vogt, P. K. Oncogenic PI3K deregulates transcription and translation. *Nat. Rev. Cancer* **5**, 921–929 (2005).
95. Sewell, A. et al. Reverse-phase protein array profiling of oropharyngeal cancer and significance of PIK3CA mutations in HPV-associated head and neck cancer. *Clin. Cancer Res.* **20**, 2300–2311 (2014).
96. Henken, F. E. et al. PIK3CA-mediated PI3-kinase signalling is essential for HPV-induced transformation in vitro. *Mol. Cancer* **10**, 71 (2011).
97. Du, L. et al. Overexpression of PIK3CA in murine head and neck epithelium drives tumor invasion and metastasis through PDK1 and enhanced TGF beta signaling. *Oncogene* **35**, 4641–4652 (2016).
98. Woenckhaus, J. et al. Genomic gain of PIK3CA and increased expression of p110alpha are associated with progression of dysplasia into invasive squamous cell carcinoma. *J. Pathol.* **198**, 335–342 (2002).
99. Verlaat, W. et al. Somatic mutation in PIK3CA is a late event in cervical carcinogenesis. *J. Pathol. Clin. Res.* **1**, 207–211 (2015).
100. Roberts, S. A. et al. An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers. *Nat. Genet.* **45**, 970–976 (2013).
101. Zhang, Y. X. 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.* **22**, 4735–4745 (2016).
102. Andor, N. et al. Pan-cancer analysis of the extent and consequences of intratumor heterogeneity. *Nat. Med.* **22**, 105–113 (2016).
103. Prince, M. E. et al. Identification of a subpopulation of cells with cancer stem cell properties in head and neck squamous cell carcinoma. *Proc. Natl Acad. Sci. USA* **104**, 973–978 (2007).
This is the first report on the existence of specific cell populations in head and neck cancer with stem cell characteristics.
104. Martens-de Kemp, S. R. et al. CD98 marks a subpopulation of head and neck squamous cell carcinoma cells with stem cell properties. *Stem Cell Res.* **10**, 477–488 (2013).
105. Rietbergen, M. M. et al. Cancer stem cell enrichment marker CD98: a prognostic factor for survival in patients with human papillomavirus-positive oropharyngeal cancer. *Eur. J. Cancer* **50**, 765–773 (2014).
106. de Jong, M. C. et al. CD44 expression predicts local recurrence after radiotherapy in larynx cancer. *Clin. Cancer Res.* **16**, 5329–5338 (2010).
107. Mroz, E. A. & Rocco, J. W. MATH, a novel measure of intratumor genetic heterogeneity, is high in poor-outcome classes of head and neck squamous cell carcinoma. *Oral Oncol.* **49**, 211–215 (2013).
108. Mroz, E. A. et al. High intratumor genetic heterogeneity is related to worse outcome in patients with head and neck squamous cell carcinoma. *Cancer* **119**, 3034–3042 (2013).
109. Mroz, E. A., Tward, A. M., Hammon, R. J., Ren, Y. & Rocco, J. W. Intra-tumor genetic heterogeneity and mortality in head and neck cancer: analysis of data from The Cancer Genome Atlas. *PLOS Med.* **10**, 1371 (2015).
110. Puram, S. V. et al. Single-cell transcriptomic analysis of primary and metastatic tumor ecosystems in head and neck cancer. *Cell* **171**, 1611–1624 (2017).
111. Biddle, A. et al. Cancer stem cells in squamous cell carcinoma switch between two distinct phenotypes that are preferentially migratory or proliferative. *Cancer Res.* **71**, 5317–5326 (2011).
112. Bettegowda, C. et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. *Sci. Transl Med.* **6**, 224ra24 (2014).
113. Wang, Y. X. et al. Detection of somatic mutations and HPV in the saliva and plasma of patients with head and neck squamous cell carcinomas. *Sci. Transl Med.* **7**, 293ra104 (2015).
114. Principe, S. et al. Tumor-derived exosomes and microvesicles in head and neck cancer: implications for tumor biology and biomarker discovery. *Proteomics* **13**, 1608–1623 (2013).
115. Braakhuis, B. J. M. et al. Expression signature in peripheral blood cells for molecular diagnosis of head and neck squamous cell carcinoma. *Oral Dis.* **19**, 452–455 (2013).
116. Best, M. G. et al. RNA-seq of tumor-educated platelets enables blood-based pan-cancer, multiclass, and molecular pathway cancer diagnostics. *Cancer Cell* **28**, 666–676 (2015).
117. Best, M. G. et al. Swarm intelligence-enhanced detection of non-small-cell lung cancer using tumor-educated platelets. *Cancer Cell* **32**, 238–252 (2017).
118. Ferris, R. L. Immunology and immunotherapy of head and neck cancer. *J. Clin. Oncol.* **33**, 3293–3305 (2015).
119. Ferris, R. L. et al. Nivolumab for recurrent squamous-cell carcinoma of the head and neck. *N. Engl. J. Med.* **375**, 1856–1867 (2016).
120. Seiwert, T. Y. et al. Safety and clinical activity of pembrolizumab for treatment of recurrent or metastatic squamous cell carcinoma of the head and neck (KEYNOTE-012): an open-label, multicentre, phase 1b trial. *Lancet Oncol.* **17**, 956–965 (2016).
References 119 and 120 are the first reports to demonstrate the clinical efficacy of immune checkpoint inhibitors for head and neck cancer.
121. Topalian, S. L., Taube, J. M., Anders, R. A. & Pardoll, D. M. Mechanism-driven biomarkers to guide immune checkpoint blockade in cancer therapy. *Nat. Rev. Cancer* **16**, 275–287 (2016).
122. Budach, W. et al. Induction chemotherapy followed by concurrent radio-chemotherapy versus concurrent radio-chemotherapy alone as treatment of locally advanced squamous cell carcinoma of the head and neck (HNSCC): a meta-analysis of randomized trials. *Radiother. Oncol.* **118**, 238–243 (2016).
123. Bonner, J. A. et al. Radiotherapy plus cetuximab for squamous-cell carcinoma of the head and neck. *N. Engl. J. Med.* **354**, 567–578 (2006).
124. Schilling, C. et al. Sentinel European Node Trial (SENT): 3-year results of sentinel node biopsy in oral cancer. *Eur. J. Cancer* **51**, 2777–2784 (2015).
125. Gan, S. J. et al. Incidence and pattern of second primary malignancies in patients with index oropharyngeal cancers versus index nonoropharyngeal head and neck cancers. *Cancer* **119**, 2593–2601 (2013).
126. Wood, L. D. et al. The genomic landscapes of human breast and colorectal cancers. *Science* **318**, 1108–1113 (2007).
127. Lawrence, M. S. et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* **499**, 214–218 (2013).
This article excellently demonstrates the challenges with identifying passenger and driver mutations in cancer, their underlying origins and bioinformatic solutions.
128. Parsons, D. W. et al. An integrated genomic analysis of human glioblastoma multiforme. *Science* **321**, 1807–1812 (2008).
129. Boscolo-Rizzo, P., Pawlita, M. & Holzinger, D. From HPV-positive towards HPV-driven oropharyngeal squamous cell carcinomas. *Cancer Treat. Rev.* **42**, 24–29 (2016).
130. Rischin, D. Oropharyngeal cancer, human papillomavirus, and clinical trials. *J. Clin Oncol.* **28**, 1–3 (2010).
131. van Houten, V. M. M. et al. Biological evidence that human papillomaviruses are etiologically involved in a subgroup of head and neck squamous cell carcinomas. *Int. J. Cancer* **93**, 232–235 (2001).
132. Smeets, S. J. et al. A novel algorithm for reliable detection of human papillomavirus in paraffin embedded head and neck cancer specimen. *Int. J. Cancer* **121**, 2465–2472 (2007).
This study demonstrates simple solutions for the false-positive findings with HPV DNA testing and for the use of test algorithms with surrogate markers to assess HPV involvement in head and neck cancer.
133. Holzinger, D. et al. Viral RNA patterns and high viral load reliably define oropharynx carcinomas with active HPV16 involvement. *Cancer Res.* **72**, 4993–5003 (2012).
134. McLaughlin-Drubin, M. E., Crum, C. P. & Mungera, K. Human papillomavirus E7 oncoprotein induces KDM6A and KDM6B histone demethylase expression and causes epigenetic reprogramming. *Proc. Natl Acad. Sci. USA* **108**, 2130–2135 (2011).
135. Rietbergen, M. M. et al. Molecular characterization of p16-immunopositive but HPV DNA- negative oropharyngeal carcinomas. *Int. J. Cancer* **134**, 2366–2372 (2014).
136. Prigge, E. S., Arbyn, M., Doeberitz, M. V. & Reuschenbach, M. Diagnostic accuracy of p16(INK4a) immunohistochemistry in oropharyngeal squamous cell carcinomas: A systematic review and meta-analysis. *Int. J. Cancer* **140**, 1186–1198 (2017).
137. Rietbergen, M. M. et al. Increasing prevalence rates of HPV attributable oropharyngeal squamous cell carcinomas in the Netherlands as assessed by a validated test algorithm. *Int. J. Cancer* **132**, 1565–1571 (2013).
138. Holzinger, D. et al. Sensitivity and specificity of antibodies against HPV16 E6 and other early proteins for the detection of HPV16-driven oropharyngeal squamous cell carcinoma. *Int. J. Cancer* **140**, 2748–2757 (2017).
139. Kreimer, A. R. et al. Evaluation of human papillomavirus antibodies and risk of subsequent head and neck cancer. *J. Clin. Oncol.* **31**, 2708–2714 (2013).

Acknowledgements
The authors' research summarized here is supported by Cancer Center Amsterdam, the Dutch Cancer Society, the European Commission, The Netherlands Organization for Scientific Research (NWO), The Fanconi Anaemia Research Fund and The German Fanconi Support Group.

Author contributions
C.R.L, P.J.F.S and R.H.B contributed to the design, concepts and writing of the manuscript.

Competing interests
R.H.B has a longstanding collaboration with and receives support from InteRNA Technologies BV and received support from Agilent Technologies Netherlands BV and AbbVie. C.R.L has a collaboration with and receives support from Genmab and InteRNA Technologies BV. C.R.L. also participates in the advisory boards of Merck and MSD. P.J.F.S. is a minority stakeholder and the Chief Science Officer of Self-Screen BV, a spin-off company of VU University Medical Center. P.J.F.S. has been on the Speaker Bureau of Roche Diagnostics, Gen-Probe, Abbott, Qiagen and Seegene and has been a consultant for Crucell BV.

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

Reviewer information
Nature Reviews Cancer thanks R. Ferris and D. Hayes for their contribution to the peer review of this work.

RELATED LINKS
cBioPortal: http://www.cbioportal.org/
The Cancer Genome Atlas: http://www.cancergenome.nih.gov
WNT homepage: http://web.stanford.edu/group/nusselab/cgi-bin/wnt/