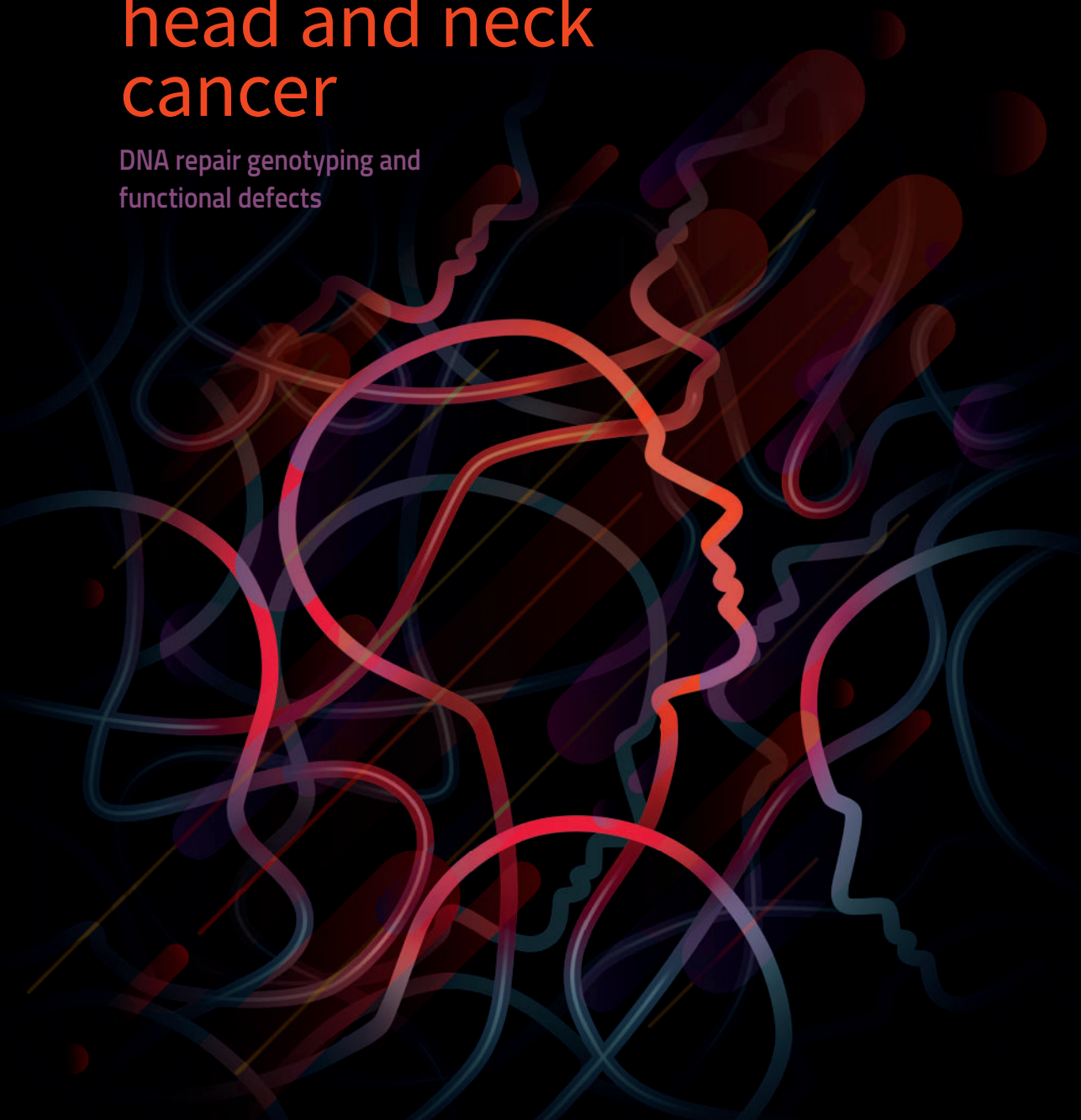
Precision medicine in head and neck cancer

DNA repair genotyping and functional defects



Caroline Verhagen

Precision medicine in head and neck cancer

DNA repair genotyping and functional defects

Caroline Vera Maria Verhagen

PRECISION MEDICINE IN HEAD AND NECK CANCER
DNA repair genotyping and functional defects

ISBN: 9789491688140

Cover design and Chapterpages
Promotie in zicht (www.promotie-inzicht.nl)

Lay-out
Nauka, Amsterdam

Precision medicine in head and neck cancer

DNA repair genotyping and functional defects

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. K.I.J. Maex
ten overstaan van een door het College
voor Promoties ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op woensdag 2 juni 2021, te 13.00 uur

door
Caroline Vera Maria Verhagen
geboren te Nijmegen

Promotiecommissie

Promotores:	prof. dr. M.W.M. van den Brekel	Universiteit van Amsterdam
	prof. dr. M. Verheij	Vrije Universiteit Amsterdam
Copromotores:	dr. C. Vens	NKI-AvL
	dr. V.B. Wreesmann	Portsmouth Hospitals University NHS Trust
Overige leden:	prof. dr. A.J.M. Balm	Universiteit van Amsterdam
	prof. dr. L.E. Smeele	Universiteit van Amsterdam
	prof. dr. C.R.N. Rasch	Universiteit Leiden
	prof. dr. H.A.M. Marres	Radboud Universiteit Nijmegen
	prof. dr. M.J. van de Vijver	Universiteit van Amsterdam
	prof. dr. J.G. Borst	Universiteit Leiden

Faculteit der Tandheelkunde

TABLE OF CONTENTS

CHAPTER 1

Introduction

Personalized medicine (or precision medicine) is defined as an emerging approach for disease treatment and prevention that takes into account individual variability in genes, environment and lifestyle for each person. It aims to give the best treatment possible based on the individual hallmarks of a patient and its tumor, rather than the 'one-size-fits-all' approach.

The human genome comprises approximately three billion nucleotides of deoxyribonucleic acid (DNA) base pairs. During the past decade, the tremendous advances in next generation sequencing (NGS) together with a spectacular drop in costs, led to an exponential gain of information about the genetic landscape. Precision medicine progressed enormously as a result of this development.

Although most of the variation in the genome between individuals has no effect on whether or not a disease develops, individual health is greatly influenced by genetic variation in combination with behavior and influences from the environment. In the field of oncology, the impact of personalized medicine is evident. Variations among individuals are present, but the variations observed in cancer cells from these individuals dictate treatment. Precision medicine in HNSCC is still in the developmental stage. The research presented in this study is intended as a step toward precision medicine in head and neck cancer.

1. Head and Neck cancer

Definition and epidemiology

Head and neck cancer, defined as squamous cell carcinoma of the upper aerodigestive tract, is the 9th most common cancer in the Netherlands, with approximately 3000 new cases annually (Nederlandse Kanker Registratie, NKR). Most often people over the age of 50 are affected and the rates in men are more than twice as high as the rates in women. Tumors of the head and neck are most commonly squamous cell carcinomas (HNSCC) and arise in the mucosal linings of the upper aerodigestive tract. The upper aerodigestive tract begins where the skin meets the mucosa at the nasal vestibule and the vermillion borders of the lips, and continues to the junction of the cricoid and cervical trachea and at the level where the hypopharynx meets the cervical esophagus. It is subdivided in several anatomic subsites, which include the oral cavity (mouth), nose and paranasal sinuses, naso-, oro- and hypopharynx and larynx (Figure 1). From the different subsites of head and neck cancer, cancer of the lip and oral cavity is the most common (44%), followed by cancer of the hypo-/oropharynx and larynx. Since HNSCC of the oral cavity, hypo-/oropharynx and larynx account for 80% of all head and neck tumors in the Netherlands, HNSCC will refer to tumors of these subsites for the remainder of this thesis. In this thesis, head and neck malignancies arising from the skin, salivary glands and thyroid gland will not be discussed.

Etiology

Smoking, alcohol and HPV

Since the aerodigestive tract is exposed to inhaled and ingested carcinogens, it is not surprising that heavy tobacco use and alcohol consumption are the most important risk factors for HNSCC. These two factors have a synergistic effect. An estimated 75% of all head and neck tumors can be attributed to those risk factors [2, 3].

In addition, Epstein Barr virus (EBV) is related to nasopharyngeal carcinomas and more importantly human papilloma virus (HPV) has been recognized as another important risk factor in a subset of HNSCC [4]. The presence of HPV is most frequently detected in tumors arising in the oropharynx and is associated with a favorable outcome [5]. In comparison to HPV-negative oropharyngeal cancer, HPV-positive oropharyngeal cancer is typically found in younger and healthier patients with little history of smoking. The incidence of these tumors is increasing in most European countries, including the Netherlands as well as in the United States [6-8]. The observed increase could be related to a change in sexual behavior (e.g. increase in practice of oral sex), since HPV also has an important role in the carcinogenesis of cervical, anal

and penile cancer. HPV-positive tumors have a far more favorable outcome compared to HPV-negative tumors. This is so substantial that the TNM staging in 2018 was adapted to include p16 immunostaining as a surrogate for HPV status.

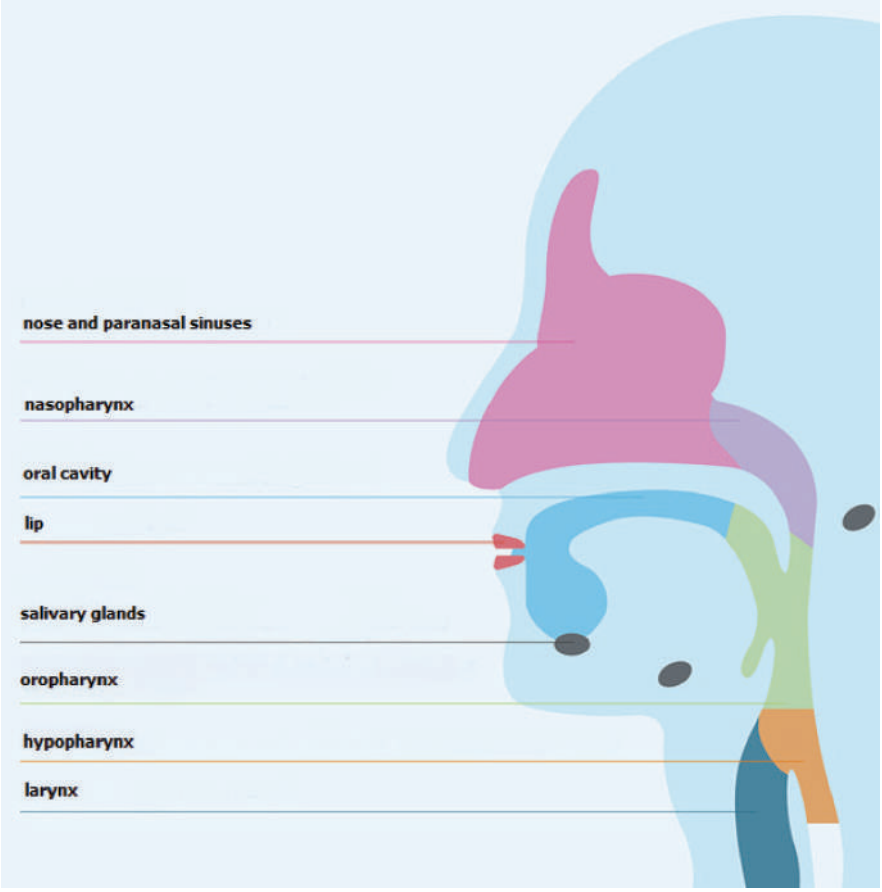


Figure 1. Head and neck cancer regions.
Figure adapted from IKNL [1].

Genetic predisposition to head and neck cancer

Genetics in hereditary syndromes

Individual genetic differences have shown to be of influence in the development of HNSCC. The important role for genetic susceptibility factors is demonstrated by the observation of an increased risk for HNSCC in patients with inherited mutations,

predisposing them to cancer [9, 10]. For example, individuals with a mutation in any of the genes associated with the rare genomic instability syndrome Fanconi anemia (FA), which is characterized by a broad variety of congenital malformations, progressive bone marrow failure, have a 500- to 700-fold higher chance to develop HNSCC than the unaffected population [11, 12]. Patients with the telomere maintenance disease Dyskeratosis congenital (DC), an inherited bone marrow failure syndrome, have a more than 1100-fold risk for HNSCC [13, 14]. Another syndrome with an increased risk of HNSCC, as well as melanoma and pancreatic cancer is Familial Atypical Multiple Mole Melanoma (FAMMM) [15]. Furthermore, the Li Fraumeni syndrome with p53 mutations, and also the Bloom syndrome with mutations in the BLM gene, predispose to development of HNSCC [16, 17].

Genetics in sporadic cancer

Genetic susceptibility also has a role in sporadic cancer as implicated by the fact that carcinogens have a strong association with the development of HNSCC, but only a minority develop a disease even at frequent/intense exposure. These interindividual differences imply a variation in genetic background. Interest has grown in the molecular aspects of susceptibility to develop cancer.

Cancer arises as an accumulation of genetic and epigenetic changes, whose encoded proteins act in a variety of signaling pathways, such as cell cycle control, cellular growth and survival, DNA repair, carcinogen metabolism, WNT-β-catenin signaling, and epigenetics. Studies on the genomic alterations in cancer were profoundly changed and accelerated by the introduction of commercial DNA sequencing machines in the mid 2000s [18]. These sequencers allowed affordable and fast sequencing of (parts of) genomes, commonly referred to as 'next-generation sequencing' (NGS). Soon after the introduction of NGS, numerous studies were undertaken to sequence and map the genomic alterations of cancer [19-23]. The Cancer Genome Atlas (TCGA) [24] is probably the best known example of such projects. TCGA performed molecular characterization at the DNA, RNA, protein, and epigenetic levels and produced the most comprehensive molecular characterization [25].

Most genomes of head and neck tumors turn out to be highly unstable. There are genetic changes at the base-pair level, stretching from a single nucleotide to confined insertions and deletions, and alterations encompassing gains and losses of sometimes entire chromosomes or chromosome arms (copy number alterations (CNA)). Chromosome regions 3p and 8p are frequently lost in HNSCC, whereas 3q, 5p and 8q are often gained [25]. CNAs of whole chromosome arms, as well as changes in copy number of small regions are observed. In addition, numerous genes are found mutated. A median of five mutations per megabase are found [26]. However, not all of these mutations are involved in carcinogenesis. Between 50 and 100 genes are indicated as substantially mutated in HNSCC and are therefore candidate driver genes. Candidate driver genes per definition carry at least one driver mutation that increases cell growth advantage to cancer cells. However, many mutations occur at

low frequencies and functional consequences of these aberrations remain unclear. They therefore remain candidates. Frequently mutated genes with proven functional impact are summarized in Table 1 [3].

Table 1. Frequently mutated genes in HPV-negative HNSCC.
Figure copied from Leemans et al. [3]

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

Some of the most commonly described defects are in the p53, CDKN2A (p16), CCND1 (cyclin D1), epidermal growth factor (EGFR), PIK3CA and NOTCH pathways [25]. Of note, many mutated genes are located in regions of CNAs often leading to loss of heterozygosity. Deregulation of signaling pathways and transcription factors are shown in Figure 2. As shown, the frequency (%) of genetic alterations differs enormously between HPV-negative and HPV-positive tumors. For example, the tumor suppressor protein p53 (TP53) on chromosomal location 17p13 is the most commonly mutated gene in HNSCC (50-80%), but mutations are rare in HPV-positive tumors [25, 27]. Loss or mutation of TP53 can lead to decreased apoptosis and increased proliferation. Inactivation of TP53 in HPV infected cells is achieved by the expression of viral oncoprotein E6, which binds p53 and targets the protein for degradation [28]. Likewise, mutations, mainly deletions, in the cyclin-dependent kinase inhibitor 2A

(CDKN2A) and amplifications CCND1 are common aberrations in HPV-negative tumors [29-31]. These aberrations enable the transition from G1 to S phase. In HPV tumors these aberrations are uncommon, because the viral oncogene E7 destabilizes the retinoblastoma protein (pRb). Besides the genetic differences, HPV-positive tumors have distinctive prognostic implications and might therefore be considered a different entity from tumors induced by smoking and drinking [32].

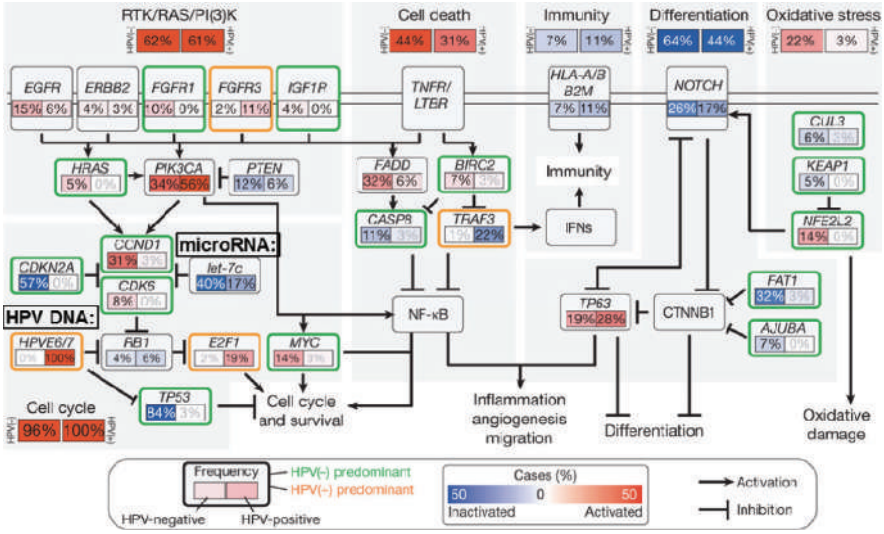


Figure 2. Deregulation of signaling pathways and transcription factors in HNSCC.
The frequency (%) of genetic alterations for HPV-negative and HPV-positive tumors are shown separately within sub-panels and highlighted. Activated and inactivated pathways/genes, and activating or inhibitory symbols are based on predicted effects of genome alterations and/or pathway functions. *Figure and figure legend adapted from TCGA [25].*

Treatment of HNSCC: current practice

Staging

Surgery, radiotherapy and concomitant chemoradiation (CCRT) are the mainstays of treatment of HNSCC. Disease stage at diagnoses is an important parameter to select treatment modality and counsel on prognosis. Staging is performed according to the TNM system which consists of three categories: T - the characteristics of the tumor at the primary site (this may be based on size, location or both); N - the degree of involvement of regional lymph nodes; M - the absence or presence of distant metastasis. In this thesis staging was performed according to the 7th edition TNM

staging system published by the American Joint Committee on Cancer (AJCC) [33]. The system was last updated in 2017 [34]. The most significant change is a separate staging algorithm for high-risk human papillomavirus-associated cancer of the oropharynx, distinguishing it from oropharyngeal cancer with other causes, thereby recognizing its favorable treatment response and prognosis. Secondly extranodal extension for the N-stage of HPV-negative HNSCC as well as depth of infiltration was considered, which is associated with poor outcome.

Based on this TNM system, clinical stages I, II, III and IV can be categorized (Table 2). Specific subdivisions may exist for each stage and denoted with an a, b, or c status. Stage I and II comprise tumors localized within the organ of origin, stage III represents local extension or with cervical lymph node metastasis and stage IV means more advanced cancer, i.e. primary tumor extension beyond the organ of origin and extensive nodal disease or distant metastasis. About one third of patients presents with early staged disease (stage I, II) and generally receive either single modality surgery or radiotherapy (RT) [30]. Patients with advanced disease stages (stage III, IV) represent about two third of HNSCC patients and often require a multi-modality treatment approach, in which surgery is frequently combined with postoperative radiotherapy or concurrent chemoradiation is employed. Non-surgical protocols are applied in the treatment regimen of resectable tumors as well, intending to preserve organ function and quality of life.

Table 2. Clinical stage grouping by T and N status.
Table copied from Deschler et al. [35]

	T				
N	**T1**	**T2**	**T3**	**T4a**	**T4b**
N0	I	II	III	IVa	IVb
N1	III	III	III	IVa	IVb
N2	IVa	IVa	IVa	IVa	IVb
N3	IVb	IVb	IVb	IVb	IVb

Radiotherapy and chemotherapy

External beam radiation therapy (RT) alone or in combination with chemotherapy has a well-established role in the curative treatment of HNSCC as a definitive therapy or adjuvant to primary surgical treatment. The development of three-dimensional (3D) conformal RT improved treatment over the conventional 2D, by limited dosing to normal tissue while adequately treating the tumor. However due to beam constraints this is still not an optimal shielding. Intensity-modulated radiation therapy (IMRT) allows for better sparing of these critical normal tissues by modulating the radiation beam in multiple small beamlets, while covering the tumor volume. Volumetric modulated arc therapy (VMAT) was introduced in 2007. This is a form of rotational

(arc) therapy, which has the advantage of delivering radiation with a varying intensity from a continuous rotation of the radiation source, thereby allowing the patient to be treated from a full 360° beam angle [36].

Chemotherapy in contrast is not an advocated regimen to cure HNSCC as a sole treatment modality. However, there are studies which point towards a role for chemotherapy exclusively in the curative treatment of T1-4N0M0 SCC of the larynx and pharynx [37]. Currently it is commonly used through one of the following approaches: *concurrent* (given along with RT), *neo-adjuvant / induction* (before surgery or (chemo) radiation), *concomitant adjuvant* (given along with RT in the postoperative setting); *adjuvant* (given alone after the completion of surgery, RT or both) or *palliative* (given to patients with incurable cancer to postpone death and/or improve quality of life).

Concurrent chemotherapy is the most commonly used of the chemotherapeutic options, and is utilized to potentiate the effects of RT in order to achieve improved locoregional control, organ preservation and survival. As mentioned above, this application is widely used in the treatment of advanced HNSCC (stage III and IV). Meta-analytic data have shown improved locoregional control and significant (4-6.5%) survival advantage for CCRT protocols compared to single modality RT [38]. Platinum-based agents, such as cisplatin and carboplatin, are typically the compounds of choice used in these regimens. At present, in the Netherlands the most common used CCRT comprises seven weeks of radiotherapy (70 Gray in 35 fractions) combined with three courses of intravenous cisplatin on day 1, 22 and 43 of RT. A cumulative dose of 200 mg/m2 or more is considered optimal, since it has been shown that overall survival was significantly better compared to those who received <200 mg/m2 (74 vs 51% at three years) [39]. Cisplatin is a crosslinking agent. When a crosslinking agent reacts with two nucleotides, the building blocks of DNA, they form a bond between them called a crosslink. These crosslinks can occur within the same strand (intrastrand) or between opposite stands of double-stranded DNA (interstrand) [40]. The formed, complex crosslink-adducts interfere with normal cell metabolism, such as DNA replication an transcription. They are complementary to the single strand breaks formed upon radiation, so cisplatin can enhance the effect of radiotherapy alone, thereby enhancing tumor response. Therefore it is called a radiosensitizer [41]. Cisplatin is very toxic and causes nausea, vomiting, renal failure, myelosuppression and hearing loss. When cisplatin is given concomitantly with radiation, a substantial increase in acute and late toxic effects is observed, compared to radiotherapy given as a single treatment modality, such as severe mucositis and feeding tube dependence due to swallowing problems and therefore it is only appropriate for patients with a good health performance and normal kidney function [42, 43]. Beside the side effects, the prognosis of patients with advanced disease remains poor; only 40-50% will survive for 5 years after treatment [30]. So over 50% of patients suffer from severe toxicities, but their prognosis is not improved. This underlines the need for personalized treatment.

2. Precision medicine

The poor prognosis of locoregionally advanced HNSCC illustrates that both the current diagnostic and the therapeutic approach are insufficient. Since a high number of patients receiving CCRT suffer from severe side effects and have no benefit from the treatment, an effort has been made and research is ongoing to design alternative approaches to treat this group of patients with higher tumor specificity and less toxicity. Until recently, cetuximab, a monoclonal antibody that inhibits EGFR, was the only molecular targeting agent that could fulfil these requirements. The concomitant application of cetuximab with radiation is well endured, but increased acute and chronic skin lesions do occur [44]. It has been shown that patients with HPV-positive oropharyngeal cancer show an inferior overall survival and higher locoregional recurrence rates compared with radiotherapy plus cisplatin [45]. The combination of cetuximab and radiation is therefore sometimes an alternative for patient unable to receive cisplatin. Although extensive research has been done, only few other druggable targets in HNSCC have been identified. An overview will be given below. Despite advances in surgical techniques, chemo- and radiotherapy and the incorporation of targeted agents such as cetuximab, the five-year survival rate has increased to a very limited extent. Approximately 40-50% still succumb to their disease. Current therapy cannot be intensified due to unacceptable toxicity and morbidity.

The TNM staging system classifies tumors into disease stages with distinct prognoses. However, it only explains 10-30% of the variance in survival rates [46, 47]. As cancer is largely a genetic disease, variance in the genetic makeup of tumors is thought to explain part of the clinical variance uncaptured by the TNM staging system. As a result, there is a great scientific effort ongoing to discover prognostic and predictive biomarkers and new targets based on genetic information as revealed by whole genome sequencing [48-50]. In the Netherlands the ongoing multicenter DRUP (Drug Rediscovery Protocol) study is an example of this approach, looking for new targets in advanced and metastasized cancer for whom standard treatment is no longer available. Biomarkers are necessary to select patients that benefit most from different treatment options.

Large scale sequencing studies have elucidated the genetic changes that characterize HNSCC [51]. It has led to the development of therapeutic strategies to target key components of aberrant signaling pathways, DNA repair pathways and immunotherapy. Evidence suggests that these agents in combination with radiotherapy may have synergistic effects [52, 53].

Novel therapeutic strategies based on the HNSCC genomic landscape

Aberrant signaling pathways

The genomic characteristics of HNSCC point to opportunities for targeted inhibitors. Only three of the frequently mutated (candidate) driver genes in HNSCC are oncogenes: EGFR, PIK3CA, and CCND1 (Table 1) [3, 54]. Driver mutations in oncogenes can cause a dependency of cancer cells on these affected pathways [55]. Amplification of EGFR is already targeted in routine clinical practice with cetuximab. The other two frequently mutated oncogenes, PIK3CA and CCND1, are promising candidates for targeted therapy [54].

PIK3CA

The PI3K signaling pathway is the most commonly mutated pathway in HNSCC [21]. Of these, PIK3CA is the most frequently mutated component of this pathway [56]. Constitutive activation of the PI3K/AKT/mTOR pathway by gain-of-function mutations makes it an important oncogenic signaling pathway. Therefore, multiple PI3K/AKT/mTOR pathway inhibitors have been developed. In preclinical studies, some of these inhibitors have demonstrated efficacy in HNSCC [57], but these agents have shown less promise in clinical trials to date [56]. Several reasons may explain this lack of efficacy, among them are acquired resistance, inadequate dosing and patient selection. PI3K/AKT/mTOR pathway mutation status or predictive biomarkers may be essential to select patients that can benefit from targeted treatment. Since single inhibitors have shown to induce compensatory feedback, the majority of clinical trials use the agents in combination with another agent to achieve a synergistic effect [54, 56, 58]. To date, the cumulative result from *in vitro, in vivo* and clinical trials analyzing the PI3K/AKT/mTOR pathway inhibition have been inconsistent. Additional studies are ongoing.

CCND1

Amplification of CCND1 is often found to be combined with loss of CDKN2A (Table 1). The protein product of CCND1 (cyclin D1) binds to either CDK4 or CDK6. In normal physiology the resulting complex contributes to cell cycle progression and is inhibited by the protein product of CDKN2A (p16). Either CCND1 gain or CDKN2A loss could thus promote cell cycle progression. Normally, unscheduled DNA replication leads to DNA damage and p53 activation [59], but TP53 is also frequently inactivated in HNSCC [30]. As both of these mutations are frequent in cancer, CDK4/6 inhibitors have been developed. Such inhibitors have demonstrated value in breast cancer [60] and are currently being evaluated for HNSCC in clinical trials [54, 61] with promising results.

DNA repair

Each cell in the human body contains the same three billion deoxyribonucleic acid (DNA) base pairs. DNA replication in these cells is an ongoing process. Although the system of replication is very accurate, DNA lesions are induced continuously from endogenous sources. Endogenous lesions are a by-product of normal cell metabolism [9, 10, 62]. Examples of exogenous factors causing DNA lesions include exposure to cigarette smoke, alcohol, but also ionizing radiation and UV exposure. Cells contain effective mechanisms to detect and repair DNA damage. Nevertheless, throughout a person's lifetime most cells will accumulate a significant amount of lesions in their DNA. Only infrequently, these alterations enable the cell to acquire the properties needed to escape control mechanisms and become a cancer cell. Interestingly, inherited DNA repair deficiencies commonly predispose to cancer. The deficiencies in the DNA repair pathways enable cancer cells to accumulate genetic alterations that attribute to their aggressive phenotype. On the other hand, in these cases, cancer cells rely on the remaining proficient DNA repair pathways to survive. DNA repair inhibiting agents try to target these back up DNA repair mechanisms to generate synthetic lethality (Figure 3).

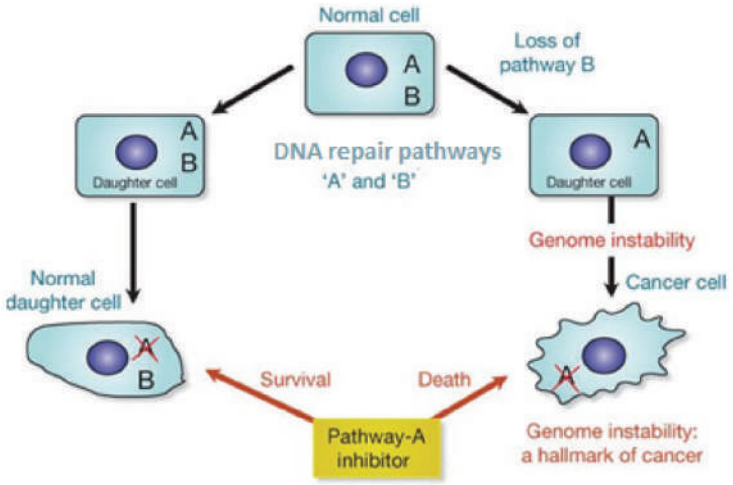


Figure 3. Synthetic lethality.

Normal cells have two DNA repair pathways, A and B. If one pathway is eliminated (B), genomic instability results, which can foster the evolution of a cancer cell. Addition of an inhibitor targeting the second pathway (A) leads to cell death. This is called synthetic lethality. *Figure and figure legend adapted from Jackson et al. [63]*

DNA repair pathways

The wide diversity of DNA lesions necessitates the engagement of multiple repair mechanisms. The key DNA repair pathways are base excision repair (BER), nucleotide excision repair (NER) and mismatch repair (MMR) for single strand DNA breaks (SSBs)

and homologous recombination (HR) and end joining (EJ) pathways for double strand breaks (DSBs) (Figure 4).

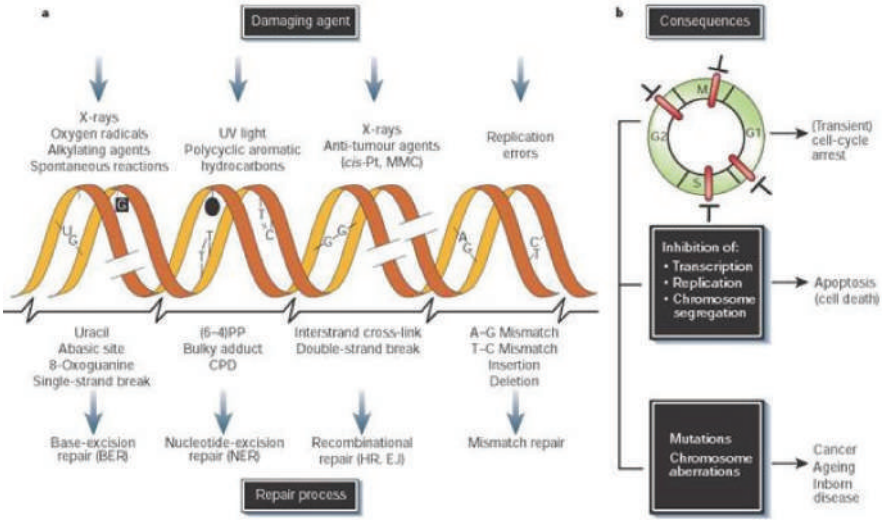


Figure 4. DNA damage, repair mechanisms and consequences.
(A) Common DNA damaging agents (top); examples of DNA lesions induced by these agents (middle); and most relevant DNA repair mechanisms responsible for the removal of the lesions (bottom). **(B)** Acute effects of DNA damage on cell-cycle progression, leading to transient arrest in the G1, S and G2 phases (top); and on DNA metabolism (middle). Long-term consequences of DNA injury (bottom) include permanent changes in the DNA sequence (point mutations affecting single genes or chromosome aberrations which may involve multiple genes) and their biological effects. *Figure and figure legend copied from Hoeijmakers et al. [64]*

In BER a damaged base is often recognized by a DNA glycosylase enzyme that mediates base removal before nuclease, polymerase and ligase proteins complete repair in processes overlapping those used in SSB repair. Whereas BER is mainly involved in endogenous damage, NER lesions predominantly arise from exogenous causes (except for some oxidative lesions). The NER system recognizes helix-distorting base lesions that interfere with base pairing and generally obstruct transcription and normal replication. During NER, the damage is excised as a 22-30-base oligonucleotide, producing ssDNA which activates DNA polymerases and associated factors before ligation ensues. MMR removes nucleotides, mispaired by DNA polymerases and insertion/deletion loops (ranging from one to ten or more bases) that result from slippage during replication of repetitive sequences or during recombination [10, 63, 65].

DSBs are more problematic since both strands are affected and therefore impact on chromosomal integrity. Most DSBs have two incompatible DNA ends that preclude direct ligation. The ends are incompatible due to mismatching overhangs or chemical

modification. EJ pathways can operate in any phase of the cell cycle and are error prone. These pathways are subdivided into the non-homologous end-joining (NHEJ) pathway, and a group of less well elucidated alternative end joining pathways (A-EJ) pathways. NHEJ repairs DNA DSBs by juxtaposing and ligating DNA ends, using very little or no complementary base pairing. In NHEJ, DSB repair is initiated by the binding of the Ku proteins to the ends, which recruits other NHEJ factors, such as the DNA-dependent protein kinase catalytic subunit (DNA-PKcs), X-ray repair cross complementing 4 (XRCC4), XRCC4-like factor (XLF) and DNA ligase IV, to catalyze DNA ligation across the DNA breaks. If the DNA ends are not compatible for ligation, DNA end processing, such as trimming, filling-in, or blocking-end removal, ensues by one of many processing enzymes such as Artemis or DNA polymerase lambda (Polλ) or DNA polymerase mu (Polμ) (Figure 5) [66]. A-EJ pathway, which includes microhomology mediated repair, repairs DNA DSBs by annealing stretches of 2-20 basepairs of overlapping bases flanking the DSB. Not all A-EJ uses microhomology to repair the DNA, but all processes require the MRN (MRE11-RAD50-NBS1 (Nijmegen breakage syndrome protein 1)) complex, carboxy-terminal binding protein interacting protein (CtIP), Poly (ADP-ribose) polymerase 1 (PARP1) and DNA polymerase θ (Polθ). The pathway provides a backup mechanism for both NHEJ and HR [66]. In contrast to EJ pathways, HR and SSA is generally restricted to S and G2 when the DNA is replicated, providing a second copy of the sequence (from the sister chromatid) as the template to mediate faithful repair [63]. These pathways are initiated after ssDNA generation. ssDNA is a result of a process termed DNA end resection, that requires, for most types of ends CtIP. CtIP is believed to stimulate the endonuclease activity of MRE11 in the MRN complex. After endonucleolytic cleavage, resection proceeds bi-directionally with the exonuclease activity of MRN in 3-5’ direction and the redundant activities of exonuclease 1 (EXO1) and the BLM/DNA2 nucleases resecting in the 5’-3’ direction. The ensuing ssDNA is first bound by replication protein A (RPA), which must then be replaced by RAD51 to form a nucleofilament competent for HR through the action of the breast cancer susceptibility genes BRCA1, BRCA2 and PALB2. The ssDNA then invades the undamaged template and DNA polymerization follows, leading to resynthesizing of the sequence surrounding the break (Figure 5) [67]. SSA is in contrast to HR not dependent on RAD51. In SSA, the complementary strands of the homologous regions, flanking a DSB anneal, producing an intermediate producing an intermediate with wo non-homologous 3’ -ended tails that must be removed for new DNA synthesis and ligation to occur [66]. Besides the repair of DSBs, HR is also used to restart stalled replications forks and to repair interstrand DNA crosslinks (ICL). Crosslink repair also involves the Fanconi anemia protein complex [63, 68]. The HR and FA pathway work in close conjunction to repair DSBs and ICLs.

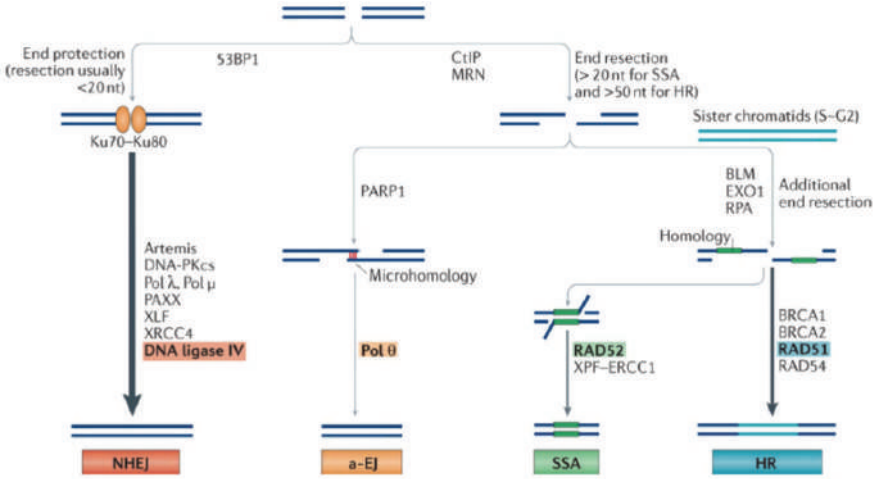


Figure 5. Double strand break repair pathway choice.
DNA DSBs can be repaired by the classical NHEJ pathway, the alternative end joining (A-EJ) pathway, the single-strand annealing (SSA) pathway or by HR. *Figure and figure legend adapted from Chang et al.* [69]

As mentioned previously, Fanconi is a rare hereditary syndrome. Patients with the FA syndrome, carry a mutation in one of the FA genes. A gene is designated FA when its bi-allelic germline mutation is found in at least two FA patients. Due to this requirement for FA nomenclature, many HR genes have received a FA name in retrospect [70]. For example, BRCA1, BRCA2 and RAD51, are also known as FANCS, FANCD1 and FANCR, respectively. Cells unable to use HR to repair the lesions mentioned above, lack repair opportunities or rely on less accurate repair mechanisms, which result in the accumulation of DNA damage and/or cell death. Consequently, defects in the FA/HR pathways can result in genomic instability and increased sensitivity to DNA crosslinking agents such as cisplatin and radiation [71].

PARPi

PARP inhibitors are an important group of DNA damage response agents. PARP1 is a prime sensor for SSBs end plays a critical role in BER [72]. PARP inhibitors (PARPi) can selectively kill HR-deficient cancer cells by synthetic lethality. Three PARPi, of which olaparib is one, are FDA approved for ovarian cancer [57]. The precise mechanism of action is under debate, although it is thought that the cytotoxicity to HR deficient cells is due to the accumulation of SSBs or the trapping of PARP due to the prevention of release of PARP1 on these lesions by PARPi. This leads to replication fork collapse and DSBs, which in turn need HR to be repaired [73]. As BRCA1 and BRCA2 are essential for HR, these DSBs cannot be repaired, therefore leading to cell death (Figure 6). PARPi are under investigation for use against HNSCC [57, 74].

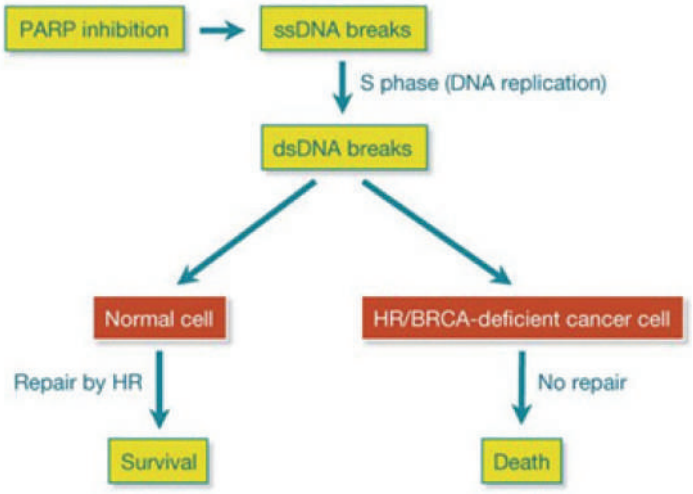


Figure 6. Exploitation of the DNA repair pathway by PARPi.
Treatment with a PARPi selectively kills HR/BRCA-deficient cells. PARP inhibition impairs the repair of SSBs, which are converted to DSBs in S-phase. Such DSBs are effectively repaired by HR in normal cells, but not in BRCA/HR-deficient cells. *Figure and figure legend adapted from Jackson et al. [63]*

Immune dysregulation

The immune system recognizes cancer cells and is entrusted with their elimination. It is held in check by inhibitory receptors and ligands. These immune checkpoint pathways, which normally maintain self-tolerance and limit collateral tissue damage during anti-microbial immune responses, can be co-opted by cancer to evade destruction [75]. The alterations in the immune cell dysfunction within the tumor microenvironment, may allow the tumor to evade immune recognition and elimination, but also provide opportunities for therapeutic intervention. Cytotoxic T-lymphocyte antigen 4 (CTLA-4) and PD-1 were the first checkpoint receptors to be discovered. Drugs interrupting these immune checkpoints, by anti-CTLA-4 antibody ipilimumab and anti-PD1 antibodies pembrolizumab and nivolumab, and others unleash anti-tumor immunity and mediate cancer regression [76]. In 2016 the use of programmed death receptor-1 (PD-1) blocking antibodies nivolumab and pembrolizumab were FDA approved as well for second line treatment. These immune checkpoint inhibitors have been found to be of benefit for patients with recurrent or metastatic HNSCC after platinum-based chemotherapy [77, 78]. Currently, several trials are running to determine optimal timing, dosage and combinations with these immunotherapy agents in the curative setting. Since these drugs are a major clinical success in many cancer types, there are prospects for HNSCC cancer therapy using several checkpoint ligands and receptors.

The immune system plays a key role in the development, establishment and progression of HNSCC [79]. The immune microenvironment of HNSCC is characterized by changes in immune cell populations, immune checkpoints, as well as tumor or microenvironmental factors that alter the balance of the immune milieu in favor of immunosuppression, allowing tumor evasion and escape from immune surveillance [80]. The first publication of immune checkpoint inhibitors in recurrent or metastatic HNSCC emerged from a trial with pembrolizumab [81, 82]. Rapidly more studies followed upon this success. In addition to the known inhibitors described above, new drugs are under evaluation in HNSCC, often as combination therapy [80]. Comprehensive characterization of the diverse immune tumor microenvironments across HNSCC and incorporation of immunologic with genomic analyses are likely to provide prognostic information about the behavior of tumors but also to guide strategies to modulate the immune system for therapeutic benefit.

Radiation response

Currently, radiation is administered, irrespective of tumor heterogeneity or radiosensitivity. HNSCC radiotherapy outcomes are determined by parameters, such as hypoxia, proliferation and intrinsic radiosensitivity, but also the presence of stem cells and microenvironmental factors have a role. A recent analysis using different gene-expression based signatures of biological factors shown to determine cellular radiation response, discovered an important role for acute and chronic hypoxia and immune cell (T cell) markers in locoregional control. Moreover, CD44 and SCL3A2, both stem cell related markers are determinants of radiation response [83].

Adding targeted agents and molecular inhibitors to radiotherapy shows promising results. In fact, PARP inhibitors are effective radiosensitizers [84]. There is also evidence that immune checkpoint inhibitors may have synergistic effects with radiotherapy in that it enhances antitumor immunity by inducing immunogenic cell death and promoting recruitment of T-cells within the tumor microenvironment [85]. Preclinical data support this synergy [79].

In short, there are various classes of drugs that are being evaluated for HNSCC treatment. Each of these novel drugs target specific characteristics of tumor cells. Therefore, predictive biomarkers and careful patient enrolment are required to demonstrate efficacy of these drugs in clinical trials. Once approved, such biomarkers will also be crucial to cost-effectively apply these drugs in the clinic.

Biomarkers

"Biomarker" is a portmanteau of "biological marker". Biomarkers are a group of objective signs of the medical state of a patient, which can be measured and reproduced. Prognostic biomarkers inform about disease outcome irrespective of the treatment administered. In contrast, biomarkers are predictive if the treatment

outcome is different between patients that are positive and negative for the biomarker [86].

Management of HNSCC is based on the TNM stage and sometimes predictive models [87]. Treatment strategies were adapted to improve the cure rate, taking the disease stage into account (eg the addition of chemo- to radiotherapy to improve the locoregional failure rate [38]). However, the differences in outcome between the stages did not change. This underlines the need to change the current treatment, which is uniform for all patients with tumors with the same histological and clinical features, to a treatment which is selected on the properties of the individual tumors, both genetically and biologically. The knowledge about molecular changes that drive carcinogenesis and tumor growth in HNSCC has increased and dozens of biomarkers have been described in literature, however only few are used in clinic [88].

Classical clinical prognostic factors

Classical prognostic factors include tumor subsite, size, volume, involvement of lymph nodes, distant metastasis and tumor grading. In addition to these, patient factors such as age, sex and comorbidity have a role. Moreover, surgery has an important position in the treatment of HNSCC. There is high post-operative risk when there are positive resection margins or when there is extranodal extension. These are factors that are currently used in clinical practice [88]. Lastly, as mentioned above, TNM staging was recently adapted for oropharyngeal HPV-positive tumors, since HPV status, determined by immunohistochemistry and PCR, became an established prognostic biomarker.

Non-genetic biomarkers

The classical clinical prognostic factors can be supported by non-genetic prognostic factors. The broad use is not yet justified since they are usually performing worse than classical prognostic factors and are not cost effective. Tumor extension is one of these important factors. Noninvasive preoperative diagnostic measures on tumor spread, for example infiltration in vital structures or perineural spread, can determine if a lesion is still operable [88]. The extent of neck disease can be determined more reliably with upcoming imaging techniques like the 99m technetium-labeled nanobodies, which target the macrophage mannose receptor (MMR) to predict the involvement of lymph nodes [89].

Radiomics might be another non-genetic biomarker. It uses the information of biomedical images that cannot be seen by eye [90]. Radiomics quantifies textural information from features, such as spatial distribution of signal intensities and pixel interrelationships, which are digitally extracted from imaging methods like CT, MRI and PET. In addition, visual differences, such as shape, image intensity and texture can be quantified, overcoming the subjective nature of image interpretation. As an example, the metabolic tumor volume was shown to be an independent prognostic factor for patients who underwent surgery as their primary treatment [91]. In another study it was shown that a radiomic signature from CT images was able to predict HPV (p16)

status [92]. Another example is the uptake of fluorodeoxyglucose after 10-20 Gy dose to monitor response by PET [93].

Novel prognostic biomarkers and tissue-based markers

The majority of biomarker studies involves tissue-based biomarkers, using various techniques like immunohistochemistry or genomic DNA and mRNA analysis. The molecular biomarkers reflect a wide variety of biological functions: cell migration, metastasis, immunity, cell cycle control, hypoxia or other biological processes. PD-L1 might be a promising biomarker for immune response. Overexpression of EGFR, CD44 or amplification of CCND1 are other biomarkers of high clinical potential.

Liquid biopsy is the analysis of circulating tumor cells (CTCs) or tumor cell products, such as tumor-derived cell-free nucleic acids, exomes and platelets, in blood samples. It has the advantage to be a minimally invasive detection method and to reflect the broad range of the malignancy's properties, including metastases and biochemical changes acquired during blood-borne dissemination, which are not present in the primary tumor. In contrast, there is a chance that tissue biopsies do not show all of the tumor characteristics, since tumors tend to be heterogenous [94]. Studies have been looking for CTCs and since an association was found between inflammation and cancer progression the interest in hematological markers was also raised [95, 96].

Challenges

Many of the published cohorts looking for biomarkers comprise tumors from various disease stages and anatomic sites, that were treated with different combinations of multiple treatment modalities. In such studies, it is uncertain to which tumors or treatment modalities the identified biomarkers apply. More homogenous cohorts are a solution for this problem. However, they are typically small (~50-150 patients) owing to HNSCC's heterogeneity and relatively low incidence. From a statistical perspective, a sample that is small relative to the number of possible predictors increases the risk of overfitting. This is a concern in cancer biomarker research, given the abundance of genomic alterations in cancer genomes, that each are possible predictors.

3. Aim and outline of this thesis

In this thesis we focus on precision medicine for locoregionally advanced HNSCC. The poor prognosis underscores the need for improvement of the treatment of these tumors. As genomic alterations lie at the heart of carcinogenesis they are interesting biomarker candidates and targets. Since FA/HR pathway maintains genomic stability and repairs DNA crosslinks, FA/HR pathway deficiency might be a useful biomarker in HNSCC. Cisplatin is a chemotherapeutical agent, frequently used in the treatment of HNSCC, which causes crosslinks. Cells with alterations causing FA/HR deficiency show increased sensitivity to those agents since they cannot be repaired. This thesis therefore first focused on exploring the occurrence of DNA repair defects in HNSCC and their functional impact as described in **chapter 2**. A panel of 29 HNSCC cell lines was tested for functional outcome measurements which define DNA repair defects. After this, functional outcome was associated with the presence of variations in genes involved in DNA repair pathways as assessed by next generation sequencing in these cell lines. Next, these variants were tested in a set of tumors of patients treated with chemoradiotherapy to test a possible link with treatment outcome.

The work in **chapter 3** elaborates on the selection methods applied in chapter 2 to identify relevant variations in the DNA repair pathway. Bi-allelic mutation of FA/HR genes is a way to lose FA/HR pathway activity and was therefore used as a variant selection criterion. In addition, variant minor allele frequency was used as a selection criterion. In breast and ovarian cancer research, FA/HR variants are often selected based on known involvement in these cancer types. The high prevalence of breast and ovarian cancer in carriers of germline FA/HR mutations suggests an interaction with sex hormones. In contrast, carcinogens from tobacco smoke and alcohol induce the DNA damage that possibly puts stress on the FA/HR pathway in HNSCC. Another subset of FA/HR variants might therefore be pathogenic in HNSCC than in breast and ovarian cancer. Discovering these variants therefore requires a different variant selection approach than one guided by breast or ovarian cancer incidence.

Since the prognosis of advanced HNSCC is poor, and current treatment has many side effects, new treatment methods are needed. To this end, the PARP inhibitor olaparib was tested. Olaparib is involved in SSB repair and proven to be lethal in BRCA-deficient cells. Cells with defects in their FA/HR pathway as described in chapter 2, were tested for their vulnerability to PARPi as monotherapy and for the radiosensitizing potential of PARP. This is described in **chapter 4**.

As discussed above, HPV is nowadays tested on a regular basis as it has shown to be a key prognostic factor in oropharyngeal cancer. It is used as a prognostic biomarker with the potential to define treatment choice. A requirement of transcriptional activity or a role of silent HPV in patients' outcome has not been well-defined. In the study described in **chapter 5** we define the proportion of HNSCC with evidence of HPV DNA and show the prognosis of the group of HNSCC with non-expressing "silent" HPV. It further compares and discusses the sensitivity and significance of the different HPV detection methods.

Finally, in **chapter 6** the results of chapters 2-5 are discussed. These results are placed in the context of contemporary HNSCC research, cancer research in general and possible future directions.

References

[1] Integraal Kankercentrum Nederland. Hoofd-halskanker. 2020, December 21; Available from: https://www.iknl.nl/kankersoorten/hoofd-halskanker.

[2] Blot, W.J., et al., Smoking and drinking in relation to oral and pharyngeal cancer. Cancer Res, 1988. **48**(11): p. 3282-7.

[3] Leemans, C.R., P.J.F. Snijders, and R.H. Brakenhoff, The molecular landscape of head and neck cancer. Nat Rev Cancer, 2018. **18**(5): p. 269-282.

[4] Syrjanen, S., Human papillomavirus (HPV) in head and neck cancer. J Clin Virol, 2005. **32 Suppl 1**: p. S59-66.

[5] Ragin, C.C. and E. Taioli, Survival of squamous cell carcinoma of the head and neck in relation to human papillomavirus infection: review and meta-analysis. Int J Cancer, 2007. **121**(8): p. 1813-20.

[6] Henneman, R., et al., Incidence Changes of Human Papillomavirus in Oropharyngeal Squamous Cell Carcinoma and Effects on Survival in the Netherlands Cancer Institute, 1980-2009. Anticancer Res, 2015. **35**(7): p. 4015-22.

[7] 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, 2013. 24(11): p. 2740-5.

[8] Chaturvedi, A.K., et al., Human papillomavirus and rising oropharyngeal cancer incidence in the United States. J Clin Oncol, 2011. **29**(32): p. 4294-301.

[9] Ciccia, A. and S.J. Elledge, The DNA damage response: making it safe to play with knives. Mol Cell, 2010. **40**(2): p. 179-204.

[10] Hoeijmakers, J.H., DNA damage, aging, and cancer. N Engl J Med, 2009. **361**(15): p. 1475-85.

[11] Kutler, D.I., et al., High incidence of head and neck squamous cell carcinoma in patients with Fanconi anemia. Arch Otolaryngol Head Neck Surg, 2003. **129**(1): p. 106-12.

[12] Auerbach, A.D., Fanconi anemia and its diagnosis. Mutat Res, 2009. **668**(1-2): p. 4-10.

[13] Alter, B.P., et al., Malignancies and survival patterns in the National Cancer Institute inherited bone marrow failure syndromes cohort study. Br J Haematol, 2010. **150**(2): p. 179-88.

[14] Alter, B.P., et al., Cancer in dyskeratosis congenita. Blood, 2009. **113**(26): p. 6549-57.

[15] Vinarsky, V., et al., Head and neck squamous cell carcinoma in FAMMM syndrome. Head Neck, 2009. **31**(11): p. 1524-7.

[16] Friedlander, P.L., Genomic instability in head and neck cancer patients. Head Neck, 2001. **23**(8): p. 683-91.

[17] Berkower, A.S. and H.F. Biller, Head and neck cancer associated with Bloom's syndrome. Laryngoscope, 1988. **98**(7): p. 746-8.

[18] Heather, J.M. and B. Chain, The sequence of sequencers: The history of sequencing DNA. Genomics, 2016. **107**(1): p. 1-8.

[19] Agrawal, N., et al., Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1. Science, 2011. **333**(6046): p. 1154-7.

[20] Stransky, N., et al., The mutational landscape of head and neck squamous cell carcinoma. Science, 2011. **333**(6046): p. 1157-60.

[21] Lui, V.W., et al., Frequent mutation of the PI3K pathway in head and neck cancer defines predictive biomarkers. Cancer Discov, 2013. **3**(7): p. 761-9.

[22] 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, 2015. **21**(3): p. 632-41.

[23] Lechner, M., et al., Targeted next-generation sequencing of head and neck squamous cell carcinoma identifies novel genetic alterations in HPV+ and HPV- tumors. Genome Med, 2013. **5**(5): p. 49.

[24] Cancer Genome Atlas Research, N., et al., The Cancer Genome Atlas Pan-Cancer analysis project. Nat Genet, 2013. **45**(10): p. 1113-20.

[25] Cancer Genome Atlas, N., Comprehensive genomic characterization of head and neck squamous cell carcinomas. Nature, 2015. **517**(7536): p. 576-82.

[26] Alexandrov, L.B., et al., Signatures of mutational processes in human cancer. Nature, 2013. **500**(7463): p. 415-21.

[27] Poeta, M.L., et al., TP53 mutations and survival in squamous-cell carcinoma of the head and neck. N Engl J Med, 2007. **357**(25): p. 2552-61.
[28] Scheffner, M., et al., The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. Cell, 1990. **63**(6): p. 1129-36.
[29] van der Riet, P., et al., Frequent loss of chromosome 9p21-22 early in head and neck cancer progression. Cancer Res, 1994. **54**(5): p. 1156-8.
[30] Leemans, C.R., B.J. Braakhuis, and R.H. Brakenhoff, The molecular biology of head and neck cancer. Nat Rev Cancer, 2011. **11**(1): p. 9-22.
[31] 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, 2006. **25**(17): p. 2558-64.
[32] 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, 2016. **17**(4): p. 440-451.
[33] Edge, S.B. and C.C. Compton, The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. Ann Surg Oncol, 2010. **17**(6): p. 1471-4.
[34] Lydiatt, W.M., et al., Head and Neck cancers-major changes in the American Joint Committee on cancer eighth edition cancer staging manual. CA Cancer J Clin, 2017. **67**(2): p. 122-137.
[35] Deschler DG, M.M., Smith RV, eds, Quick Reference Guide to TNM Staging of Head and Neck Cancer and Neck Dissection Classification. 2014, American Academy of Otolaryngology–Head and Neck Surgery Foundation.
[36] Teoh, M., et al., Volumetric modulated arc therapy: a review of current literature and clinical use in practice. Br J Radiol, 2011. **84**(1007): p. 967-96.
[37] Laccourreye, O., et al., Ten-year outcome of curative "exclusive" chemotherapy in N0M0 squamous cell carcinoma of the larynx and pharynx with complete clinical response. Head Neck, 2019. **41**(7): p. 2190-2196.
[38] Pignon, J.P., et al., Meta-analysis of chemotherapy in head and neck cancer (MACH-NC): an update on 93 randomised trials and 17,346 patients. Radiother Oncol, 2009. **92**(1): p. 4-14.
[39] Al-Mamgani, A., et al., The impact of cumulative dose of cisplatin on outcome of patients with head and neck squamous cell carcinoma. Eur Arch Otorhinolaryngol, 2017. **274**(10): p. 3757-3765.
[40] Kelland, L., The resurgence of platinum-based cancer chemotherapy. Nat Rev Cancer, 2007. **7**(8): p. 573-84.
[41] Sharma, V.M. and W.R. Wilson, Radiosensitization of advanced squamous cell carcinoma of the head and neck with cisplatin during concomitant radiation therapy. Eur Arch Otorhinolaryngol, 1999. **256**(9): p. 462-5.
[42] Bourhis, J., et al., Concomitant chemoradiotherapy versus acceleration of radiotherapy with or without concomitant chemotherapy in locally advanced head and neck carcinoma (GORTEC 99-02): an open-label phase 3 randomised trial. Lancet Oncol, 2012. **13**(2): p. 145-53.
[43] Heukelom, J., et al., Late follow-up of the randomized radiation and concomitant high-dose intra-arterial or intravenous cisplatin (RADPLAT) trial for advanced head and neck cancer. Head Neck, 2016. **38 Suppl 1**: p. E488-93.
[44] Bonner, J.A., et al., Radiotherapy plus cetuximab for locoregionally advanced head and neck cancer: 5-year survival data from a phase 3 randomised trial, and relation between cetuximab-induced rash and survival. Lancet Oncol, 2010. **11**(1): p. 21-8.
[45] Suton, P., et al., Cisplatin-based chemoradiotherapy vs. cetuximab-based bioradiotherapy for p16-positive oropharyngeal cancer: an updated meta-analysis including trials RTOG 1016 and De-ESCALaTE. Eur Arch Otorhinolaryngol, 2019. **276**(5): p. 1275-1281.
[46] Groome, P.A., et al., A comparison of published head and neck stage groupings in carcinomas of the tonsillar region. Cancer, 2001. **92**(6): p. 1484-94.
[47] Groome, P.A., et al., A comparison of published head and neck stage groupings in carcinomas of the oral cavity. Head Neck, 2001. **23**(8): p. 613-24.
[48] Willemsen, A., et al., Molecular tumour boards and molecular diagnostics for patients with cancer in the Netherlands: experiences, challenges, and aspirations. Br J Cancer, 2019. **121**(1): p. 34-36.

[49] Lee, Y.S., D.E. Johnson, and J.R. Grandis, An update: emerging drugs to treat squamous cell carcinomas of the head and neck. Expert Opin Emerg Drugs, 2018. **23**(4): p. 283-299.
[50] Subbiah, V. and K. Sircar, Personalizing Precision Oncology: From Light Microscope to Molecular Microscope. J Natl Compr Canc Netw, 2019. **17**(7): p. 886-888.
[51] Lim, S.M., et al., Investigating the Feasibility of Targeted Next-Generation Sequencing to Guide the Treatment of Head and Neck Squamous Cell Carcinoma. Cancer Res Treat, 2019. **51**(1): p. 300-312.
[52] Frederick, B.A., et al., Combined EGFR1 and PARP1 Inhibition Enhances the Effect of Radiation in Head and Neck Squamous Cell Carcinoma Models. Radiat Res, 2020. **194**(5): p. 519-531.
[53] Devaraja, K., Current Prospects of Molecular Therapeutics in Head and Neck Squamous Cell Carcinoma. Pharmaceut Med, 2019. **33**(4): p. 269-289.
[54] Cho, J., D.E. Johnson, and J.R. Grandis, Therapeutic Implications of the Genetic Landscape of Head and Neck Cancer. Semin Radiat Oncol, 2018. **28**(1): p. 2-11.
[55] Weinstein, I.B., Cancer. Addiction to oncogenes--the Achilles heal of cancer. Science, 2002. **297**(5578): p. 63-4.
[56] Cai, Y., S. Dodhia, and G.H. Su, Dysregulations in the PI3K pathway and targeted therapies for head and neck squamous cell carcinoma. Oncotarget, 2017. **8**(13): p. 22203-22217.
[57] Glorieux, M., R. Dok, and S. Nuyts, Novel DNA targeted therapies for head and neck cancers: clinical potential and biomarkers. Oncotarget, 2017. **8**(46): p. 81662-81678.
[58] Wang, L. and R. Bernards, Taking advantage of drug resistance, a new approach in the war on cancer. Front Med, 2018. **12**(4): p. 490-495.
[59] Toledo, L., K.J. Neelsen, and J. Lukas, Replication Catastrophe: When a Checkpoint Fails because of Exhaustion. Mol Cell, 2017. **66**(6): p. 735-749.
[60] Finn, R.S., et al., The cyclin-dependent kinase 4/6 inhibitor palbociclib in combination with letrozole versus letrozole alone as first-line treatment of oestrogen receptor-positive, HER2-negative, advanced breast cancer (PALOMA-1/TRIO-18): a randomised phase 2 study. Lancet Oncol, 2015. **16**(1): p. 25-35.
[61] Michel, L., et al., Phase I trial of palbociclib, a selective cyclin dependent kinase 4/6 inhibitor, in combination with cetuximab in patients with recurrent/metastatic head and neck squamous cell carcinoma. Oral Oncol, 2016. **58**: p. 41-8.
[62] Lindahl, T. and D.E. Barnes, Repair of endogenous DNA damage. Cold Spring Harb Symp Quant Biol, 2000. **65**: p. 127-33.
[63] Jackson, S.P. and J. Bartek, The DNA-damage response in human biology and disease. Nature, 2009. **461**(7267): p. 1071-8.
[64] Hoeijmakers, J.H., Genome maintenance mechanisms for preventing cancer. Nature, 2001. **411**(6835): p. 366-74.
[65] Hoeijmakers, J.H., DNA repair mechanisms. Maturitas, 2001. **38**(1): p. 17-22; discussion 22-3.
[66] Seol, J.H., E.Y. Shim, and S.E. Lee, Microhomology-mediated end joining: Good, bad and ugly. Mutat Res, 2018. **809**: p. 81-87.
[67] Hustedt, N. and D. Durocher, The control of DNA repair by the cell cycle. Nat Cell Biol, 2016. **19**(1): p. 1-9.
[68] Kennedy, R.D. and A.D. D'Andrea, The Fanconi Anemia/BRCA pathway: new faces in the crowd. Genes Dev, 2005. **19**(24): p. 2925-40.
[69] Chang, H.H.Y., et al., Non-homologous DNA end joining and alternative pathways to double-strand break repair. Nat Rev Mol Cell Biol, 2017. **18**(8): p. 495-506.
[70] Nalepa, G. and D.W. Clapp, Fanconi anaemia and cancer: an intricate relationship. Nat Rev Cancer, 2018. **18**(3): p. 168-185.
[71] Stecklein, S.R. and R.A. Jensen, Identifying and exploiting defects in the Fanconi anemia/BRCA pathway in oncology. Transl Res, 2012. **160**(3): p. 178-97.
[72] Dietlein, F., L. Thelen, and H.C. Reinhardt, Cancer-specific defects in DNA repair pathways as targets for personalized therapeutic approaches. Trends Genet, 2014. **30**(8): p. 326-39.
[73] Wurster, S., et al., PARP1 inhibition radiosensitizes HNSCC cells deficient in homologous recombination by disabling the DNA replication fork elongation response. Oncotarget, 2016. **7**(9): p. 9732-41.

[74] de Haan, R., et al., Study protocols of three parallel phase 1 trials combining radical radiotherapy with the PARP inhibitor olaparib. BMC Cancer, 2019. **19**(1): p. 901.
[75] Topalian, S.L., C.G. Drake, and D.M. Pardoll, Immune checkpoint blockade: a common denominator approach to cancer therapy. Cancer Cell, 2015. **27**(4): p. 450-61.
[76] Hodi, F.S., et al., Improved survival with ipilimumab in patients with metastatic melanoma. N Engl J Med, 2010. **363**(8): p. 711-23.
[77] Ferris, R.L., et al., Nivolumab for Recurrent Squamous-Cell Carcinoma of the Head and Neck. N Engl J Med, 2016. **375**(19): p. 1856-1867.
[78] Larkins, E., et al., FDA Approval: Alectinib for the Treatment of Metastatic, ALK-Positive Non-Small Cell Lung Cancer Following Crizotinib. Clin Cancer Res, 2016. **22**(21): p. 5171-5176.
[79] Ferris, R.L., Immunology and Immunotherapy of Head and Neck Cancer. J Clin Oncol, 2015. **33**(29): p. 3293-304.
[80] Solomon, B., R.J. Young, and D. Rischin, Head and neck squamous cell carcinoma: Genomics and emerging biomarkers for immunomodulatory cancer treatments. Semin Cancer Biol, 2018. **52**(Pt 2): p. 228-240.
[81] 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, 2016. **17**(7): p. 956-965.
[82] Chow, L.Q.M., et al., Antitumor Activity of Pembrolizumab in Biomarker-Unselected Patients With Recurrent and/or Metastatic Head and Neck Squamous Cell Carcinoma: Results From the Phase Ib KEYNOTE-012 Expansion Cohort. J Clin Oncol, 2016. **34**(32): p. 3838-3845.
[83] van der Heijden, M., et al., Biological Determinants of Chemo-Radiotherapy Response in HPV-Negative Head and Neck Cancer: A Multicentric External Validation. Front Oncol, 2019. **9**: p. 1470.
[84] Verhagen, C.V., et al., Extent of radiosensitization by the PARP inhibitor olaparib depends on its dose, the radiation dose and the integrity of the homologous recombination pathway of tumor cells. Radiother Oncol, 2015. **116**(3): p. 358-65.
[85] Pilones, K.A., C. Vanpouille-Box, and S. Demaria, Combination of radiotherapy and immune checkpoint inhibitors. Semin Radiat Oncol, 2015. **25**(1): p. 28-33.
[86] Ballman, K.V., Biomarker: Predictive or Prognostic? J Clin Oncol, 2015. **33**(33): p. 3968-71.
[87] Batenburg, R. and E. Van den Broek, Pharmacy information systems: the experience and user satisfaction within a chain of Dutch pharmacies. Int J Electron Healthc, 2008. **4**(2): p. 119-31.
[88] Budach, V. and I. Tinhofer, Novel prognostic clinical factors and biomarkers for outcome prediction in head and neck cancer: a systematic review. Lancet Oncol, 2019. **20**(6): p. e313-e326.
[89] Agrawal, A., et al., [(99m)Tc]Tilmanocept Accurately Detects Sentinel Lymph Nodes and Predicts Node Pathology Status in Patients with Oral Squamous Cell Carcinoma of the Head and Neck: Results of a Phase III Multi-institutional Trial. Ann Surg Oncol, 2015. **22**(11): p. 3708-15.
[90] van Timmeren, J.E., et al., Radiomics in medical imaging-"how-to" guide and critical reflection. Insights Imaging, 2020. **11**(1): p. 91.
[91] Zhang, H., et al., Metabolic tumour volume as a prognostic factor for oral cavity squamous cell carcinoma treated with primary surgery. J Otolaryngol Head Neck Surg, 2014. **43**: p. 33.
[92] Leijenaar, R.T., et al., Development and validation of a radiomic signature to predict HPV (p16) status from standard CT imaging: a multicenter study. Br J Radiol, 2018. **91**(1086): p. 20170498.
[93] Hentschel, M., et al., Early FDG PET at 10 or 20 Gy under chemoradiotherapy is prognostic for locoregional control and overall survival in patients with head and neck cancer. Eur J Nucl Med Mol Imaging, 2011. **38**(7): p. 1203-11.
[94] Mader, S. and K. Pantel, Liquid Biopsy: Current Status and Future Perspectives. Oncol Res Treat, 2017. **40**(7-8): p. 404-408.
[95] Nakashima, H., et al., Pre-treatment neutrophil to lymphocyte ratio predicts the chemoradiotherapy outcome and survival in patients with oral squamous cell carcinoma: a retrospective study. BMC Cancer, 2016. **16**: p. 41.
[96] Nakahira, M., et al., Prognostic role of the combination of platelet count and neutrophil-lymphocyte ratio in patients with hypopharyngeal squamous cell carcinoma. Eur Arch Otorhinolaryngol, 2016. **273**(11): p. 3863-3867.

CHAPTER 2

Fanconi anemia and homologous recombination gene variants are associated with functional DNA repair defects in vitro and poor outcome in patients with advanced head and neck squamous cell carcinoma

Caroline V.M. Verhagen*, David M. Vossen*, Kerstin Borgmann, Floor Hageman, Reidar Grénman, Manon Verwijs-Janssen, Lisanne Mout, Roel J.C. Kluin, Marja Nieuwland, Tesa M. Severson, Arno Velds, Ron Kerkhoven, Mark J. O'Connor, Martijn van der Heijden, Marie-Louise van Velthuysen, Marcel Verheij, Volkert B. Wreesmann, Lodewyk F.A. Wessels, Michiel W.M. van den Brekel, Conchita Vens.

* Shared first authorship

Oncotarget. 2018 Apr; 9(26): 18198-18213.

Abstract

Mutations in Fanconi Anemia or Homologous Recombination (FA/HR) genes can cause DNA repair defects and could therefore impact cancer treatment response and patient outcome. Their functional impact and clinical relevance in head and neck squamous cell carcinoma (HNSCC) is unknown. We therefore questioned whether functional FA/HR defects occurred in HNSCC and whether they are associated with FA/HR variants. We assayed a panel of 29 patient-derived HNSCC cell lines and found that a considerable fraction is hypersensitive to the crosslinker Mitomycin C and PARP inhibitors, a functional measure of FA/HR defects. DNA sequencing showed that these hypersensitivities are associated with the presence of bi-allelic rare germline and somatic FA/HR gene variants. We next questioned whether such variants are associated with prognosis and treatment response in HNSCC patients. DNA sequencing of 77 advanced stage HNSCC tumors revealed a 19% incidence of such variants. Importantly, these variants were associated with a poor prognosis (p = 0.027; HR = 2.6, 1.1-6.0) but favorable response to high cumulative cisplatin dose. We show how an integrated *in vitro* functional repair and genomic analysis can improve the prognostic value of genetic biomarkers. We conclude that repair defects are marked and frequent in HNSCC and are associated with clinical outcome.

Introduction

Chromosomal stability is governed by DNA damage response and repair processes such as the homologous recombination (HR) and Fanconi Anemia (FA) pathways. The FA-pathway is essential for the repair of DNA interstrand crosslinks and together with elements of the homologous recombination (HR) repair pathway they also strongly determine cellular survival upon exposure to crosslinking agents [1]. Aberrations in the FA/HR-pathway have been reported in multiple cancer types and their therapeutic exploitation has been described [2,3]. The breast cancer susceptibility genes *BRCA1* and *BRCA2* (*BRCA1/2*), well-known members of the FA/HR-pathway, have a well-described role in hereditary breast and ovarian cancer. Recent DNA sequencing studies highlight the high occurrence of DNA repair gene aberrations; however, the assessment of a functional impact lags behind and their clinical relevance remains poorly defined. Notably, the underlying DNA repair defects in *BRCA1/2* mutated breast and ovarian tumors can be exploited with PARP inhibitors [4–6] further stressing the importance of functional DNA repair defect studies.

Fanconi anemia patients suffer from a condition caused by germline mutations in the Fanconi anemia (FA) genes and have an increased susceptibility to cancer. Head and neck squamous cell carcinoma (HNSCC) is the most common solid cancer in these patients, with a 700-fold increased risk [7,8]. Sporadic HNSCC is the sixth most common cancer worldwide and its incidence is strongly associated with alcohol consumption, smoking and HPV infection [9,10]. A considerable proportion of patients is diagnosed at an advanced stage, at which patients are often treated with surgery or a combination of radiotherapy and cisplatin. This combination is effective, although not all patients benefit and less than half of the patients will be cured [11]. In addition, many suffer severe side effects without possibly benefiting from the treatment. New treatment decision aids and alternative therapeutic approaches are therefore urgently needed [12–14].

The strong impact of smoking and alcohol in the development of HNSCC, both likely based on the DNA crosslinking nature of these mutagens [15,16], suggests a protective role of the FA/HR repair pathway. Meta-analysis has shown the benefit of the addition of crosslinking agents to radiotherapy to improve outcome in HNSCC [11] and may further indicate tumor DNA repair defects to be involved in crosslinker sensitivity. Together these data point to a role of crosslink repair defects, particularly those of the FA/HR pathway, in the etiology and treatment of HNSCC.

In sporadic HNSCC, downregulation of FA gene expression [17] and frequent *FANCF* silencing by methylation was found [18]. Furthermore, copy number alterations [19] and somatic mutations of individual FA genes have been described in HNSCC [20,21]. A recent study found FA gene variants in HNSCC cell lines that were responsive to a chromosomal breakage assay [22]. Comprehensive genomic analysis of the FA/HR pathway are rare and it is unknown whether these alterations compromise cellular crosslink repair activity, as functional analyses are lacking [23]. Importantly however,

the clinical relevance of functional or genetic FA/HR tumor defects has not been elucidated.

In this study we therefore investigate the incidence and properties of functional DNA repair defects in HNSCC by applying multiple functional assays to a large HNSCC cell line panel. We then integrate data from these functional assays and DNA sequencing to improve the selection of functionally relevant genetic alterations. Finally, we probe the association of such FA/HR aberrations with clinical outcome in a well-defined homogenous HNSCC patient cohort (n = 77) treated with radiotherapy and cisplatin to test their prognostic value.

Results

Hypersensitivity to the DNA crosslinking agent mitomycin C reveals functional crosslink repair defects in HNSCC

Hypersensitivity to the crosslinking agent mitomycin C (MMC) and a strong G2 cell cycle block in response to MMC are hallmarks of FA-pathway disruption [24,25]. To test whether sporadic HNSCCs have such DNA repair defects, we treated 29 HNSCC cell lines with MMC and assessed their survival in long term growth assays. The HNSCC cell lines showed a broad spectrum of sensitivities to MMC (Figure 1A) with IC_{50} values ranging over 50-fold from 5-250nM (Figure 1B, Supplementary Table 1). MMC-hypersensitivity, in particular if as pronounced as in the FA-patient derived cells, strongly suggests a functional crosslink repair defect in a significant proportion of the cell lines.

To confirm this FA-like phenotype additional functional endpoints were analyzed. We first examined cell cycle progression after treatment with MMC (Figure 1C, Supplementary Figure 1, Supplementary Table 1). Exponentially growing cells were treated with MMC and analyzed by flow-cytometry for cell cycle phase distribution 48hrs after treatment. Both FA-positive controls showed a strong G2-block. Consistent with the prior analysis, the MMC-hypersensitive UT-SCC-12A, UT-SCC-20A, UT-SCC-24B, UT-SCC-45 and UT-SCC-60B cell lines showed a G2-block that was comparable to the FA-positive controls. The IC_{50} for MMC-induced cell killing correlated strongly with induction of a G2-block ($p < 0.0005$) (Supplementary Figure 1B). Proliferation ultimately exposes the cytotoxicity of MMC-induced DNA crosslinks through replication attempts. This could therefore affect drug sensitivity values in a manner that is unrelated to repair efficiency. However, we found no such association between the MMC-induced cytotoxicity or the G2-block and S-phase content (Supplementary Figure 2), further illustrating the value of individual cell doubling time adaptation and choice of long term survival assay. Taken together, both tests show functional defects in cellular DNA crosslink repair in a significant proportion of HNSCC cell lines.

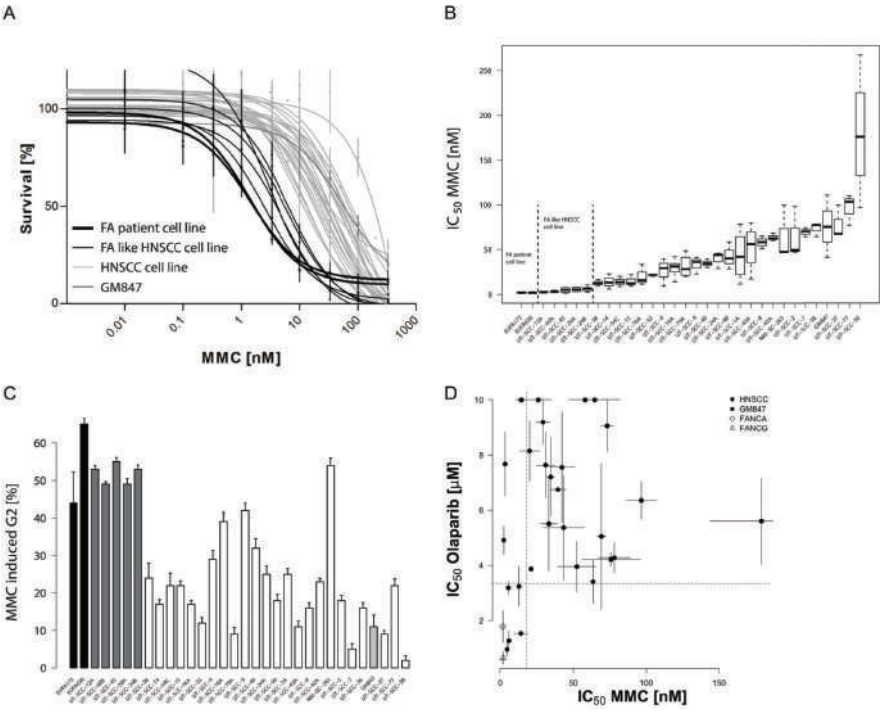


Figure 1. Sensitivity of HNSCC cell lines to mitomycin C and PARP inhibition.
(A) MMC sensitivity as measured by a prolonged growth assay. The average surviving fraction derived from three to five independent experiments per cell line. Errors are SEM. Note, MMC concentrations are log-transformed. A non-linear fit on the log-transformed data is shown. **(B)** Boxplot with MMC IC50 values in the cell line panel. Values are calculated from the curve fits on the individual experiment data and are the average of three to five independent experiments. **(C)** G2/M cell cycle phase arrest 48hrs after 1µM MMC treatment. Cell lines are ranked according to their MMC IC50. MMC-induced G2 values are corrected for the untreated. Errors are SEM. **(D)** Comparison of MMC and olaparib sensitivity in the HNSCC cell line panel. The graph demonstrates the lack of MMC-resistant but olaparib hypersensitive cell lines. Olaparib IC50 values were determined on the individual curve-fits of three to five independent experiments. Errors are SEM.
"FA-like" have been highlighted for presentation and cross-comparison purposes and depicts HNSCC cell lines with MMC IC50 values that are not significantly different from those of the FA-patient cell lines (EUFA173 and EUFA 636) that served as positive controls.

Hypersensitivity to the PARP inhibitor olaparib supports functional DNA repair defects

PARP inhibitors have been shown to reveal FA/HR-defects, by inducing kill in repair-defective cells [5,26]. We therefore tested our HNSCC panel for sensitivity to the PARP inhibitor olaparib (AstraZeneca). Olaparib response varied highly (Supplementary Figure 3). IC_{50} was not reached at the highest tested dose in eight of the cell lines. As reported previously, FA-cells were hypersensitive to olaparib [26] and define the lower limit in this sensitivity range. Several HNSCC cell lines were indistinguishable from these FA-patient derived fibroblasts, strongly indicating functional FA/HR-pathway defects in these lines. No correlation was found between S-phase content or doubling time and olaparib (Supplementary Figure 4A).

We did not find a strong correlation between MMC and olaparib response (Figure 1D, Supplementary Figure 4B). Consistent with the existence of olaparib resistance mechanisms unrelated to or bypassing HR-mediated repair, some MMC-hypersensitive cell lines, the UT-SCC-12A and UT-SCC-60B, did not exhibit a concomitant olaparib hypersensitivity [27]. Notably, all highly olaparib-sensitive cells were also MMC-hypersensitive, further confirming a functional FA/HR-pathway defects in these cell lines.

Aberrant FANCD2 expression and mono-ubiquitylation point to FA-pathway defects

Mono-ubiquitylation of FANCD2 is an essential step in FA-pathway activation upon DNA damage [27,28]. A lack of MMC-induced FANCD2-L demonstrates FA-pathway defects upstream of the ubiquitylation event. We tested the FANCD2-ubiquitylation capacity in 18 HNSCC cell lines and found that anomalies are common. FANCD2 expression varied and was very high in UT-SCC-43A (Figure 2, Supplementary Figure 5). Only three cell lines responded to MMC treatment with an increase in FANCD2-mono-ubiquitylation. Despite a strong MMC cell survival assay response, UT-SCC-45 and UT-SCC-43A lack efficient FANCD2 mono-ubiquitylation at any condition. These data further support DNA damage response irregularities and FA-complex defects leading to disrupted FA-pathway activation in HNSCC.

FA expression analysis reveals lack of FANCF expression in one HNSCC cell line

The observed FA phenotype-like properties of a proportion of the HNSCC prompted us to search for the genetic cause. FANCF down-regulation mediated by promoter methylation has been reported in HNSCC [18]. We therefore determined FANCF expression by PCR and tested whether it was associated with the observed functional defects. No correlation or cut-off analysis supported a role of FANCF expression in defining MMC sensitivity (Figure 3A). However, the FANCF expression level was undetectable in UT-SCC-43A and is consistent with its lack of MMC-induced FANCD2-ubiquitylation, thereby revealing the likely cause of the observed FA-pathway defect

in this cell line. RNA-sequencing analysis in all cell lines confirmed the lack of FANCF expression in UT-SCC-43A but did not reveal additional hits.

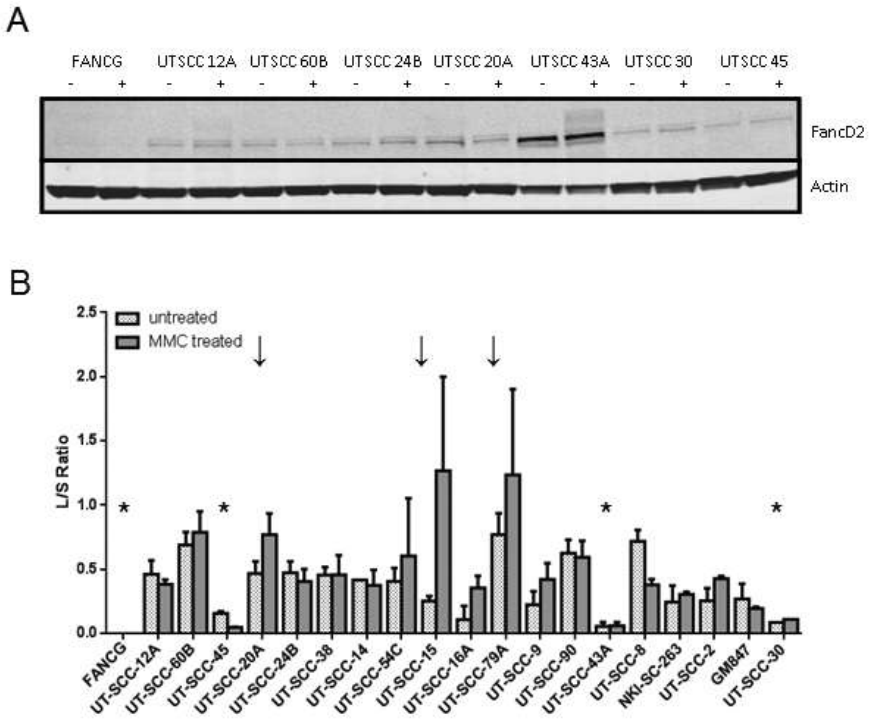


Figure 2. FANCD2-mono-ubiquitylation capacity in the HNSCC panel.
FANCD2-mono-ubiquitylation ability was assessed by exposure to MMC**. (A)** Representative example of FANCD2-ubiquitylation western blot analyses are shown. Lysates were prepared from untreated (-) or MMC-treated cells (+) 6h after treatment. The lack of the upper band indicates a lack of the mono-ubiquitylated form of FANCD2 (FANC2-L) and a defect upstream in the Fanconi pathway. Actin served as a loading control. **(B)** Quantification of MMC-induced FANCD2-mono-ubiquitylation in the HNSCC panel. Quantified FANCD2-L/S values in untreated (dotted bars) and MMC-treated (solid bars) samples of each analyzed HNSCC cell line are shown. HNSCC values are ranked according to their MMC sensitivity. Errors are SEM. Stars (*) indicate examples with an overall lack of FANCD2-ubiquitylation, arrows (↓) in contrast depict HNSCC with a pronounced MMC-induced FANCD2 mono-ubiquitylation as expected by a fully functional pathway.

DNA sequencing identifies FA/HR gene variants with a functional association

To further uncover genetic defects that may explain the above observed crosslink repair defects, we performed capture-based sequencing including 27 canonical FA/HR-pathway genes (Supplementary Table 2). Copy number analysis on the sequencing data did not reveal homozygous deletions of the FA/HR genes. We next called single nucleotide variants and small indels and applied a variant selection protocol that was designed to enrich for functional alterations (Supplementary Table 3). In brief, considering the requirement for loss of heterozygosity (LOH) for most mutated DNA repair genes to affect cellular function, it selects for homozygous non-synonymous variants with a low or lacking minor allele frequency with the aim to enrich for variants and mutations with a potential functional association. Confirming the specificity of this approach, the known *FANCA* and *FANCG* germline mutations in the FA-patient derived fibroblasts were identified and no additional variants were selected. No FA/HR gene variants, as defined by our selection criteria, were found in the GM847 normal fibroblasts.

The analysis revealed seven FA/HR-variants in seven HNSCC lines: three in *BRCA1*, two in *FANCD2* and one in *BRCA2* and *BRIP1* (Figure 3B). With the exception of UT-SCC-30, these variants exclusively occurred in the MMC-hypersensitive cell lines, supporting their potential role in identifying crosslinker sensitivity (Figure 3C). Consistent with an FA/HR defect and the MMC and olaparib hypersensitivity endpoints, the suspected BRCA1 mutation carrier UT-SCC-60B showed impaired radiation-induced rad51 foci formation (Supplementary Figure 6). In contrast, the resistant UT-SCC-30 did not show any apparent HR pathway defects (Supplementary Figure 6), suggesting a false positive assignment by the variant selection protocol. We next used the *in silico* algorithms PolyPhen [29] and SIFT [30] to predict deleteriousness of the variants. Five of these seven variants were predicted to be damaging or deleterious (Supplementary Table 4). While among those variants some are rare or moderately rare SNPs, two of these variants, those in *BRIP1* and *FANCD2*, are unreported in the 1000Genomes database. The location and nature of these variants strongly suggests an impact on protein function. The *BRIP1* variant Gly690Arg is located within the helicase domain and is strongly predicted to affect *BRIP1* protein function (Figure 3B, Supplementary Table 4). The *FANCD2* mutation affects a proline flanking a highly conserved region that encompasses the heterodimer interface, resulting in a deleterious prediction by PolyPhen. The other *FANCD2* missense variant is reported in the 1000 Genomes database and is located within highly conserved regions of the DNA binding domain. It was detected in the FANCD2-mono-ubiquitylation defected UT-SCC-45, indicating a causative link.

LOH is common in tumors of carriers of pathogenic BRCA1/2 germline variants that predispose to breast cancer. This prompted us to investigate LOH in the genes in which the seven variants are located. We therefore assessed the zygosity of all sequenced SNPs in these genes (Supplementary Figure 7). We found that all 19 *BRCA1* SNPs detected in UT-SCC-38 are homozygous, strongly suggesting LOH of the potentially mutated *BRCA1*. Similar evidence of full or partial LOH of *BRIP1* and *BRCA2* is present

in the UT-SCC-12A and UT-SCC-76A respectively. Two out of three *BRCA1* variants in UT-SCC-60B are homozygous and have a minor allele frequency (MAF) below 5%, also pointing to potential LOH events. We were not able to assess LOH events using SNPs in the other three cell lines due to low SNP density, however, copy number data point to LOH in the respective genes in two of them, UT-SCC-15 and UT-SCC-45 Supplementary Figure 7).

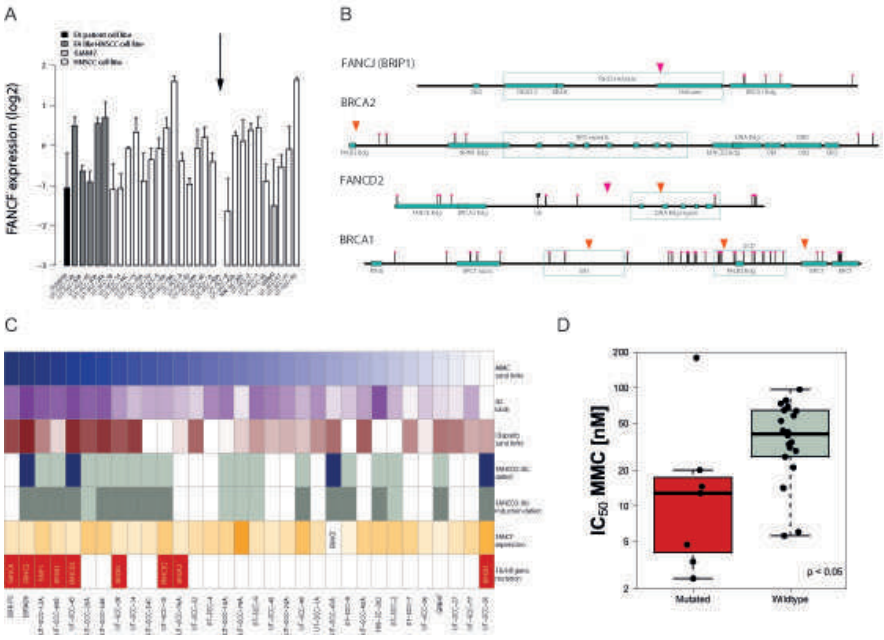


Figure 3. Identification of genetic FA and HR pathway alterations and their association with a functional repair defect.

(A) FANCF expression in the HNSCC cell line panel. HNSCC cell lines are ordered according to their MMC sensitivity. The relative FANCF expression in the individual HNSCC cell lines is shown as a deviation from average (log2-transformed) after normalization to the two housekeeping genes. Arrow (↓) depicts lack of FANCF expression. Errors are SD on the means of 3 to 5 independent PCR reactions. **(B)** Identification of potential FA and HR gene mutations in HNSCC. Homozygous rare sequence variants were found in BRCA1, two in FANCD2 and one in BRIP1 (FANCJ) and BRCA2 in 7 of the 29 HNSCC. Rare SNPs are depicted in orange, unreported non-synonymous variants in pink. **(C)** Comprehensive summary of the HNSCC DNA repair defect data. HNSCC cell lines are ordered according to their ranking in MMC sensitivity (top panel). MMC sensitivity, MMC-induced G2 block and olaparib sensitivity are represented by a color grading with darker colors representing defects in those parameters. Blue bars display defects in FANCD2 mono-ubiquitylation and grey bars represent a lack of induction by MMC (white bars = not determined). FANCF bars are color-ranked according to their expression values. Red bars demonstrate the identification of DNA sequence variants (as shown in B). **(D)** MMC sensitivity of HNSCC with FA/HR gene variants (in red) compared to HNSCC in which such variants could not be found ($p < 0.05$).

Some of the selected FA/HR-variants, due to the selection criteria, may just expose LOH events, rather than being causative. Their incidence in the repair-defected cell lines, however, suggests a functional link and therefore a potential role as repair defect markers. Therefore, we evaluated whether all the variants (including possible false positive, i.e. the UT-SCC-30) were associated with MMC-sensitivity. The 'FA/HR-gene affected' HNSCC cell lines (i.e. variant-positive) were significantly more sensitive than the non-mutated (Figure 3D, $p < 0.05$). Likewise, MMC-hypersensitive were significantly enriched for FA/HR-pathway variants (Supplementary Figure 8A, $p < 0.005$). This association did not originate from a higher mutation load in these cell lines and was specific to the FA/HR-pathway further supporting the bioinformatics selection approach (manuscript in preparation). We analyzed 17 additional crosslink repair and DNA damage response genes that act in the periphery of the FA/HR-pathways and found one additional rare variant in *ATR* (Supplementary Figure 8B,C) in the MMC-sensitive UT-SCC-14, hence improving the MMC-sensitivity association ($p < 0.01$; Supplementary Figure 8D).

Taken together, we identified genetic variants in the FA/HR-pathway that are associated with functional repair defects.

FA/HR gene variants are present in a considerable proportion of HNSCC tumors

Encouraged by the strong association of the FA/HR-variants with the functional crosslink repair defect *in vitro*, we investigated the presence of such variants in a cohort of 77 advanced stage oro- and hypopharyngeal HNSCC tumors from chemo-radiated patients (Table 1). Since associated with function, we applied the variant selection protocol that was used for the cell lines, while correcting for stromal contribution in order to enrich for homozygous variants. 19% (15/77) of the HNSCC, a fraction similar to the cell line panel (24%), possessed such FA/HR-variants (herein termed 'FA/HR-affected' tumors). Figure 4A and Supplementary Table 5 summarize and list these variants, their predicted alterations and distribution over the patient cohort. Half of these variants were predicted to be deleterious or damaging by PolyPhen and SIFT or were referenced in COSMIC. We further found a higher prevalence of rs17885240 in our patient cohort than in the general population (4/77 versus MAF = 0.0129; $p < 0.05$). Others have shown a high prevalence of rs17885240 in childhood AML (Supplementary Table 5).

FA/HR gene variants with in vitro functional association are associated with poor outcome in HNSCC

Due to the functional repair defect association of the FA/HR gene variants *in vitro*, we next questioned whether the identified patients had a different prognosis. The known and reported HNSCC prognostic variables tumor site, tumor volume and HPV-status affect outcome in our patient cohort (Supplementary Figure 9A-F) [12]. We next compared the clinical outcome of patients with 'FA/HR-affected' tumors to the others,

while adjusting for these known prognostic variables in a Cox proportional hazards model. We find that the overall survival (OS) of patients with 'FA/HR-affected' tumors is lower (multivariate HR 2.6; 95% CI 1.1-6.0; $p < 0.05$) (Figure 4B, Supplementary Table 6). Locoregional control was worse in these patients (univariate HR 2.1; 95% CI 0.5-8.1; $p = 0.21$ and multivariate HR 3.5; 95% CI 0.8-15.2; $p = 0.09$) (Supplementary Figure 10), but this did not reach significance. This was likely due to the small number of locoregional events (n = 10) and also prevents meaningful conclusions from multivariate Cox model analyses that aim to account for the three major clinical variables. There was no apparent association between HPV-status and the presence of FA/HR-variants (Supplementary Table 7). Standard somatic FA/HR-variant selection did not reveal any outcome association, further highlighting the value of FA/HR-variant selection criteria and the functional *in vitro* analysis (Supplementary Figure 11).

Next, we evaluated indirect or FA/HR unrelated elements that may have driven the observed association of the FA/HR-affected tumors with poor patient outcome. Mutation load was not of influence, since these tumors have similar variant and mutation frequencies to the other tumors (Supplementary Figure 12). 529 additional HNSCC and cancer related genes were sequenced in these tumors. To confirm specificity to the FA/HR pathways, we then assessed how likely the HR of 2.6 that was found in the 'FA/HR-affected', is, by applying the FA/HR gene variant selection protocol to 10.000 random gene sets of similar total base coverage (Figure 4C). Only 304/10.000 random gene sets have a HR > 2.6, confirming the significance of the association between HR/FA-variants and the poor outcome. Finally, our variant selection approach may have marked LOH of larger chromosomal segments associated with poor outcome. We therefore repeated our analysis on other genes in the same chromosomal location and did not find a significant association with overall survival or locoregional control (Figure 4D, Supplementary Figure 10B, Supplementary Table 6 and 8) further confirming specificity to the FA/HR-pathway genes.

Not all patients received equally high cisplatin doses, either due to discontinuation or a planned low dose cisplatin chemo-radiation scheme. True (functional) FA/HR-pathway disruption would imply an increased tumor sensitivity to crosslinking agents. We therefore further grouped patients according to their received cumulative cisplatin dose (Figure 4E). In line with the prediction, only FA/HR variants harboring tumors responded to a change in cumulative dose and our data show a benefit from high cumulative cisplatin doses for these patients while failing to show this in the non-FA/HR affected. In summary, we find that FA/HR-pathway variants and mutations mark FA/HR repair defects *in vitro*. Identifying such variants in HNSCC patients allowed us to reveal a worse prognosis group, suggesting that such tumor repair defects have an impact on patient outcome.

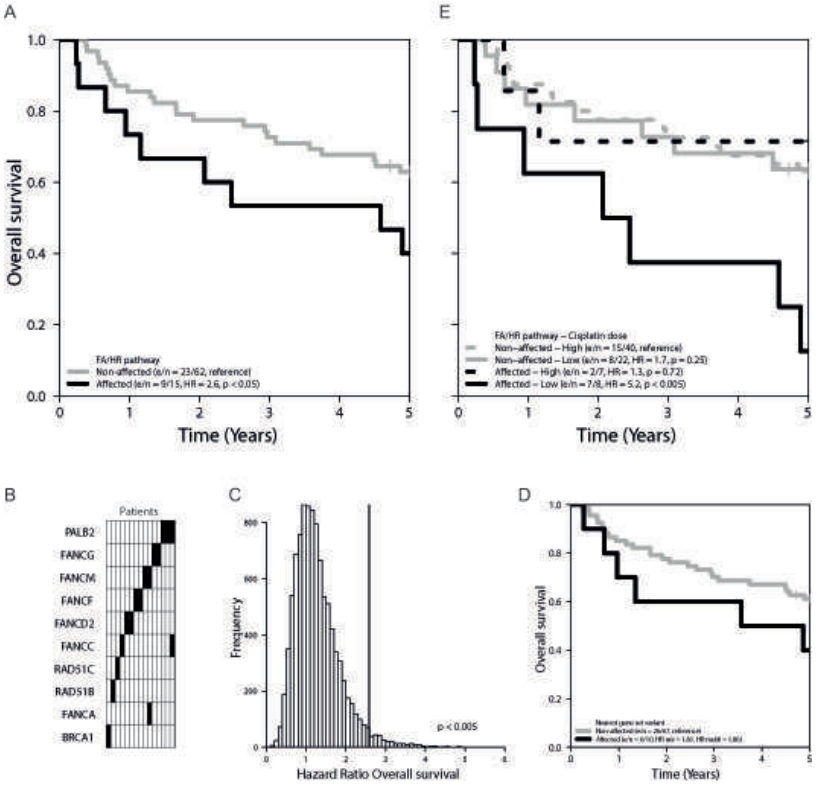


Figure 4. FA/HR gene variants in tumors of HNSCC patients and their prognostic value.

(A) Nineteen FA/HR gene variants were found in fifteen patient samples (columns), resulting in a 19.5% incidence rate (Supplementary Table 5). **(B)** Oro- and hypopharyngeal tumor samples of seventy-seven chemoradiated HNSCC patients were analysed for variants in canonical FA/HR genes by applying the variant selection criteria that returned a functional DNA repair defect association *in vitro*. HNSCC patients with FA/HR gene variants in the tumors (= FA/HR-affected) had a worse overall survival (OS) ($p < 0.05$ in multivariate analysis), demonstrating the impact of potential functional FA/HR repair defects in this patient population. **(C)** Distribution of OS hazard ratios, obtained by repeating the analysis of Figure 4A on ten-thousand randomly selected genes sets of similar base coverage, highlights the significance and specificity of the FA/HR gene variants HR finding. Gene sets were sampled from 529 sequenced cancer-related genes. **(D)** Kaplan-Meier graph showing OS of patients with variants in genes in FA/HR gene chromosomal locations. No significant association was found. **(E)** OS is worst in patients with functionally associating FA/HR gene variants and low cumulative cisplatin dose (HR 5.2, $p < 0.005$). In-figure: multivariate hazard ratios (HR) with confidence intervals (95% CI) from multivariate Cox proportional hazard models that include the tested gene set, tumor site, HPV-status and tumor volume.

Discussion

Prompted by the relevance of genetic HR/FA repair defects in breast cancer and the high incidence of HNSCC in FA-patients [7,8], we investigated the role of such defects in sporadic HNSCC. As similar genetic markers were missing, we first determined the incidence and properties of functional DNA repair defects by performing multiple functional assays on 29 HNSCC cell lines. We find that a significant proportion exhibit DNA crosslink repair defects comparable to FA-patient derived fibroblasts, as determined by the functional endpoints MMC-hypersensitivity, G2-blocks, olaparib-hypersensitivity and FANCD2 mono-ubiquitylation. Our comprehensive large cell line panel data therefore shows the high prevalence of functional DNA crosslink repair defects in HNSCC in vitro. We next determined possible genetic causes with the intention to translate these findings to the clinic and find multiple FA/HR gene variants that are associated with the functional outcome parameters. When retrospectively determining FA/HR gene variants with such an *in vitro* confirmed functional relevance in clinical samples, we find an incidence of 19%. Importantly, patients with tumors that harbor such genetic markers, do worse. No other genetic pathway analysis was able to depict this worse prognosis patient population, thereby further supporting the important role of such potential repair defects in these cases.

Other *in vitro* studies indicated crosslink repair defects in HNSCC. A large variation in response to the crosslinker cisplatin in ten cell lines tested by Snyder et al [31] supports our conclusion. We find at least two HNSCC cell lines that are incapable of mono-ubiquitylating FANCD2 (UT-SCC-45 and UT-SCC-43A) indicating a defect upstream or in FANCD2 itself. In both cell lines this was consistent with the genetic defect. This low incidence is consistent with other smaller studies and explains the lack of disrupted FANCD2 mono-ubiquitylation or FA-gene expression deregulation in the study of Snyder et al. Also Burkitt [32] observed defective FANCD2-foci formation in three cisplatin sensitive HNSCC. The authors found decreased BRCA1 expression and showed this caused cisplatin sensitivity in one cell line. Furthermore, MMC and cisplatin-responsive HNSCC cell lines have been found by a chromosomal breakage assay [22]. No other study has comprehensively tested for both functional and genetic FA/HR-defects in such a large panel of HNSCC cell lines. Some studies sought after genetic defects in functional affected only, others employed endpoints, such as FANCD2-ubiquitylation, that do however not capture repair defects downstream of the event.

The discovery of frequent FA/HR defects may have also been impeded by weak IC50 values that were determined by short term metabolic assays in other studies. Our HNSCC cell line panel exhibited a large range of doubling times (17h to 79h, Supplementary Table 1) providing sensitivity and specificity issues in standard short term (72h) survival assays. Our long term and doubling time adjusted survival assay was able to uncover MMC sensitivity in slow HNSCC with DTs of over 48h. Yet, other HR/FA unrelated factors can affect some of the endpoints on an individual basis and therefore demands multiple endpoint analysis. The response to MMC partly

depends on metabolic activation by cellular reductases such as the NAD(P)H:Quinone oxireductase-1 (NQO1) and the extent of glutathione detoxification [33,34]. However, these processes rarely affect the response by an enhancement factor of more than three [35–37]. It should be noted that such factors could also affect the apparent FANCD2-ubiquitylation response at a given MMC concentration. Rad51 foci induction by radiation, however, reflects the ability to engage HR [38]. Consistent with its resistance to MMC and olaparib, this response therefore excludes a BRCA defect in UTSCC-30 (Supplementary Figure 6). Apparent olaparib resistance is possible in the presence of FA/HR defects. This can be the result of an increased expression of the drug efflux transporters *Abcb1a/b* genes, which encode the P-glycoprotein [27]. Moreover, recent studies show a role for p53-binding protein 1 (53BP1) and REV7 for olaparib resistance in BRCA1 deficient cells [39,40]. Notably and consistent with our observation, the authors report olaparib resistant tumors which maintained crosslinker cisplatin sensitivity [40]. While FA/HR-defects can therefore not be excluded in olaparib resistant HNSCC or in HNSCC with somewhat higher MMC IC50 values, the manifestation of the combined olaparib and MMC hypersensitivity is a strong indicator of functional crosslink repair defects. By our multiple endpoint analysis we were able to unmask multiple FA/HR defects. At least a quarter of HNSCC exhibited strong FA/HR defects as determined by multiple endpoints. Any cut-off for classification would be however arbitrary, clearly evident from the distribution shown in Figure 1B or Figure 3C and many more HNSCC show signs of repair defects.

Aided by the quantity of different cell lines and endpoint analyses in our study, we were also able to define genetic and repair defect associations ultimately providing a variant selection protocol that was designed to enrich for variants with a potential functional association. The selection of homozygous variants and rare SNPs are noteworthy elements of our variant selection protocol. First, we reasoned that functional DNA repair defects will generally be enabled or caused by the loss of the functional allele and therefore only selected homozygous variants. LOH of three FA genes has recently been reported in sporadic HNSCC (33). As shown here, our approach can also capture and enrich for potentially relevant LOH events in the case of rare SNPs. Thus we acknowledge that these particular variants may not have a direct impact on the gene function. A related challenge in evaluating the consequence of individual missense variants was the absence of matched normal samples. Selection of homozygous variants does also not exclude an effect on repair of potential heterozygous compound mutations which are difficult to evaluate. Second, and different from other studies that discard SNPs, we included SNPs with a MAF < 2.5%. This allows for increased LOH event detection and is also based on our hypothesis that a fraction of HNSCC patients may bear hypomorphic FA/HR germline variants [41] that are pathogenic only when exposed to high levels of crosslinkers. HNSCC patients are often heavy drinkers and smokers and exposed to DNA crosslinking chemicals in both. From demographic data we estimated that 5-10% of all heavy alcohol and tobacco users will develop HNSCC [42,43]. In breast cancer, a single pathogenic BRCA1/2 variant can reach a penetrance of more than 50% [44]. Hence, we estimate that if

HNSCC would be promoted by few SNPs, these SNPs could be present in up to 2.5-5% of the general population, thus used this value as cut-off. These estimates highlight how SNPs with such a relative high MAF could contribute to a HNSCC predisposition but remain masked as they would not harm the majority of carriers with little exposure to the carcinogens. Due to the heavy exposure to crosslinking agents we also expect that the type and nature of "pathogenic" FA/HR variants is largely different from those causing breast cancer or Fanconi Anemia. Indeed, we find FA/HR gene variants that have not been classified as pathogenic in hereditary breast cancer or Fanconi Anemia. Yet, some show indications of a potential disruptive nature (FANCD2 P834A, RAD51C G264S). Other variants may have simply marked LOH events in these cells and patients. Importantly, the identified variants by the chosen selection criteria were strongly associated with MMC-sensitivity, thus providing a functional link.

Homozygous variant detection is challenging in patient tumor samples due to unknown tumor sample purity, ploidy and intra-tumor heterogeneity [45]. Corrections for stromal contributions were made, using the pathologist's tumor purity estimate. Yet, in tumor samples with high stroma contribution of 40-50%, concessions had to be made ultimately allowing for a higher degree of false positives with regards to the homozygous status of the selected variants. However, when analyzing genetic data of 77 HNSCC tumors, and classifying patients according to tumor FA/HR-variant presence as FA/HR-affected we found that patients with such tumors had a worse prognosis. This worse outcome is analogous to reports in breast cancer studies, in which confirmed pathogenic BRCA1/2 mutations were associated with a more malignant phenotype and a worse prognosis [46,47]. Chromosomal and genetic instability promoted by the DNA repair defect and the associated tumor heterogeneity could be the driving force for the malignant phenotype and the apparent treatment failure in such patients. Since such tumors are hypersensitive to crosslinking agents *in vitro*, future studies will have to determine whether this specific patient group do benefit from crosslinker-based treatment. HNSCC patients benefit from high cumulative cisplatin doses [48]. Our preliminary analysis on our initial cohort data shows that this benefit may very well be based on tumor DNA repair defects (Figure 4E). Although belonging to a poor prognosis group, the patients with FA/HR variants containing tumors were the only ones who appeared to benefit from cisplatin dose intensification (high dose vs low dose) thereby establishing survival rates as high as in the other patients.

The accomplishment of personalized medicine requires the discovery and identification, by functional and genomic approaches, of processes that could define treatment options. We assessed, quantified and characterized the FA/HR DNA repair pathway defects and probed their clinical relevance. FA/HR genomic variant selection that was supported by the functional *in vitro* analysis helped to reveal an association with poor survival in a cohort of chemo-radiated HNSCC patients and thereby points to the prognostic value of DNA repair defect identification in cancer. HNSCC has a dismal prognosis, particularly when diagnosed at an advanced state. Our data stress the relevance of repair defects in establishing such bad prognosis and reveal clinical

treatment options at the same time since these defects are associated with the benefit of high cumulative doses of crosslinking agents.

Together, our data suggest a novel role of FA/HR-pathway aberrations in both sporadic HNSCC etiology and prognosis. To our knowledge this is also the first report that shows how a comprehensive *in vitro* functional and genetic DNA sequencing-based pathway analysis can reveal or enrich for functional and therefore relevant genetic alterations, thereby educating clinical marker selection strategies for biomarker development. The study enabled us to depict a subpopulation of patients with a bad prognosis that might be associated with genomic instability features through the depicted DNA repair defects. In the context of precision medicine, these defects can be exploited for tumor-targeted therapy options (e.g. PARP inhibitors) or personalized cancer treatment with crosslinking agents [6,49–51].

Material and Methods

Cell lines and cell culture

Head and neck cancer cell lines were established at the University Hospital in Turku (UT-SCC; listed in Supplementary Table 1), Finland and at the Netherlands Cancer Institute (NKI-SCC-263) [52–54]. 23 HNSCC are TP53 mutated and only UT-SCC-45 is confirmed TP53 wildtype. Cells were grown and assayed under low oxygen (4%) conditions and were cultured under standard culture condition at 37°C and 5% CO_2 in DMEM supplemented with 10% fetal bovine serum, 1% penicillin/streptomycin and 1% non-essential amino acids (Invitrogen).

Patients and clinical cohort

Pre-treatment tumor samples were obtained by biopsy from patients enrolled in our hospital from 2001 to 2010 and after documented informed consent. Patient and tumor characteristics are listed in Table 1. All patients were treated with concurrent cisplatin-based chemoradiotherapy (70Gy in 35 fractions). Different cisplatin regimens were administered: daily (6mg/m^2 intravenously), three-weekly (100mg/m^2 intravenously), or weekly for the first 4 weeks of radiotherapy (150mg/m^2 intra-arterial). Final cumulative cisplatin doses were recorded and patients were categorized into low (<300mg/m^2) and high (≥300mg/m^2) dose categories (Table 1). Tumor volumes were assessed on RT treatment planning CTs. Two categories (0-30 and >30cc) were established and considered in multivariate analyses. Time to locoregional recurrence was calculated from the start of treatment until the first of the following events: local or regional recurrence (event), death or last follow-up (censored). Overall survival time was calculated until death (event) or until the last follow-up (censored). HPV status was determined by capture-based sequencing and validated using p16 and p53 IHC and PCR on suspected positive cases.

Table 1: Demographics of HNSCC patient cohort.

Patient characteristics		N (%)
Gender	M	55 (71)
	F	22 (29)
Primary site	oropharynx	49 (64)
	hypopharynx	28 (36)
T-stage	T1	1 (1)
	T2	13 (17)
	T3	35 (46)
	T4	28 (36)
N-stage	N0	10 (13)
	N1	7 (9)
	N2	52 (68)
	N3	8 (10)
T-volumes	0-30 cc	38 (49)
	>30	39 (51)
Events	Death	32 (42)
	Locoregional Recurrence	10 (13)
HPV	positive	21 (27)
	negative	56 (73)
Smoker	current	46 (60)
	former	19 (25)
	never	4 (5)
	unknown	8 (10)
Alcohol consumption	yes	47 (61)
	former-alcoholic	13 (17)
	never	8 (10)
	unknown	9 (12)
Cisplatin regimen	daily (6mg/m^2, 5 weeks)	17 (23)
	3-weekly (100mg/m^2, 3x)	46 (59)
	weekly (150 mg/m^2, 4x)	14 (18)
Cumulative cisplatin dose	low (< 300 mg/m^2)	30 (39)
	high (≥ 300 mg/m^2)	47 (61)
		time
Median age	at diagnosis	58 years (SD=9,6)
Median survival	Overall survival	63 months (SD=39)
	Locoregional control	63 months (SD=41)

Pretreatment biopsy material from HNSCC tumors of 77 patients was sequenced and tested for functional repair associated FA/HR-variants, as determined in the cell line panel. All patients received concurrent cisplatin-based chemoradiotherapy, with some patients reaching a high cumulative cisplatin dose of ≥ 300 mg/m2.

G2 block analysis

Cells were cultured in 6-well plates for three days and treated with the indicated doses of MMC (Sigma-Aldrich) for 2 hours. After MMC treatment, cells were washed with PBS and incubated with fresh non-drug-containing medium. 48hrs later cells were prepared for flow-cytometry analysis (FAC Scan, Becton Dickinson, San Jose California USA) and re-suspended in PBS containing PI (propidium iodide, 10µg/ml) and RNAse (0,02mg/ml). Cell cycle phase distributions were measured on the PI histograms using the software CellQuest (Becton Dickinson).

MMC sensitivity analysis

Cultures were exposed to different concentrations of MMC (Sigma-Aldrich) after 1 day of culture. Cultures grew until the untreated cells had undergone at least five population doublings. Sub-confluent status was tested with internal linearity controls. Live cells from 3-6 sub-confluent wells were counted with a CASY cell counter (Schärfe system, Scotch plains, New Jersey). Survival is the fraction of the number of treated to untreated cells in %. IC_{50} values (MMC concentration with 50% growth inhibition) were calculated from third-order polynomial curve-fits on the growth inhibition values of the individual experiments.

PARP inhibitor sensitivity analysis

PARP inhibitor olaparib (formerly AZD2281/KU-0059436; provided by AstraZeneca) sensitivity was determined by a long term growth assay using the CyQuant-Cell-Proliferation Assay Kit (Invitrogen) that measures DNA content after cell lysis. Cells were cultured for one doubling-time before exposure to various doses of olaparib while adjusting DMSO concentration in all. Cells were then cultured for a period of at least 5 population doublings. Internal plate controls test linearity and assured sub-confluent status within each experiment. Survival was determined as the fraction of the value in the treated samples compared to the untreated using a TECAN-infinite fluorescence plate reader. IC_{50} values were determined from polynomial curve-fits on the data.

FANCD2 mono-ubiquitylation analysis

Cells were treated for six hours with or without 100nM MMC two days after plating and harvested for lysate preparation. Proteins were extracted, resolved by SDS-PAGE using a gradient gel (4-15%; Bio-Rad Laboratories) and detected by anti-FANCD2 (1:500 (FI17) Santa Cruz Biotechnology, Santa Cruz) or anti-β-actin IgG (Sigma, 1:50.000), IRDYE 680/800 conjugated anti-mouse or anti-rabbit IgG (Licor, 1:7500). Quantitative analysis of FANCD2-L (long and mono-ubiquitylated) and FANCD2-S (short) of two to three independent experiments on multiple blots each was done on the LICOR platform (Biosciences) and expressed as L/S ratio.

FANCF expression analysis

RNA from exponentially growing cells was isolated according to standard protocols. FANCF RT-PCR analysis was performed in the 7500 Fast Real-Time PCR system (Applied Biosystems) using the primers listed in Supplementary Table 9. Samples were assayed in triplicate in at least 3 independent RT-PCR reactions and runs. FANCF expression values were normalized to two housekeeping genes. The 7500 Fast system SDS software and the 2-(ddCt) method was used for data analysis. FANCF expression in UT-SCC-43 was undetectable in all independent samples despite high RNA quality values (RIN10) and average housekeeping gene expression. The lower limit for detection was assigned to this line for calculation purposes.

HNSCC material and DNA isolation

Cellular DNA of the HNSCC cell lines and from fresh frozen tumor material was isolated using the Qiagen AllPrep DNA/RNA Mini Kit. Only material with an average tumor content of 50% and higher, as determined by a pathologist on H&E sections adjacent to and in the midst of the sections collected for DNA sequencing, was included. Researchers and bioinformaticians were blinded to patient information and outcome data. Clinicians were blinded to sequencing data. Clinical variables data and anonymous outcome data were applied after sequencing and variant selection.

DNA capture and sequencing

Paired-end (PE) fragment libraries were prepared using a genomic DNA library preparation kit (Illumina). The libraries were hybridized to a SureSelect custom-based bait library (Agilent) designed to capture exonic regions (with a 50bp extension on both sides). After washing, the captured DNA was amplified. Enriched libraries were barcoded, pooled and sequenced on a GAII (Illumina Hiseq-2000) using a 2x75bp PE protocol. Sequencing reads were aligned to the GRCh37.55 Ensembl human reference genome using the Burrows-Wheeler Aligner 0.5.10 backtrack algorithm. Potential PCR duplicates were removed using picard-tools MarkDuplicates. An average read depth of 247 in the cell line samples and 255 in the tumor samples was achieved. Copy numbers were inferred from the DNA-seq data using PropSeg on the cell line panel and CNVkit in the tumor cohort [55,56].

Variant and mutation calling

Variants, single nucleotide variants (SNVs) and indels, were called with VarScan 2.3.9 using Samtools mpileup 0.1.19 [57]. Next, we annotated these variants with the RefSeq and 1000 Genomes august 2015 databases using Annovar version date 11-05-2016 [58]. CADD [59], PolyPhen [29], REVEL [60] and SIFT [30] were used for *in silico* variant effect predictions. Supplementary Table 3 describes the filtering steps of the variant selection protocol that was designed to enrich for variants with a functional impact. In brief, variant selection considered variant allele frequency (VAF > 0.8) and rare SNPs (MAF < 2.5%). While maintaining VAF criteria for selection, VAF values in the tumor

samples were adjusted to correct for the stromal contribution. Taking into account the pathologist's tumor fraction assessment (TF), the VAF values from the sequencing data were corrected as follows to compute a VAF value for the tumor (VAF_T):

$$VAF_T = \frac{VAF - VAF_S \times (1 - TF)}{TF}$$

Matched normal tissue blood samples were largely unavailable. Stromal VAF (VAF_S) was therefore set to 0.2, thereby providing a minimal VAF value for heterozygosity.

Statistics

All analyses were performed in the R environment for statistical computing. The MMC IC_{50} values of FA/HR-variant positive versus negative cell lines were compared using a Mann-Whitney U test. A two-by-two Table holding the FA/HR-variant status of the third most MMC sensitive cell lines (n = 10) versus the other cell lines (n = 19) was analyzed with Fisher's exact test. Overall survival, locoregional control and Kaplan-Meier estimators were calculated. Multivariate Cox proportional hazard models included the tested gene set of interest and the clinical covariates tumor site, HPV-status and tumor volume. These clinical covariates have prognostic significance in HNSCC and reached significance in univariate analyses in our cohort. Further statistical analyses and values are accordingly specified in the Supplementary data.

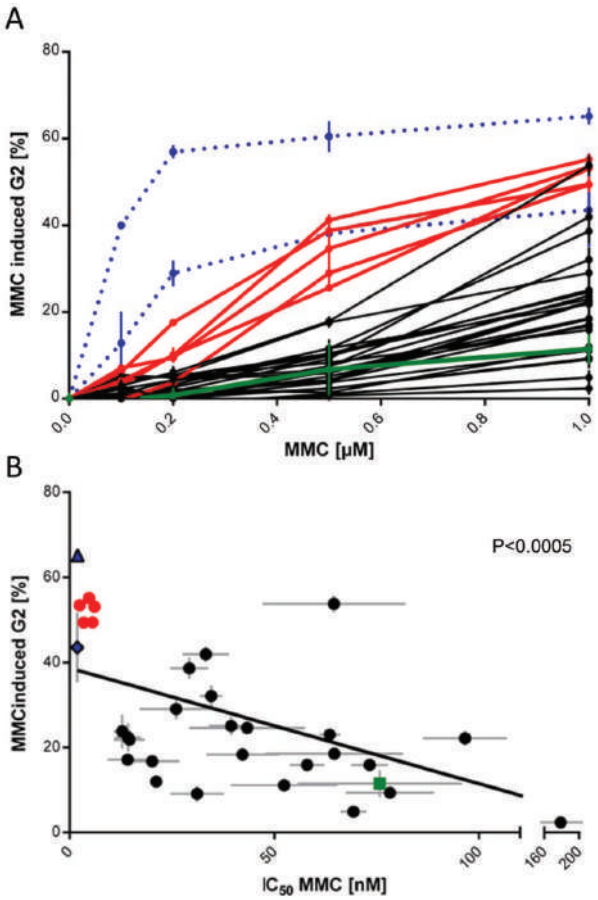
Abbreviations

BRCA1/2: Breast Cancer Type 1 Susceptibility Protein 1 or 2; CT: computed tomography scans; FA: Fanconi Anemia; H&E: hematoxylin and eosin; HNSCC: Head and Neck squamous cell carcinoma; HPV: Human papillomavirus; HR: hazard ratio; HR: Homologous recombination; IC50: half maximal inhibitory concentration; IHC: Immunohistochemistry; MAF: Minor allele frequency; MMC: Mitomycin C; OS: Overall Survival; PARP: Poly(ADP-Ribose) Polymerase; PE: paired-end; RT: radiotherapy; SDS-PAGE: Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis; SNV: single nucleotide variant; TF: tumor fraction; UT-SCC: University Turku - Squamous Cell Carcinoma; VAF: variant allele frequency;

Acknowledgements

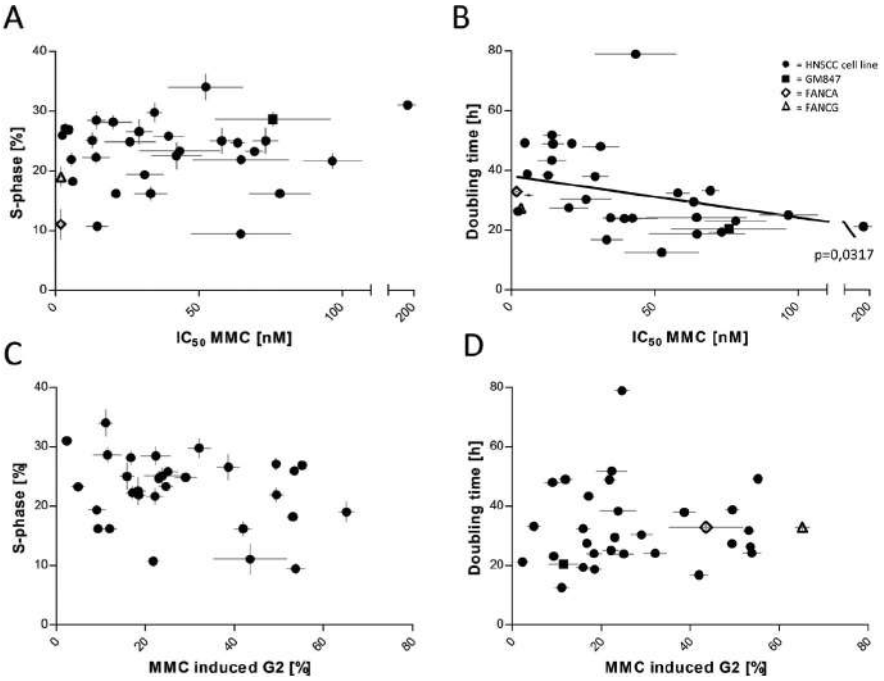
We thank Michael Hauptman and Katarzyna Jozwiak for statistical advice. A. Pfauth and F. van Diepen from the NKI flow-cytometry facility for their technical assistance and Monique de Jong and Magali Michaut for their support in expression and genomic

analysis. Olaparib was provided by M.J.O and AstraZeneca. We thank Hein te Riele and Ruud Brakenhoff for critical reading of the manuscript. Our dear colleague Prof. Dr. Adrian Begg had an integral role in the initiation and early execution of this study. He sadly passed away in 2014. We would like to dedicate this work to him.

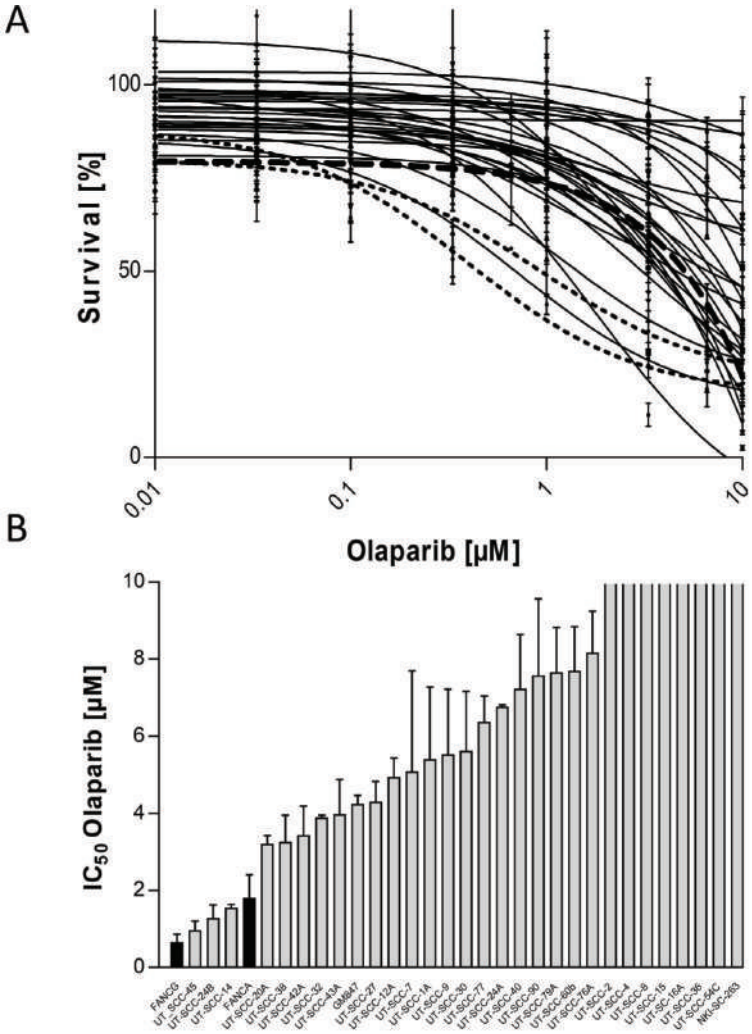


Supplementary Figure 1. MMC-induced G2/M cell cycle phase arrest in HNSCC cell lines and G2 block association with MMC survival.

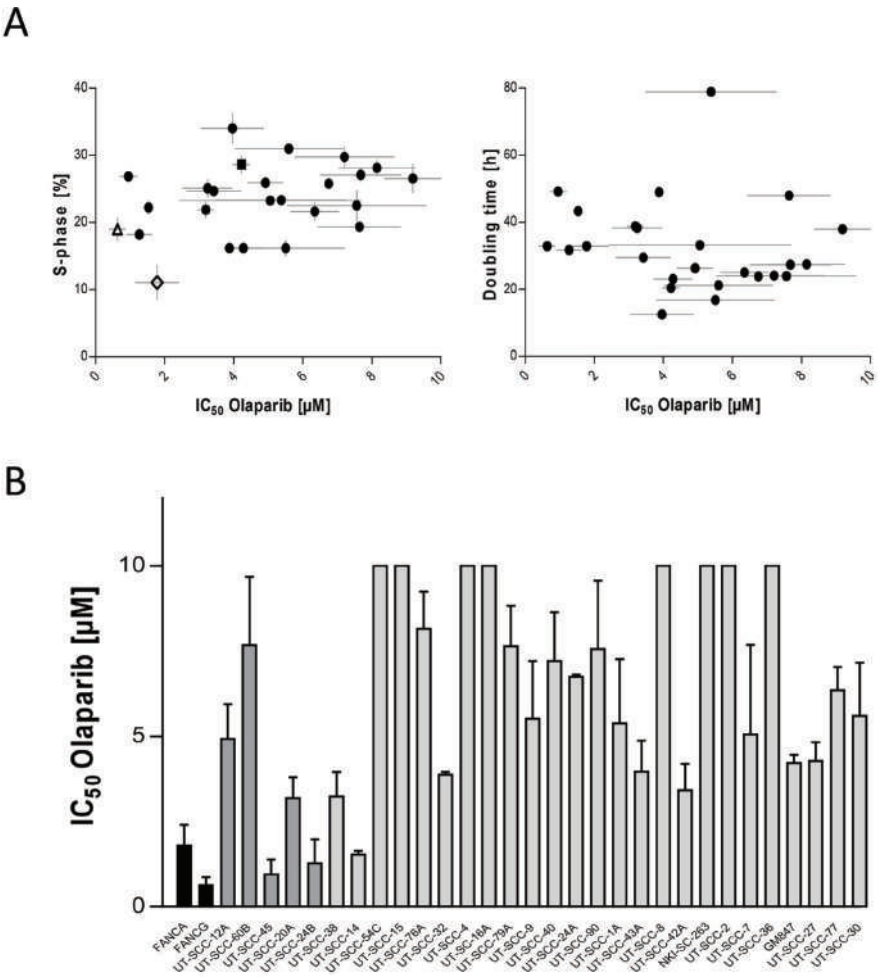
(A) G2 phase cell values in % 48hrs after MMC treatment at indicated MMC concentrations. MMC-induced G2 block values are corrected for the untreated cells. Errors are SEM. **(B)** Correlation of MMC sensitivity and MMC-induced G2 block within the HNSCC panel. The IC50 values of the MMC sensitivity correlated significantly with the MMC-induced G2 block (percentage of cells in G2 after 1 µM MMC) with $p < 0.0005$. Initial experiments at 24, 48 and 72 h defined the optimal time point for MMC-induced G2 block analysis in these cell lines. Considering the doubling times, 48 h was chosen for comparison throughout the cell line panel. Lack of G2 blocks in cell lines with considerable long doubling times has been confirmed by analyzing later time points. Untreated, the cell lines varied in the G2 cell cycle phase fraction from 10 to 20%. Two Fanconi Anemia patient-derived fibroblast cell lines with mutations in FANCA and FANCG (EUFA173 and EUFA 636, in blue) serve as positive controls and are compared to a human fibroblast cell line (GM847, green). HNSCC cell lines defined as most FA-like as described in the manuscript are depicted in red.



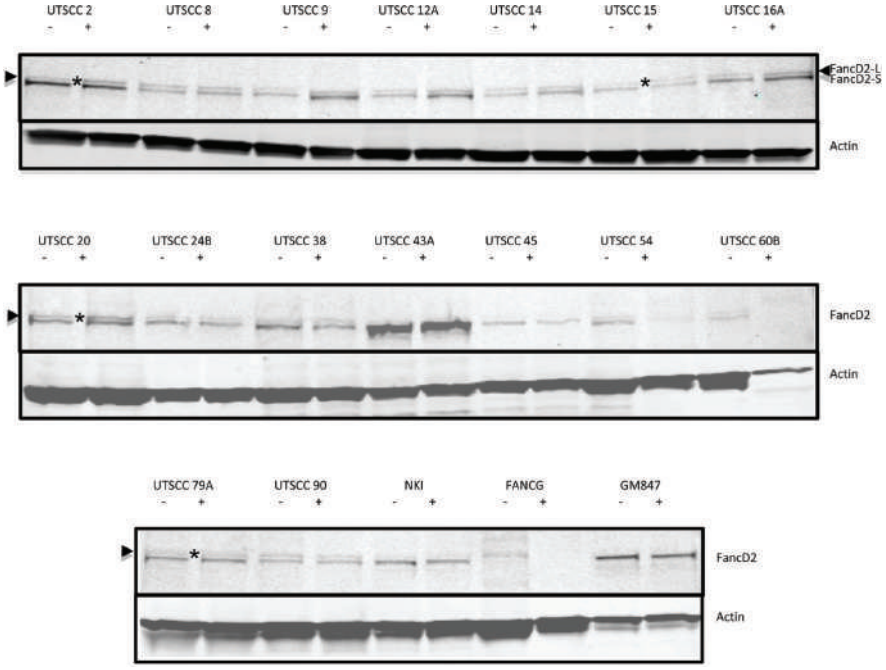
Supplementary Figure 2. Association of MMC response with cell line characteristics. Comparison of MMC sensitivity and MMC-induced G2 block to S-phase content in the cell lines (**A** and **C**, respectively) or to the doubling times (**B** and **D**, respectively). MMC-induced G2 at 1 μM, MMC IC50 and olaparib IC50 values were compared to S-phase content and doubling times in the HNSCC (filled symbols) and assessed for correlation. The controls (open symbols) have been excluded for regression and significance analyses. Error bars are SEM.



Supplementary Figure 3. Sensitivity of HNSCC cell lines to PARP inhibition by olaparib. **(A)** Olaparib sensitivity in the HNSCC cell line panel as determined by a prolonged growth assay. The average surviving fraction (in %) of three to five independent experiments per cell line is shownas a dose response to the PARP inhibitor olaparib. Olaparib concentrations are log-transformed and nonlinear fits on the log-transformed survival data are shown; errors are SEM. The FA control cell lines (EUFA173 and EUFA 636) are depicted by the dotted lines. **(B)** Olaparib IC_{50} values of the individual 29 HNSCC cell lines and FA controls as in A. The average IC_{50} values of each cell line are shown and were determined on the curve fi ts of the individual experiments; errors are SEM. The highest olaparib concentration that was tested is 10 µM and this value was assigned to olaparib insensitive cell lines in which less than 30% kill was achieved by this dose. FA reference controls are depicted in black.

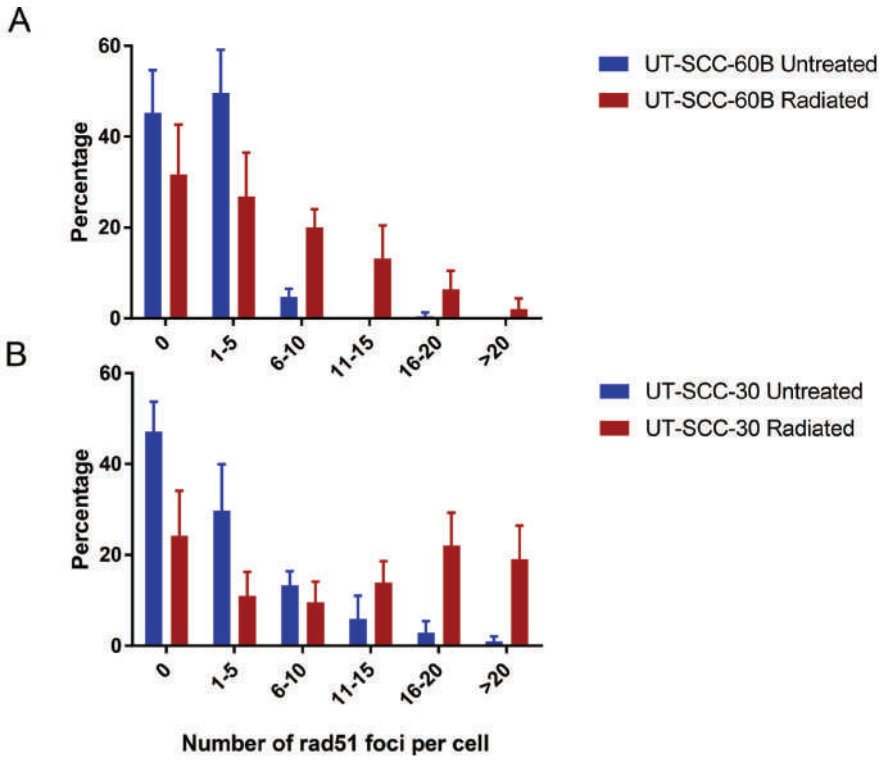


Supplementary Figure 4. Sensitivity of HNSCC cell lines to PARP inhibition by olaparib. **(A)** Comparison of PARP inhibitor olaparib sensitivity to S-phase content in the cell line or to the doubling times as indicated. Olaparib IC_{50} values were compared to S-phase content and doubling times in the HNSCC (filled symbols) and assessed for correlation. The controls (open symbols) and HNSCC cell lines with undetermined IC_{50} values have been excluded for regression and significance analyses. Errors are SEM. **(B)** Olaparib sensitivity of the 29 HNSCC cell lines and controls. The average IC_{50} values of three to five independent experiments are shown for each cell line and were determined on the curve fits of the individual experiments; errors are SEM. HNSCC are ranked according to their MMC sensitivity. The highest olaparib concentration that was tested is 10µM and this value was assigned to olaparib insensitive cell lines.



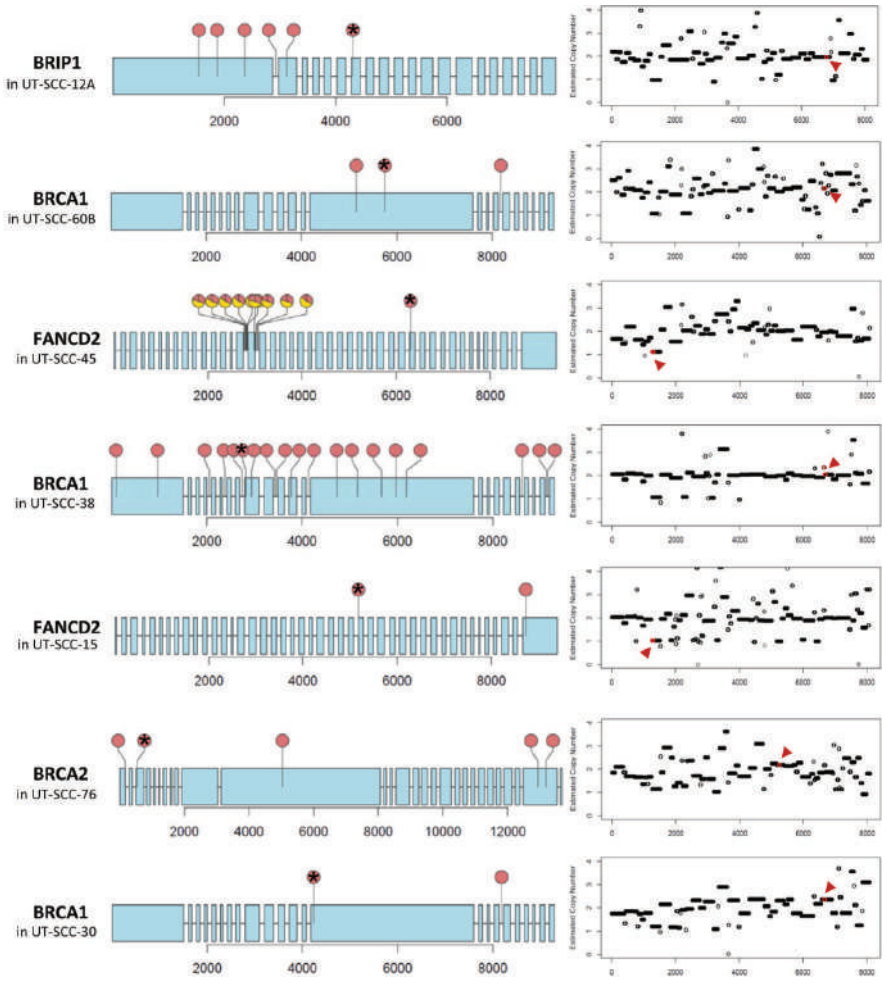
Supplementary Figure 5. FANCD2-monoubiquitylation.

Additional representative examples of MMC induced FANCD2-monoubiquitylation in the analyzed HNSCC cell lines. Western images on untreated and MMC treated (36 nM) cell lysates, used for FANCD2-L(ong), indicated by a black arrow head, and FANCD2-S(hort), indicated by a grey arrow head, determinations. Stars (*) indicate positive controls with examples of FANCD2-monoubiquitylation induction by MMC exposure. This is seen by an increase in the intensity in the upper band (FANCD2-L) after MMC treatment. Other HNSCC display signs of FANCD2-monoubiquitylation, even in untreated conditions, and shows FA activation by endogenous sources (replication stress) and is typical for tumor cells. Other HNSCC lack FANCD2-monoubiquitylation (FANCD2-L band). If however sensitive to MMC, one can assume efficient drug activation and this therefore shows the inability to monoubiquitylate FANCD2, such as in the UT-SCC-43A.

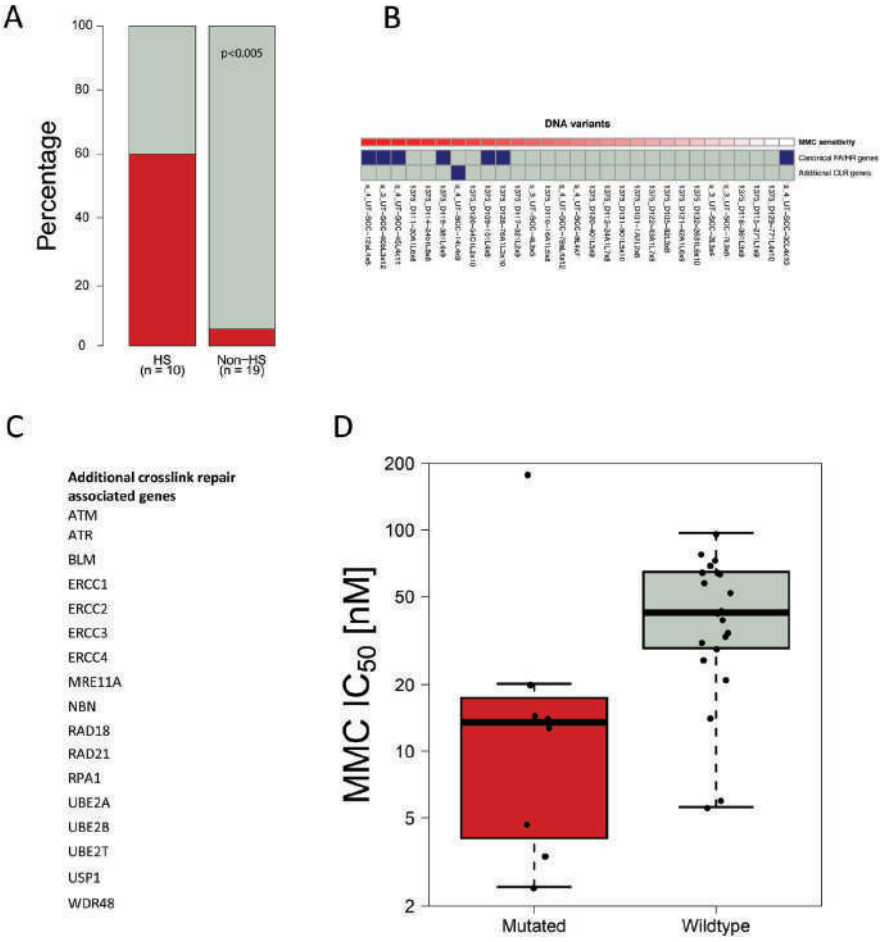


Supplementary Figure 6. Rad51 foci formation after radiation.

Radiation-induced RAD51 foci histograms in UT-SCC-60B **(A)** and UT-SCC-30 **(B)**. The percentage of cells with rad51foci counts in the indicated bins are shown. In HR-proficient cells radiation (6 Gy) induces the formation of nuclear rad51foci. The data show a pronounced induction of such rad51 foci in the UT-SCC-30 and a greatly dampened response in the UT-SCC-60B. This is consistent with the MMC-hypersensitivity and indicates a compromised FA/HR pathway in the UT-SCC-60B. The average of data of 3-5 independent experiments are shown (error = SD). Cells were seeded in 6-well plates containing sterilized microscope cover slips and irradiated with 6Gy. After 5 hours, cells were fixed with 4% paraformaldehyde, permeablized with 0.15% Triton-X and incubated in 1% BSA. Rad51 labelling was performed by overnight incubation with rabbit anti-RAD51 primary antibody (polyclonal H29, Santa Cruz), followed by a FITC labelled anti-rabbit secondary antibody (Jackson ImmunoResearch). Images were collected for counting with the Leica TCS SP5 Confocal microscope at 63× or 100× magnification. As shown, UT-SCC-30 which contains a variant in BRCA1, is capable of RAD51 foci formation, indicating an intact repair pathway. This is in accordance with functional BRCA.

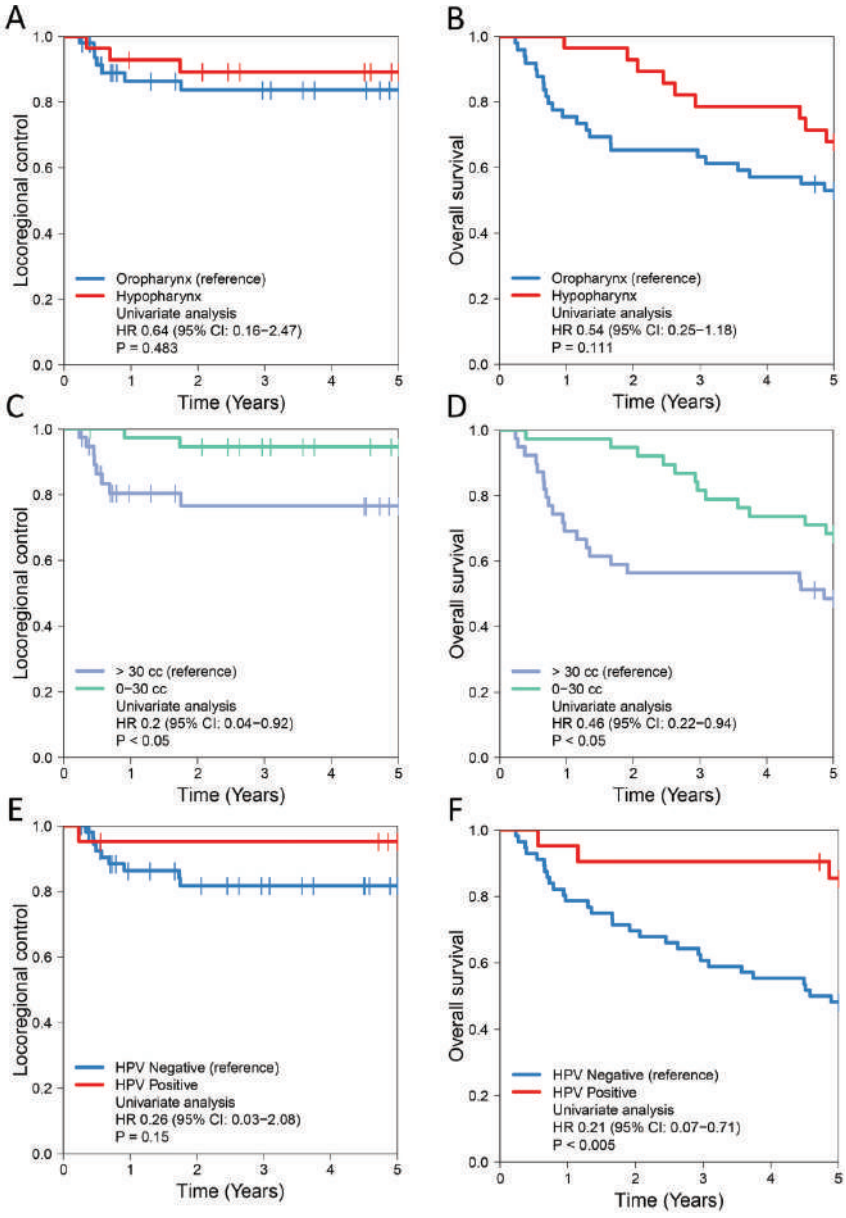


Supplementary Figure 7. indications for loss of heterozygosity in selected genes.
Lolliplot diagrams to the left show the variant zygosity distribution in genes of cell lines that were classified as having a FA/HR variant (marked with *) according to our variant selection protocol. In addition to these variants, all other SNV variants (originally discarded by the variant selection protocol), but with MAF below 60%, are shown. Homozygous variants are shown in red, heterozygous variants in red and yellow. Estimated copy number across all baits in the capture and plotted in chromosomal order (bait No. on the X-axis) are shown to the right for the respective HNSCC line. Red circle and arrow highlight the position of the affected gene as indicated to the left.



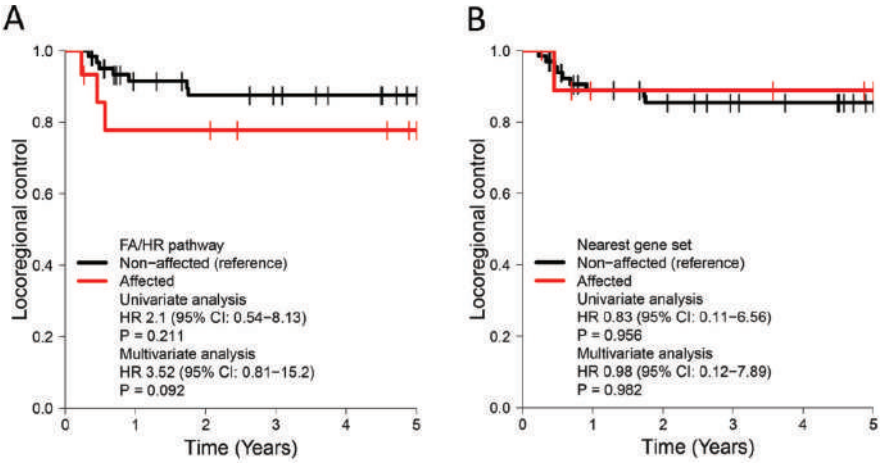
Supplementary Figure 8. FA/HR-variant enrichment in MMC sensitive HNSCC and extension of crosslink repair gene list to peripheral genes for DNA repair defect sequence variant selection.

(A) FA/HR variants are significantly enriched in the MMC sensitive HNSCC cell lines (HS; top 1/3 most MMC sensitive cell lines) compared to non-HS ($p < 0.005$). **(B)** The variant selection protocol as in Sup-Table 3 was applied to 17 additional genes that act at the periphery of the FA/HR pathway and that could have an impact on crosslink repair (listed under **(C)**). This approach revealed one additional cell line with a potential ATR gene mutation. **(D)** HNSCC cell lines with a variant in any of the 27 FA/HR plus 17 peripheral genes ("mutated") are significantly more sensitive to MMC (Wilcoxon Rank Sum test: $p < 0.01$; Fisher exact test: $p < 0.005$).



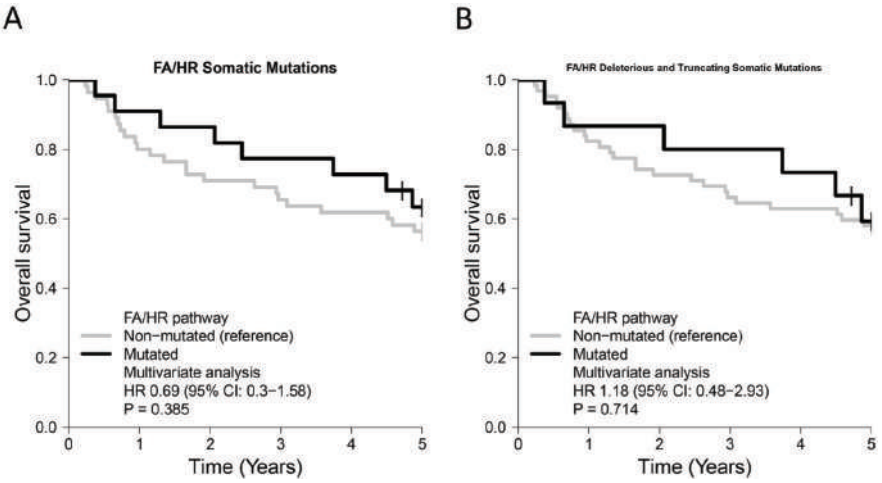
Supplementary Figure 9. Major clinical variables.

Locoregional control (**A, C, E**) and overall survival (**B, D, F**) in the HNSCC patient cohort showing the influence of (A,B) tumor site (oropharynx vs hypopharynx), (C, D) tumor size (below and above 30 cubic centimeters) and (E, F) HPV status. In-figure: hazard ratios (HR) from univariate Cox models and p-values from the exact log-rank test (see also Supplementary Table 6).



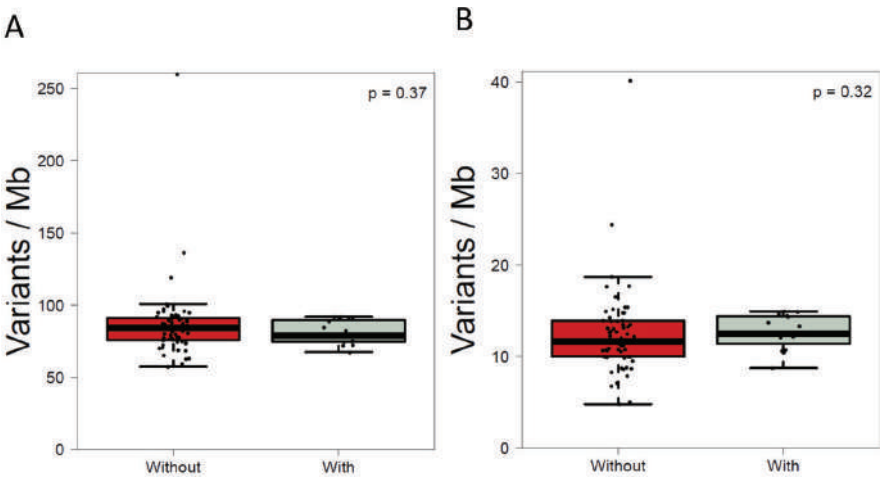
Supplementary Figure 10. Association of functional FA/HR repair defect associated variants with locoregional control in HNSCC patients.

FA/HR gene variant selection in the seventy-seven strong chemo-radiated patient cohort with oropharyngeal or hypopharyngeal HNSCC returned 15 patients among the 77 with such tumors. **(A)** Kaplan-Meier graph shows locoregional control in the patients with tumors with and without FA/HR variants as in Figure 4. Data shows a trend to worse prognosis of patients with 'FA/HR affected' tumors. **(B)** Kaplan-Meier graph shows locoregional control in patients with tumors with and without variants in the nearest gene set as in Figure 4D. In-figure: hazard ratios (HR) from Cox proportional hazards models with univariate *p*-values from the exact log-rank test. Multivariate HR and *p*-values from Cox models that include the tested gene set, tumor site, HPV-status and tumor volume are also shown for comparison purposes, however it should be noted that small number of events per variable result in inaccurate parameter estimates and restricts meaningful multivariate analyses. The number of LRC events were 3 for affected and 7 for non-affected.



Supplementary Figure 11. Association of somatic FA/HR gene variants with clinical outcome.

(A) Somatic variant calling in our patient cohort identified 22/77 patient tumors with somatic FA/HR-variants. The somatic variants were not associated with overall survival, showing that common approach of selecting somatic variants would have missed the clinical outcome associations presented in this study. Only variants in exonic and splice site regions were considered. SNP databases were used to remove germline variants, because matched normal samples were unavailable. Variants were considered somatic if they (1) were absent in all three of the 1000 Genomes Project, Exome Sequencing Project or Exome Aggregation Consortium databases, or (2) if they occurred >20 times in the COSMIC database. **(B)** From the somatic FA/HR-variants identified in (A), only those predicted to be deleterious (by SIFT or polyphen) or truncating (frameshift and non-frameshift indels, stop gain and loss variants) were selected. This analysis identified 15/77 patient tumors with deleterious or truncating somatic FA/HR variants. These patient tumors did not have a different outcome than the others. In-figure multivariate hazard ratios (HR) come from multivariate Cox proportional hazard models that include the tested gene set, tumor site, HPV-status and tumor volume.



Supplementary Figure 12. Mutational load in tumors with and without a FA/HR-variant.

A higher mutational load in tumors (due to inherent mutational processes) can contribute to a higher chance to select sequence variants. Tumors that had FA/HR-variants (i.e were selected after applying the FA/HR variant selection criteria) were compared to those without (Figure 4). **(A)** Non-SNPs sequence variants (synonymous and non-synonymous) representing de novo mutations within all the captured genes (556) were counted. This value was normalized to the individual sample base coverage and is shown as number of variants per sequenced Mb bases for the different tumors. **(B)** Using selection criteria similar to the FA-HR gene variant calling that select for non-synonymous variants and applies a MAF cut off of 2.5% did also not expose any increased variant burden in the sequenced genes in the identified cell lines. Variant load was normalized to the base coverage of the individual samples and is shown as number of variants per sequenced Mb bases. Boxplots are shown with individual tumor sample values as dots.

Supplementary Table 1: Characteristics and sensitivity values of the analysed head and neck cancer cell lines

HNSCC	MMC [nM]			MMC ind. G2 [%]			Olaparib [µM]			AUC		SF2 [Gy]	S-phase content [%]			Doubling time [h]
	IC50	sd	*N*	1 µm	sd	*N*	IC50	sd	*N*		sd		average	sd	*N*	
UT-SCC-1A	43	28,3	4	25	2,7	3	5,4	4,2	5	1.7	0.3	0,38	23	1,0	3	79
UT-SCC-2	65	29,1	3	18	2,6	3	10*		4	1.8	0.2	0,35	22	1,5	3	19
UT-SCC-4	26	15,4	3	29	4,2	3	10*		4	1.7	0.2	0,35	25	1,0	3	30
UT-SCC-7	69	5,5	3	5	2,5	3	5,1	4,6	3	2.0	0.2	0,42	23	1,0	3	33
UT-SCC-8	58	7,3	3	16	2,1	2	10*		3	1.9	0.1	0,37	25	3,8	3	32
UT-SCC-9	33	9,8	3	42	3,3	3	5,5	4,5	7	1.4	0.1	0,25	16	2,0	3	17
UT-SCC-12A	2	0,7	4	53	1,7	3	4,9	1,0	4	2.1	0.1	0,38	26	1,4	3	26
UT-SCC-14	14	8,5	3	17	2,2	3	1,5	0,2	3	1.7	0.3	0,34	22	1,7	3	43
UT-SCC-15	15	6,6	3	22	2,2	3	10*		3	2.1	0.1	0,42	11	1,5	3	49
UT-SCC-16A	29	10,4	5	39	4,3	3	9,2	1,6	4	1.8	0.1	0,33	27	3,7	3	38
UT-SCC-20A	6	2,8	4	49	2,6	3	3,2	0,6	7	2.1	0.2	0,46	22	4,0	12	39
UT-SCC-24A	39	9,4	3	25	3,8	3	6,7	0,1	3	2.6	0.3	0,51	26	1,3	3	24
UT-SCC-24B	6	3,1	5	53	2,0	3	1,3	0,7	4	2.3	0.1	0,43	18	0,0	3	32
UT-SCC-27	78	18,7	3	9	1,7	3	4,3	1,0	3	1.9	0.1	0,37	16	0,0	3	23
UT-SCC-30	179	70,0	4	2	2,1	3	5,6	2,7	3	2.0	0.1	0,39	31	0,6	3	21
UT-SCC-32	21	1,2	3	12	2,6	3	3,9	0,1	3	1.7	0.3	NDT	16	1,0	3	49
UT-SCC-36	73	7,8	3	16	2,1	2	10*		3	2.2	0.2	NDT	25	3,8	3	19
UT-SCC-38	13	3,6	3	24	6,9	3	3,2	1,2	3	2.3	0.3	0,45	25	2,2	3	38
UT-SCC-40	35	4,8	3	32	4,3	3	7,2	2,5	3	2.3	0.2	0,45	30	2,7	3	24
UT-SCC-42A	64	4,4	3	23	1,6	3	3,4	1,9	6	2.1	0.1	0,44	25	1,0	3	29
UT-SCC-43A	52	26,1	4	11	2,6	3	4,0	1,8	4	1.8	0.2	0,34	34	3,9	3	12
UT-SCC-45	5	3,1	4	55	1,9	3	1,0	0,4	3	2.0	0.1	0,37	27	0,6	3	49
UT-SCC-54C	14	5,5	3	22	5,7	3	10*		3	2.3	0.1	0,42	28	2,7	3	52
UT-SCC-60B	3	1,2	3	49	1,2	3	7,7	2,0	3	2.2	0.3	NDT	27	1,7	3	27
UT-SCC-76A	20	11,6	3	17	1,8	3	8,2	2,2	4	2.5	0.2	0,51	28	2,0	3	27
UT-SCC-77	97	17,8	3	22	3,1	3	6,4	1,2	3	2.5	0.2	NDT	22	2,3	3	25
UT-SCC-79A	31	13,1	4	9	3,2	3	7,6	2,1	3	2.4	0.2	NDT	19	1,1	3	48
UT-SCC-90	42	15,2	3	18	2,9	3	7,6	3,5	3	2.2	0.2	NDT	22	3,9	3	24
FANCA	2	0,9	4	44	11,7	2	1,8	1,6	7	NDT	NDT	NDT	11	5,7	5	33
FANCG	2	0,6	4	65	2,8	3	0,6	0,5	5	NDT	NDT	NDT	19	5,9	12	33
GM847	76	34,9	3	11	5,5	3	4,2	0,4	3	NDT	NDT	NDT	29	4,7	13	20
NKI-SC-263	65	30,3	3	54	3,5	3	10*		3	2.51	NDT	0,52	9	1,5	3	24

List of used HNSCC cell lines generated and provided by R. Grenamn and the determined cellular characteristics and response parameters. The FA patient-derived fibroblast cell lines with known mutations in FANCG (EUFA 636)(25) and FANCA (VU0173-F; EUFA 173)(24) were kindly provided by Dr. H. Joenje and served as positive controls, whereas an hTERT transformed human fibroblast cell line (GM847), was included for comparison in a non-tumor derived cell line. The FA gene mutation status in the FANCA and FANCG positive control cell lines was confirmed by PCR and subsequent conventional Sanger sequencing and by capture based sequencing that replicated these results.
Doubling times of all cell lines were assessed by plating different cell numbers into 6-well plates in triplicates and counting live cells after harvesting them by trypsinization at different time points thereafter. Population doubling times (DT) were calculated from the exponential fits on the steepest part of the growth curves. Radiation survival data were kindly provided by R. Grènman.
* = IC_{50} was not achieved within the tested dose range; the maximum tested olaparib concentration was assigned to these cell lines. NDT = not determined, sd = standard deviation, AUC = Area under the radiation survival curve, SF2 = surviving fraction at 2Gy, *N* = number of independent experiments. # all cell lines were regularly tested for mycoplasma (Janetzko et al, Transfus Med Hemother 2014; 41(1):83-89) twice a month when in culture. Passage number never exceeded 20. Photographs, morphology, cell doubling, PE, characteristics and later DNA-Seq served as authentication method.

Supplementary Table 2: FA/HR gene list

HGNC	ENSEMBL
canonical:	
FANCA	ENSG00000187741
FANCB	ENSG00000181544
FANCC	ENSG00000158169
BRCA2	ENSG00000139618
FANCD2	ENSG00000144554
FANCE	ENSG00000112039
FANCF	ENSG00000183161
FANCG	ENSG00000221829
FANCI	ENSG00000140525
BRIP1	ENSG00000136492
FANCL	ENSG00000115392
FANCM	ENSG00000187790
PALB2	ENSG00000083093
RAD51C	ENSG00000108384
SLX4	ENSG00000188827
RAD51	ENSG00000051180
TP53BP1	ENSG00000067369
BRCA1	ENSG00000012048
FAAP100	ENSG00000185504
RAD50	ENSG00000113522
RAD51B	ENSG00000182185
RAD51D	ENSG00000185379
RAD52	ENSG00000002016
RAD54B	ENSG00000197275
RAD54L	ENSG00000085999
XRCC3	ENSG00000126215
XRCC2	ENSG00000196584

List of canonical FA/HR pathway genes with core activities in the Fanconi anemia and homologous recombination DNA repair pathways that were chosen for the genetic analyses in this study. Canonical FA and HR pathway genes that were available within the sequencing capture set were selected and are listed with their ENSEMBL reference IDs.

Supplementary Table 3: Filtering steps in bio-informatics analysis for variant selection

		Exclusion criteria:
Annovar refSeq annotation	Region	not annotated as 'exonic' or 'splicing'
	Consequence	Synonymous SNV
Coverage criteria	Total reads	<10
	Variant allele containing reads	<4
Healthy blood samples		>1
Minor allele frequency	1000 Genomes Project (August 2015, European population)	>0.025*
Variant allele fraction	Cell lines	<0.80
	Patient samples	<0.80**

Variants, single nucleotide variants (SNVs) and indels, were called with VarScan. Only variants that are annotated by RefSeq to be in (ncRNA) exonic or splicing regions were selected. Variant calls were filtered against the following criteria: at least 10 total reads and 4 reads containing the variant allele pass the quality controls, and at least 80% of all reads need to contain the variant allele (variant allele frequency, VAF, of > 0.8). Since assuming loss of gene function to primarily cause the observed DNA repair defects we focused on variants with such a high read fraction, as this is thought to reflect homozygosity. Only rare polymorphisms were retained. This was done by removing variants with a minor allele frequency (MAF) higher than 2.5% in the 1000 Genomes database. Variants that were not present in this database were considered rare or de novo, hence retained. Splicing variants were removed if they were not in COSMIC70. Finally, variants that occurred in two or more of four healthy control subjects that were included in the sequencing analysis were removed. Identified variants were individually confirmed using IGV viewer.

*Cut offs were chosen based on HNSCC risk estimates as described in discussion.

**After applying sample specific corrections that considered the pathologist's tumour percentage estimates in the sample as described in Materials and Methods.

Supplementary Table 4: References and characteristics of selected canonical FA/HR gene set variants in HNSCC cell line panel

Cell line	Gene	Protein change	SNP ID	cosmic70	MAF 1000G	CADD	REVEL	SIFT	PolyPhen2	ClinVar	Comments to potential functional effects	PMID
Control FANCA EUFA173	FANCA	R951Q				16.36	0.731		Probably Damaging		described as affecting function and referenced in LOVD; FANCA mutation found in FA patient	17924555
Control FANCG EUFA636	FANCG	T550fs									Frameshift deletion resulting in truncation, FANCG mutation in FA patient	12552564
Control FANCG EUFA636	FANCG	S7F	rs35984312	ID = COSM 753539	0,002	12.71	0.062	Deleterious	B	Benign (Fanconi Anemia)	FANCG mutation found in FA patient	
UT-SCC-12A	BRIP1	G690R				16.23	0.697		Probably Damaging		No report found. The G to R change at the border of the helicase domain could affect the protein function.	
UT-SCC-60B	BRCA1	R841W	rs1800709	ID = COSM 1246204	0,005	5.12	0.355	Deleterious	Possibly Damaging	Benign (hereditary breast cancer)	Referenced in LOVD; Predicted to affect function in evolutionary conservation analysis Likely pathogenic through bayesian analysis considering breast cancer family frequencies 840S is phosphorylation site	12531920 18415037 9585599 8968716
UT-SCC-45	FANCD2	G1043S	rs142238966			15.38	0.169	T	Probably Damaging		Located in DNA Binding region	
UT-SCC-38	BRCA1	M1652I	rs1799967		0,0159	11.83	0.426	T	B	Benign (breast cancer)	Referenced in LOVD: Evolutionary conservation analysis predicts deleteriousness within BRCT domain	18694767 12955716

Supplementary Table 4: Continued

Cell line	Gene	Protein change	SNP ID	cosmic70	MAF 1000G	CADD	REVEL	SIFT	PolyPhen2	ClinVar	Comments to potential functional effects	PMID
UT-SCC-15	FANCD2	P834A				19.66	0.414	T	Probably Damaging		Proline to alanine change likely to affect protein structure at DNA binding site	
UT-SCC-76A	BRCA2	Y42C	rs4987046		0,003	4.474	0.089	T	B	Benign (breast cancer)	Alters PALB2 binding site. Variant reported to affect RPA Binding in co-immunoprecipitation assay	12527904
UT-SCC-30	BRCA1	R1347G	rs28897689		0,001	10.36	0.527	Deleterious	B	Benign (breast cancer)	Located at poorly conserved region, predicted to affect function by Grantham score	18415037

PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/) is a tool that predicts a possible impact of an amino acid change on the protein structure. SIFT (http://sift.jcvi.org/) prediction is based on the degree of conversation of aminoacids in the sequence. CADD (http://cadd.gs.washington.edu/) and REVEL (https://sites.google.com/site/revelgenomics/) are variant effect prediction algorithms that combine the knowledge of a multitude of algorithms such as PolyPhen-2 and SIFT. CADD and REVEL scores above 15 and 0.5, respectively, are considered pathogenic.
T = Tolerated; B = Benign; PMID = reference PubMed ID number. Minor allele frequencies (MAF) from 1000 genome (1000G) are listed.

Supplementary Table 5A: References and characteristics of selected canonical FA/HR gene set variants in the tumor samples of the patients in the study

Pts	Gene	Protein change	SNP ID	cosmic70	MAF 1000G	CADD	REVEL	SIFT	PolyPhen2	ClinVar	Comments to potential functional effects	REF PMID
1	PALB2	L337S	rs45494092		0,0249	8.918	0.041	Deleterious	B	Conflicting interpretations	Possibly enriched in familial cutaneous malignant melanoma. Similar frequency in breast cancer than controls.	24949998 21618343 26283626
	FANCC	V60I	rs138629441			0.185	0.101	T	B	Likely benign (Fanconi Anemia)	Unlikely to affect function	14695169
	FANCG	R513Q	rs17885240		0,0129	15.01	0.016	T	B	Likely benign (Fanconi Anemia)	Increased frequency in children with AML	16643430
2	PALB2	T1099R	rs142132127	ID=COSM1666745		16	0.214	Deleterious	Damaging	Uncertain significance	Semiconservative AS substitution in conserved 5th WD repeat, region of BRCA2, RAD51 and POLH interaction. Possible enrichment in individuals with breast or ovarian cancer.	25186627 26315354
3	PALB2	L337S	rs45494092		0,0249	8.918	0.041	Deleterious	B	*as above*	*see above*	*see above*
4	FANCG	R513Q	rs17885240		0,0129	15.01	0.016	T	B	*as above*	*see above*	*see above*
5	FANCG	R513Q	rs17885240		0,0129	15.01	0.016	T	B	*as above*	*see above*	*see above*
6	FANCM	K953N	rs142864437			14.93	0.085	Deleterious	Damaging	Uncertain significance	No reports found	
	FANCA	A554V				19.95	0.636	T	Damaging		No reports found	
7	FANCM	T77A	rs61746895		0,0129	2.724	0.041	T	B	Likely benign (Fanconi Anemia)	Neighboring MPH1 (ERCC4-related helicase) region	
8	FANCF	P320L	rs45451294		0,0119	16.83	0.096	T	Possibly Damaging	Benign (Fanconi Anemia)	No reports found	

Supplementary Table 5A: Continued

Pts	Gene	Protein change	SNP ID	cosmic70	MAF 1000G	CADD	REVEL	SIFT	PolyPhen2	ClinVar	Comments to potential functional effects	REF PMID
	FANCF	R38H				23.2	0.14	Deleterious	Damaging		No reports found	
9	FANCF	P320L	rs45451294		0,0119	16.83	0.096	T	Possibly Damaging	*as above*	*see above*	*see above*
10	FANCD2	N545S	rs145522204		0,0089	0.887	0.053	T	B	Benign	No reports found	
11	FANCD2	R997Q				15.1	0.091	T	B		No reports found	
12	FANCC	H256R				16.51	0.058	T	B		Reported in one individual in controls in pancreatic cancer study	15695377
13	RAD51C	G264S	rs147241704			23	0.202		B	Uncertain significance	Non-conservative amino acid substitution in conserved region (ATPase domain) LOVD: hypomorph feature: partial complementation in cells, normal Rad51 foci; Increased frequency in breast / ovarian Moderate penetrance suggestion in ovarian cancer	20400964 21990120
14	RAD51B	K243R	rs34594234		0,0089	25.4	0.186	T	Damaging			
15	BRCA1	R841W	rs1800709	ID=COSM1246204	0,005	5.12	0.355	Deleterious	Possibly Damaging	Benign (Hereditary Can)	Referenced in LOVD; Predicted to affect function in evolutionary conservation analysis Likely pathogenic through bayesian analysis considering breast cancer family frequencies neighboring 840S is phosphorylation site	22753008 21520273 18415037 8968716

Abbreviations and listing as in Supplementary Table 4.

Supplementary Table 5B: References and characteristics of selected canonical FA/HR gene set variants in the tumor samples of the patients in the study

Pts	Gene	Protein change	Age at diagnosis (years)	Primary site	HPV-status	Tumor volume	Smoker	Alcohol consumption
1	PALB2	L337S	75	Hypopharynx	Negative	0-30 cc	yes	yes
	FANCC	V60I						
	FANCG	R513Q						
2	PALB2	T1099R	62	Oropharynx	Positive	0-30 cc	never	yes
3	PALB2	L337S	60	Oropharynx	Negative	> 30 cc	yes	yes
4	FANCG	R513Q	64	Hypopharynx	Negative	> 30 cc	yes	former-alcoholic
5	FANCG	R513Q	54	Oropharynx	Negative	0-30 cc	yes	yes
6	FANCM	K953N	55	Hypopharynx	Negative	0-30 cc	yes	yes
	FANCA	A554V						
7	FANCM	T77A	57	Oropharynx	Positive	> 30 cc	former-smoker	yes
8	FANCF	P320L	56	Oropharynx	Negative	> 30 cc	unknown	unknown
	FANCF	R38H						
9	FANCF	P320L	59	Hypopharynx	Negative	0-30 cc	yes	yes
10	FANCD2	N545S	48	Oropharynx	Negative	0-30 cc	yes	yes
11	FANCD2	R997Q	65	Hypopharynx	Negative	0-30 cc	former-smoker	never
12	FANCC	H256R	64	Hypopharynx	Negative	0-30 cc	former-smoker	yes
13	RAD51C	G264S	54	Oropharynx	Negative	> 30 cc	yes	yes
14	RAD51B	K243R	73	Hypopharynx	Negative	0-30 cc	yes	yes
15	BRCA1	R841W	37	Oropharynx	Negative	> 30 cc	former-smoker	former -alcoholic

Continuation of Supplementary Table 5 listing patient characteristics.

Supplementary Table 6: Statistical values from overall survival analysis

				Univariate Analysis			Multivariate Analysis		
Variables		**Patients (*n*)**	**Events (*n*)**	**HR**	**95% CI**	***P*-Value**	**HR**	**95% CI**	***P*-Value**
Tumor site									
	Oropharynx	49	23	1.0 (ref)			1.0 (ref)		
	Hypopharynx	28	9	**0.54**	(0.25–1.18)	*0.1113*	**0.27**	(0.12–0.6)	***0.0014*
Tumor volume									
	> 30 cc	39	20	1.0 (ref)			1.0 (ref)		
	0-30 cc	38	12	**0.46**	(0.22–0.94)	**0.03173*	**0.4**	(0.18–0.87)	**0.0205*
HPV status									
	Negative	56	29	1.0 (ref)			1.0 (ref)		
	Positive	21	3	**0.21**	(0.07–0.71)	***0.00319*	**0.13**	(0.04–0.45)	***0.0012*
Canonical FA/HR									
	No variant	62	23	1.0 (ref)			1.0 (ref)		
	Variant	15	9	**1.96**	(0.9–4.23)	*0.1035*	**2.59**	(1.11–6.01)	**0.0271*
Nearest genes									
	No variant	67	26	1.0 (ref)			1.0 (ref)		
	Variant	10	6	**1.81**	(0.75–4.41)	*0.2388*	**1.86**	(0.74–4.68)	*0.1889*

Table shows results of the univariate and multivariate analyses. Hazard ratios (HR) and corresponding 95% confi dence intervals (95% CI) derived from Cox proportional hazards models are shown for the univariate analysis. The groups were of uneven size, with one group being small, resulting in an inappropriate asymptotic distribution. Therefore exact log-rank test *p*-values which are listed in the Table are based on Monte Carlo simulations with 100,000 iterations. Hazard ratios (HR), 95% CI and *p*-values are derived from Cox proportional hazards models in the multivariate analysis. The multivariate model considered tumor site, tumor volume, HPV status, and the genetic variant status as fourth variable. Genetic variant status addressed either the canonical FA/HR variant status or the presence of variants in the nearest genes.

Supplementary Table 7: FA/HR gene mutation and HPV status

A

		HPV status	
		–	+
FA/HR Var	–	43	19
	+	13	2

B

		HPV status	
		–	+
FA/HR Var	–	22	19
	+	6	2

Table shows the overlap of HPV positivity and FA/HR-variant presence. (**A**) All tumors (n = 77), no signifi cant association between FA/HR variant and HPV status was found. Fishers exact test p = 0.215. (**B**) Only oropharynx tumors (n = 49), no association between FA/HR variant and HPV status was found. Fishers exact test p = 0.4384.

Supplementary Table 8: Matched control gene sets

Canonical Fa/Hr	matched nearest
FANCA	CDT1
FANCB	POLA1
FANCC	PTCH1
BRCA2	RFC3
FANCD2	VHL
FANCE	CDKN1A
FANCF	ABCC8
FANCG	APTX
FANCI	POLG
BRIP1	PPM1D
FANCL	PSME4
FANCM	POLE2
PALB2	PLK1
RAD51C	AKAP1
SLX4	ABCA3
RAD51	LTK
TP53BP1	LTK
BRCA1	TOP2A
FAAP100	CSNK1D
RAD50	PPP2CA
RAD51B	HIF1A
RAD51D	LIG3
RAD52	FOXM1
RAD54B	NBN
RAD54L	MUTYH
XRCC3	AKT1
XRCC2	SMARCD3

List of matched genes which are close to the canonical FA/HR pathway genes (Supplementary Table 2). Genes within the > 500 genes strong DNA sequencing capture set that best fulfi lled the requirements for matching (nearest) were selected for control analyses.

Supplementary Table 9: RT-PCR primers and kits

Kit/primer	Cat.No:
AllPrep DNA/RNA Mini Kit	Qiagen, Cat. No.80004
QuantiFast SYBR RT-PCR kit	Qiagen, Cat. No.204154
Hs_FANCF_1_SG QuantiTect Primer Assay	Qiagen, QT00241136 (NM_022724)
Hs_GAPDH_2_SG QuantiTect Primer Assay	Qiagen, QT01192646 (NM_002046)
Hs_B2M_1_SG QuantiTect Primer Assay	Qiagen, QT00088935 (NM_004048)

List of primers and kits used for FANCF RT-PCR analysis.
RNA quality was assessed by the Agilent 2100 Bio-analyzer (Agilent Technologies). All RNA samples used for the FANC-F RT-PCR reached maximal RIN 10 values.

References

[1] Kim H, D'Andrea AD. Regulation of DNA cross-link repair by the Fanconi anemia/BRCA pathway. Genes Dev 2012;26:1393–408. doi:10.1101/gad.195248.112.

[2] Mateo J, Carreira S, Sandhu S, Miranda S, Mossop H, Perez-Lopez R, et al. DNA-Repair Defects and Olaparib in Metastatic Prostate Cancer. N Engl J Med 2015;373:1697–708. doi:10.1056/NEJMoa1506859.

[3] Moschetta M, George A, Kaye SB, Banerjee S. BRCA somatic mutations and epigenetic BRCA modifications in serous ovarian cancer. Ann Oncol Off J Eur Soc Med Oncol 2016;27:1449–55. doi:10.1093/annonc/mdw142.

[4] Helleday T. Homologous recombination in cancer development, treatment and development of drug resistance. Carcinogenesis 2010;31:955–60. doi:10.1093/carcin/bgq064.

[5] Farmer H, McCabe N, Lord CJ, Tutt ANJ, Johnson DA, Richardson TB, et al. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. Nature 2005;434:917–21. doi:10.1038/nature03445.

[6] O'Connor MJ. Targeting the DNA Damage Response in Cancer. Mol Cell 2015;60:547–60. doi:10.1016/j.molcel.2015.10.040.

[7] Rosenberg PS, Alter BP, Ebell W, Rosenberg PS. Cancer risks in Fanconi anemia : findings from the German Fanconi Anemia Registry 2008;93:511–7. doi:10.3324/haematol.12234.

[8] Kutler DI, Auerbach AD, Satagopan J, Giampietro PF, Batish SD, Huvos AG, et al. High incidence of head and neck squamous cell carcinoma in patients with Fanconi anemia. Arch Otolaryngol Head Neck Surg 2003;129:106–12.

[9] Kamangar F, Dores GM, Anderson WF. Patterns of cancer incidence, mortality, and prevalence across five continents: defining priorities to reduce cancer disparities in different geographic regions of the world. J Clin Oncol 2006;24:2137–50. doi:10.1200/JCO.2005.05.2308.

[10] Leemans CR, Braakhuis BJM, Brakenhoff RH. The molecular biology of head and neck cancer. Nat Rev Cancer 2011;11:9–22. doi:10.1038/nrc2982.

[11] Pignon JP, Maître A le, Maillard E, Bourhis J. Meta-analysis of chemotherapy in head and neck cancer (MACH-NC): An update on 93 randomised trials and 17,346 patients. Radiother Oncol 2009;92:4–14. doi:10.1016/j.radonc.2009.04.014.

[12] Baumann M, Krause M, Overgaard J, Debus J, Bentzen SM, Daartz J, et al. Radiation oncology in the era of precision medicine. Nat Rev Cancer 2016;16:234–49. doi:10.1038/nrc.2016.18.

[13] Linge A, Löck S, Gudziol V, Nowak A, Lohaus F, von Neubeck C, et al. Low Cancer Stem Cell Marker Expression and Low Hypoxia Identify Good Prognosis Subgroups in HPV(-) HNSCC after Postoperative Radiochemotherapy: A Multicenter Study of the DKTK-ROG. Clin Cancer Res 2016;22:2639–49. doi:10.1158/1078-0432.CCR-15-1990.

[14] Begg AC, Stewart FA, Vens C. Strategies to improve radiotherapy with targeted drugs. Nat Rev Cancer 2011;11:239–53. doi:10.1038/nrc3007.

[15] Brooks PJ, Theruvathu JA. DNA adducts from acetaldehyde: implications for alcohol-related carcinogenesis. Alcohol 2005;35:187–93. doi:10.1016/j.alcohol.2005.03.009.

[16] Barón AE, Franceschi S, Barra S, Talamini R, La Vecchia C. A comparison of the joint effects of alcohol and smoking on the risk of cancer across sites in the upper aerodigestive tract. Cancer Epidemiol Biomarkers Prev n.d.;2:519–23.

[17] Wreesmann VB, Estilo C, Eisele DW, Singh B, Wang SJ. Downregulation of Fanconi anemia genes in sporadic head and neck squamous cell carcinoma. ORL J Otorhinolaryngol Relat Spec 2007;69:218–25. doi:10.1159/000101542.

[18] Marsit CJ, Liu M, Nelson HH, Posner M, Suzuki M, Kelsey KT. Inactivation of the Fanconi anemia/BRCA pathway in lung and oral cancers: implications for treatment and survival. Oncogene 2004;23:1000–4. doi:10.1038/sj.onc.1207256.

[19] Sparano A, Quesnelle KM, Kumar MS, Wang Y, Sylvester AJ, Feldman M, et al. Genome-wide profiling of oral squamous cell carcinoma by array-based comparative genomic hybridization. Laryngoscope 2006;116:735–41. doi:10.1097/01.mlg.0000205141.54471.7f.
[20] Cancer Genome Atlas Network. Comprehensive genomic characterization of head and neck squamous cell carcinomas. Nature 2015;517:576–82. doi:10.1038/nature14129.
[21] Morris LGT, Chandramohan R, West L, Zehir A, Chakravarty D, Pfister DG, et al. The Molecular Landscape of Recurrent and Metastatic Head and Neck Cancers. JAMA Oncol 2017;3:244. doi:10.1001/jamaoncol.2016.1790.
[22] Stoepker C, Ameziane N, van der Lelij P, Kooi IE, Oostra AB, Rooimans MA, et al. Defects in the Fanconi Anemia Pathway and Chromatid Cohesion in Head and Neck Cancer. Cancer Res 2015;75:3543–53. doi:10.1158/0008-5472.CAN-15-0528.
[23] Romick-Rosendale LE, Lui VWY, Grandis JR, Wells SI. The Fanconi anemia pathway: repairing the link between DNA damage and squamous cell carcinoma. Mutat Res n.d.;743–744:78–88. doi:10.1016/j.mrfmmm.2013.01.001.
[24] Akkari YM, Bateman RL, Reifsteck CA, D'Andrea AD, Olson SB, Grompe M. The 4N cell cycle delay in Fanconi anemia reflects growth arrest in late S phase. Mol Genet Metab 2001;74:403–12. doi:10.1006/mgme.2001.3259.
[25] Sasaki MS, Tonomura A. A high susceptibility of Fanconi's anemia to chromosome breakage by DNA cross-linking agents. Cancer Res 1973;33:1829–36.
[26] McCabe N, Turner NC, Lord CJ, Kluzek K, Bialkowska A, Swift S, et al. Deficiency in the repair of DNA damage by homologous recombination and sensitivity to poly(ADP-ribose) polymerase inhibition. Cancer Res 2006;66:8109–15. doi:10.1158/0008-5472.CAN-06-0140.
[27] Rottenberg S, Jaspers JE, Kersbergen A, van der Burg E, Nygren AOH, Zander SAL, et al. High sensitivity of BRCA1-deficient mammary tumors to the PARP inhibitor AZD2281 alone and in combination with platinum drugs. Proc Natl Acad Sci U S A 2008;105:17079–84. doi:10.1073/pnas.0806092105.
[28] Garcia-Higuera I, Taniguchi T, Ganesan S, Meyn MS, Timmers C, Hejna J, et al. Interaction of the Fanconi anemia proteins and BRCA1 in a common pathway. Mol Cell 2001;7:249–62. doi:10.1016/s1097-2765(01)00173-3.
[29] Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nat Methods 2010;7:248–9. doi:10.1038/nmeth0410-248.
[30] Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res 2003;31:3812–4. doi:10.1093/nar/gkg509.
[31] Snyder ER, Ricker JL, Chen Z, Waes C Van. Variation in cisplatinum sensitivity is not associated with Fanconi Anemia/BRCA pathway inactivation in head and neck squamous cell carcinoma cell lines. Cancer Lett 2007;245:75–80. doi:10.1016/j.canlet.2005.12.028.
[32] Burkitt K, Ljungman M. Compromised Fanconi anemia response due to BRCA1 deficiency in cisplatin-sensitive head and neck cancer cell lines. Cancer Lett 2007;253:131–7. doi:10.1016/j.canlet.2007.01.017.
[33] Seow HA, Penketh PG, Baumann RP, Sartorelli AC. Bioactivation and resistance to mitomycin C. Methods Enzymol 2004;382:221–33. doi:10.1016/S0076-6879(04)82012-3.
[34] Tomasz M. Mitomycin C: small, fast and deadly (but very selective). Chem Biol 1995;2:575–9. doi:10.1016/1074-5521(95)90120-5.
[35] Volpato M, Phillips RM. Tailoring targeted therapy to individual patients: lessons to be learnt from the development of mitomycin C. Cancer Genomics Proteomics n.d.;4:175–86.
[36] Sawamura AO, Aoyama T, Tamakoshi K, Mizuno K, Suganuma N, Kikkawa F, et al. Transfection of human cytochrome P-450 reductase cDNA and its effect on the sensitivity to toxins. Oncology n.d.;53:406–11. doi:10.1159/000227596.
[37] Belcourt MF, Hodnick WF, Rockwell S, Sartorelli AC. Differential toxicity of mitomycin C and porfiromycin to aerobic and hypoxic Chinese hamster ovary cells overexpressing human NADPH:cytochrome c (P-450) reductase. Proc Natl Acad Sci U S A 1996;93:456–60. doi:10.1073/pnas.93.1.456.

[38] Naipal KAT, Verkaik NS, Ameziane N, van Deurzen CHM, Ter Brugge P, Meijers M, et al. Functional ex vivo assay to select homologous recombination-deficient breast tumors for PARP inhibitor treatment. Clin Cancer Res 2014;20:4816–26. doi:10.1158/1078-0432.CCR-14-0571.

[39] Xu G, Chapman JR, Brandsma I, Yuan J, Mistrik M, Bouwman P, et al. REV7 counteracts DNA double-strand break resection and affects PARP inhibition. Nature 2015;521:541–4. doi:10.1038/nature14328.

[40] Jaspers JE, Kersbergen A, Boon U, Sol W, van Deemter L, Zander SA, et al. Loss of 53BP1 causes PARP inhibitor resistance in Brca1-mutated mouse mammary tumors. Cancer Discov 2013;3:68–81. doi:10.1158/2159-8290.CD-12-0049.

[41] Chandrasekharappa SC, Chinn SB, Donovan FX, Chowdhury NI, Kamat A, Adeyemo AA, et al. Assessing the spectrum of germline variation in Fanconi anemia genes among patients with head and neck carcinoma before age 50. Cancer 2017;123:3943–54. doi:10.1002/cncr.30802.

[42] Ferreira Antunes JL, Toporcov TN, Biazevic MGH, Boing AF, Scully C, Petti S. Joint and independent effects of alcohol drinking and tobacco smoking on oral cancer: a large case-control study. PLoS One 2013;8:e68132. doi:10.1371/journal.pone.0068132.

[43] Hashibe M, Brennan P, Benhamou S, Castellsague X, Chen C, Curado MP, et al. Alcohol drinking in never users of tobacco, cigarette smoking in never drinkers, and the risk of head and neck cancer: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium. J Natl Cancer Inst 2007;99:777–89. doi:10.1093/jnci/djk179.

[44] Prevalence and penetrance of BRCA1 and BRCA2 mutations in a population-based series of breast cancer cases. Anglian Breast Cancer Study Group. Br J Cancer 2000;83:1301–8. doi:10.1054/bjoc.2000.1407.

[45] Carter SL, Cibulskis K, Helman E, McKenna A, Shen H, Zack T, et al. Absolute quantification of somatic DNA alterations in human cancer. Nat Biotechnol 2012;30:413–21. doi:10.1038/nbt.2203.

[46] Goodwin PJ, Phillips K-A, West DW, Ennis M, Hopper JL, John EM, et al. Breast cancer prognosis in BRCA1 and BRCA2 mutation carriers: an International Prospective Breast Cancer Family Registry population-based cohort study. J Clin Oncol 2012;30:19–26. doi:10.1200/JCO.2010.33.0068.

[47] van den Broek AJ, Schmidt MK, van 't Veer LJ, Tollenaar RAEM, van Leeuwen FE. Worse breast cancer prognosis of BRCA1/BRCA2 mutation carriers: what's the evidence? A systematic review with meta-analysis. PLoS One 2015;10:e0120189. doi:10.1371/journal.pone.0120189.

[48] Strojan P, Vermorken JB, Beitler JJ, Saba NF, Haigentz M, Bossi P, et al. Cumulative cisplatin dose in concurrent chemoradiotherapy for head and neck cancer: A systematic review. Head Neck 2016;38 Suppl 1:E2151-8. doi:10.1002/hed.24026.

[49] Verhagen CVM, de Haan R, Hageman F, Oostendorp TPD, Carli ALE, O'Connor MJ, et al. Extent of radiosensitization by the PARP inhibitor olaparib depends on its dose, the radiation dose and the integrity of the homologous recombination pathway of tumor cells. Radiother Oncol 2015;116:358–65. doi:10.1016/j.radonc.2015.03.028.

[50] Kirova YM, Savignoni A, Sigal-Zafrani B, de La Rochefordiere A, Salmon RJ, This P, et al. Is the breast-conserving treatment with radiotherapy appropriate in BRCA1/2 mutation carriers? Long-term results and review of the literature. Breast Cancer Res Treat 2010;120:119–26. doi:10.1007/s10549-009-0685-6.

[51] Vollebergh MA, Lips EH, Nederlof PM, Wessels LFA, Wesseling J, Vd Vijver MJ, et al. Genomic patterns resembling BRCA1- and BRCA2-mutated breast cancers predict benefit of intensified carboplatin-based chemotherapy. Breast Cancer Res 2014;16:R47. doi:10.1186/bcr3655.

[52] Joenje H, Levitus M, Waisfisz Q, D'Andrea A, Garcia-Higuera I, Pearson T, et al. Complementation analysis in Fanconi anemia: assignment of the reference FA-H patient to group A. Am J Hum Genet 2000;67:759–62. doi:10.1086/303067.

[53] Demuth I, Wlodarski M, Tipping AJ, Morgan N V, de Winter JP, Thiel M, et al. Spectrum of mutations in the Fanconi anaemia group G gene, FANCG/XRCC9. Eur J Hum Genet 2000;8:861–8. doi:10.1038/sj.ejhg.5200552.

[54] Lansford CD, Grenman R, Bier H, Somers KD, Kim SY, Whiteside TL, et al. Head and Neck Cancers. Hum. Cell Cult., Dordrecht: Kluwer Academic Publishers; 1999, p. 185–255. doi:10.1007/0-306-46861-1_12.

[55] Rigaill GJ, Cadot S, Kluin RJC, Xue Z, Bernards R, Majewski IJ, et al. A regression model for estimating DNA copy number applied to capture sequencing data. Bioinformatics 2012;28:2357–65. doi:10.1093/bioinformatics/bts448.
[56] Talevich E, Shain AH, Botton T, Bastian BC. CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. PLoS Comput Biol 2016;12:e1004873. doi:10.1371/journal.pcbi.1004873.
[57] Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009;25:2078–9. doi:10.1093/bioinformatics/btp352.
[58] Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 2010;38:e164. doi:10.1093/nar/gkq603.
[59] Kircher M, Witten DM, Jain P, O'roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet 2014;46:310–5. doi:10.1038/ng.2892.
[60] Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, et al. REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. Am J Hum Genet 2016;99:877–85. doi:10.1016/j.ajhg.2016.08.016.

CHAPTER 3

Role of variant allele fraction and rare SNP filtering to improve cellular DNA repair endpoint association

David M. Vossen, Caroline V.M. Verhagen, Reidar Grénman, Roel J.C. Kluin, Marcel Verheij, Michiel W.M. van den Brekel, Lodewyk F.A. Wessels, Conchita Vens.

PLoS One. 2018 Nov; 13(11): e0206632.

Abstract

Background

Large cancer genome studies continue to reveal new players in treatment response and tumorigenesis. The discrimination of functional alterations from the abundance of passenger genetic alterations still poses challenges and determines DNA sequence variant selection procedures. Here we evaluate variant selection strategies that select homozygous variants and rare SNPs and assess its value in detecting tumor cells with DNA repair defects.

Methods

To this end we employed a panel of 29 patient-derived head and neck squamous cell carcinoma (HNSCC) cell lines, of which a subset harbors DNA repair defects. Mitomycin C (MMC) sensitivity was used as functional endpoint of DNA crosslink repair deficiency. 556 genes including the Fanconi anemia (FA) and homologous recombination (HR) genes, whose products strongly determine MMC response, were capture-sequenced.

Results

We show a strong association between MMC sensitivity, thus loss of DNA repair function, and the presence of homozygous and rare SNPs in the relevant FA/HR genes. Excluding such selection criteria impedes the discrimination of crosslink repair status by mutation analysis. Applied to all KEGG pathways, we find that the association with MMC sensitivity is strongest in the KEGG FA pathway, therefore also demonstrating the value of such selection strategies for exploratory analyses. Variant analyses in 56 clinical samples demonstrate that homozygous variants occur more frequently in tumor suppressor genes than oncogenes further supporting the role of a homozygosity criterion to improve gene function association or tumor suppressor gene identification studies.

Conclusion

Together our data show that the detection of relevant genes or of repair pathway defected tumor cells can be improved by the consideration of allele zygosity and SNP allele frequencies.

Introduction

Recent large-scale sequencing efforts stimulated oncology research and revealed a multitude of novel genomic alterations and somatic mutations in various tumor types [1–3]. These genetic studies are driven by the need to understand the processes related to tumor development and treatment response with the ultimate goal to find therapeutic biomarkers and targets for novel therapeutic approaches [4,5]. A major challenge in the clinical translation of these results is the discrimination of genetic alterations with a functional impact from the vast number of detected alterations.

While high gene mutation frequencies in tumors can point to potential oncogenes, defining a role for tumor suppressor genes (TSG) can be challenging. This is because TSG variants with a functional impact likely affect pathway performance only when the wild-type allele is lost. Variants in DNA repair genes, archetypal TSGs, are particularly difficult to evaluate. Many genes are involved in the repair of DNA damage and determine cellular survival following DNA damage. Mutations in any of these genes could influence cellular outcome following DNA damage. The multitude of genes and variants therefore hampers gene mutation detection as it is difficult to identify among them the affected gene or the pathway disrupting variant. Attempts to deduct DNA repair defects from genetic data for example for treatment response analyses suffer from the ignorance of the functional impact of many of the gene variants. Experimental validation of the functional impact of the individual variants is however time-consuming and costly. Rarity may also discourage thorough characterizations of the individual variant. An important role of genetic DNA repair defects in tumors could therefore been masked by a multitude of different, rare but functionally important, variants across many genes [6]. Computational tools that help to prioritize the potential functional relevance of genetic alterations are therefore of great importance. Many different tools, comprehensively reviewed in Eilbeck et al. and compared by Mahmood et al. [7,8], have been developed. They use different tactics and are valuable tools to assess a variant's probability to affect gene function. Algorithms that predict the effect of non-synonymous variants often consider the degree of conservation of homologous sequences and how disruptive an amino acid change is based on its physical properties [9,10]. However, these approaches may not suffice and in the presence of wild-type alleles such annotations are inapt to predict cellular pathway performance such when applied to tumor suppressor or DNA repair genes.

Here we propose a simple, complementary strategy that combines the zygosity status and allele frequency of variants in order to enrich for variants that label genes or tumor cells with a functional DNA repair defect. Germline *BRCA1/2* variants that predispose to breast cancer illustrate the potential value of such approaches. Pathogenic *BRCA1/2* variants manifest their deleterious potential after loss of the wild-type allele, a process known as loss of heterozygosity (LOH) [11]. Homozygosity can thus be a selection criterion to enrich for variants with a functional impact, in particular those causing a loss of function or that mark LOH events. Pathogenic *BRCA1/2* variants occur at very low allele frequencies and are therefore often annotated in

dbSNP. Variant selection protocols that remove dbSNP annotated variants in order to identify somatic mutations, also remove rare pathogenic variants [12]. To prevent this some studies select only variants that are very likely to impair gene function, such as nonsense and frameshift mutations. Other selection criteria question whether variants have a confirmed association with familial breast cancer. Such associations do, however, not reflect cellular DNA repair performance well; an endpoint relevant to treatment response association studies, in particular to targeted drugs such as to PARP inhibitors [13–15]. This is partly due to the rarity of these pathogenic variants. Also, hypomorphic gene mutations among rare single-nucleotide polymorphisms (SNPs) can influence DNA repair and response to a degree that affects drug response but not cancer incidence rates. These hypomorphic gene mutations as well as low penetrance variants will be removed by variant selection protocols that solely retain confirmed pathogenic variants. The retention of variants with low population frequency (rare SNPs) could address these issues and improve repair outcome association.

Here we set out to test the benefit of allele zygosity and SNP allele frequencies filters. Selection of homozygous variants and rare SNPs as selection criteria have been considered previously. These studies were either restricted to retrieving known cancer-associated variants [12] or performed exploratory analyses in multidrug screens without verifying causality [16]. In an effort to validate these variant selection criteria, we apply them in a setup that provides a functional link to the phenotypic effect of the selected, i.e. variant-marked, tumor cell lines. We focus on genes in the Fanconi anemia (FA) and homologous recombination (HR) pathways that are required to repair DNA crosslinks. Sensitivity to the DNA crosslinker Mitomycin C (MMC) drug is a hallmark of FA pathway defects and resulted in the identification of multiple FA pathway genes [17–20]. DNA interstrand crosslinks (ICL) pose critical obstacles in replication and this activates the FA core complex that is composed by ten FA pathway proteins. Together they function as a ubiquitin E3 ligase to mono-ubiquitylate FANCD2-I. This activation recruits endonucleases to cleave the DNA and allow translesion synthesis at the ICL affected DNA, ultimately however requiring members of the homologous recombination repair pathway to finalize ICL repair. Together, the FA and HR pathway has an important role in resolving mitomycinC induced ICLs. We therefore performed capture sequencing and variant calling on FA/HR genes in a panel of 29 patient derived head and neck squamous cell carcinoma (HNSCC) cell lines, of which a proportion was previously shown to have FA/HR pathway defects [21]. The defects were revealed by MMC response data and confirmed by additional crosslink repair function parameters [21]. The MMC response data provide a robust functional readout for crosslink repair that allowed us to assess the value of combined allele zygosity and SNP allele frequency filters to detect cellular repair defects and/or relevant genes.

Methods

Cell line panel

The following HNSCC patient derived cell lines were generated at the Turku University Hospital Finland between 1990-2002: UT-SCC-1A, UT-SCC-2, UT-SCC-4, UT-SCC-7, UT-SCC-8, UT-SCC-9, UT-SCC-12A, UT-SCC-14, UT-SCC-15, UTSCC-16A, UT-SCC-20A, UT-SCC-24A, UT-SCC-24B, UT-SCC-27, UT-SCC-30, UT-SCC-32, UT-SCC-36, UT-SCC-38, UT-SCC-40, UT-SCC-42A, UT-SCC-43A, UT-SCC-45, UTSCC-54C, UT-SCC-60B, UT-SCC-76A, UT-SCC-77, UT-SCC-79A and UT-SCC-90. The conditions under which these cell lines were cultured have been described previously [22,23]. The HNSCC cell line NKI-SCC-263 was established at the Netherlands Cancer Institute. Two FA patient derived fibroblast cell lines were provided by Dr. H. Joenje (Vrije University Amsterdam). Our sequence analyses confirmed the reported FANCG 1649delC and FANCA Arg951Gln mutations in these FA patient fibroblast line (EUFA636 [24] and EUFA173 [25]). These FA pathway mutated cell lines served as positive controls and provided the reference values in the MMC response data. These cell lines were not considered in the variant selection criteria assessment studies. Notably, all cell lines were cultured and tested under low oxygen (5%) conditions since high oxygen conditions affect cellular growth and fitness of repair defected cell lines. The negative control cell line was an hTERT transformed human fibroblast cell line (GM847), provided by Roderick Beijersbergen (the Netherlands Cancer Institute). Mitomycin C (MMC; Sigma Aldrich) sensitivities have been determined as described previously [21]. In brief, cell doublings were assessed and cellular survival after treatment with different concentrations of MMC was determined by live cell counting after multiple divisions (minimum 5) in a long-term growth assay. Controls were included with lower cell densities that tested and assured linearity in these assays. Survival was determined relative to untreated cells and the MMC concentration resulting in 50% survival (MMC IC_{50} value) was calculated for each cell line from third order polynomial curve fits on the growth inhibition curves of the individual independent experiments [14,21]. Data are from three to five independent experiments per cell line with 3 to 6 replicates each. Mean MMC IC_{50} values were used in analyses. MMC concentrations were adapted in the individual cell lines to assure a good coverage of data points in particular in the IC50 to IC90 inducing MMC dose range.

Patient samples

Tumor samples were obtained after documented informed consent. Consent forms were approved by the medical ethical committee of the Netherlands Cancer Institute. Use of the material for this genetic study was approved by the institutional ART-CFMP biobank review board. Tumor samples are from 56 patients with advanced head and neck squamous cell carcinoma (HNSCC) that were enrolled in our hospital between 2001 and 2010 and obtained from fresh frozen pretreatment biopsies. Matched blood samples were not available for these samples.

DNA capture and sequencing

Genomic DNA was isolated using the AllPrep DNA/RNA Mini Kit (Qiagen). Paired-end fragment libraries were prepared with the TruSeq DNA library preparation kit (Illumina) and target-captured with a SureSelect custom-based bait library (Agilent) targeting 556 genes (Supplementary Table A). The capture covers 1.9 Mb with about 24000 probes. Baits covered all exons (average coverage 1.5) and UTR and TSS were additionally covered in the canonical FA/HR pathway genes. DNA was washed, amplified, barcoded, pooled and sequenced on the Illumina Hiseq 2000 using a 2x75bp paired-end protocol. Sequencing reads were aligned to the human reference genome (GRCh37.55/hg19) with the backtrack algorithm of the Burrows-Wheeler Aligner 0.5.10 [26]. Potential polymerase chain reaction duplicates were removed with Picard Tools (http://picard.sourceforge.net). The average read coverage was 247. Blood samples from four healthy volunteers were sequenced to depict potential capture errors and bias and to monitor sequencing noise.

Variant calling

Variants, both single nucleotide variants and small insertions and deletions, were called with VarScan 2.3.9 [27] in conjunction with Samtools mpileup 0.1.19 [28]. VarScan's VarFreq (divided by 100) provided the variant allele fraction (VAF) values used in this study and represent the fraction of reads that show the individual variants. Single nucleotide variants were called when allele coverage was at least ten, the number of variant reads at least four, and the VAF at least 0.10. The same parameters were used for calling small insertions and deletions, except the minimum number of variant reads was set to ten with no VAF restriction. Single nucleotide variants adjacent to small insertions and deletions were removed. We used Annovar (version date 11-05-2016) [29] to annotate variants with the RefSeq and 1000 Genomes European august 2015 (Phase 3) databases [30]. SNPs refer to variants with an associated minor allele frequency (MAF) in the 1000 Genomes European database. All analyses and figures were restricted to non-synonymous exonic variants and essential splice site variants throughout this paper. Variants that occurred in all four healthy volunteer blood samples were removed as they were considered potential artifacts.

Statistical analysis

In order to assess the value of the selection criteria to detect cellular crosslink repair defects, we compared 'pathway mutated' with 'non-mutated' cell lines in various pathways and gene sets. 'Pathway mutated' cell lines had one or more variants (potential mutations) in any of the pathway genes, the 'non-mutated' had none (Supplementary Figure A). The association between pathway mutation status and MMC response was evaluated with the Wilcoxon rank-sum test and visualized by the Positive Predictive Value (PPV). The Wilcoxon rank-sum test compares the (ranked) MMC IC_{50} values of 'pathway mutated' and 'non-mutated' cell lines. The PPV was calculated as the fraction of 'FA/HR pathway mutated' cell lines present among the ten

most MMC sensitive cell lines (Supplementary Figure A). The number of LOH events across TSGs and oncogenes was compared with the Wilcoxon rank-sum test. Fisher's exact test was used to compare the proportion of heterozygous and homozygous variants between TSGs and oncogenes. All statistical analyses were performed in the R environment for statistical computing.

Results

Variant allele fraction filtering improves functional outcome association

Here we set out to investigate the association between the zygosity status or allele frequency of a variant and its ability to mark a phenotypic effect. To this end we employed a panel of 29 patient-derived HNSCC cell lines. We employed response to the DNA crosslinker MMC as a hallmark and readout of cellular DNA repair proficiency (Supplementary Figure B). Tumor cell DNA repair capacity is the essential functional endpoint in cancer treatment response or DNA repair targeting strategies studies. Cellular MMC response is strongly influenced by the activity of the FA and HR gene products [17–20]. Assay performance and MMC sensitivity to repair relation was confirmed by the revealed MMC hypersensitivity of FA patient fibroblast controls with confirmed FA gene mutations and lack thereof in wild-type fibroblasts (Supplementary Figure B). Capture sequencing data on 556 genes, which included the canonical FA/HR pathway genes (Supplementary Table A), was available and we performed variant calling with different parameters on those genes.

We first factored out the allele frequency by removing all SNPs present in the 1000 Genomes database and focused on the role of the zygosity status. To this end we employed variant allele fraction (VAF) as a measure of allele zygosity, with high VAF corresponding to homozygosity. The VAF distribution of all non-synonymous variants reflected allele zygosity well in these cell lines (Supplementary Figure C). We varied VAF thresholds from low to high and at each threshold value removed variants with a VAF below the threshold. Thus, each VAF threshold resulted in a unique set of variants being retained. A cell line was assigned as 'FA/HR pathway mutated' if any FA/HR gene variant in that cell line was retained under the tested threshold. Increasing VAF thresholds gradually reduced the number of retained variants and number of cell lines with a 'FA/HR pathway mutated' assignment (Supplementary Figure D). We then tested whether the MMC IC_{50} values showed an association with the 'FA/HR pathway mutated' status. Specifically, we performed the Wilcoxon rank-sum test and calculated the Positive Predictive Value (PPV), i.e. the fraction of 'FA/HR pathway mutated' cell lines present among the ten most MMC sensitive cell lines. As control we repeated our analysis on 10.000 random gene sets, each having cumulative sequence lengths similar to the FA/HR gene set ('Similar sized gene sets'). The association between 'FA/HR pathway mutated' and MMC sensitivity increased and reached significance with

higher VAF thresholds (Figure 1A and Supplementary Figure D). In contrast, the PPV of the control sets remained constant at 0.3, indicating a random assignment hence a lack of discrimination power (Figure 1A and Supplementary Figure E). This indicates that the specificity of FA/HR variants to detect MMC hypersensitive, and therefore repair-defected, cell lines improves with increasing homozygosity.

Next, we assessed indirect factors that could have caused or confounded the here observed association between homozygosity of FA/HR variants and MMC sensitivity. The apparent association with homozygosity of the FA/HR variants could be an indirect consequence of an increased total or homozygous variant load in these MMC sensitive cell lines. However, neither the total nor the homozygous variant load was associated with MMC sensitivity (Supplementary Figures F and G), regardless whether SNPs were included or excluded. Homozygous FA/HR variants could act as 'tags' for other homozygous variants in genes in the proximity of the FA/HR genes (Supplementary Table A). However, we find that variants in neighboring genes were not associated with MMC sensitivity (Figure 1A and Supplementary Figure E, "Nearest genes"). LOH events, across a larger region neighboring and containing the FA/HR genes, are thus unlikely to have caused the association between the homozygosity of FA/HR variants and MMC sensitivity. These results show that the improved MMC response association by these filters is FA/HR gene-specific. Together, our results support homozygosity as a variant selection criterion.

Rare SNP filtering improves functional outcome association

We next investigated the benefit of considering the allele frequency of FA/HR variants for DNA repair defect marking. We factored out the variable zygosity status by only using homozygous variants. This analysis was similar to the previous, except that variants were now filtered according to their allele frequency in the 1000 Genomes database instead of their VAF in the cell line sample. We did so by varying the allele frequency, also known as the MAF, threshold (Supplementary Figure H). The number of variants and cell lines with a 'FA/HR pathway mutated' assignment decreased pronouncedly at the MAF threshold value of 0.1. Notably, the association between FA/HR variant marked cell lines and MMC sensitivity increased and reached significance with lower MAF thresholds (Figure 1B, Supplementary Figures H and I). This effect was neither seen in the 'nearest genes' nor in the 10.000 similar sized gene sets.

We further studied the interaction between zygosity and allele frequency by filtering on both the VAF and the MAF. We then evaluated the association of the cell lines classified as repair defected by the remaining variants with MMC sensitivity. This analysis shows that the PPV increases proportionally to the VAF threshold, though only after common SNPs have been removed, i.e. at lower MAF thresholds (Figs 1C and 1D). Multiple of the VAF and MAF threshold combinations show a significant association with MMC sensitivity, further illustrating the value of a combined approach (Supplementary Figures J and K). In addition, when assessed in the two FA patient derived fibroblast cell lines, these filters identified the reported FANCG 1649delC and

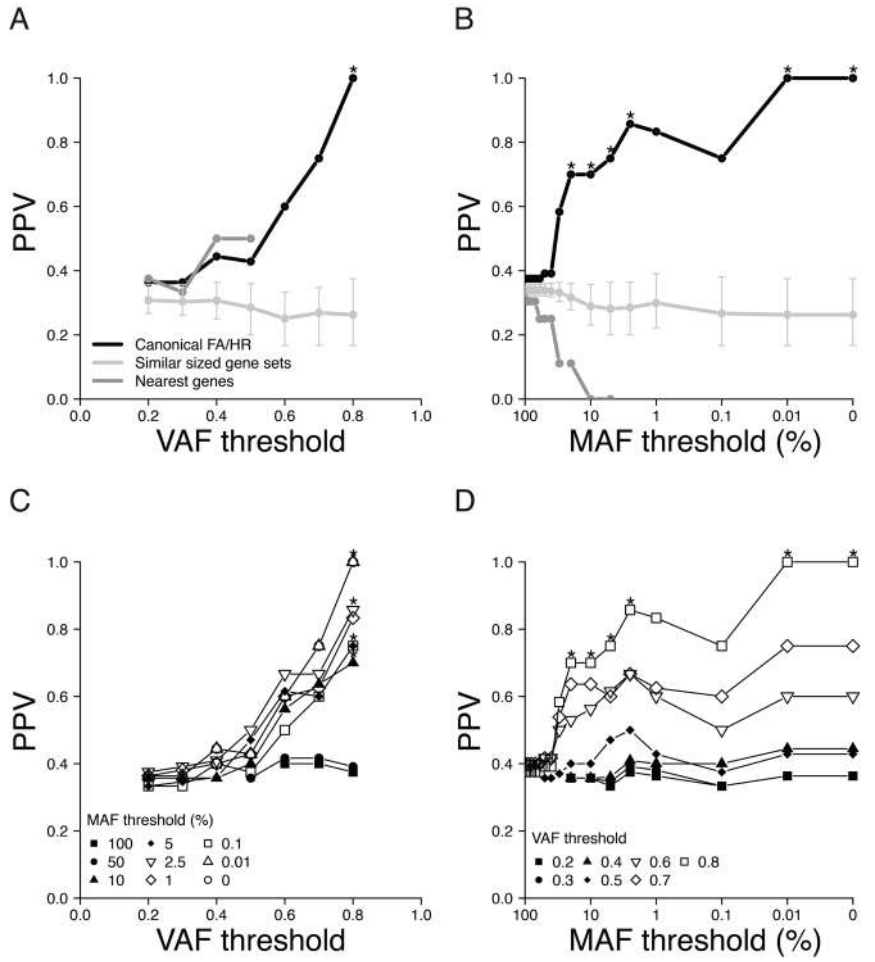


Figure 1. High VAF and low MAF selection criteria improve functional association.
The positive predictive value (PPV) for MMC sensitivity was used to quantify the ability of variants to mark repair defected cell lines, i.e. ten most MMC sensitive, by the different variant selection criteria. Statistical analyses were omitted at sample sizes of a group of less than two and these data points and lines have been excluded in the figures. Asterisks mark a significant association with MMC response. **(A)** PPV values at each VAF threshold that was applied for variant selection are shown. Lines show the results for the canonical FA/HR genes and of the two controls: the "nearest genes" and the 10.000 similar sized randomly grouped gene sets. Error bars delineate the first and third quartile from the median in the latter. **(B)** PPV with progressively decreasing maximum MAF thresholds. Line coloring is identical to Fig 1A. **(C)** Impact of canonical FA/HR gene variant selection filters on PPV when combining multiple maximum MAF thresholds (as indicated in in-figure legend) with increasing minimum VAF thresholds (x-axis). **(D)** Influence of filters on PPVs after applying decreasing maximum MAF thresholds for multiple minimum VAF thresholds for canonical FA/HR gene variant selection as indicated.

FANCA Arg951Gln mutations. In summary, these results show that the variant selection criteria allele zygosity and allele frequency can be combined to enrich for variants that mark cell lines with a loss of cellular repair function. Selecting homozygous variants and rare SNPs detects variants with a link to loss of pathway function in our dataset, and thus also helps to prioritize variants for further experimental validation.

VAF and rare SNP filtering improve pathway identification

The previous analyses focused on the FA/HR pathway, because defects in FA/HR genes are well-known to impair the repair of crosslinks induced by MMC. In practice, the aim of genetic analysis is often to reveal novel associations with drug response, e.g. in pharmacogenomic interaction studies [5,31,32]. Given the importance of the FA/HR pathway for the removal of MMC induced crosslinks, the prioritization of this pathway over others by variant selection strategies that consider homozygosity and rare SNPs would confirm their value.

To this end we broadened the analysis to include all genes annotated in all Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. Based on the results of the previous section we selected homozygous variants (VAF ≥ 0.8) and included rare SNPs (MAF ≤ 1%, a threshold commonly used to define rare variants). We scored a pathway as 'mutated' when any of the genes annotated to that pathway were 'mutated', i.e. has a variant as selected by the above criteria. For each pathway, the MMC IC_{50} values of 'pathway-mutated' cell lines were compared to those of 'non-mutated' cell lines. Some mutational profiles of pathways correlated strongly (Pearson correlation $\rho \geq 0.75$) because they were only represented by a few overlapping genes in our capture-set. In these cases the pathway with the weakest MMC response association was removed. The results of this analysis are represented in Figure 2A. We find that the Fanconi anemia (FA) pathway is the only pathway that is significantly associated with MMC sensitivity.

We then compared this analysis to an approach that includes all non-synonymous variants regardless of zygosity status, i.e. performs no VAF filtering. At the same time, variants and SNPs with a MAF ≤ 1% were retained. This approach was not able to reveal an association between the Fanconi anemia pathway and MMC sensitivity. This result is consistent with the requirement of loss of the wild-type, the functional DNA repair pathway gene allele, to impact cellular repair. It thereby confirms the benefit of using zygosity as a variant selection criterion to depict repair defected cells. The analysis did, however, reveal significant associations between four other pathways and MMC resistance (Supplementary Figure L). Upon further examination we found that two of these associations were mainly driven by falsely assigned *TP53* mutation status in the cell lines. Two *TP53* mutations were missed in MMC sensitive cell lines; one point mutation remained undetected because of low read depth and one known medium sized deletion was not picked up due to limitations of short-read sequencing data (Supplementary Table B) [22,33–37]. When correcting the *TP53* mutation status these two associations disappear (Figure 2B). The other two pathways associated with

MMC resistance were the Hippo signaling and mineral absorption pathways, while the Rap1 signaling pathway was associated with MMC sensitivity (Figure 2B).

We next compared homozygous and rare SNP variant selection strategies to standardly used variant selection methods such as REVEL and CADD. Variants in FA/HR genes that were selected by these methods did not show an association with the functional endpoint (MMC response) at any threshold (Supplementary Figure M). These variant selection methods did also not reveal a FA pathway link in the broader pathway analysis (Supplementary Figure N). Our proposed homozygous and rare SNP variant selection strategies however did prioritize the FA pathway over all other KEGG pathways (Figure 2A), since it had the strongest association with MMC response. As the FA pathway has an established role in cellular MMC response, this supports the value of our variant selection strategy for exploratory analyses.

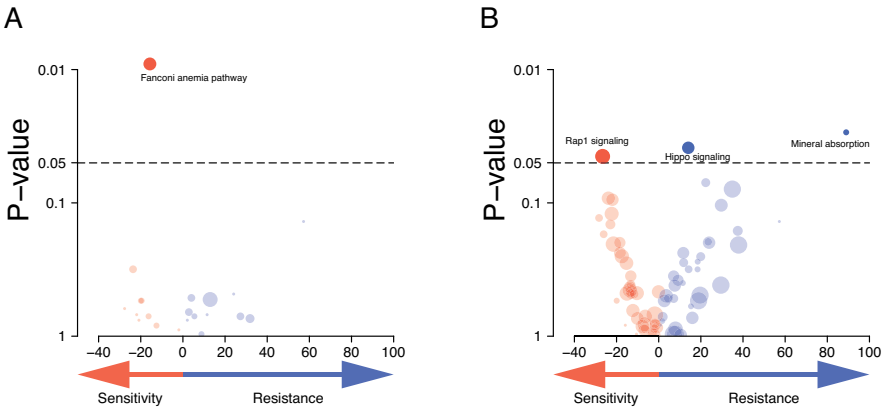


Figure 2. Homozygous and rare SNP variant selection criteria identify the KEGG Fanconi anemia pathway to be strongly associated with MMC response.

Volcano plot showing the significance of the associations between MMC response and 'pathway-mutation' classification in the individual KEGG pathways as determined by the presence of homozygous and rare SNP variants in these pathway genes. The x-axis shows the difference in mean MMC IC_{50} between 'pathway mutated' and 'pathway non-mutated' cell lines. Pathway mutated cell lines are those with one or more variants in any gene of the individual KEGG pathway. The y-axis shows the significance values (Wilcoxon rank-sum test *p*-value) of the difference that were found in the MMC IC_{50} between 'pathway-mutated' and 'non-mutated' cell lines. Dot size is proportional to the number of 'pathway-mutated' cell lines, dot color intensity proportional to significance value. **(A)** Results after applying the variant selection strategy that selects homozygous variants (VAF ≥ 0.8) and includes rare SNPs (MAF ≤ 1%). **(B)** Results of analyses that select variants regardless of zygosity status, including rare SNPs (MAF ≤ 1%).

Tumor suppressor genes are enriched for LOH and homozygous variants

After focusing on DNA repair genes, we next investigated TSGs in a broader sense and assessed the potential value of the homozygosity criterion for tumor suppressor gene identification in patient tumor samples. Sample tumor purity, intra-tumor heterogeneity and (aneu)ploidy hamper the ability to distinguish between heterozygosity and homozygosity based on the VAF values in tumor samples [38]. We therefore questioned whether selecting homozygous variants in tumor samples would be able to enrich for tumor suppressor genes. Homozygosity, often through the loss of a functional wild-type allele, is one of the means to accomplish the loss of the cellular function of a gene, a requirement common to TSGs. Conversely, the activating mutations in oncogenes do not necessarily require homozygosity to have a functional impact. We therefore tested whether known TSGs were enriched for LOH events/homozygous variants in patient tumor samples.

To this end, we applied the same variant calling pipeline that was used for the HNSCC cell line panel to capture sequencing data from 56 HNSCC tumor samples. Rare non-synonymous SNPs were included (MAF ≤ 1%). We find that the VAF distribution has three peaks in the tumor samples, in contrast to the bimodal VAF distribution of the HNSCC cell line panel (Supplementary Figure O). It illustrates how the contribution of tumor purity, intra-tumor heterogeneity and ploidy to the VAF values complicates the distinction between heterozygous and homozygous variants based on VAF values. We therefore used the algorithm PureCN [39] to identify genes and variants that underwent LOH. PureCN employs VAF and read coverage to produce an improved estimate of the allele specific copy number and thus LOH. Six documented HNSCC TSGs and ten oncogenes (Supplementary Table C) were used to narrow the analysis [40]. First, we determined the number of samples with LOH per gene (Figure 3A). LOH occurred five times more often in TSGs than oncogenes (average of 27 LOH events across TSGs versus 5 across oncogenes; $p < 0.005$; Figure 3B). Second, we pooled the variants of all samples in the TSGs and oncogenes and calculated the fraction of gene variants that was homozygous. The fraction of homozygous variants was six times higher in the TSGs than in the oncogenes (37 / 68 versus 1 / 11; $p < 0.005$).

The results show that in the HNSCC tumor samples, LOH and homozygous variants are more common in TSGs than oncogenes. This is in line with the requirement of loss of the wild-type allele for TSGs. These data indicate that selecting homozygous variants and rare SNPs could indeed enrich for genes or variants linked to a loss of function event. This variant selection strategy is therefore expected to improve associations with functional endpoints.

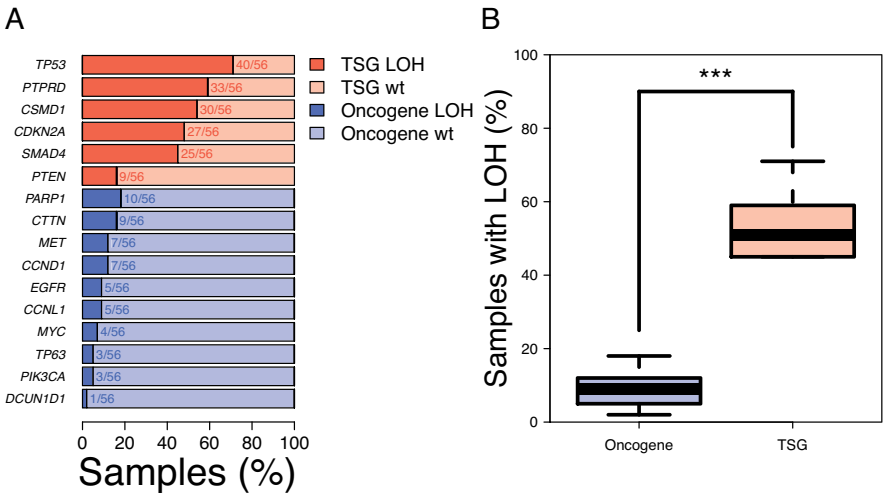


Figure 3. Loss of heterozygosity events are more common in tumor suppressor genes. **(A)** The percentage of HNSCC samples with LOH per TSG (dark red) and OG (dark blue). HNSCC TSG and oncogenes (OG) are as reported by Leemans et al. [40]. Light colors represent the percentage of samples without LOH (wt). Fraction of samples with LOH are indicated by numbers and dark colors. **(B)** Boxplot representation of the percentage of HNSCC tumor samples with LOH in any TSGs or OGs as shown in Fig 3A. TSGs are enriched for LOH events ($p < 0.005$).

Discussion

Here we tested whether variant selection strategies that consider allele zygosity and allele frequency help to mark tumor cells with functional defects. To this end we used 29 HNSCC cell lines with confirmed DNA crosslink repair defects as assessed by functional assays. We applied these strategies to the selection of variants in FA/HR pathway genes that govern crosslink repair and found a benefit of selecting homozygous variants and of retaining rare SNPs. The results show that the presence of homozygous variants and rare SNPs in the FA/HR genes is associated with functional outcome, i.e. repair defect, in the cell lines. This highlighted the potential to mark such defects in tumor cells. To demonstrate the value of our approach for exploratory analyses, we extended it to all KEGG pathways. We found that the functional outcome association is strongest in the KEGG Fanconi anemia pathway. This exemplifies the value of the variant selection criteria to identify genes or pathways relevant to a given functional endpoint. Finally, clinical application feasibility was demonstrated in the HNSCC patient tumor analysis. The enrichment of LOH events and homozygous variants in TSGs indicates improved functional association, further highlighting the potential benefit of such variant selection strategies.

Cancer genomics studies often use SNP databases to remove germline variants, but seldom use them to incorporate rare SNPs in their analysis [12]. Guiding variant selection by allele zygosity is even less customary in cancer genomics, even though this approach has successfully been used to identify pathogenic germline variants in congenital disorders [41,42]. A possible reason why this approach was underused in cancer research, is the complexity of accurately calling homozygous variants caused by LOH in tumors due to intra-tumor heterogeneity, aneuploidy and normal cell admixture [38]. Recently developed algorithms have made progress in meeting this challenge, and the identification of homozygous variants caused by LOH in tumors is now possible [38,39]. Indeed recent studies further point to the importance of considering LOH events and/or homozygosity for marking tumors with HR defects [43,44].

The selection of rare SNPs has been proposed before and was applied previously in order to capture known tumor-associated variants [12]. Since the authors acknowledged the importance of LOH, zygosity and rare SNP criteria have been also applied in attempts to identify genetic DNA repair defects [16]. However, their validity or effectiveness in depicting repair defects has formally never been tested nor confirmed. Here we therefore investigated whether such selection criteria would improve the detection of cell lines with a functional defect or the identification of the relevant genes in repair defected cell lines. In comparison to previous pharmacogenomic projects [5,31,32,45] our cell line panel data provides a unique and reliable model system for studying the effect of variant selection methods on functional outcome association. This is due to a relatively large number of uniformly treated cell lines of the same cancer type and the choice of a robust functional read out, i.e. long-term growth assays to determine MMC IC_{50} values. The commonly used short-term survival assays often reflect a mixture of growth delay and kill. They are also affected by the apoptosis proficiency of a cell line. In contrast, long-term assays reflect the overall repair proficiency, and therefore survival, better. Our chosen drug-doses also ensured that MMC IC_{50} values were reached experimentally rather than estimated by extrapolation. This is a possible improvement to the approach of others [31,32], whose IC_{50} values are discordant [46]. Unfortunately, it was not possible to extend our analysis to the publically available cell line data due to various reasons: lack of publicly available HNSCC cell line data [45], robust MMC response data [32], or DNA sequence information of the canonical FA/HR pathway genes [31]. Two resistant cell lines showed ATP7A variants, while the Hippo pathway depiction was mostly driven by APC and CTNNB1 gene variants and this points to a wnt signaling pathway contribution. A link to platinum drugs has been reported for ATP7A [47]. The validation of the possible role of the wnt signaling pathway in determining MMC resistance is however inhibited by the lack of long-term MMC survival assay data in independent HNSCC cell line panels.

While our analyses showed how the selection of non-synonymous variants, homozygous and rare as SNPs, can identify the genes and pathways responsible for the functional endpoint in question (here the FA/HR pathway for MMC response), it does not show that the selected variants cause the repair defect. This may be the case for a fraction of the selected variants. Others, however, may have simply marked LOH events in the respective genes. As rare SNPs are unlikely to be present

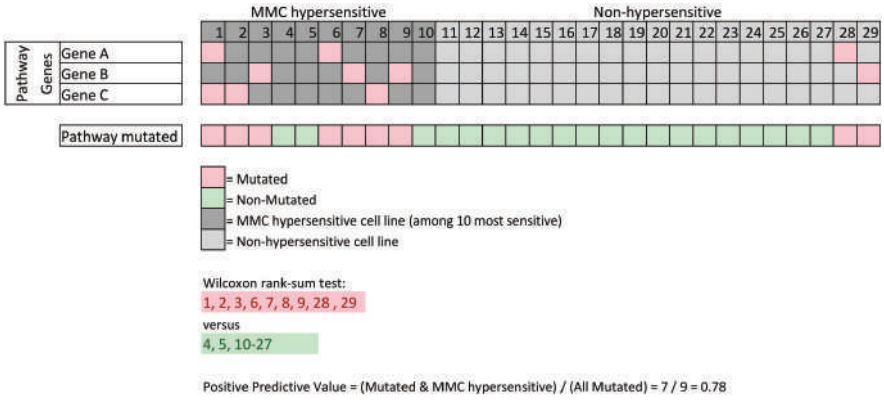
in a homozygous state, selection of homozygosity will therefore mark LOH events and point to gene alterations that are more directly related to the repair defect / functional endpoint. Yet, with respect to the identification of variants with a functional impact, a similar selection procedure did enrich for variants with an annotation in COSMIC or predicted to alter protein function by other algorithms in our previous study [21]. After accounting for the individual stroma component in patient tumor samples, it also enabled a patient outcome association that was not evident with regular selection criteria as shown by our data in Verhagen et al. [21]. Together our data also suggest that variant selection should not solely be guided by pathogenicity criteria (derived from cancer incidence data as by COSMIC annotation) since this results in a lack of hits in our repair defected HNSCC cell line panel. The potential impact on protein function of the selected variants will have to be evaluated in future studies. This is a massive endeavor considering the multitude of variants and questionable in value due to the infrequency of the individual variants. This points to the importance of variant selection strategies that provide an increased link to functional endpoints on a cellular level so to be able to perform valuable outcome association studies.

Conclusion

Assuming to be a prerequisite but without proof of concept, homozygocity of DNA repair gene variants has been used to mark tumor samples with potential defects. Here we tested this in the genetic context of tumor cells and show its benefit and association with repair endpoints. The inclusion of rare homozygous SNPs to mark potentially repair defected cell lines further improves this association and highlights their potential role in genomic studies. We conclude that allele zygosity and SNP allele frequency selection criteria can be used to identify FA/HR repair-defected samples and relevant genes.

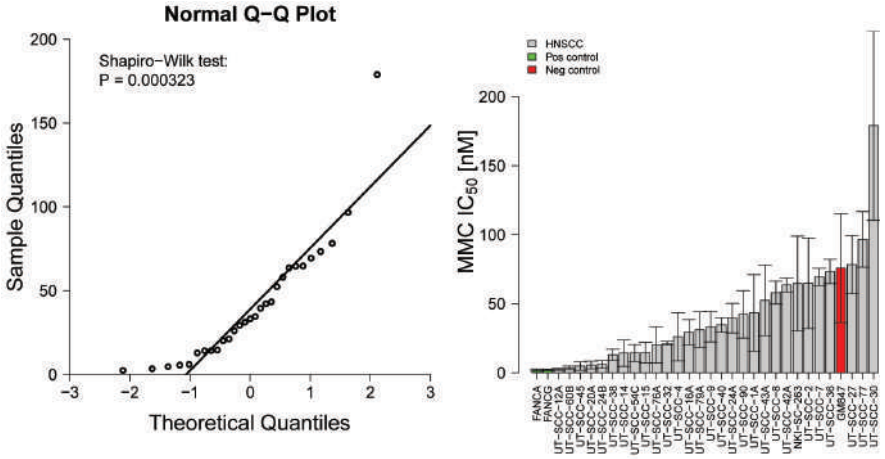
Acknowledgements

We are grateful for advice and support by Floor Hageman, Magali Michaut and Paul Essers. We would like to acknowledge the NKI- AVL Core Facility Molecular Pathology & Biobanking (CFMPB) for supplying NKI-AVL Biobank material.

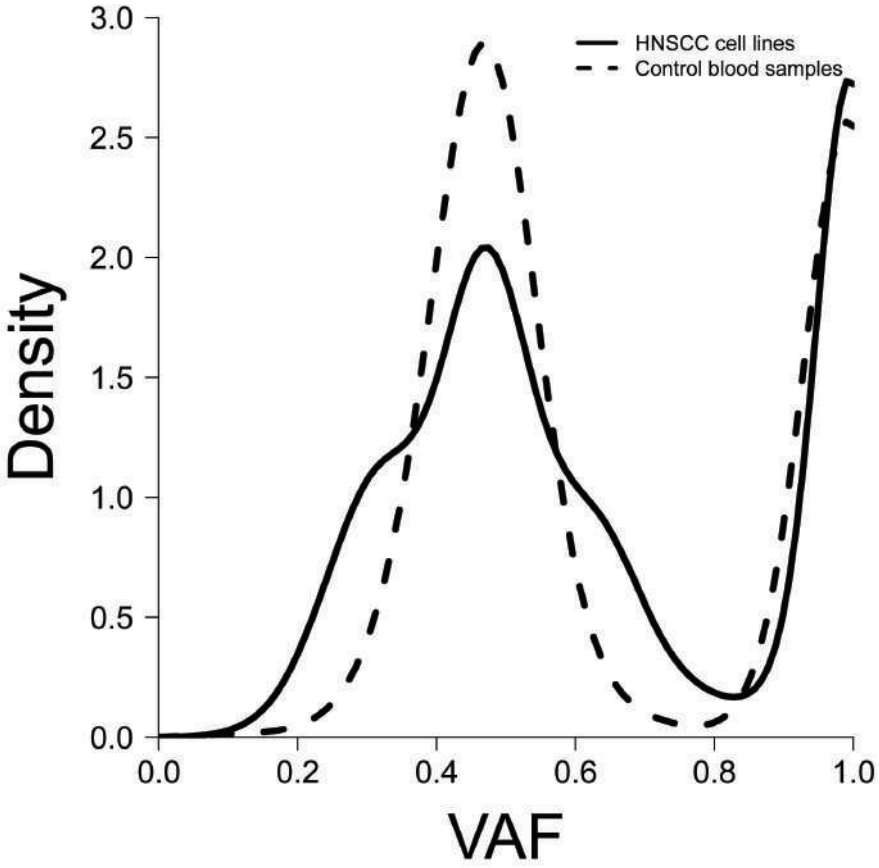


Supplementary Figure A. Illustration of the association analysis between pathway variants and MMC response.

We compared 'pathway-mutated' with 'non-mutated' cell lines. 'Pathway-mutated' cell lines had one or more variants in any of the pathway genes under investigation (pink), the 'nonmutated' had none (green). The Wilcoxon rank-sum test was used to compare the (ranked) MMC IC_{50} values of 'pathway mutated' and 'non-mutated' cell lines. The PPV was calculated as the ratio of true positive calls over all positive calls. Positive calls are 'pathway-mutated' cell lines as marked by the different variant selection criteria; true positive calls were 'pathway-mutated' cell lines that were also among the ten most MMC sensitive cell lines (hypersensitive cell lines in dark grey).

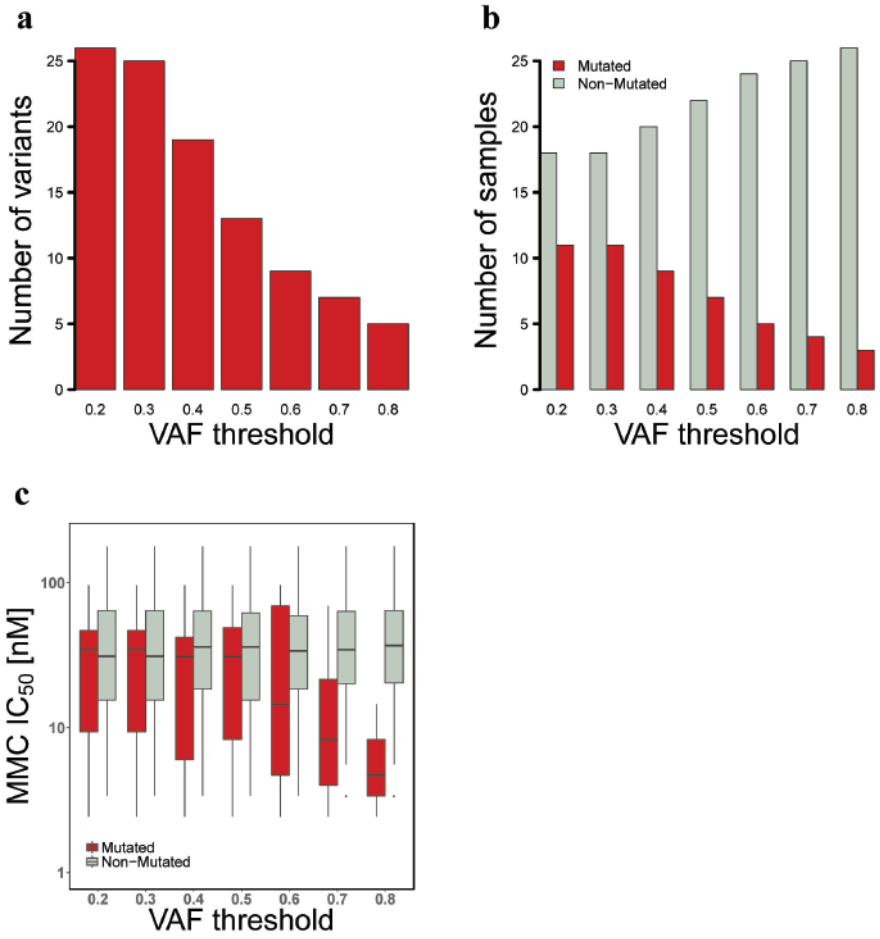


Supplementary Figure B. MMC IC_{50} values distribution in the HNSCC cell line panel. **(A)** The distribution of the MMC IC_{50} values (as previously reported by us in Verhagen et al. [23] was assessed. The inlet Q-Q plot and stated significant Shapiro-Wilk test illustrates the non-normal distribution of the HNSCC cell line MMC IC_{50} values. **(B)** MMC IC_{50} values. The positive control cell lines with confirmed FA gene defects are depicted in green (FANCA, FANCG), the negative control normal fibroblast line in red. MMC response of the HNSCC cell lines (grey bars) varied from hypersensitive (similar sensitivity as the positive control) to normal sensitivity (similar sensitivity as the negative control). Values are calculated from the curve-fits on the individual experiment data and are the average of three to five independent experiments; bars are SEM.



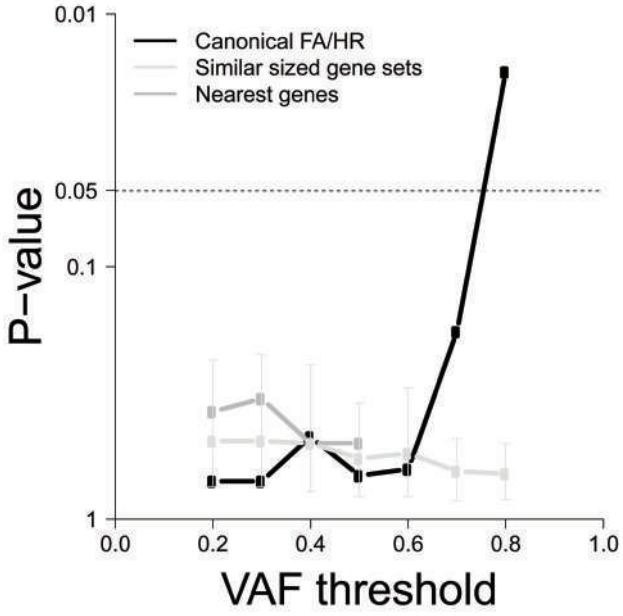
Supplementary Figure C. Average allele zygosity is reflected in the VAF distribution of variants.

Density plot of the VAF distribution of all non-synonymous exonic and splice site variants in all captured genes (both SNP and non-SNP). The VAF distributions of the HNSCC cell line panel and control blood samples both show a bimodal distribution of hetero- and homozygous variants; the latter shows a narrower distribution. The minimum in the curve appears around 0.8. The threshold for homozygous variants was therefore set at 0.8.



Supplementary Figure D. The effect of VAF filtering on the number of FA/HR gene variants and 'mutated' cell line calling.

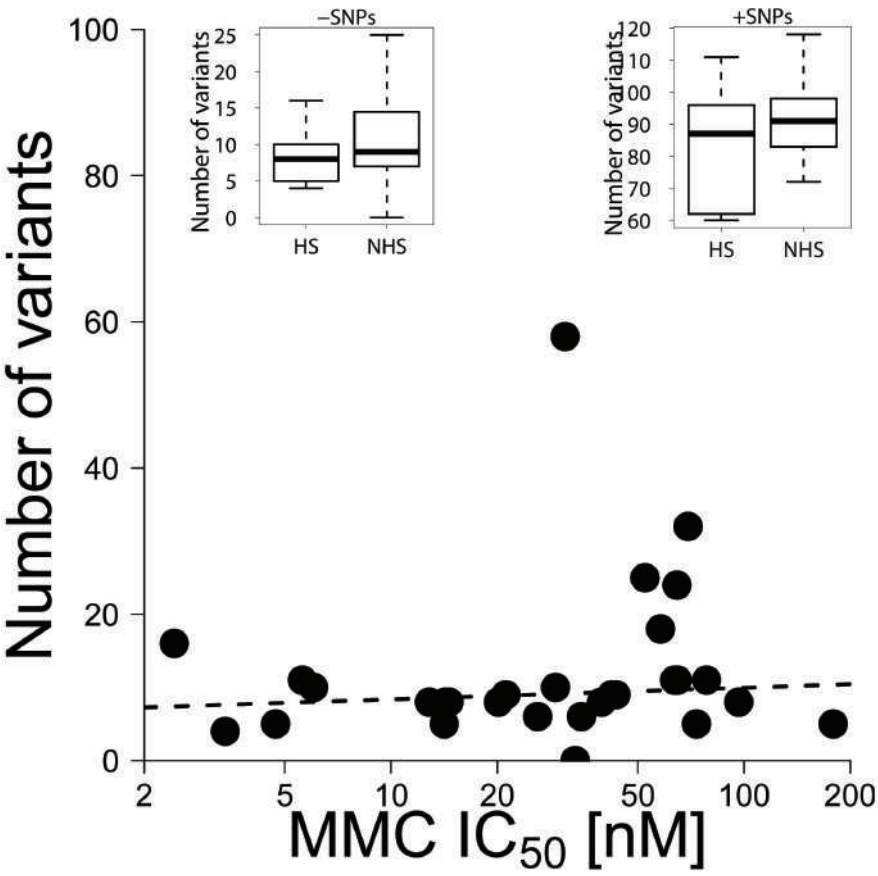
We varied VAF thresholds from low to high and at each threshold value removed variants with a VAF below the threshold. At each VAF threshold we depict **(A)** the total number of variants exceeding the threshold, **(B)** the number of cell lines with a FA/HR variant exceeding the threshold ('mutated'), and **(C)** the MMC IC_{50}s of the FA/HR 'mutated' and 'non-mutated' cell lines at the indicated thresholds.



Supplementary Figure E. Selection of high VAF variants improves functional, MMC response, association of selected cell lines.

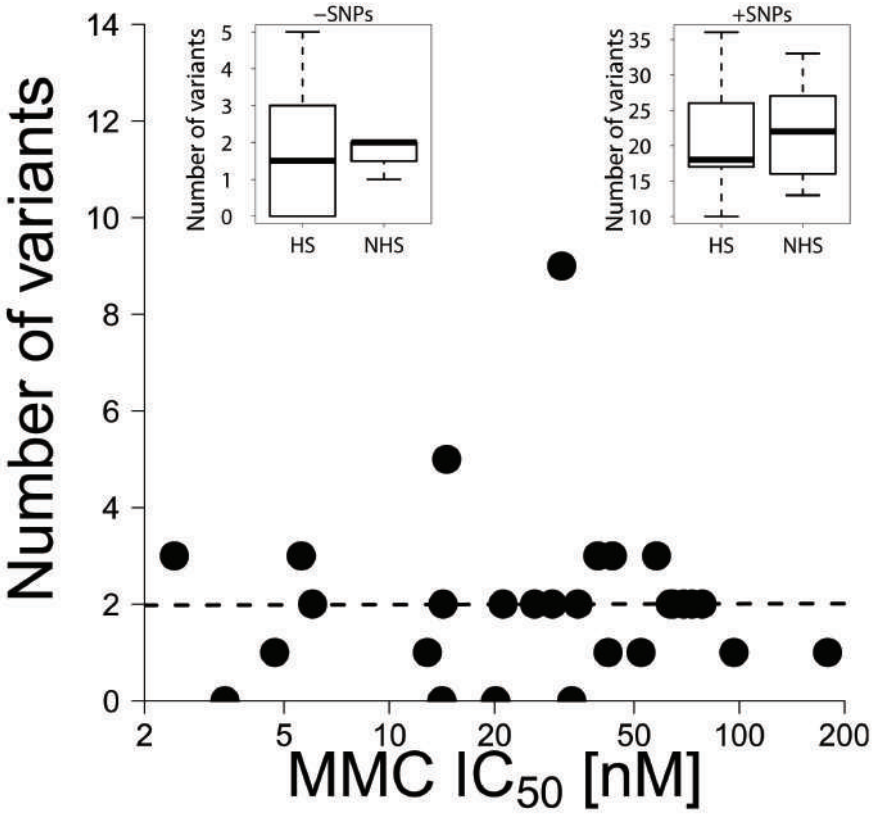
Association of MMC sensitivity with increasing VAF thresholds for variant selection. Wilcoxon rank-sum test *p*-values are shown*. MMC response associations are shown for cell lines marked by canonical FA/HR gene set variants (excluding SNPs black solid line, including SNPs black dashed line) and control gene sets. The control gene sets are 10.000 randomly grouped similar-sized gene sets (light grey solid line). The with error bars extend from the median to the first and third quartiles. Dark grey solid line depicts the results for the nearest gene set.

* Statistical tests were not performed at thresholds that resulted in less than two or all samples to be marked by a variant. Thin dashed horizontal line shows significance level of 0.05.



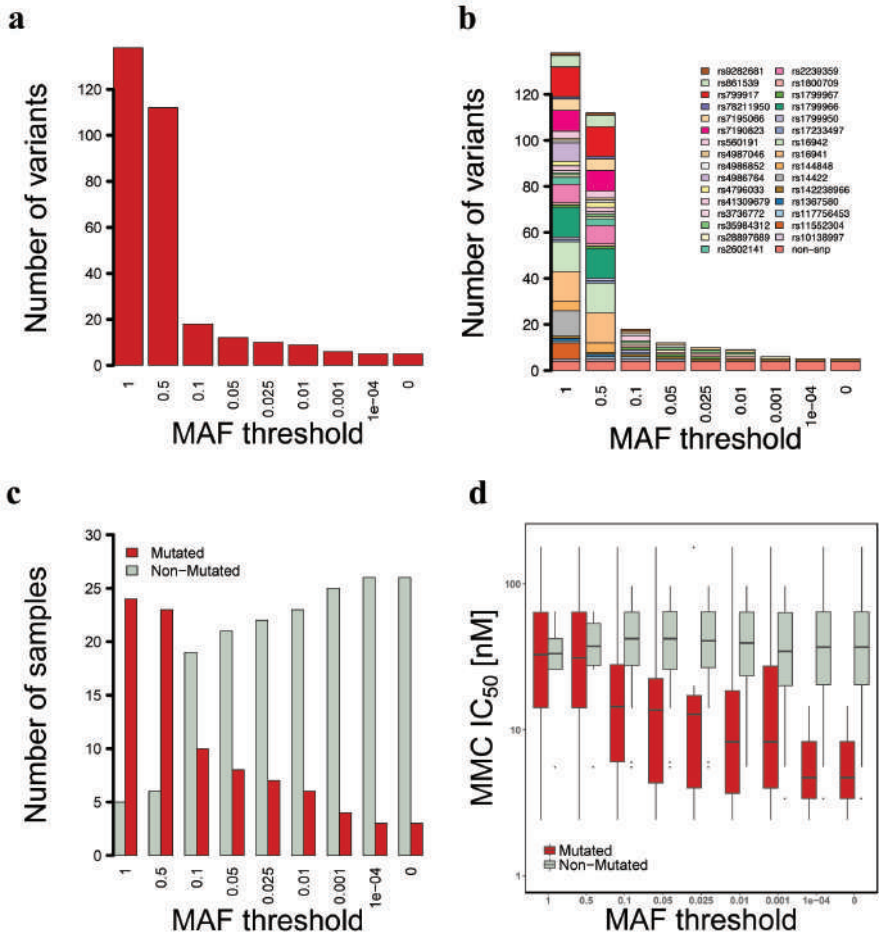
Supplementary Figure F. Variant load in the cell lines is not associated with MMC sensitivity.

The total number of variants in all captured and sequenced genes, as obtained by applying the bioinformatics pipeline described in Materials and Methods and after removing all SNPs, were calculated for each cell line and plotted against the MMC IC_{50} value of the respective cell line. Dashed black line shows the regression model with MMC sensitivity as a covariate ($p = 0.81$). The inlets show the average variant load as determined above in the hypersensitive cell lines (HS) and non-hypersensitive lines (NHS) both excluding (-) or including (+) SNPs.



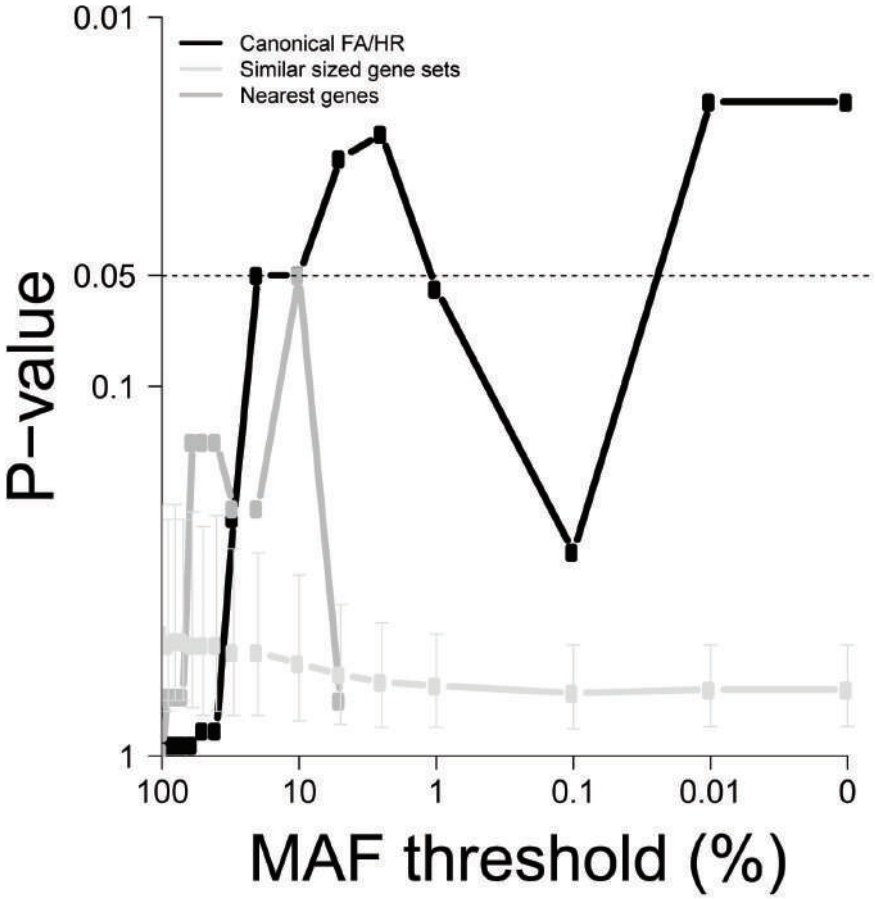
Supplementary Figure G. Homozygous variant load in the cell lines is not associated with MMC sensitivity.

The number of homozygous variants in all captured and sequenced genes as obtained by applying the bioinformatics pipeline described in Materials and Methods and after removing all SNPs. Dashed black line shows the regression model with MMC sensitivity as a covariate ($p = 0.65$). The inlets show the average homozygous variant load as determined above in the hypersensitive cell lines (HS) and non-hypersensitive lines (NHS) either excluding (-) or including (+) SNPs.



Supplementary Figure H. The effect of MAF filtering on the number of selected FA/HR gene variants and mutated cell lines.

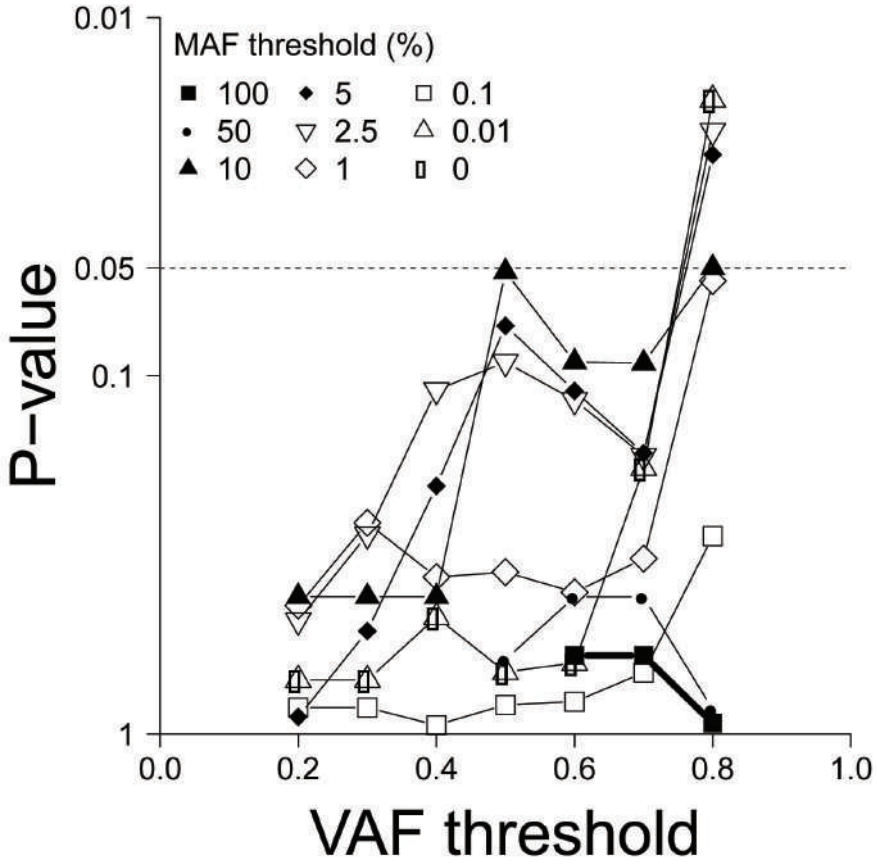
Focusing on all homozygous FA/HR variants (VAF ≥ 0.8), MAF thresholds were varied and at each threshold value SNPs with a MAF below the threshold were removed. The total number of variants and SNPs with MAF below the threshold **(A)**, the split according to dbSNP identifier **(B)**, the number of cell lines with a FA/HR gene variant or SNP with MAF below the threshold ('mutated') **(C)**, and **(D)** the MMC IC_{50}s of the cell lines assigned to be FA/HR 'mutated' and 'non-mutated' according to the presence of a variant selected by these criteria are shown at each MAF threshold. Stars indicate significance levels.



Supplementary Figure I. Selecting low MAF SNPs and non-SNP variants improves functional, MMC response, association of selected cell lines.

Association of variants with MMC sensitivity as determined by the Wilcoxon rank-sum test *p*-value* in the canonical FA/HR gene set (black solid line) and control gene sets (light grey solid line for the 10.000 similar sized gene sets with errors bars extending from the median to the first and third quartiles, dark grey solid line for the nearest genes) with decreasing MAF thresholds.

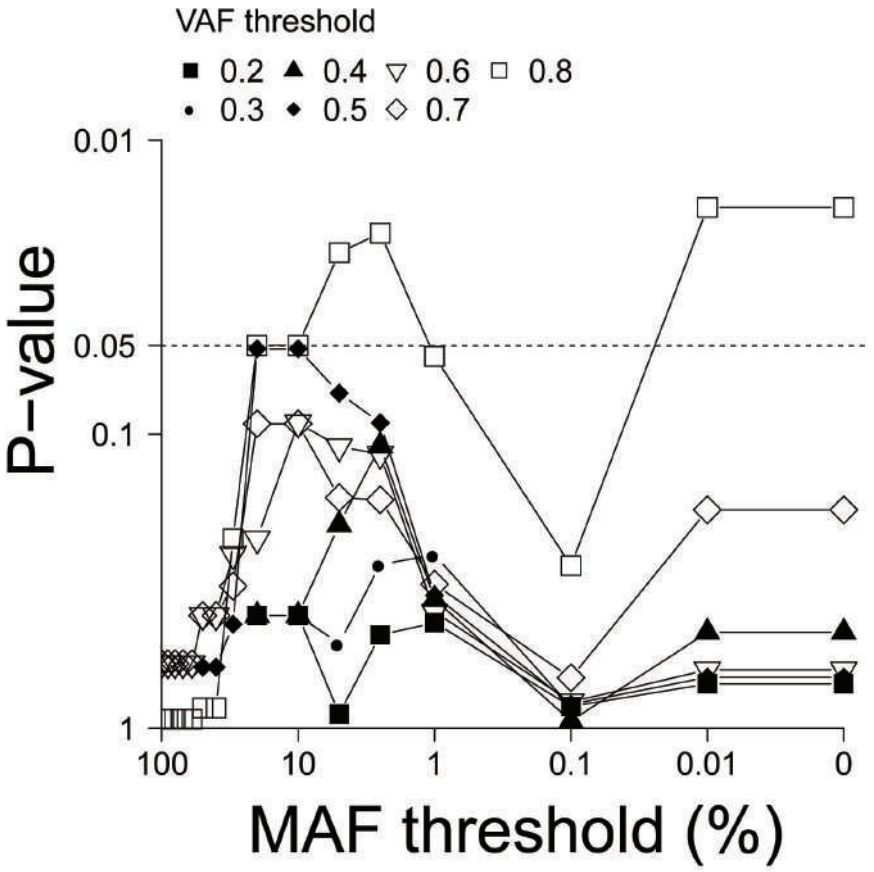
* Statistical tests were not performed at thresholds that resulted in less than two or all samples to be marked by a variant. Thin dashed horizontal line shows significance level of 0.05.



Supplementary Figure J. Selection of high VAF variants results in a significant functional endpoint association with multiple MAF thresholds.

Association of canonical FA/HR gene set variants with MMC sensitivity. Wilcoxon rank-sum test *p*-values* with increasing VAF thresholds are shown. Figure shows this analysis for different MAF thresholds above which SNPs were excluded.

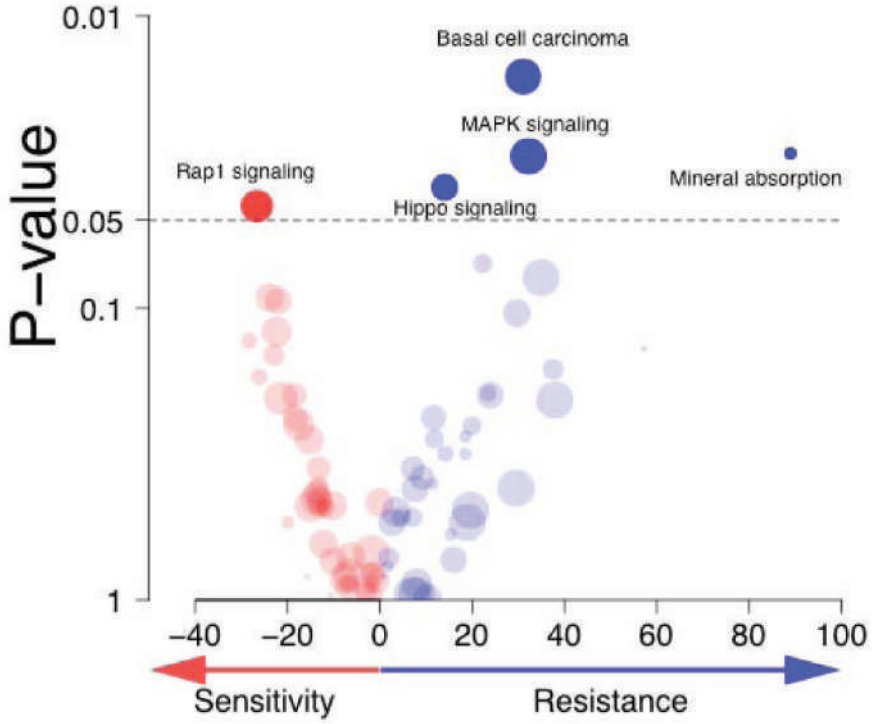
* Statistical tests were not performed at thresholds that resulted in less than two or all samples to be marked by a variant. Thin dashed horizontal line shows significance level of 0.05.



Supplementary Figure K. Selection of low MAF SNPs and non-SNP variants results in a significant functional endpoint association when focusing on those with a high VAF.

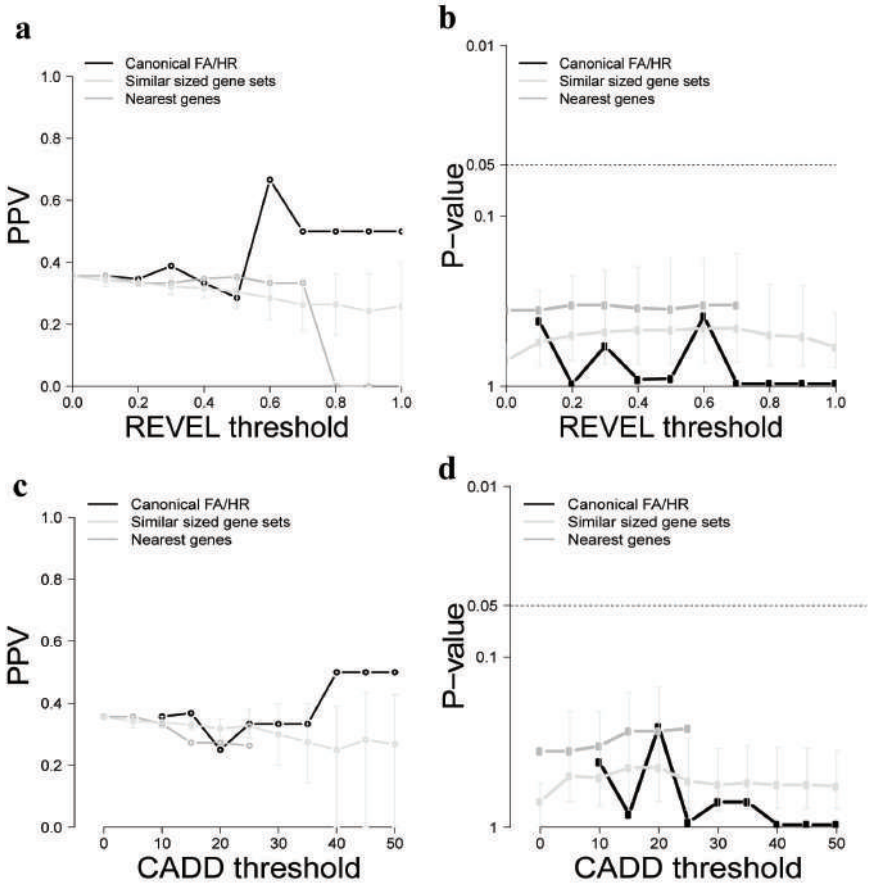
Association of canonical FA/HR gene set variants with MMC sensitivity. Wilcoxon rank-sum test *p*-values* with decreasing MAF thresholds are shown. Different VAF thresholds, excluding variants with a VAF below these thresholds, as indicated above were tested. Significant associations are found at multiple MAF thresholds, but only when the analysis is restricted to homozygous variants (VAF ≥ 0.8).

* Statistical tests were not performed at thresholds that resulted in less than two or all samples to be marked by a variant. Thin dashed horizontal line shows significance level of 0.05.



Supplementary Figure L. Volcano plot MMC response and individual KEGG pathways associations (without *TP53* mutation status correction).

The x-axis is the difference in mean MMC IC_{50} between the cell lines that were called 'pathway mutated' and 'non-mutated'. The y-axis shows the significance of this difference (Wilcoxon rank-sum test). 'Pathway-mutated' cell lines are those with one or more variants in any gene of the individual KEGG pathway. Dot size is proportional to the number of 'pathway mutated' cell lines. Note: the association of MMC with the MAPK signaling and basal cell carcinoma pathways was erroneous due to missed *TP53* mutations in 2 out of 10 MMC hypersensitive cell lines (see main text and Table C).

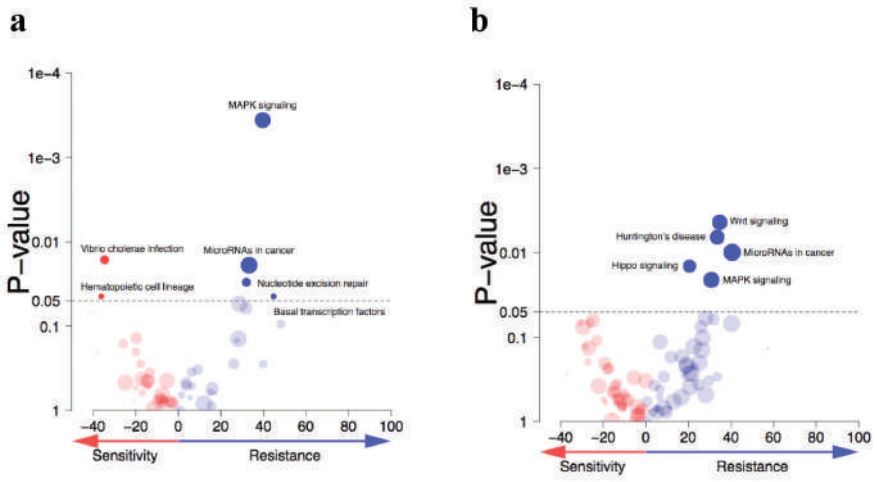


Supplementary Figure M. Variant selection based on REVEL and CADD deleteriousness scores and association with MMC sensitivity.

We varied REVEL and CADD score thresholds from low to high and at each threshold value removed variants with a score below the threshold. Variants with REVEL scores > 0.5 and CADD scores > 15 are generally considered deleterious. Although in-frame and frameshift indels and nonsense mutations aren't scored by REVEL or CADD, these were considered as potentially pathogenic and therefore included at each threshold.

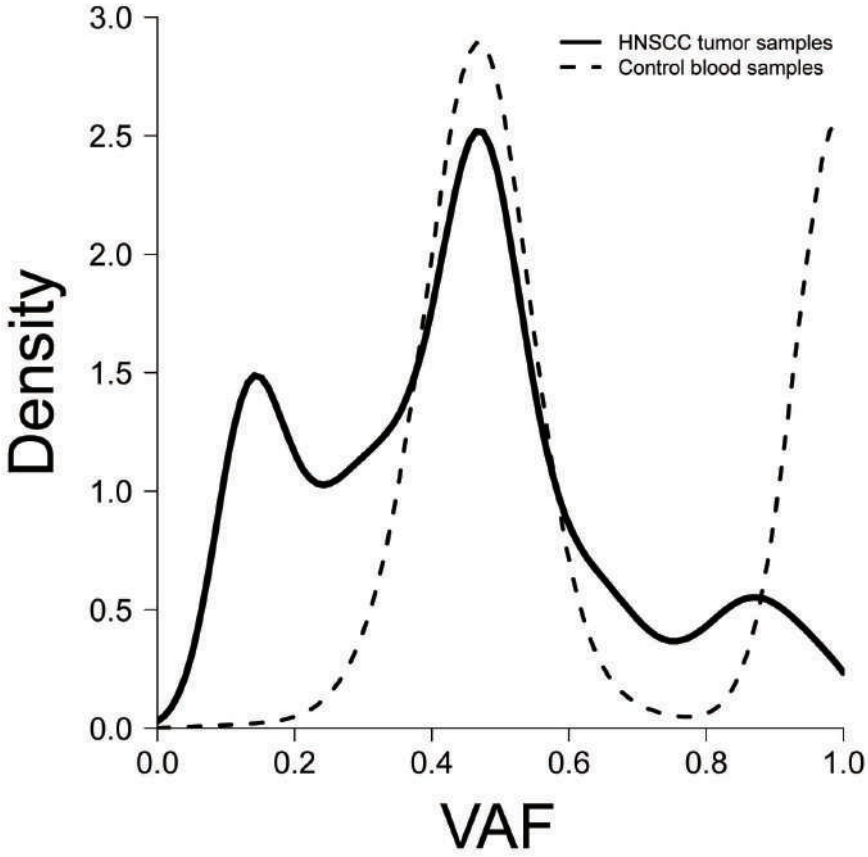
The positive predictive value (PPV) for MMC sensitivity was used to quantify the ability of variants to label repair defected cell lines (i.e. ten most MMC sensitive) *. **(A)** and **(C)** PPV values at each REVEL and CADD threshold. **(B)** and **(D)** association of variants with MMC response as determined by the Wilcoxon rank-sum test P-value.

* Statistical analyses were omitted at sample sizes of a group of less than two and these data points and lines have been excluded in the figures.



Supplementary Figure N. Volcano plot MMC response and individual KEGG pathways associations after REVEL (A) and CADD (B) variant selection and based on 'deleteriousness' prediction.

The x-axis is the difference in mean MMC IC_{50} between 'pathway mutated' and 'non-mutated' cell lines. The y-axis shows the significance of the difference in MMC IC_{50} between pathway mutated and non-mutated cell lines (Wilcoxon rank-sum test). Pathway mutated cell lines are those with one or more variants in any gene of the individual KEGG pathway. Dot size is proportional to the number of 'pathway mutated' cell lines. Variants with **(A)** a REVEL score > 0.5 and **(B)** a CADD score > 15 were retained, as these are generally considered pathogenic. Although in-frame and frameshift indels and nonsense mutations aren't scored by REVEL or CADD, we considered these pathogenic and included them at each threshold.



Supplementary Figure O. VAF density plot of variants in 56 HNSCC tumor samples. Variants were called with the same variant calling pipeline that was used for the HNSCC cell line panel. Only rare SNPs were included (MAF ≤ 1%). The VAF density of variants in the HNSCC tumor samples has three peaks. In contrast, the density function of variants in control blood samples has a bimodal distribution reflecting the hetero- and homozygous states.

Supplementary Table A

I		II	
Capture composition		**Canonical FA/HR gene set**	**Matched nearest genes**
Genes involved in DNA damage response		*FANCA*	*CDT1*
Homologous Recombination		*FANCB*	*POLA1*
Fanconi Anemia pathway		*FANCC*	*PTCH1*
Non Homologous End Joining		*BRCA2*	*RFC3*
Nucleotide Excision Repair		*FANCD2*	*VHL*
Base Excision Repair		*FANCE*	*CDKN1A*
Mismatch Repair		*FANCF*	*ABCC8*
Translesion Synthesis		*FANCG*	*APTX*
Genes involved in cell cycle regulation and proliferation		*FANCI*	*POLG*
Genes involved in drug activation and transport		*BRIP1*	*PPM1D*
Genes involved in drug sensitivity		*FANCL*	*PSME4*
cisplatin		*FANCM*	*POLE2*
MMC		*PALB2*	*PLK1*
PARP inhibitors		*RAD51C*	*AKAP1*
Frequently mutated genes in cancer, in particular:		*SLX4*	*ABCA3*
Head and Neck Cancers (HNSCC)		*RAD51*	*LTK*
Lung Cancer (NSCLC)		*TP53BP1*	*LTK*
Breast Cancer		*BRCA1*	*TOP2A*
Colon Cancer		*C17orf70*	*CSNK1D*
		RAD50	*PPP2CA*
		RAD51B	*HIF1A*
		RAD51D	*LIG3*
		RAD52	*FOXM1*
		RAD54B	*NBN*
		RAD54L	*MUTYH*
		XRCC3	*AKT1*
		XRCC2	*SMARCD3*

Composition of 556 gene capture set (I) and list of genes comprising the canonical FA/HR gene set (II) used in this study. **(I)** The primary objective in the selection of the capture genes was to cover DNA repair genes and genes reported to be relevant in head and neck squamous cell carcinoma (HNSCC). Cell cycle regulation and drug activation and transporter genes were selected since they influence crosslinker response, even though to a lesser degree. Genes mediating sensitivity to drugs such as cisplatin and PARPi, unrelated to DNA repair, were added. Moreover, we included frequently mutated genes in HNSCC (as of 2014) to be able to assess those in the HNSCC cell line panel. We broadened this selection with additional frequently mutated genes, as reported in other human cancers. **(II)** The matched nearest genes were selected from the 556 gene capture set to be closest to each canonical FA/HR gene, as measured by base pair distance.

Supplementary Table B

Cell line	*TP53* literature	*TP53* DNA-seq	Reference	Comment
UT-SCC-12A	Mut	Wildtype	[33]	Read coverage of 4, therefore discarded by pipeline due to low coverage
UT-SCC-60b	Mut	Wildtype	[34]	Deletion of 25 base pairs
UT-SCC-45	Wildtype	Wildtype	[35]	Cell line from HPV positive tumor
UT-SCC-20A	Mut	Mut	[36]	
UT-SCC-24B	Mut	Mut	[33]	
UT-SCC-38	Unknown	Wildtype		
UT-SCC-14	Mut	Mut	[36]	Deletion of 30 base pairs documented, splice site variant detected
UT-SCC-54C	Mut	Mut	[34]	
UT-SCC-15	Mut	Mut	[33]	
UT-SCC-76A	Unknown	Mut		
UT-SCC-32	Mut	Mut	[33]	
UT-SCC-4	Mut	Mut	[33]	
UT-SCC-16A	Mut	Mut	[36]	
UT-SCC-79A	Unknown	Mut		
UT-SCC-9	Mut	Wildtype	[33]	Deletion of exons 2-9
UT-SCC-40	Mut	Wildtype	[22]	Deletion of 30 base pairs
UT-SCC-24A	Mut	Mut	[33]	
UT-SCC-90	Unknown	Wildtype		
UT-SCC-1A	Mut	Mut	[36]	
UT-SCC-43A	Mut	Mut	[34]	
UT-SCC-8	Mut	Mut	[22]	
UT-SCC-42A	Mut	Mut	[34]	
NKI-SCC-263	Unknown	Mut		
UT-SCC-2	Mut	Mut	[36]	
UT-SCC-7	Mut	Mut	[33]	
UT-SCC-36	Mut	Mut	[22]	
UT-SCC-27	Mut	Mut	[22]	
UT-SCC-77	Mut	Mut	[37]	
UT-SCC-30	Mut	Mut	[36]	

List of cell lines and their *TP53* mutation status in the literature and as called by our DNAseq pipeline. Literature references and a comment explaining the discrepancy are given between the literature. Five reported *TP53* mutations were missed. The mutation in UT-SCC-12A was present in the raw reads, but not called by VarScan due to low sequencing coverage. The other four comprised medium-sized deletions and were missed due to the short-read DNA-sequencing approach that is not well suited to detect deletions of this size. Importantly, two of these missed mutations were among the most MMC sensitive cell lines. This resulted in false associations between *TP53* mutation status and MMC response and a false association with the MAPK signaling and basal cell carcinoma pathways in the pathway analyses (Figure L). These associations disappeared when correcting *TP53* mutation status to the reported (Fig 2B).

Supplementary Table C

Gene	Function	Capture sequenced
FHIT	Tumor suppressor	No
RASSF1A	Tumor suppressor	No
CSMD1	Tumor suppressor	Yes
CDKN2A	Tumor suppressor	Yes
PTPRD	Tumor suppressor	Yes
PTEN	Tumor suppressor	Yes
TP53	Tumor suppressor	Yes
SMAD4	Tumor suppressor	Yes
CCNL1	Oncogene	Yes
PARP1	Oncogene	Yes
PIK3CA	Oncogene	Yes
TP63	Oncogene	Yes
DCUN1D1	Oncogene	Yes
EGFR	Oncogene	Yes
MET	Oncogene	Yes
MYC	Oncogene	Yes
PTK2	Oncogene	No
CCND1	Oncogene	Yes
CTTN	Oncogene	Yes
FADD	Oncogene	No

List of established and candidate cancer genes in HNSCC as described and taken from Leemans et al. 2011 [40]. The analysis was narrowed to the genes that were captured sequenced (third column: 'Yes').

References

[1] Cancer Genome Atlas Research Network, Weinstein JN, Collisson EA, Mills GB, Shaw KRM, Ozenberger BA, et al. The Cancer Genome Atlas Pan-Cancer analysis project. Nat Genet 2013;45:1113–20. doi:10.1038/ng.2764.

[2] Zack TI, Schumacher SE, Carter SL, Cherniack AD, Saksena G, Tabak B, et al. Pan-cancer patterns of somatic copy number alteration. Nat Genet 2013;45:1134–40. doi:10.1038/ng.2760.

[3] Lawrence MS, Stojanov P, Mermel CH, Robinson JT, Garraway LA, Golub TR, et al. Discovery and saturation analysis of cancer genes across 21 tumour types. Nature 2014;505:495–501. doi:10.1038/nature12912.

[4] Gao H, Korn JM, Ferretti S, Monahan JE, Wang Y, Singh M, et al. High-throughput screening using patient-derived tumor xenografts to predict clinical trial drug response. Nat Med 2015;21:1318–25. doi:10.1038/nm.3954.

[5] Iorio F, Knijnenburg TA, Vis DJ, Bignell GR, Menden MP, Schubert M, et al. A Landscape of Pharmacogenomic Interactions in Cancer. Cell 2016;166:740–54. doi:10.1016/j.cell.2016.06.017.

[6] Hofree M, Carter H, Kreisberg JF, Bandyopadhyay S, Mischel PS, Friend S, et al. Challenges in identifying cancer genes by analysis of exome sequencing data. Nat Commun 2016;7:12096. doi:10.1038/ncomms12096.

[7] Eilbeck K, Quinlan A, Yandell M. Settling the score: Variant prioritization and Mendelian disease. Nat Rev Genet 2017;18:599–612. doi:10.1038/nrg.2017.52.

[8] Mahmood K, Jung C, Philip G, Georgeson P, Chung J, Pope BJ, et al. Variant effect prediction tools assessed using independent, functional assay-based datasets: implications for discovery and diagnostics. Hum Genomics 2017;11:10. doi:10.1186/s40246-017-0104-8.

[9] Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc 2009;4:1073–81. doi:10.1038/nprot.2009.86.

[10] Ramensky V, Bork P, Sunyaev S. Human non-synonymous SNPs: server and survey. Nucleic Acids Res 2002;30:3894–900.

[11] Turner N, Tutt A, Ashworth A. Hallmarks of “BRCAness” in sporadic cancers. Nat Rev Cancer 2004;4:814–9. doi:10.1038/nrc1457.

[12] Jung H, Bleazard T, Lee J, Hong D. Systematic investigation of cancer-associated somatic point mutations in SNP databases. Nat Biotechnol 2013;31:787–9. doi:10.1038/nbt.2681.

[13] O’Connor MJ. Targeting the DNA Damage Response in Cancer. Mol Cell 2015;60:547–60. doi:10.1016/j.molcel.2015.10.040.

[14] Verhagen CVM, de Haan R, Hageman F, Oostendorp TPD, Carli ALE, O’Connor MJ, et al. Extent of radiosensitization by the PARP inhibitor olaparib depends on its dose, the radiation dose and the integrity of the homologous recombination pathway of tumor cells. Radiother Oncol 2015;116:358–65. doi:10.1016/j.radonc.2015.03.028.

[15] de Haan R, Pluim D, van Triest B, van den Heuvel M, Peulen H, van Berlo D, et al. Improved pharmacodynamic (PD) assessment of low dose PARP inhibitor PD activity for radiotherapy and chemotherapy combination trials. Radiother Oncol 2018;126:443–9. doi:10.1016/j.radonc.2017.10.017.

[16] Sousa FG, Matuo R, Tang S-W, Rajapakse VN, Luna A, Sander C, et al. Alterations of DNA repair genes in the NCI-60 cell lines and their predictive value for anticancer drug activity. DNA Repair (Amst) 2015;28:107–15. doi:10.1016/j.dnarep.2015.01.011.

[17] Sasaki MS, Tonomura A. A high susceptibility of Fanconi’s anemia to chromosome breakage by DNA cross-linking agents. Cancer Res 1973;33:1829–36.

[18] Kim H, D’Andrea AD. Regulation of DNA cross-link repair by the Fanconi anemia/BRCA pathway. Genes Dev 2012;26:1393–408. doi:10.1101/gad.195248.112.

[19] Knipscheer P, Räschle M, Smogorzewska A, Enoiu M, Ho TV, Schärer OD, et al. The Fanconi anemia pathway promotes replication-dependent DNA interstrand cross-link repair. Science 2009;326:1698–701. doi:10.1126/science.1182372.

[20] Wang AT, Smogorzewska A. SnapShot: Fanconi anemia and associated proteins. Cell 2015;160:354-354.e1. doi:10.1016/j.cell.2014.12.031.
[21] Verhagen CVM, Vossen DM, Borgmann K, Hageman F, Grénman R, Verwijs-Janssen M, et al. Fanconi anemia and homologous recombination gene variants are associated with functional DNA repair defects in vitro and poor outcome in patients with advanced head and neck squamous cell carcinoma. Oncotarget 2018;9:18198–213. doi:10.18632/oncotarget.24797.
[22] Lansford CD, Grenman R, Bier H, Somers KD, Kim SY, Whiteside TL, et al. Head and Neck Cancers. Hum. Cell Cult., Dordrecht: Kluwer Academic Publishers; 1999, p. 185–255. doi:10.1007/0-306-46861-1_12.
[23] Verhagen CVM, Vossen DM, Borgmann K, Hageman F, Grénman R, Verwijs-Janssen M, et al. Fanconi anemia and homologous recombination gene variants are associated with functional DNA repair defects in vitro and poor outcome in patients with advanced head and neck squamous cell carcinoma. Oncotarget 2018;9. doi:10.18632/oncotarget.24797.
[24] Demuth I, Wlodarski M, Tipping AJ, Morgan N V, de Winter JP, Thiel M, et al. Spectrum of mutations in the Fanconi anaemia group G gene, FANCG/XRCC9. Eur J Hum Genet 2000;8:861–8. doi:10.1038/sj.ejhg.5200552.
[25] Joenje H, Levitus M, Waisfisz Q, D'Andrea A, Garcia-Higuera I, Pearson T, et al. Complementation analysis in Fanconi anemia: assignment of the reference FA-H patient to group A. Am J Hum Genet 2000;67:759–62. doi:10.1086/303067.
[26] Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009;25:1754–60. doi:10.1093/bioinformatics/btp324.
[27] Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res 2012;22:568–76. doi:10.1101/gr.129684.111.
[28] Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009;25:2078–9. doi:10.1093/bioinformatics/btp352.
[29] Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 2010;38:e164. doi:10.1093/nar/gkq603.
[30] 1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. Nature 2015;526:68–74. doi:10.1038/nature15393.
[31] Garnett MJ, Edelman EJ, Heidorn SJ, Greenman CD, Dastur A, Lau KW, et al. Systematic identification of genomic markers of drug sensitivity in cancer cells. Nature 2012;483:570–5. doi:10.1038/nature11005.
[32] Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature 2012;483:603–7. doi:10.1038/nature11003.
[33] Hauser U, Balz V, Carey TE, Grénman R, Van Lierop A, Scheckenbach K, et al. Reliable detection of p53 aberrations in squamous cell carcinomas of the head and neck requires transcript analysis of the entire coding region. Head Neck 2002;24:868–73. doi:10.1002/hed.10128.
[34] Takebayashi S, Hickson A, Ogawa T, Jung K-Y, Mineta H, Ueda Y, et al. Loss of chromosome arm 18q with tumor progression in head and neck squamous cancer. Genes, Chromosom Cancer 2004;41:145–54. doi:10.1002/gcc.20066.
[35] Rieckmann T, Tribius S, Grob TJ, Meyer F, Busch C-J, Petersen C, et al. HNSCC cell lines positive for HPV and p16 possess higher cellular radiosensitivity due to an impaired DSB repair capacity. Radiother Oncol 2013;107:242–6. doi:10.1016/j.radonc.2013.03.013.
[36] Pekkola-Heino K, Servomaa K, Kiuru A, Grenman R. Increased radiosensitivity is associated with p53 mutations in cell lines derived from oral cavity carcinoma. Acta Otolaryngol 1996;116:341–4.
[37] Farnebo L, Jerhammar F, Vainikka L, Grénman R, Norberg-Spaak L, Roberg K. Number of negative points: a novel method for predicting radiosensitivity in head and neck tumor cell lines. Oncol Rep 2008;20:453–61.
[38] Carter SL, Cibulskis K, Helman E, McKenna A, Shen H, Zack T, et al. Absolute quantification of somatic DNA alterations in human cancer. Nat Biotechnol 2012;30:413–21. doi:10.1038/nbt.2203.

[39] Riester M, Singh AP, Brannon AR, Yu K, Campbell CD, Chiang DY, et al. PureCN: copy number calling and SNV classification using targeted short read sequencing. Source Code Biol Med 2016;11:13. doi:10.1186/s13029-016-0060-z.
[40] Leemans CR, Braakhuis BJM, Brakenhoff RH. The molecular biology of head and neck cancer. Nat Rev Cancer 2011;11:9–22. doi:10.1038/nrc2982.
[41] Audo I, Bujakowska KM, Léveillard T, Mohand-Saïd S, Lancelot M-E, Germain A, et al. Development and application of a next-generation-sequencing (NGS) approach to detect known and novel gene defects underlying retinal diseases. Orphanet J Rare Dis 2012;7:8. doi:10.1186/1750-1172-7-8.
[42] Zeitz C, Jacobson SG, Hamel CP, Bujakowska K, Neuillé M, Orhan E, et al. Whole-exome sequencing identifies LRIT3 mutations as a cause of autosomal-recessive complete congenital stationary night blindness. Am J Hum Genet 2013;92:67–75. doi:10.1016/j.ajhg.2012.10.023.
[43] Riaz N, Blecua P, Lim RS, Shen R, Higginson DS, Weinhold N, et al. Pan-cancer analysis of bi-allelic alterations in homologous recombination DNA repair genes. Nat Commun 2017;8:857. doi:10.1038/s41467-017-00921-w.
[44] Maxwell KN, Wubbenhorst B, Wenz BM, De Sloover D, Pluta J, Emery L, et al. BRCA locus-specific loss of heterozygosity in germline BRCA1 and BRCA2 carriers. Nat Commun 2017;8:319. doi:10.1038/s41467-017-00388-9.
[45] Abaan OD, Polley EC, Davis SR, Zhu YJ, Bilke S, Walker RL, et al. The Exomes of the NCI-60 Panel: A Genomic Resource for Cancer Biology and Systems Pharmacology. Cancer Res 2013;73:4372–82. doi:10.1158/0008-5472.CAN-12-3342.
[46] Haibe-Kains B, El-Hachem N, Birkbak NJ, Jin AC, Beck AH, Aerts HJWL, et al. Inconsistency in large pharmacogenomic studies. Nature 2013;504:389–93. doi:10.1038/nature12831.
[47] Safaei R, Howell SB. Copper transporters regulate the cellular pharmacology and sensitivity to Pt drugs. Crit Rev Oncol Hematol 2005;53:13–23. doi:10.1016/j.critrevonc.2004.09.007.

CHAPTER 4

Extent of radiosensitization by the PARP inhibitor olaparib depends on its dose, the radiation dose and the integrity of the homologous recombination pathway of tumor cells

Caroline V.M. Verhagen*, Rosemarie de Haan*, Floor Hageman,
Tim P.D. Oostendorp, Annalisa Carli, Mark J. O'Connor, Jos Jonkers, Marcel Verheij,
Michiel van den Brekel and Conchita Vens.

* Shared first authorship

Radiotherapy & Oncology. 2015 Sep; 116(3): 358-65.

Abstract

Background and Purpose

The PARP inhibitor olaparib is currently tested in clinical phase 1 trials to define safe dose levels in combination with RT. However, certain clinically relevant insights are still lacking. Here we test, while comparing to single agent activity, the olaparib dose and genetic background dependence of olaparib-mediated radiosensitization.

Materials and Methods

Long-term growth inhibition and clonogenic assays were used to assess radiosensitization in BRCA2-deficient and BRCA2-complemented cells and in a panel of human head and neck squamous cell carcinoma cell lines.

Results

The extent of radiosensitization greatly depended on the olaparib dose, the radiation dose and the homologous recombination status of cells. Olaparib concentrations that resulted in radiosensitization prevented PAR induction by irradiation. Seven hours olaparib exposures were sufficient for radiosensitization. Importantly, the radiosensitizing effects can be observed at much lower olaparib doses than the single agent effects.

Conclusions

Extrapolation of this data to the clinic suggests that low olaparib doses are sufficient to cause radiosensitization, underlining the potential of the treatment. Here we show that drug doses achieving radiosensitization can greatly differ from those achieving single agent activities, an important consideration when developing combined radiotherapy strategies with novel targeted agents.

Introduction

New strategies to enhance the efficacy of radiotherapy focus on the combination with novel targeted agents that modulate radiosensitivity[1-4]. These strategies aim to increase the therapeutic window by radiosensitizing tumors to a greater extent than normal tissue. One such strategy exploits the difference in the capacity to repair radiation induced DNA damage between tumor and normal tissue cells. An example of such a strategy that is currently tested in clinical trials, is the combination of radiotherapy with PARP-1 inhibitors such as veliparib [5, 6] and olaparib.

Olaparib is an oral drug that inhibits poly-(ADP-ribose)-polymerase-1 (PARP-1), an enzyme involved in base excision repair and DNA single strand break (SSB) repair. As single agent, olaparib is well tolerated and has proven anti-tumor activity in BRCA1 and BRCA2 mutation carriers [7-10]. Maximal tolerated doses for monotherapy at bi-daily 400mg capsules provided clinical plasma olaparib concentrations of 5 to 14μM [8]. Single agent activity has been proposed to result from the accumulation of unrepaired endogenous lesions such as SSBs. These unrepaired SSBs can convert into cytotoxic double strand breaks (DSB) upon cell replication. Alternative mechanisms have also been proposed focusing on the cytotoxicity through replication fork stalling by the DNA-trapped PARP molecule when inhibited. Cells with homologous recombination (HR) defects, e.g. BRCA1 or BRCA2 deficient cells, have lost their replication-associated opportunity to repair such DSBs and stalled forks and are therefore particularly sensitive to PARP-1 inhibition.

The clinical development of PARP-1 inhibitors such as olaparib in combination with radiotherapy is less advanced than monotherapy strategies. Safety and tolerability of such combinations are currently tested in several clinical phase 1 trials. A number of pre-clinical studies have shown that PARP-1 inhibition leads to effective radiosensitization in multiple models [11-14]. Radiation induces DNA lesions that require PARP-1 activity for repair. Consistent with a cytotoxicity that is mediated by replication stalling or DSB conversion upon replication, radiosensitization by PARP-1 inhibitors is greater in proliferating cells [15, 16] and in cells with vulnerabilities associated with DSB repair [17]. PARP-1 inhibition therefore provides an opportunity for tumor-specific radiosensitization in repair-defected and highly proliferating tumor cells.

A challenge in the clinical development of combination strategies with targeted agents, such as olaparib, concerns the optimal dose of the targeted agent [3, 18, 19]. The optimal dose of a radiosensitizer is the dose resulting in maximum tumor-specific radiosensitization. It therefore depends on both tolerability and efficacy of the combination treatment. Notably, as the underlying mechanisms of the activity as radiosensitizer or as single agent may differ, also tolerability and effectivity, and thus the optimal dose, may differ.

Despite extensive studies showing the radiosensitizing potential of PARP inhibitors, clinically relevant questions remained unanswered with respect to the effective dose as radiosensitizer: which olaparib dose is needed for radiosensitization, is it different in

different genetic backgrounds, and how does the dose compare to the effective dose as single agent? In this study, we evaluated the influence of olaparib dose, radiation dose and the HR status of tumor cells on the extent of olaparib-mediated radiosensitization and compared effective doses of olaparib as radiosensitizer to effective doses as single agent.

Materials and Methods

Reagents

Olaparib was provided by AstraZeneca. Stock solutions were prepared in dimethylsulfoxide (DMSO) at a concentration of 10mM. DMSO concentrations were equalized in all treatments and doses including the control cells without olaparib.

Cell culture

The mouse mammary tumor cell lines KB2P3.4 with homozygous deletion of the BRCA2-gene (here referred to as BRCA2-deficient) and KB2P3.4R3 with complemented BRCA2 (BRCA2-complemented) [20] were cultured at 37°C, 5% CO_2 and 5% O_2 in medium as described previously [20].

The human HNSCC cell lines UT-SCC-12A, UT-SCC-20A, UT-SCC-24B, UT-SCC-30, UT-SCC-45 and UT-SCC-60B were kindly provided by Dr. Grenman (Turku University Central Hospital, Finland) and cultured as described previously under low oxygen (4%) in DMEM medium (Invitrogen) [21]. Note, that these lower than generally applied but more relevant oxygen levels may have an impact on PARP inhibitor sensitivity.

Radiation

Cells were exposed to single radiation doses (0.5-6Gy) at a dose rate of 1 Gy/min using the Gammacell® 40 Exactor (Best Theratronics Ltd., Ottawa, Ontario, Canada). Control cells were sham-irradiated.

Long term growth inhibition assay

Proliferating cells were treated with various concentrations of olaparib 1hr prior irradiation and further cultured for at least five doubling times in the continuous presence of olaparib. As PARP inhibition can affect metabolic activity a non-metabolic growth assay read out was chosen. Cell numbers and survival were determined by quantifying the DNA content using the CyQuant Cell Proliferation Assay Kit (Invitrogen). DMSO treated controls were plated at different cell densities to evaluate linearity of the assay i.e. the sub-confluent state of cell cultures when assayed. Growth inhibition was monitored as a decrease in “viability” determined by dividing the cell number values of the treated samples by those of the controls.

Clonogenic survival assay

Proliferating cells were irradiated after an 1hr olaparib pre-incubation. Six hours thereafter, cells were trypsinized, plated in appropriate cell densities and cultured with or without olaparib for at least six doubling times to allow colony formation. Colonies were fixed and stained with 0.5% crystal violet/6.0% glutaraldehyde and counted. Surviving fractions (SF) were calculated by dividing the plating efficiency of treated samples by the plating efficiency of control samples.

PARP pharmacodynamic assay

Cellular levels of PAR (poly-ADP-ribose, the direct product of PARP-1) were measured by using the HT PARP in vivo Pharmacodynamic Assay II (Trevigen®) following the manufactures protocol. Cells were cultured and irradiated on ice at a sub-confluent state after 1hr of olaparib pre-incubation. After cell extract preparation and protein content determination (BCA protein assay, ThermoScientific), PAR levels were measured by chemiluminescent readings using a Tecan plate reader.

Data analysis and statistics

Clonogenic survival values (SFs) and growth inhibition values (viability) were normalized to either non-radiated controls or to DMSO controls (without olaparib treatment) as indicated. SFs on radiation survival curves were fitted using the linear quadratic (LQ)-model *($SF=exp(-\alpha d-\beta d^2)$; d = radiation dose)*. For growth inhibition assays, the area under the curve (reflecting mean inactivation doses) of these fits from 0 to 6Gy ($AUC_{0\text{-}6Gy}$) were calculated for each individual experiment. The radiation dose resulting in 37% survival was calculated from the fits for each olaparib dose in each individual experiment. The radiation dose enhancement factor (DEF) of olaparib is the ratio of the radiation doses resulting in 37% survival in control samples to olaparib treated samples.

Statistical analyses were performed in GraphPad Prism using the Student's t-test (two-sided) and considered significant if p-value <0.05.

Results

BRCA2-deficient cells are hypersensitive to olaparib and to radiation

To evaluate all main factors that potentially affect the ability of olaparib to radiosensitize cells, we used two isogenic cell lines that differ in their HR status. The BRCA2-deficient cells derived from a spontaneous tumor mouse model that drives BRCA2 loss in mammary cells and are HR-deficient and (p53-/-) [2]. As its HR-proficient counterpart we used the same cell line in which BRCA2 function was restored by genetic complementation (BRCA2-complemented) [20]. Firstly, we assessed sensitivity to radiation and to olaparib in this model. Consistent with previous reports [20, 23-25], the BRCA2-deficient cells were hypersensitive to olaparib with an IC_{50} (i.e. concentration

resulting in 50% survival) of 16nM, whereas the IC_{50} was not yet reached at 3.3µM olaparib in the BRCA2-complemented cells (Figure 1A). BRCA2-deficient cells also confirmed to be more sensitive to radiation than BRCA2-complemented cells by a DEF_{37} of 2.1 [26,27] (Figure 1B, Supplemental Table 1).

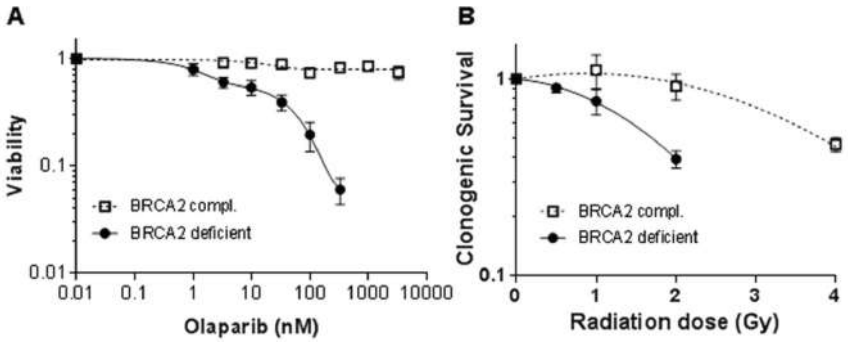


Figure 1. BRCA2-deficient cells are more sensitive to olaparib and radiation than BRCA2-complemented.

(A) Growth inhibition in BRCA2-complemented and BRCA2-deficient cells after continuous olaparib exposure as determined by long-term growth inhibition assays (Cyquant). Note, DMSO-treated (0nM) control samples are plotted at 0.01nM on the log scale X-axis. **(B)** SF of these cells after radiation as determined by clonogenic survival assays. Data present the mean ±SEM of 3-6 independent experiments.

Olaparib radiosensitizes BRCA2-proficient and deficient cells to different extent

We wished to test a wide range of olaparib doses combined with different radiation doses. To do so, we performed long-term growth inhibition assays in which BRCA2-deficient and BRCA2-complemented cells were continuously exposed to olaparib. We acknowledged the potential interference of PARP inhibition in cellular metabolic activity and hence applied an assay based on a DNA content read out. In both cell lines, the combination of olaparib and radiation killed cells more effectively than either alone (Figure 2A-B, Supplemental Figure 1A-B). The combination resulted in an additive or more than additive kill, hence radiosensitization, as shown in the survival curves derived after normalization to non-radiated controls at the respective olaparib doses (Figure 2C-D, Supplemental Figure 1C-D). To compare the extent of radiosensitization in the two cell lines, the area under the curve (AUC) and dose enhancement factors (DEF_{37}) were calculated on the linear quadratic fits (Figure 2E-F). The overall lower AUC values reflect the increased intrinsic radiosensitivity of the BRCA2-deficient cells (Figure 1B). In both cell lines, the AUC decreased and the DEF increased with increasing olaparib dose, showing the radiosensitization dependence on olaparib dose. Notably, BRCA2-deficient cells were radiosensitized at more than ten-fold lower concentrations than BRCA2-complemented cells (Figure 2F and Supplemental Table 2; with 8nM

and 220nM for a DEF of 1.3 respectively). We conclude that olaparib effectively radiosensitizes both HR-deficient and HR-proficient cells, however, the olaparib dose needed for radiosensitization is lower in HR-impaired cells.

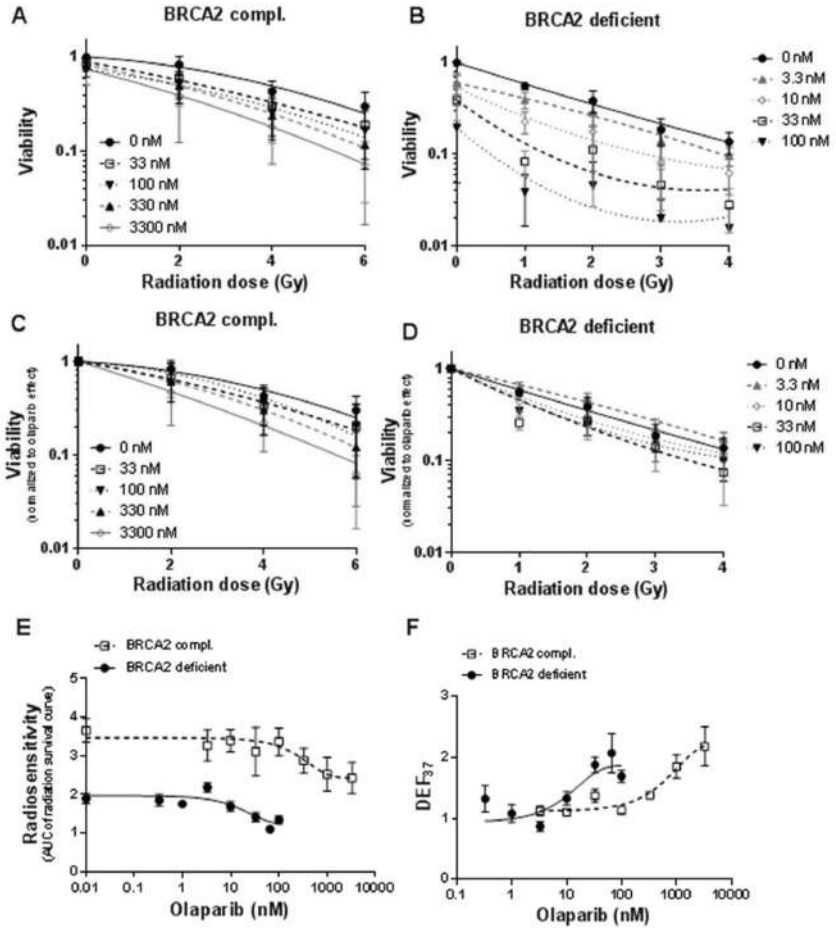


Figure 2. Olaparib radiosensitizes BRCA2-deficient and BRCA2-complemented cells.
(A-B) Survival in BRCA2-complemented and BRCA2-deficient cells after radiation and continuous olaparib exposure (for clarity only representative olaparib doses are shown, for full data see Supplemental Figure 1) as determined by long term growth inhibition assays. **(C-D)** Data as in Figures A-B respectively, now normalized to un-radiated values at the different olaparib doses. Datapoints present mean ±SD of 3-7 independent experiments (A-D) **(E-F)** Area under the curve (AUC) values and dose enhancement factors (DEF37) of BRCA2-complemented and BRCA2-deficient cells at different olaparib doses from radiation survival curve fits that were normalized to the respective olaparib effects. Datapoints present mean ±SEM of 3-7 independent experiments (except for 0.33nM and 66nM BRCA2-deficient datapoints n=2).

Seven hours exposure to olaparib is sufficient to radiosensitize

As the highly cytotoxic effect of olaparib as single agent in HR-deficient cells may mask the full radiosensitization potential of olaparib, we reduced olaparib exposure time from continuous to 7 hours (1hr before and 6hrs after irradiation). This reduction in olaparib exposure significantly reduced cell kill (single agent effect)in the BRCA2-deficient cells (Supplemental Figure 2). In combination with radiation, a short exposure time to olaparib killed cells more effectively than either alone in both cell lines (Supplemental Figure 2, Figure 3A). At the lower radiation doses radiosensitization was more pronounced in the short than in the continuous exposure schedule (Supplemental Figure 3), thereby partly compensating for the loss of single agent activity (olaparib cytotoxicity) in the short exposure schedule. We conclude that seven hours exposure to olaparib is sufficient to radiosensitize.

Extent of radiosensitization depends on olaparib dose and BRCA2 status

In both, BRCA2-deficient and BRCA2-complemented, cells radiosensitization increased with increasing olaparib dose (Figure 3A, Supplemental Figure 3A-B). To compare radiosensitization effects at different olaparib doses, dose enhancement factors (DEF) at 37% survival were calculated from the clonogenic assay data. These data show that BRCA2-deficient cells are radiosensitized by much lower doses of olaparib than BRCA2-complemented cells, after both, short and continuous, exposures to olaparib (Figure 3B, Supplemental Figure 3C-E, Table 1).

Effective doses for radiosensitization are lower than for single agent activity

To compare effective doses of olaparib as radiosensitizer with effective doses as single agent directly , surviving fractions (SF) were normalized to radiated controls without olaparib. Lower doses of olaparib were needed for radiosensitization than for single agent activity in both cell lines and for short and long exposures (Figure 3C-D, Supplemental Figure 4). In the 7h exposure schedule, effective single agent activity doses in BRCA2-deficient cells are around 1000nM (with an IC50 of 3500nM, Supplemental Table 2), whereas doses as low as 3 or 33nM achieve significant radiosensitization. BRCA2-complemented cells showed a radiation dose dependent increase in the extent of radiosensitization by olaparib (Figure 3C). In contrast, BRCA2-deficient cells were sensitized by just 0.5Gy. Increasing the radiation dose did not increase the extent of radiosensitization further (Figure 3D). Together, we conclude that the extent of radiosensitization by olaparib depends on the olaparib and radiation dose and the cellular repair status. As we show here, this is the case at schedules that result in high single agent activity, but also when olaparib is applied shortly, (with little single agent activity). Importantly, effective doses of olaparib for radiosensitization are much lower than for single agent activity.

Table 1. Dose Enhancement Factors

	Cell line	BRCA2-compl.			BRCA2-deficient			UT-SCC-12A	UT-SCC-20A	UT-SCC-24B	UT-SCC-30	UT-SCC-45	UT-SCC-60B
	Assay	CA		LT-GI	CA		LT-GI	LT-GI	LT-GI	LT-GI	LT-GI	LT-GI	LT-GI
	Ola exposure	7 h	cont.	cont.	7 h	cont.	cont.	cont.	cont.	cont.	cont.	cont.	cont.
DEF_{37}	**3.3 nM**	-	-	1.14	1.25	1.12	1.11	1.07	0.96	1.15	1.08	1.01	1.09
	33 nM	0.97	-	1.17	1.67	1.64	1.74	1.08	0.97	1.16	1.08	1.02	1.09
	100 nM	-	1.80	1.21	-	-	1.86	1.08	1.00	1.19	1.08	1.03	1.10
	1000 nM	1.32	2.06	1.72	2.56	-	-	1.17	1.26	1.61	1.11	1.16	1.21
	3300 nM	1.77	-	2.27	-	-	-	1.28	1.71	2.66	1.16	1.45	1.38
Ola-EF	**0,5 Gy**	-	-	-	20	0.8	-	-	-	-	-	-	-
	1 Gy	-	-	-	33	1.1	2.0	-	-	-	-	-	-
	2 Gy	(>2)	1	9	55	1.5	1.6	-	-	-	-	-	-
	3 Gy	-	-	-	-	-	1.7	1.8	70	2.9	1.1	1.4	6.4
	4 Gy	(>5)	32	47	-	-	1.6	-	-	-	-	-	-
	6 Gy	-	-	214	-	-	-	4.6	224	6.3	3.2	13	49

The Table summarizes the radiation dose enhancement factors at 37% survival (DEF37) achieved at the given olaparib doses in the analyzed cell lines. Values derived from colony assays (CA) that have typically less drug doses tested are average DEF on all experiments at the given dose. Values derived from long term growth inhibition (LT-GI) assays were calculated from the fits considering all tested olaparib doses (Figure 2F and Supplemental Figure 6). Ola-EF indicates the "olaparib enhancement factor" by comparing the IC50 values in the unirradiated controls to the radiation normalized data at the indicated radiation doses (e.g. a value of 20 indicates a 20fold lower olaparib IC50 value). Ola-EF could not be calculated in the BRCA2-complemented 7h experiments (indication are given based on the highest tested non-toxic olaparib concentration).

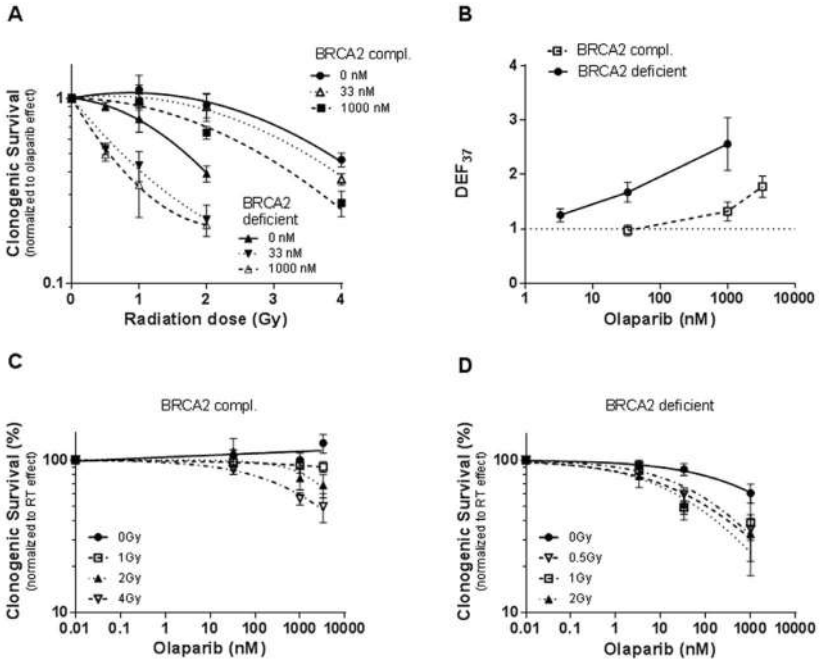


Figure 3. Effective radiosensitization by short olaparib exposure is dependent on olaparib dose and BRCA2 status and occurs at lower than single agent activity dose. **(A)** Surviving fractions (SF) of BRCA2-complemented and BRCA2-deficient cells after radiation and 7 hours olaparib exposure as determined by clonogenic assays. SF are normalized to the values of the respective non-radiated olaparib treated samples. For comparison, only olaparib doses tested in both are shown, see Supplemental Figure 3 for all tested olaparib doses. **(B)** Dose enhancement factors at 37% survival (DEF_{37}) at different olaparib doses. Data present mean ±SEM of 3-6 independent experiments (BRCA2-deficient 2Gy-1000nM datapoint n=2). **(C-D)** Survival (in %) of BRCA2-complemented and BRCA2-deficient cells after radiation and 7 hours olaparib exposure as determined by clonogenic survival assays. SF are normalized to DMSO-treated controls at the respective radiation dose. Note, DMSO-treated control-samples are plotted at 0.01nM. Datapoints present mean ±SEM of 4-6 independent experiments (BRCA2-deficient 2Gy-1000nM datapoint n=2).

Olaparib prevents PAR induction upon irradiation at radiosensitizing doses

Radiation effectively induces PAR, the direct product of PARP-1[28]. A different PAR response could be the underlying cause of the different radiosensitization doses in the two cell lines. To date, it is not known at which olaparib concentration parylation (i.e. PAR-formation), induced by radiation, can be prevented. In order to test PAR induction in the context of radiation/PARP inhibitor combinations, we applied a quantitative ELISA based assay on lysates of treated cells. Radiation alone induced PAR in both cell lines similarly (Figure 4). Despite the observed differences in the sensitivity to olaparib

as single agent, we show that olaparib resulted in a comparable dose dependent decrease in PAR levels in these cells (Figure 4, Supplemental Figure 5) indicating that olaparib activity is similar. Notably, at concentrations that radiosensitized BRCA2-deficient (but not BRCA2-complemented) cells, olaparib prevented PAR induction by radiation in both cell lines similarly (Figure 4). Thus, consistent with the observed radiosensitization potential, our results show that olaparib effectively abolished radiation-induced parylation at very low drug concentrations, whereas baseline PAR levels were not yet fully abolished.

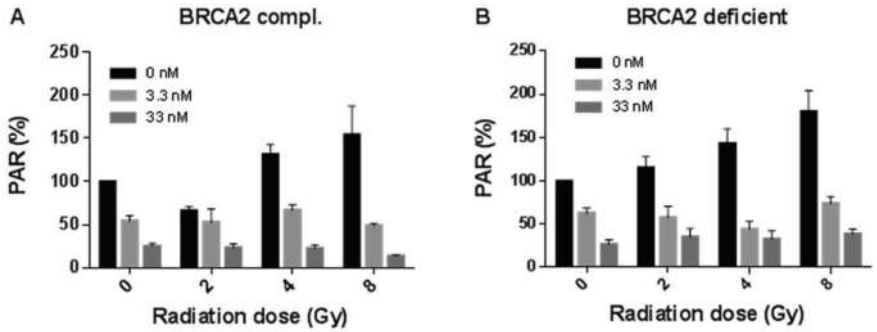


Figure 4. Olaparib prevents PAR induction upon radiation.
(A-B) PAR levels normalized to the DMSO-treated controls upon olaparib and radiation treatment as indicated. Datapoints present mean ±SEM of 3-4 independent experiments.

Radiosensitization by olaparib in a panel of human HNSCC cell lines

To further validate the results of the isogenic BRCA2 mouse tumor cell model, we next tested a panel of human head and neck squamous cell carcinoma (HNSCC) cell lines. These HNSCC cell lines differed in their sensitivity to olaparib as single agent (Figure 5, 0Gy). All HNSCC cell lines were radiosensitized by olaparib (as shown by the normalized values in Figure 5 and by the DEF_{37} in Supplemental Figure 6). Table 1 summarizes the overall radiosensitization effect by listing the DEF_{37} that were achieved at the tested olaparib concentrations in the different cell lines. Olaparib enhancement factors (Ola-EF) were calculated from the growth inhibition curves by comparing the IC50 values in the unirradiated controls to those derived from the radiation normalized curves. They demonstrate that olaparib activity was “radiation enhanced” in a radiation dose dependent manner. Consistent with our previous results, the olaparib concentration at which radiosensitization occurred, was about 2 to 100 fold (average of 30-fold in the HNSCC) lower than the concentration at which single agent activity occurred. We defined a DEF of 1.3 as a robust radiosensitization in our models and compared the olaparib dose that is required to achieve this extent of radiosensitization across the tested cell lines (Supplemental Table 2 and Supplemental Figure 7). These doses were in general much lower than the olaparib IC50s. This pattern was observed in all cell lines, with the exception of UT-SCC-45 that showed both, radiosensitization and

single agent activity, at very low doses. The extent of radiosensitization increased with increasing radiation dose. With this data we show that the effective doses of olaparib as radiosensitizer are consistently lower than the effective doses as single agent in all analyzed cell lines.

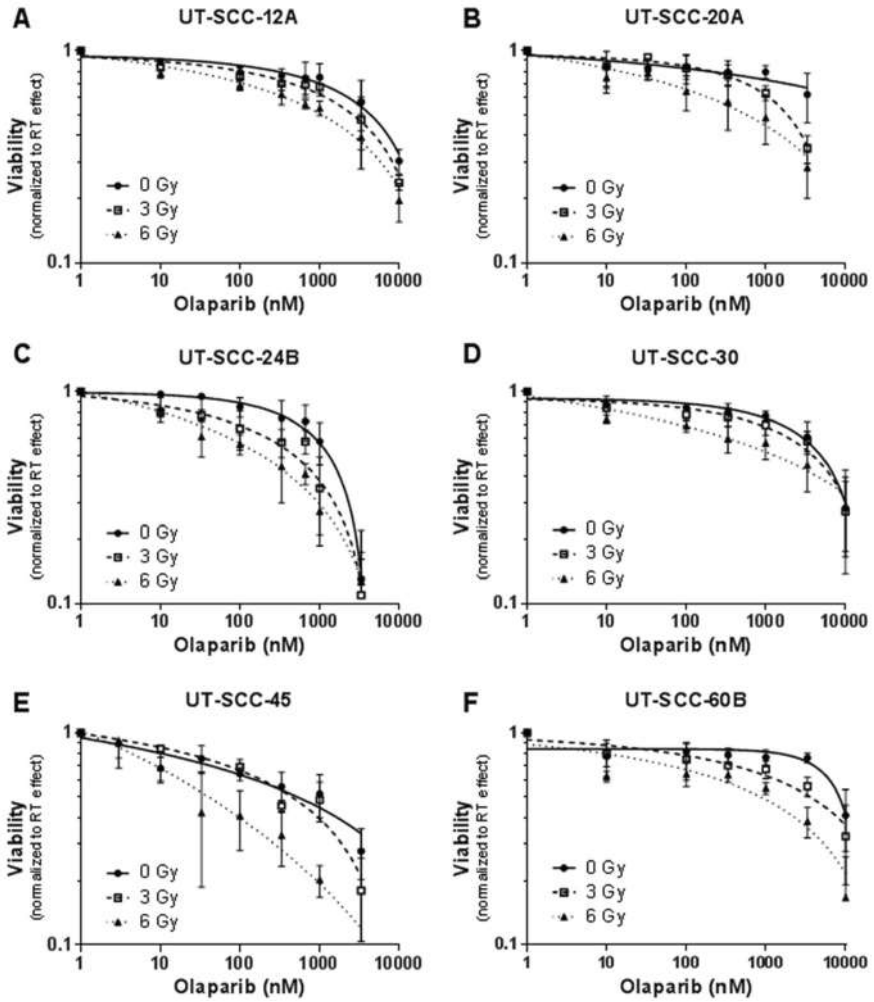


Figure 5. Olaparib dose-dependent radiosensitization in a panel of human HNSCC cell lines.

(A-F) Growth inhibition by continuous olaparib exposure in UT-SCC-12A, 20A, 24B, 30, 45, 60B cells and indicated radiation doses. Values were normalized to the DMSO-treated controls at the respective radiation doses (IR) as indicated. DMSO treated control-samples are plotted at 0.01nM. Datapoints present mean ±SD of 3-4 independent experiments (UT-SCC-24B 33nM and 660nM datapoints n=2).

Discussion

In this study we show that the extent of radiosensitization by the PARP inhibitor olaparib depends on radiation and olaparib dose and on the HR status of the cells. We further showed that olaparib radiosensitizes at much lower doses than needed for single agent activity.

This difference in effective doses (for radiosensitization and for single agent activity) has important clinical implications since this suggests radiosensitization activity at very low dose levels. Monotherapy phase I trials have defined a maximum tolerated dose (MTD) for olaparib of 400mg bi-daily (in the capsule formulation)[8], a dose applied in consecutive trials to demonstrate single agent tumor activity[7, 9, 10]. Combination treatments that assessed the safety of olaparib with chemotherapeutics have, however, been challenging and were not able to achieve such high doses [29-32]. Our findings, however, would indicate that much lower doses than administered in monotherapy are needed to achieve activity in combination with RT combination. As a reference, clinical plasma concentrations of olaparib found at monotherapy MTD would translate into a 5000-14000nM concentration range [8]. We conclude that three quarter of the tested cell lines are efficiently radiosensitized at a 3 to 10-fold lower dose than this monotherapy MTD. Considering available patient pharmacokinetic data [8], and dependent on the scheduling and the genetic tumor background one could expect radiosensitization around a 10-20mg olaparib dose level (capsule) based on our data. There is some evidence suggesting that PARP inhibitors such as olaparib accumulate in tumor cells, indicating that even lower doses may be sufficient [33, 34]. Comparing olaparib concentrations in resected breast cancer tumors and corresponding plasma samples at the time of surgery, Bundred et al. found that on average, tumour concentrations were approximately 41 % of plasma concentrations (range, 5–154 %) [35]. Notably, they found significant PARP-1 inhibition in tumors of patients receiving only 10mg olaparib bi-daily[35]. This further supports a low dose activity (also Figure 4) that may be fully exploited in the radiation setting. Often, clinical combination studies aim to achieve monotherapy doses and schedules that showed single agent efficacy, assuming that the needed target inhibition is similar. However, here we show that efficacy in combination with radiation, efficient radiosensitization, could be reached at different i.e. lower doses.

The rational of combining PARP inhibitors with radiation is based on a potential tumor-specific activity that derives from the reported pronounced radiosensitization in replicating tumor cells and in DNA repair and DNA damage response impaired cells (Figure 1-3, [13, 15-17]) and from the potential to exploit or alter hypoxia [14,36]; all attributes that are inherent to tumor cells [37]. Despite this potential tumor selectivity, however, our data further call for careful dose escalation in phase I RT combination trials. Considering the potent low olaparib dose activity one should anticipate a similar pattern in concomitantly radiation induced normal tissue toxicities.

Olaparib single agent activity on tumor cell kill requires a long-term, continuous exposure to the drug ([34], Supplemental Figure 2B). To fully take advantage of the potential single agent tumor activity, monotherapy trial designs should provide a continuous exposure at high doses of the drug. However, here we show that for the establishment of efficient radiosensitization short exposures are sufficient and at low radiation doses this may partly compensate for the lack of cytotoxicity by olaparib itself (Table 1, Supplemental Tables 1-2). With this and good tumor PK/PD kinetic data at hand, one will be able to design sensible schedules that would optimize the timing of RT and drug administration.

We observed that the low PARP inhibitor doses that decreases PAR formation only partially at unchallenged conditions (i.e. unirradiated), however, efficiently abrogate PAR formation even at higher radiation doses (Figure 4 and Supplemental Figure 5). Our data, thus, suggest that the lower dose requirement for radiosensitization is due to this increased requirement of PARP activity after radiation. Mechanistically, this may also explain the radiation dose dependence as induced PAR levels depend on radiation dose. Variability in PAR formation capacity, drug transport and drug activation efficiencies, together with genetic background differences may likely cause the wide range of radiosensitizing olaparib doses. Ultimately, however, this lack of PAR formation affects DNA damage repair and is thought to cause double strand breaks by conversion from single strand breaks or trapped PARP molecules during replication. It should also be noted that the continuous treatment schedule most likely affected replication in the olaparib sensitive cell lines and therefore partly compromised the full radiation effect.

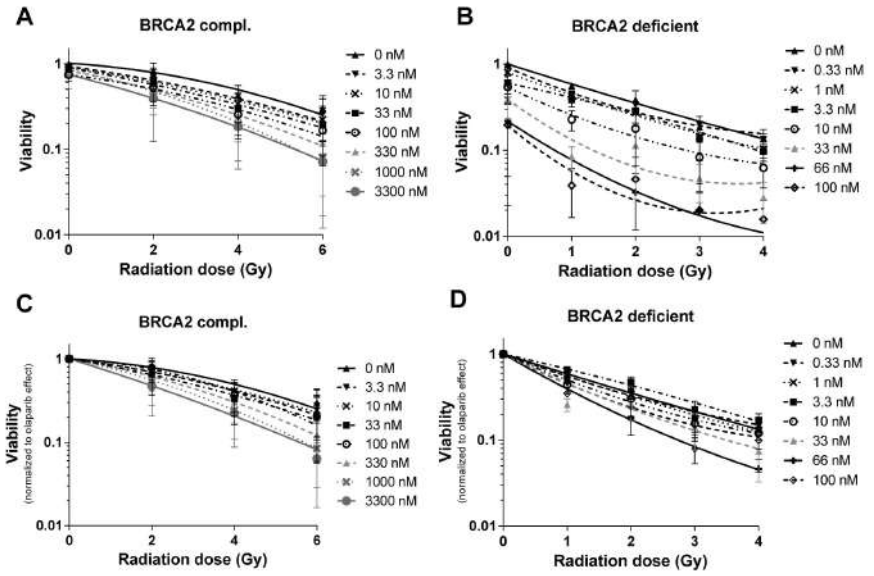
Together, our study highlights that the optimal biological dose of a targeted agent in combination with radiation can greatly differ from the optimal biological dose used as single agent and further illustrates the relevance of such investigations. Alongside with the guidelines for development of radiosensitizers [38] we thus also recommend to define optimal dosing for radiosensitization. In particular when radiosensitization occurs at very low doses as shown here, our study also stresses the need of robust and sensitive (pharmacodynamic) biomarkers in the clinic that in contrast to the monotherapy setting reflect the altered radiation response well.

Conflict of interest statement

The authors have no conflict of interest. MJC is employee of AstraZeneca.

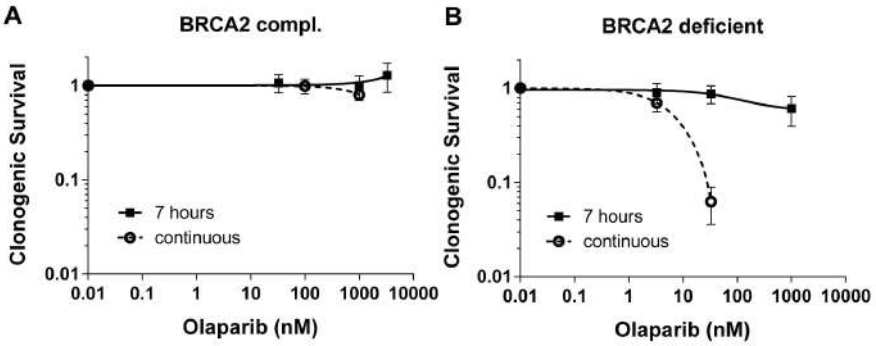
Acknowledgement

The authors wish to acknowledge the significant contribution from A.C. Begg, our esteemed colleague who passed away in January 2014. We would like to thank M. Verwijs-Janssen, B. Inanc, D. Pluim for technical assistance and F. Balm, S. Zerp and B. van Triest for discussions and careful reading of the manuscript. We are grateful for the kind gift of the HNSCC cell lines from R. Grenman. This work was supported by grants from Verwelius and Brunel.



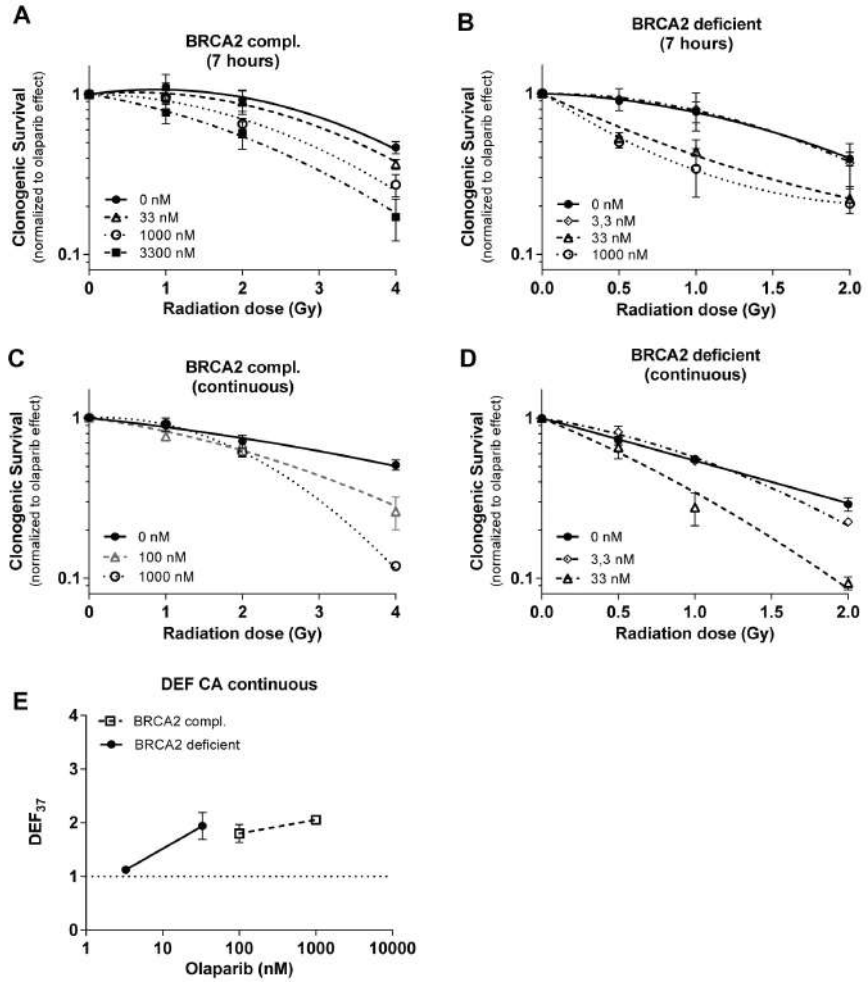
Supplemental Figure 1. Olaparib radiosensitizes both, the BRCA2-deficient and the BRCA2-complemented, cell lines.

(A-B) Viability of BRCA2-complemented and BRCA2-deficient cells after radiation and continuous olaparib exposure as determined by long-term growth inhibition assays (Cyquant). **(C-D)** Viability of respectively figures A & B normalized to olaparib effect at the respective indicated doses. Data are presented as mean ±SD of 3-7 independent experiments (except for 0.33nM and 66nM BRCA2-deficient datapoints n=2).



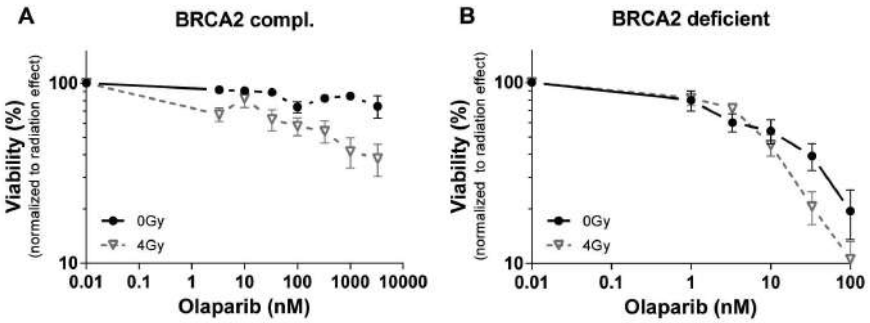
Supplemental Figure 2. Reduction in olaparib exposure time reduces cell kill by olaparib (i.e. single agent activity) in BRCA2-deficient cells.

(A-B) Surviving fractions of respectively BRCA2-complemented and BRCA2-deficient cells after continuous or short-term (7 hours) olaparib exposure as determined by clonogenic survival assay. Data are presented as mean ±SD of 4-6 independent experiments.



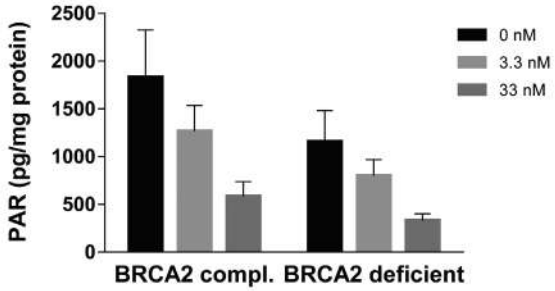
Supplemental Figure 3. Clonogenic survival after short and continuous olaparib exposure confirms that the extent of radiosensitization is dependent on olaparib dose and BRCA2 status.

(A-B) Surviving fractions (SF) of BRCA2-complemented and BRCA2-deficient cells after radiation and *7 hours* olaparib exposure as determined by clonogenic assays. SF are normalized to the values of the respective non-radiated samples. Data are presented as mean ±SEM of 3-6 independent experiments (datapoint 2Gy-1000nM BRCA2-deficient n=2). **(C-D)** SF of BRCA2-complemented and BRCA2-deficient cells after radiation and *continuous* olaparib exposure as determined by clonogenic assays. SF are normalized to the values of the respective non-radiated samples. Data are presented as mean ±SEM of 3-5 independent experiments (datapoint 1Gy-33nM BRCA2-deficient n=2). **(E)** Dose enhancement factors at 37% survival (DEF_{37}) at different olaparib doses after continuous exposure to olaparib and radiation as determined by clonogenic assays. Data are presented as mean ±SEM of 3-4 independent experiments.



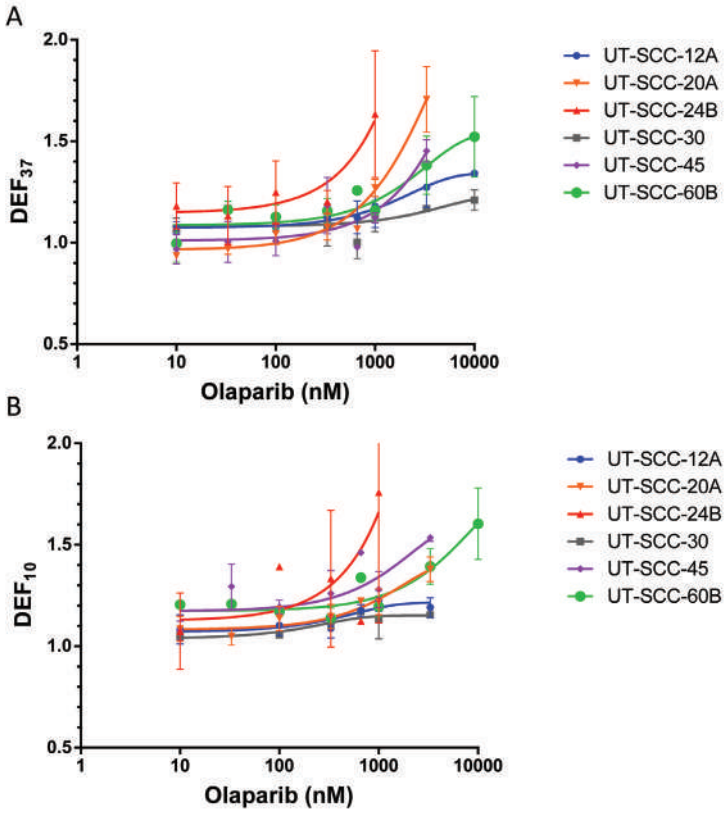
Supplemental Figure 4. Olaparib dose needed for effective radiosensitization is lower than dose needed for single agent activity.

(A-B) Viability of BRCA2-complemented and BRCA2-deficient cells after radiation and continuous olaparib exposure as determined by long-term growth inhibition assays (Cyquant). Viability values are normalized to 0nM olaparib controls at the respective radiation dose. Data are presented as mean ±SEM of 3-6 independent experiments.

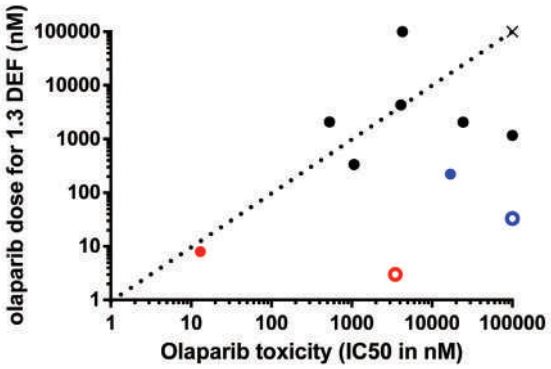


Supplemental Figure 5. Olaparib reduces PAR levels in both BRCA2-complemented and BRCA2-deficient cells.

PAR levels in pg/mg protein measured after olaparib treatment as indicated. There is no significance difference in baseline PAR levels (i.e. PAR levels of 0nM olaparib) between BRCA2-complemented and BRCA2-deficient cells (PAR levels in mean ±SD respectively 3508 ±1854 and 2034 ±823, p=0.20). Data are presented as mean ±SEM of 3-4 independent experiments.



Supplemental Figure 6. Olaparib mediated radiation dose enhancement factors
Radiation dose enhancement factors at 37% or 10% survival (DEF37, DEF10) achieved at different olaparib doses in the tested panel of human head and neck squamous cell carcinoma cell lines.



Supplemental Figure 7. Olaparib toxicity and radiosensitization

Olaparib doses achieving radiation dose enhancement factors (DEF_{37}) of 1.3 compared to olaparib caused toxicity (single agent activity) in the tested cell lines. Red circles indicate values from BRCA2-deficient mouse tumor cell lines after continuous or short olaparib exposure (filled and empty circles respectively), blue are the BRCA2-complemented. Note that, based on the non-linear dose response curves, a 50% kill (IC50) does not compare to the additional kill achieved by a 1.3 DEF. The graph shows that radio-sensitization potential does not closely correlate with olaparib sensitivity. However in general, radiosensitization was observed at lower olaparib concentrations than when exerting substantial olaparib toxicity in most cell lines and exposures. Values that could not be reached within the tested range were set to approach 100000nM for presentation purposes.

Supplemental Table 1: Olaparib toxicity and 2Gy Surviving Fractions (SF2)

	Cell line	BRCA2-compl.			BRCA2-deficient			UT-SCC-12A	UT-SCC-20A	UT-SCC-24B	UT-SCC-30	UT-SCC-45	UT-SCC-60B
	Assay	CA		LT-GI	CA		LT-GI	LT-GI	LT-GI	LT-GI	LT-GI	LT-GI	LT-GI
	Ola exposure	7 h	cont.	cont.	7 h	cont.	cont.	cont.	cont.	cont.	cont.	cont.	cont.
SF2	0nM	0.99 (±0.16)	0.76 (±0.03)	0.83 (±0.07)	0.39 (±0.05)	0.29 (±0.03)	0.38 (±0.04)	0.74 (±0.09)	-	-	-	-	-
	3.3nM	0.71 (±0.03)	-	0.75 (±0.11)	0.32 (±0.12)	0.18 (±0.02)	0.47 (±0.03)	-	-	-	-	-	-
	33nM	0.70 (±0.08)	-	0.66 (±0.16)	0.20 (±0.05)	0.13 (±0.05)	0.27 (±0.04)	-	-	-	-	0.32 (±0.14)	0.50 (±0.07)
	100nM	-	0.63 (±0.08)	0.77 (±0.08)	-	-	0.24 (±0.02)	0.59 (±0.04)	0.56 (±0.14)	0.34 (±0.03)	0.50 (±0.07)	0.29 (±0.07)	0.47 (±0.12)
	1000nM	0.53 (±0.12)	0.68 (±0.11)	0.55 (±0.13)	0.10 (±0.05)	-	-	0.54 (±0.07)	0.45 (±0.09)	0.19 (±0.05)	0.45 (±0.08)	0.22 (±0.06)	0.42 (±0.06)
Tox	3.3nM	-	-	0.92 (±0.03)	0.90 (±0.09)	0.70 (±0.07)	0.60 (±0.07)	-	-	-	-	0.89 (±0.08)	-
	33nM	1.07 (±0.10)	-	0.89 (±0.01)	0.87 (±0.08)	0.06 (±0.01)	0.39 (±0.07)	-	0.82 (±0.06)	0.95 (±0.16)	-	0.76 (±0.07)	-
	100nM	-	0.99 (±0.09)	0.74 (±0.05)	-	-	0.19 (±0.06)	0.79 (±0.02)	0.84 (±0.05)	0.85 (±0.04)	0.84 (±0.01)	0.64 (±0.02)	0.81 (±0.04)
	330nM	-	-	0.83 (±0.03)	-	-	0.06 (±0.02)	0.76 (±0.03)	0.79 (±0.05)	0.75 (±0.08)	0.80 (±0.04)	0.56 (±0.05)	0.79 (±0.03)
	1000nM	1.00 (±0.11)	0.80 (±0.05)	0.85 (±0.03)	0.61 (±0.09)	-	-	0.75 (±0.06)	0.79 (±0.03)	0.58 (±0.06)	0.77 (±0.02)	0.51 (±0.06)	0.77 (±0.04)

The surviving fraction at 2Gy and the toxicity of olaparib as surviving fraction in unradiated cells at the given olaparib dose are listed as determined by clonogenic assays (CA) or long-term growth inhibition assays (LT-GI) after continuous (cont) or 7h olaparib exposure. The SF2 are determined using the linear-quadratic fits on the surviving fraction data that were normalized to the olaparib cytotoxicity at the indicated olaparib doses. Data are presented as mean ±SEM of 4-6 independent experiments.

Supplemental Table 2: Comparison of radiation enhancement effect and olaparib cytotoxicity

Cell line	Ola Exposure	Ola con (nM) at DEF 1.3	Ola IC50 (nM)
BRCA2-deficient	7 hours	3.3	3503
	Cont.	8	16
BRCA2-compl.	7 hours	1000	> 10000
	Cont.	223	16562
UT-SCC-12A	Cont.	4317	4104
UT-SCC-20A	Cont.	1166	120498
UT-SCC-24B	Cont.	335	1068
UT-SCC-30	Cont.	>10000	4289
UT-SCC-45	Cont.	2082	528
UT-SCC-60B	Cont.	2071	24361

The Table lists and compares the olaparib dose needed to achieve a radiation dose enhancement factor (DEF_{37}) of 1.3 and the olaparib dose needed to achieve 50% kill as single agent (IC_{50}) in the tested cell lines. Note, that all cells were cultured at 5% oxygen. Values were calculated from the fits on the DEF graphs, except 7h Ola concentration values that were based on the DEF calculations from clonogenic assay data at the given dose (see Figure 2F and Supplemental Figure 6). Values that could not be reliably calculated and were not reached within the tested dose ranges are indicated as larger than the highest tested dose.

References

[1] Begg AC, Stewart FA, Vens C. Strategies to improve radiotherapy with targeted drugs. Nat Rev Cancer. 2011;11:239-53.

[2] Ree AH, Hollywood D. Design and conduct of early-phase radiotherapy trials with targeted therapeutics: lessons from the PRAVO experience. Radiother Oncol. 2013;108:3-16.

[3] Verheij M, Vens C, van Triest B. Novel therapeutics in combination with radiotherapy to improve cancer treatment: Rationale, mechanisms of action and clinical perspective. Drug ResistUpdat. 2010;13:29-43.

[4] Ataman OU, Sambrook SJ, Wilks C, Lloyd A, Taylor AE, Wedge SR. The clinical development of molecularly targeted agents in combination with radiation therapy: a pharmaceutical perspective. Int J Radiat Oncol Biol Phys. 2012;84:e447-54.

[5] Mehta MP, Wang D, Wang F, Kleinberg L, Brade A, Robins HI, et al. Veliparib in combination with whole brain radiation therapy in patients with brain metastases: results of a phase 1 study. J Neurooncol. 2015.

[6] Reiss KA, Herman JM, Zahurak M, Brade A, Dawson LA, Scardina A, et al. A Phase I Study of Veliparib (ABT-888) in Combination with Low-Dose Fractionated Whole Abdominal Radiation Therapy in Patients with Advanced Solid Malignancies and Peritoneal Carcinomatosis. Clin Cancer Res. 2015;21:68-76.

[7] Audeh MW, Carmichael J, Penson RT, Friedlander M, Powell B, Bell-McGuinn KM, et al. Oral poly(ADP-ribose) polymerase inhibitor olaparib in patients with BRCA1 or BRCA2 mutations and recurrent ovarian cancer: a proof-of-concept trial. Lancet. 2010;376:245-51.

[8] Fong PC, Boss DS, Yap TA, Tutt AN, Wu P, Mergui-Roelvink M, et al. Inhibition of poly(ADP-ribose) polymerase in tumors from BRCA mutation carriers. N Engl J Med. 2009;361:123-34.

[9] Kaufman B, Shapira-Frommer R, Schmutzler RK, Audeh MW, Friedlander M, Balmana J, et al. Olaparib Monotherapy in Patients With Advanced Cancer and a Germline BRCA1/2 Mutation. J Clin Oncol. 2014.

[10] Tutt A, Robson M, Garber JE, Domchek SM, Audeh MW, Weitzel JN, et al. Oral poly(ADP-ribose) polymerase inhibitor olaparib in patients with BRCA1 or BRCA2 mutations and advanced breast cancer: a proof-of-concept trial. Lancet. 2010;376:235-44.

[11] Chatterjee P, Choudhary GS, Sharma A, Singh K, Heston WD, Ciezki J, et al. PARP inhibition sensitizes to low dose-rate radiation TMPRSS2-ERG fusion gene-expressing and PTEN-deficient prostate cancer cells. PLoS ONE. 2013;8:e60408.

[12] Nowsheen S, Bonner JA, Yang ES. The poly(ADP-Ribose) polymerase inhibitor ABT-888 reduces radiation-induced nuclear EGFR and augments head and neck tumor response to radiotherapy. Radiother Oncol. 2011;99:331-8.

[13] Powell C, Mikropoulos C, Kaye SB, Nutting CM, Bhide SA, Newbold K, et al. Pre-clinical and clinical evaluation of PARP inhibitors as tumour-specific radiosensitisers. Cancer Treat Rev. 2010;36:566-75.

[14] Senra JM, Telfer BA, Cherry KE, McCrudden CM, Hirst DG, O'Connor MJ, et al. Inhibition of PARP-1 by olaparib (AZD2281) increases the radiosensitivity of a lung tumor xenograft. Mol Cancer Ther. 2011;10:1949-58.

[15] Dungey FA, Loser DA, Chalmers AJ. Replication-dependent radiosensitization of human glioma cells by inhibition of poly(ADP-Ribose) polymerase: mechanisms and therapeutic potential. Int J Radiat Oncol Biol Phys. 2008;72:1188-97.

[16] Noel G, Godon C, Fernet M, Giocanti N, Megnin-Chanet F, Favaudon V. Radiosensitization by the poly(ADP-ribose) polymerase inhibitor 4-amino-1,8-naphthalimide is specific of the S phase of the cell cycle and involves arrest of DNA synthesis. Mol Cancer Ther. 2006;5:564-74.

[17] Loser DA, Shibata A, Shibata AK, Woodbine LJ, Jeggo PA, Chalmers AJ. Sensitization to radiation and alkylating agents by inhibitors of poly(ADP-ribose) polymerase is enhanced in cells deficient in DNA double-strand break repair. Mol Cancer Ther. 2010;9:1775-87.

[18] Le QT, Raben D. Integrating biologically targeted therapy in head and neck squamous cell carcinomas. Semin Radiat Oncol. 2009;19:53-62.

[19] Deutsch E, Soria JC, Armand JP. New concepts for phase I trials: evaluating new drugs combined with radiation therapy. Nat Clin Pract Oncol. 2005;2:456-65.

[20] Evers B, Schut E, van der Burg E, Braumuller TM, Egan DA, Holstege H, et al. A high-throughput pharmaceutical screen identifies compounds with specific toxicity against BRCA2-deficient tumors. Clin Cancer Res. 2010;16:99-108.
[21] Lansford CD, Grenman R, Bier H, Somers KD, Kim SY, Whiteside TL, et al. Cancer Cell Lines Part 2. In: Masters JRW, Palsson B, editors. Human Cell Culture. Dordrecht: Kluwer Academic Press; 2000. p. 185-255.
[22] Jonkers J, Meuwissen R, van der Gulden H, Peterse H, van der Valk M, Berns A. Synergistic tumor suppressor activity of BRCA2 and p53 in a conditional mouse model for breast cancer. Nat Genet. 2001;29:418-25.
[23] Bryant HE, Schultz N, Thomas HD, Parker KM, Flower D, Lopez E, et al. Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. Nature. 2005;434:913-7.
[24] Evers B, Drost R, Schut E, de Bruin M, van der Burg E, Derksen PW, et al. Selective inhibition of BRCA2-deficient mammary tumor cell growth by AZD2281 and cisplatin. Clin Cancer Res. 2008;14:3916-25.
[25] Farmer H, McCabe N, Lord CJ, Tutt AN, Johnson DA, Richardson TB, et al. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. Nature. 2005;434:917-21.
[26] Verhaegh GW, Jongmans W, Morolli B, Jaspers NG, van der Schans GP, Lohman PH, et al. A novel type of X-ray-sensitive Chinese hamster cell mutant with radioresistant DNA synthesis and hampered DNA double-strand break repair. Mutat Res. 1995;337:119-29.
[27] Kraakman-van der Zwet M, Wiegant WW, Zdzienicka MZ. Brca2 (XRCC11) deficiency results in enhanced mutagenesis. Mutagenesis. 2003;18:521-5.
[28] Benjamin RC, Gill DM. ADP-ribosylation in mammalian cell ghosts. Dependence of poly(ADP-ribose) synthesis on strand breakage in DNA. J Biol Chem. 1980;255:10493-501.
[29] Balmana J, Tung NM, Isakoff SJ, Grana B, Ryan PD, Saura C, et al. Phase I trial of olaparib in combination with cisplatin for the treatment of patients with advanced breast, ovarian and other solid tumors. Ann Oncol. 2014;25:1656-63.
[30] Khan OA, Gore M, Lorigan P, Stone J, Greystoke A, Burke W, et al. A phase I study of the safety and tolerability of olaparib (AZD2281, KU0059436) and dacarbazine in patients with advanced solid tumours. Br J Cancer. 2011;104:750-5.
[31] Liu JF, Tolaney SM, Birrer M, Fleming GF, Buss MK, Dahlberg SE, et al. A Phase 1 trial of the poly(ADP-ribose) polymerase inhibitor olaparib (AZD2281) in combination with the anti-angiogenic cediranib (AZD2171) in recurrent epithelial ovarian or triple-negative breast cancer. Eur J Cancer. 2013;49:2972-8.
[32] Rajan A, Carter CA, Kelly RJ, Gutierrez M, Kummar S, Szabo E, et al. A phase I combination study of olaparib with cisplatin and gemcitabine in adults with solid tumors. Clin Cancer Res. 2012;18:2344-51.
[33] Murray J, Thomas H, Berry P, Kyle S, Patterson M, Jones C, et al. Tumour cell retention of rucaparib, sustained PARP inhibition and efficacy of weekly as well as daily schedules. Br J Cancer. 2014;110:1977-84.
[34] Rottenberg S, Jaspers JE, Kersbergen A, van der Burg E, Nygren AO, Zander SA, et al. High sensitivity of BRCA1-deficient mammary tumors to the PARP inhibitor AZD2281 alone and in combination with platinum drugs. ProcNatlAcadSciU S A. 2008;105:17079-84.
[35] Bundred N, Gardovskis J, Jaskiewicz J, Eglitis J, Paramonov V, McCormack P, et al. Evaluation of the pharmacodynamics and pharmacokinetics of the PARP inhibitor olaparib: a Phase I multicentre trial in patients scheduled for elective breast cancer surgery. Invest New Drugs. 2013.
[36] Liu SK, Coackley C, Krause M, Jalali F, Chan N, Bristow RG. A novel poly(ADP-ribose) polymerase inhibitor, ABT-888, radiosensitizes malignant human cell lines under hypoxia. RadiotherOncol. 2008;88:258-68.
[37] Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144:646-74.
[38] Harrington KJ, Billingham LJ, Brunner TB, Burnet NG, Chan CS, Hoskin P, et al. Guidelines for preclinical and early phase clinical assessment of novel radiosensitisers. Br J Cancer. 2011;105:628-39.

CHAPTER 5

Transcriptionally inactive HPV is associated with poor prognosis in chemo-radiotherapy treated oropharyngeal squamous cell carcinomas patients

Caroline V.M. Verhagen, Roelof JC Kluin, Martijn vd Heijden, David M. Vossen, Marcel Verheij, Stefan M. Willems, Michiel W.M. van den Brekel and Conchita Vens.

Submitted

Abstract

Objectives

Human papilloma virus (HPV) - positivity is an important prognostic marker in head and neck squamous cell carcinoma (HNSCC). HPV positive tumors have a favorable prognosis, especially in oropharyngeal tumors (OPSCC). Current methods are mostly based on HPV transcriptionally active tumors. However, little is known about the role of HPV transcriptionally inactive (silent) tumors. We therefore tested different methods for detection of HPV and the role of silent HPV tumors.

Materials and Methods

In a homogenously treated cohort of 71 tumors of the oropharynx and hypopharynx, the presence of HPV was detected by several methods. These consist of (1) DNA and RNA (targeted) next generation sequencing, (2) immunohistochemical staining of $p16^{ink4a}$ and p53 and (3) classic HPV specific PCR on p16 overexpressing samples.

Results

Targeted sequencing showed that read coverage for silent HPV tumors is lower than for transcriptionally active tumors; transcriptional activity correlated highly with HPV-DNA read load. Silent HPV tumors as determined by sequencing were assigned to be HPV negative by the other methods used. The presence of silent HPV was 52% in the oropharyngeal tumors and 74% in the hypopharyngeal tumors. Moreover, patients with those tumors had a survival which is significantly worse than those with true HPV positive OPSCC (p=0.0005).

Conclusion

We found a higher percentage of silent HPV in HNSCC tumors than previously reported. Patients with silent HPV tumors have a survival similar to patients with tumors in which no HPV was detected. These transcriptionally inactive tumors are significantly different from transcriptionally active OPSCC.

Introduction

Human papilloma virus (HPV) has been implicated in head and neck squamous cell carcinoma (HNSCC) since over 20 years and has been detected in a significant proportion of HNSCC [1]. The incidence of HPV-related carcinomas is increasing in most developed countries. From all HNSCC subsites, oropharynx tumors are most frequently found to be HPV positive. HNSCC patient outcome association studies revealed that HPV-associated oropharyngeal SCC (OPSCC) show better outcomes compared to HPV negative, with all treatment modalities, even at de-escalated RT doses [2-12]. This led to an independent classification in the recent TNM staging [13]. Routine testing for HPV in oropharynx tumors has become standard in most clinics [14] to support prognosis and sometimes amend treatment. Changing this treatment, however, requires a HPV detection methodology with (a) high accuracy and (b) a clinically relevant outcome association [15].

Based on this favorable prognosis, treatment de-intensification trials are being conducted for HPV-associated OPSCC that aim to minimize normal tissue toxicities from standard chemo-radiotherapy. Thus, HPV-positivity is an important prognostic biomarker in HNSCC [16]. Only a fraction of non-oropharyngeal HNSCC has been reported to be HPV positive; these tumors do not appear to have a better outcome in these specific subgroups/populations [2]. However, recent data suggests that HPV-positivity may also be associated with improved survival in oral cavity and hypopharyngeal tumors [17]. Given the clinical relevance of the presence of HPV, it is important to evaluate both the HPV detection methods and the impact of HPV activity in this distinction.

HPV is a DNA virus of the papilloma family, comprising over 120 different types that can infect mucosal and cutaneous epithelia [18].The expression of the E1 and E2 HPV genes support HPV episome amplification after infection and endocytosis. E6 and E7 expression, from episomal HPV or from HPV integrated in the host genome, however, interferes in cellular signaling to suppress cellular differentiation and to promote re-entrance into the cell cycle. Among others, it results in p53 inactivation and degradation of the tumor suppressor pRB, which leads to $p16^{INK14}$ overexpression, a prime surrogate marker for HPV in HNSCC. The elimination of damage-induced cell cycle blockade facilitates chromosomal aberrations and their retention, thereby supporting progression to invasive cancer. Together with the maintenance of cellular proliferation, also driven by the expression of viral E6 and E7, these are essential components in tumorigenesis [18]. The loss of E1 and E4/5 and parts of E2 during HPV integration has been postulated to cause the elevation of E6/E7 expression [19, 20]. Later studies, however, showed that E6/E7 expression was high in HPV-PCR positive and p16 overexpression positive OPSCC, regardless whether HPV was integrated, episomal or both [21, 22]. All analyzed HNSCC cell lines that were positive for HPV-DNA show expression of E2/E6 or E7, independent of copy number or viral load [21], supporting the notion that HPV is transcriptionally active in HNSCC.

HPV-associated HNSCC are biologically distinct from HPV negative HNSCC [23]. Recent genetic studies highlight the dissimilar gene mutation spectrum, most notably the absence of TP53 mutations and CDKN2A or CCND1 gene aberrations, copy number variations and also differences in the gene expression profile in HPV positive and HPV negative HNSCC [23-26]. To determine HPV status, most of these studies used established methods, such as assessment of p16 overexpression, by immunohistochemistry in combination with HPV-specific PCR, or in situ hybridization techniques, to detect and/or localize DNA of HNSCC-relevant HPV types, or E6 RNA detection methods. As determined by these methods, a causal link to HPV infection is most apparent in oropharyngeal cancers [18, 27]. Although HPV positivity in HNSCC has been reported outside the oropharynx, the low incidence does not warrant routine testing yet. A high proportion (>90%) of virus positive OPSCC are infected by the variant HPV16. However, also HPV18 and 33, among other types, have been reported in OPSCC [28-30].

Many studies epitomize this crucial role of HPV detection and methodology [31-34]. Immunohistochemical staining followed by a HPV-specific PCR in case of p16 positivity, is a common strategy to determine HPV status in OPSCC. Overexpression of p16 is defined as strong nuclear and/or cytoplasmic expression in >70% of invasive tumor cells [35]. P16 positivity has been reported to have a 96.8% sensitivity, but only a 83.8% specificity as a HPV surrogate marker [36], since HPV-unrelated p16 overexpression occurs. In HPV positive HNSCC the oncoprotein E6 binds to p53, causing p53 degradation and inactivation which exempts p53 aberrations during carcinogenesis. The absence of p53 alterations in p16 positive OPSCC has been therefore proposed as a good surrogate marker for HPV-associated HNSCC [37]. However, p16 high but HPV negative are linked to a particular poor prognosis [33] justifying rigorous HPV-specific detection assays, such as HPV-DNA in situ hybridization or PCR [14, 38]. HPV classification methods will vary in their sensitivity and specificity and have been the focus of recent studies [32]. DNA sequencing, facilitated by next generation sequencing technologies, has been proposed to confirm HPV status in HNSCC [39, 40]. In situ hybridization FISH, RNA-seq and other techniques further complement the spectrum of HPV tests that aim to sensitively detect either HPV-DNA or HPV-RNA, from HPV gene expression. Notably, concordance in these methods can only be expected under the condition that HPV in OPSCC is always transcriptionally active.

Despite its potential significance, only few studies investigated the occurrence of transcriptionally inactive HPV in HNSCC and their characteristics. The role of HPV transcriptional activity in OPSCC patient outcome association, in particular, remains elusive.

In this study we aimed to investigate the transcriptional activity of HPV in HNSCC since HPV status is an important marker and has a role in staging in OPSCC. However, little is known about the role of transcriptional inactive HPV tumors. We therefore

look at the impact of tumors with transcriptional inactive HPV on patient outcome in a cohort of curatively chemo-radiotherapy treated patients with advanced HNSCC.

Materials and methods

Patient and patient material

Pre-treatment pharyngeal squamous cell carcinoma tumor samples were available from 71 patients, enrolled in the Netherlands Cancer Institute for concomitant chemoradiotherapy (Supplementary Table 1). All patients received cisplatin based chemotherapy. Radiotherapy was administered according to a conventional (35 fractions of 2 Gy; overall treatment time 7 weeks) or DAHANCA scheme (35 fractions of 2 Gy; overall treatment time 6 weeks). Chemotherapy consisted of daily (25 x 6 mg/m2) or 3-weekly (3 x 100 mg/m2) intravenous administration of cisplatin. A minority (n=14) received intra-arterial cisplatin (4 x 150 mg/m2) according to the RADPLAT trial [41]. Not all patients (n=12) completed chemotherapy and therefore the individual total cumulative cisplatin dose was recorded. The collection of the biopsies of this tumor material was approved by the Institutional Review Board and all patients granted written informed consent. The retrospective 'NKI-CRAD' cohort provided pretreatment material, of patients with HNSCC, who were treated at the Netherlands Cancer Institute between 2001 and 2014 [42]. From 71 patients with oropharyngeal (n=43) and hypopharyngeal (n=28) tumors, sufficient material with the sufficient tumor percentage (see next paragraph) was available. Formalin fixed paraffin-embedded tissue samples, as well as matched frozen material were available for all these patients

Sample preparation and DNA and RNA extraction

Fresh frozen tumor material was collected after sectioning and tumor content was assessed by a pathologist from representative sections after hematoxylin and eosin (H&E)-staining. Only samples with a tumor content above 50% were used for DNA and RNA isolation. DNA and RNA were extracted with AllPrep DNA/RNA Mini Kits (Qiagen) following the manufacture's protocols.

Library preparation and DNA sequencing

Paired-end (PE) fragment libraries were prepared using a genomic DNA library preparation kit (Illumina) and hybridized to a SureSelect custom-based bait library (Agilent) which contained baits to 550 cancer related genes [43] and to the E1, E6, E7 and L1 region of HPV 16 and 18 (Figure 1a and Supplementary Table 2). After washing, the captured DNA was amplified. Enriched libraries were multiplexed using custom

adapters containing an index sequence and sequenced on Illumina GAII and Hiseq 2000 platforms using a 2x75bp PE protocol.

RNA sequencing

RNA sequencing was performed as published by van der Heijden et al. [44].

HPV classification by immunohistochemistry

Immunohistochemical staining of $p16^{ink4a}$ and p53 was used to test for HPV**.** Immunohistochemistry of the samples was performed on a BenchMark Ultra autostainer (Ventana Medical Systems). Paraffin sections were cut at 3µm, heated at 75º for 2 minutes and deparaffinized with E prep solutions (Ventana Medical Systems) and submitted to heat-induced antigen retrieval procedures with Cell Conditioning 1 (CC1, Ventana Medical Systems). The antibody clone DO-7 was used for p53 (Dako) and clone JC8 for p16 detection (Santa Cruz) while using the Optiview Amplification kit (Ventana Medical Systems) for p16 detection only. Slides were counterstained with Hematotoxylin II and Bluing Reagent (Ventana Medical Systems) after applying the Optiview DAB Detection Kit (Ventana Medical Systems).

Classification by the pathologists followed institutional and international guidelines: to be classified as p16 positive, >70% of invasive tumor cells should show strong and diffuse nuclear and cytoplasmatic p16 expression [35]. P53 immunohistochemistry was used to classify invasive tumor cells as either wildtype, or mutated. A mutated phenotype was defined as i) null mutation (complete lack of nuclear p53 staining), or ii) overexpression phenotype as previously described [45]. All p16 overexpressing samples were further tested by PCR.

HPV classification by PCR

Standard PCR-based techniques were used to identify HPV-DNA on diagnostic formalin fixed diagnostic HNSCC. Fresh frozen material was used for a total of two cases to compensate for the lack of FFPE material and an invalid result. To validate our method, PCR was performed on FFPE and fresh frozen tissue biopsies of four other samples, including two non-oropharyngeal tumors. HPV genotyping was performed using the PapilloCheck (Greiner Bio-One), that assesses for 24 HPV types simultaneously, including 18 high and 6 low risk types.

HPV classification by DNA and RNA sequencing

For HPV alignment and read count determination, several genomes were collected from the following sources after querying the European nucleotide archive for HPV (using various keywords, but leaving out all integration, diagnostic and patented sequences): The European Nucleotide Archive, Zilvinas Kucinskas bioinformatics database and the Nucleotide database (NCBI). The HPV genomes were combined in a single reference

file. After detection all samples were mapped against a reference containing only the observed HPV genomes, using BWA mem 7.12. To enable filtering of human reads in the samples, samples were also aligned with BWA mem to hg38. Reads which mapped better to the human than the HPV genome according to the number of matches, were filtered using XenofilteR [46]. The read counts per HPV-strain was taken as an indicator for its presence in the sample.

Statistics

The end points were progression free (PFS) and overall survival (OS). Progression free survival was defined as the time from the last day of the treatment to death or the first documented local or regional recurrence or distant metastases. Overall survival was defined as the time from the last day of treatment until death.

Results

Transcriptionally inactive HPV in advanced pharyngeal HNSCC

Pretreatment HNSCC samples from 71 patients (Supplementary Table 1), all advanced pharyngeal and treated with definitive cisplatin-based chemoradiotherapy, were tested for presence of HPV-DNA. The read coverage of HPV was determined after DNA sequencing using a cancer related gene capture set that also contained baits for HPV16 and 18. We found evidence for HPV in 86% of the cases with different degrees of read coverage (Figure 1b). Remarkably, and distinctively different from previous reports that used different detection techniques, only seven percent (n=3) of the oropharynx tumors showed no signs of HPV-DNA, however 93% did, while 75% of the hypopharyngeal were positive for HPV-DNA. Three different read coverage groups are discernable from the distribution: no HPV reads, low read counts, up to 20,000 reads of which the ^{2}log is 14, and samples with high read counts, up to 1.4x10E7 reads of which the ^{2}log is 24 (Figure 1c). The latter lacks hypopharyngeal cases. Overall, oropharyngeal cases show greater read numbers in the low read count group than hypopharyngeal HNSCC (2812 vs 591) indicating a higher degree of HPV-DNA in the samples (Figure 1c).

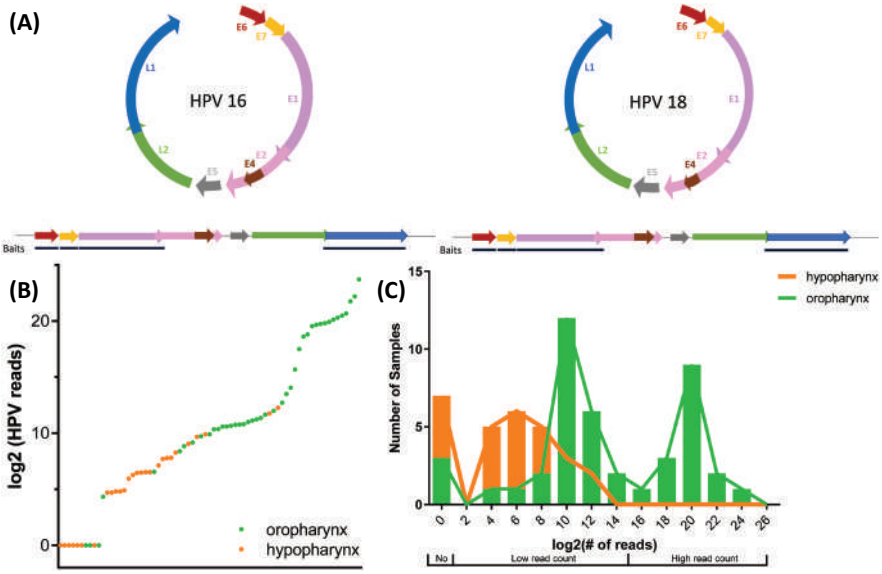


Figure 1: DNA sequencing in HNSCC

(A) Schematic of the genome of HPV types 16 and 18 designed from Ref seq entry: NC_001526 resp. AY262282. The HPV genome has a circular double-stranded DNA structure. The genomes consist of 6 early (E1, E2, E4, E5, E6 and E7) and 2 late (L1 and L2) genes. Underneath is a linear representation of the genome, which shows for which regions of the viral genome baits were designed. **(B)** Coverage distribution of HPV-DNA reads (log scale) of oropharyngeal (green) and hypopharyngeal (orange) tumor samples (n=71). **(C)** Coverage distribution in bins of HPV-DNA reads of oropharyngeal (green) and hypopharyngeal (orange) samples (n=71). The distribution shows three peaks.

We next determined transcriptional activity of HPV by RNA sequencing (Figure 2a) for 42 out of the 43 oropharyngeal tumors. RNA quality was not sufficient in one sample. Of the hypopharyngeal tumors 27 were sequenced. Figure 2b shows the range of RNA read coverage over the hypo- and oropharyngeal samples, revealing that 24% of the HNSCC samples express genes to a high degree, with 40% of oropharyngeal samples showing transcriptional active HPV. Notably, none of the hypopharyngeal samples show HPV expression (Supplementary Figure 1).

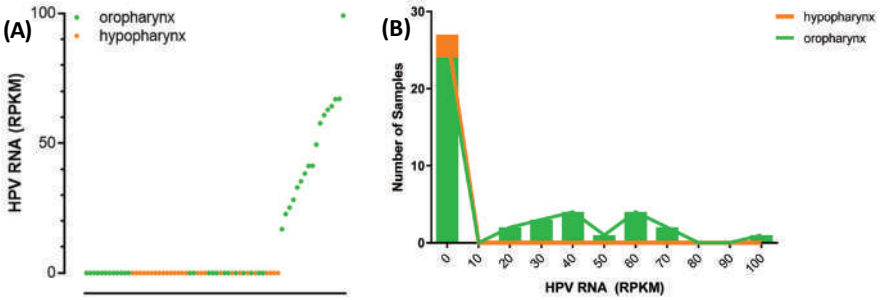


Figure 2: RNA sequencing in HNSCC

Oropharyngeal tumors are shown in green, hypopharyngeal tumors in orange (n=69). **(A)** Coverage distribution of HPV RNA reads (in Reads Per Kilobase Million (RPKM)). **(B)** Coverage distribution in bins of HPV RNA reads (in RPKM).The distribution shows two peaks.

A relatively large fraction of the HPV-DNA positive samples turn out to have transcriptionally inactive HPV in both the oro- and the hypopharyngeal tumors (Table 1 and Figure 3a). Transcriptional activity correlates highly with HPV-DNA load (Figure 3b). HPV positive, but transcriptionally inactive (termed silent HPV here) oropharyngeal cases have been reported before [47], however with a much lower incidence. Next generation DNA sequencing techniques and the inclusion of HPV baits to facilitate HPV detection, likely increased the sensitivity in detecting HPV-DNA.

Comparison with $p16^{ink4a}$ /PCR and $p16^{ink4a}$ /p53 immunohistochemistry based classifications

Next, HPV status by standard techniques, $p16^{ink4a}$ /PCR and $p16^{ink4a}$ /p53 immunohistochemistry, was assessed in this patient group. Of the 43 oropharyngeal carcinomas, 18 showed p16 overexpression. Seventeen of these tumors had a p53 wildtype status, whereas one was mutated. Hence this latter one was classified as HPV negative and the others as HPV positive (Figure 3a).

We next performed standard HPV genotyping by PCR on p16-overexpressing oropharynx tumors (n=18). In concordance with the p53 wildtype status, this technique classified 17 oropharyngeal carcinomas as HPV positive (40%) (Figure 3a). One of these 17 tumors was positive for HPV33, the others for HPV16. One of the p16-overexpressing samples was HPV negative as determined by PCR and in concordance showed a mutated p53 on IHC.

Overall, we find that, despite substantial HPV read counts, only samples which are transcriptionally active were detected as HPV positive by standard techniques using p16/p53 immunohistochemistry and a subsequent PCR validation step (Figure 3a).

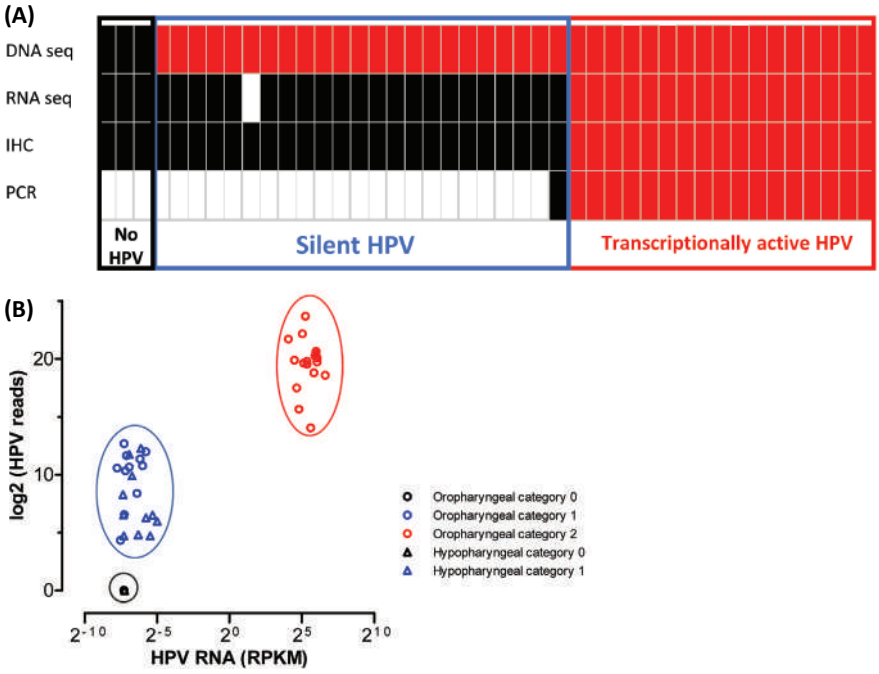


Figure 3: Comparison of methods of HPV detection

(A) Comprehensive summary of the different methods of HPV detection in oropharyngeal tumors. Tumors are ordered based on number of HPV-DNA reads. HPV positive samples as assessed by the specific method are shown in red; in black HPV negative samples. **(B)** Correlation of RNA sequencing reads versus DNA sequencing reads. Tumors in which no HPV is detected are shown in black, silent HPV tumors in blue and in red the transcriptionally active tumors.

Role of p16 status, p16 percentage and HPV16 variants

In literature it has been shown that p16+/HPV negative tumors do significantly worse than true HPV negative tumors [33, 48]. Accordingly, in our series, the patient with a p16-positive, but as defined by p53 status and PCR, HPV negative OPSCC had a local recurrence and distant metastasis after only 6 months. Moreover, we checked if patients with p16- oropharyngeal tumors with some (10-70%) p16 staining (n=4) do worse. Their overall survival is 50%, which is not significantly different from patients with p16-negative OPSCC (<10% staining), nor from p16-positive OPSCC (Data not shown). In our series, however, the numbers are too low to allow for any statistically relevant comparison.

Since we wondered if some HPV16 variants were more present in the high or low DNA read count group, we determined HPV16 variants in our cohort of oropharyngeal tumors. Sequence variants of the HPV 16 E6 ORF have been found to correctly classify HPV16-variants [49, 50]. Nucleotide positions of E6 were checked for positions 109,

131, 132, 143, 145, 178, 183, 256, 267, 269, 286, 289, 335, 350, 403 and 532. The Non European (NE) variant contains E6 changes at nucleotide 145, 286, 289 and 335 when compared to the HPV16 reference genome (E(P)), which is a German isolate and is member of the European lineage. In addition to these consensus changes, class-specific changes can be present. If changes are present, but not all the consensus changes, the sample is identified as a European variant (EV). We found a random distribution of coverage over the different HPV16 variants (Supplementary Figure 2a). In total there were 7/39 E(P), 24/39 EV and 8/39 NE HPV16 variants. In the high coverage group, 4/16 were E(P), 10/16 EV and 2/16 NE HPV16. This is not significantly different. Progression-free survival of patients with a European (E(P)), a European variant or a Non European (NE) variant of HPV16 is not different (Supplementary Figure 2b), nor when the E(P) and EV were grouped together vs the NE variant (p=0.8498; data not shown)

HPV transcriptional activity association with overall survival and progression free survival after chemoradiotherapy

The low incidence and number of samples with silent HPV, as determined by older techniques, prevented patient survival association analysis with transcriptional activity in previous studies [47].

In patients with oropharyngeal HNSCC, an association between HPV and better prognosis has been extensively proven. Patients with HPV positive oropharyngeal tumors, as determined by standard techniques, consisting of p16-staining followed by PCR on p16 positive tumors, have significantly better progression free and overall survival, confirming such an association in our patient cohort (Data similar to Figure 4a and b).

Using our combined sequencing data, we defined three HNSCC groups with i) HPV observed in both RNA-seq and DNA (transcribed HPV), ii) HPV-DNA, but missing expression (silent HPV) and iii) a small group free of HPV markers. These three groups and their characteristics are shown in Table 1. We find that patients with transcriptionally inactive but HPV-DNA positive OPSCC had a significantly worse progression free survival, compared to OPSCC cases with transcriptionally active HPV (p=0.0005). Progression free survival is not different (p=0.7294) from the third group free of HPV markers, although number of cases were too low for appropriate comparisons (Figure 4a). Overall survival showed a similar relation (Figure 4b). Locoregional control, however, is not different between groups (Figure 4c). Hypopharyngeal tumors do not show HPV gene expression (Figure 2b), and survival of patients is similar for the groups with or without HPV-DNA (Figure 4d and e). Interestingly locoregional control is significantly worse in hypopharyngeal tumors where no HPV is detected compared with the low read count group (Figure 4f).

Together, while we find a considerable fraction of HPV-DNA positive HNSCC, only HNSCC with transcriptional active HPV are associated with a good patient prognosis. Classification by transcriptional activity as determined by RNA-Sequencing corresponds with standard immunohistochemical and PCR based classifications further support its use for patient classification in survival analyses.

Table 1: Demographics of HNSCC cohort

Variable		Oropharynx (=42)			Hypopharynx (n=27)	
		No HPV detected (n=3)	Silent HPV (n=22)	Transcribed HPV (n=17)	No HPV detected (n=7)	Silent HPV (n=20)
Gender						
	Female	0 (0)	8 (36)	5 (29)	2 (29)	5 (25)
	Male	3 (100)	14 (64)	12 (71)	5 (71)	15 (75)
Age at diagnosis						
	Median (range)	70 (61-74)	57 (49-75)	58 (40-64)	61 (51-73)	59 (48-76)
	Mean (sd)	68 (7)	59 (8)	56 (7)	62 (7)	60 (8)
Disease stage						
	III	0 (0)	3 (14)	1 (6)	1 (14)	6 (30)
	IVA	3 (100)	13 (59)	12 (71)	5 (72)	13 (65)
	IVB	0 (0)	6 (27)	4 (24)	1 (14)	1 (5)
cT classification						
	T1	0 (0)	1 (5)	0 (0)	0 (0)	0 (0)
	T2	0 (0)	1 (5)	5 (29)	2 (29)	6 (30)
	T3	1 (33)	10 (45)	7 (41)	3 (43)	8 (40)
	T4a	2 (64)	7 (32)	4 (24)	1 (14)	5 (25)
	T4b	0 (0)	3 (14)	1 (6)	1 (14)	1 (5)
cN classification						
	N0	0 (0)	4 (18)	1 (6)	0 (0)	5 (25)
	N1	0 (0)	1 (5)	0 (0)	1 (14)	5 (25)
	N2	3 (3)	14 (64)	13 (76)	6 (86)	10 (50)
	N3	0 (0)	3 (14)	3 (18)	0 (0)	0 (0)
Smoker						
	Former smoker	0 (0)	2 (9)	3 (18)	1 (14)	6 (30)
	Never	0 (0)	0 (0)	2 (12)	0 (0)	1 (5)
	Unknown	0 (0)	1 (5)	3 (18)	1 (14)	1 (5)
	Yes	3 (100)	19 (86)	9 (53)	5 (72)	12 (60)
Alcohol consumption						
	Former alcoholic	1 (33)	4 (18)	1 (6)	1 (14)	4 (20)
	Never	0 (0)	1 (5)	0 (0)	0 (0)	1 (5)
	Unknown	0 (0)	1 (5)	4 (24)	1 (14)	1 (5)
	Yes	2 (64)	16 (73)	12 (71)	5 (72)	14 (70)
Number of DM events (%)		1 (33)	10 (45)	1 (6)	0 (0)	3 (15)
Number of LRC events (%)		1 (33)	3 (14)	2 (12)	2 (29)	0 (0)
Dead at 5 years (%)		2 (64)	16 (73)	3 (18)	3 (43)	7 (35)

The DNA and RNA from 69 oropharyngeal and hypopharyngeal pretreatment biopsies of patients were sequenced. Three groups are shown: 1) No HPV detected (DNA and RNA negative), 2) Silent HPV (DNA positive but RNA negative), 3) Transcribed HPV (DNA and RNA positive).

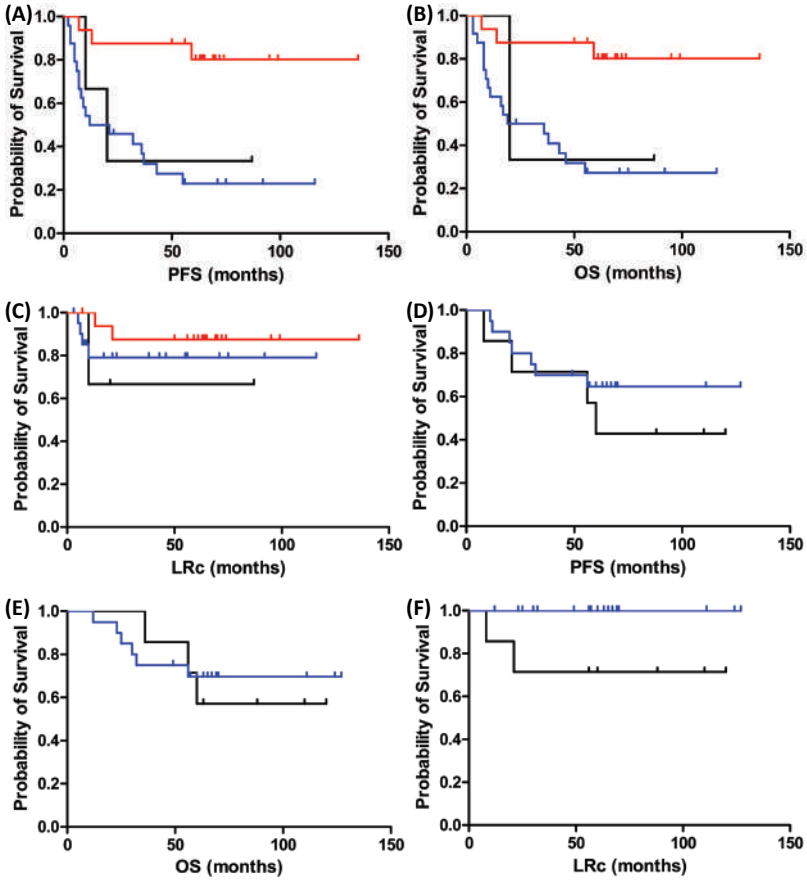


Figure 4: Survival curves of 42 oropharyngeal and 27 hypopharyngeal tumors

Tumors where no HPV is detected are shown as a black line, the silent HPV tumors in blue and in red the transcriptionally active tumors. **(A)** Progression free survival of the different subgroups of oropharyngeal tumors. Survival for silent HPV tumors is significantly worse than for transcriptionally active tumors (p=0.0005), but not from tumors where no HPV is detected (p=0.7924). **(B)** Overall survival of the different subgroups of oropharyngeal tumors. Survival for the silent HPV tumors is significantly worse than for transcriptionally active tumors (p=0.0016), but not from tumors where no HPV is detected (p=0.7301). **(C)** Locoregional control of the different subgroups of oropharyngeal tumors. Survival for silent HPV tumors is not significantly different from transcriptionally active tumors (p=0.5729), nor from tumors where no HPV is detected (p=0.6001). **(D)** Progression free survival of the different subgroups of hypopharyngeal tumors. Survival of silent HPV tumors is not significantly different from tumors where no HPV is detected (p=0.4711). **(E)** Overall survival of the different subgroups of hypopharyngeal tumors. Survival of silent HPV tumors is not significantly different from tumors where no HPV is detected (p=0.8500). **(F)** Locoregional control of the different subgroups of hypopharyngeal tumors. Survival of silent HPV tumors is significantly different from tumors where no HPV is detected (p=0.0169).

Discussion

In this study we investigated the role of transcriptional activity of HPV in HNSCC in both oropharyngeal and hypopharyngeal tumors. Using sensitive DNA sequencing technologies, we find here that a fraction of oropharyngeal tumors, larger than previously documented, contain HPV-DNA. HPV-DNA was also present in a considerable fraction of hypopharyngeal tumors. RNA-sequencing shows, however, that many are not transcriptionally active. HPV positive results from standard PCR and immunohistochemistry assays largely correspond with transcriptional activity and positive RNA-sequencing results. We further show that this group of tumors with silent HPV-DNA has a poor prognosis that is significantly different from transcriptionally active tumors.

There are several hypotheses for the presence of HPV-DNA reads, which is not transcriptionally active. It could be that the virus has not reached the basal cells, but only the non-dividing suprabasal cells, so that there is no replication of the virus. Another possibility is that a low viral load stimulates the immune system to eliminate the virus, or is a manifestation of a properly functioning immune response. Third, low viral loads reflect latent infections that do not carry a risk of neoplastic transformation because of a lack of expression of viral oncoproteins [51].

Previous studies described the presence of silent HPV-DNA in HNSCC. New technologies helped to reveal that a proportion of oropharyngeal HNSCC tumors are HPV-DNA positive but do not show E6-E7 RNA expression, suggesting that the HPV is not transcriptionally active in a number of OPSCC (14-50%) [52]. Subsequent studies demonstrated that such tumors have biological characteristics resembling tumors without evidence of HPV-DNA [48, 53]. Comparing E6 and E7 expressing HNSCC, as determined by RT-PCR, with other HNSCC, the authors showed that cellular gene expression differs significantly. Overall gene expression appeared similar among the tumors where no HPV was detected and the silent HPV tumors. Both were characterized by a high level of LOH at 3p, 9p and/or 17p, whereas there were minor abnormalities in HPV positive tumors [53]. While the authors were able to confirm the reported favorable outcome in the HPV RNA-positive HNSCC in their study, the low number of DNA-positive but RNA-negative samples made it difficult to assess the impact of HPV transcription for patient outcome.

Interestingly, Wichmann and colleagues found no difference in survival between the transcriptionally inactive tumors and transcriptionally active tumors, nor from the group where no HPV was detected [48]. Moreover, this group looked at global gene expression. No differences between silent and HPV-DNA negative tumors were found. Importantly, Wichmann looked into a heterogenous group of HNSCC tumors, both in subsite and in the way they were treated. When they looked into oropharyngeal tumors only, just 9 out of 84 oropharyngeal tumors were found to contain silent HPV, whereas 47 out of 84 were HPV-DNA negative. Similar results for the difference in global expression between HPV-DNA negative and transcriptionally inactive tumors versus transcriptionally active tumors were found for oropharyngeal tumors compared

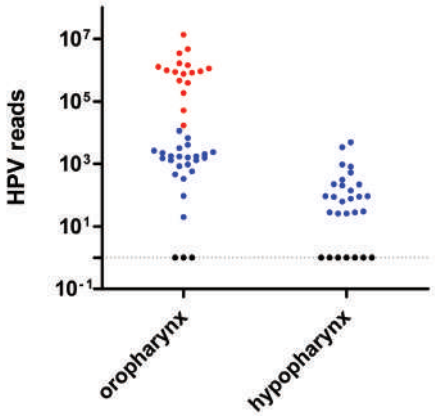
to the heterogenous group of HNSCC. However, Wichmann et al. did not look at survival of those subgroups of oropharyngeal tumors, probably because of the low numbers.

Our study consisted of only 71 patients, of which 43 had oropharyngeal tumors. HPV status by classical methods, $p16^{ink4a}$ /PCR and $p16^{ink4a}$ /p53 immunohistochemistry was compared with HPV status determined by DNA- and RNA-NGS (Figure 3b). RNA sequencing of all oropharyngeal tumors had the same outcome of HPV status as assessed by classical methods. However, there were many more HPV positive tumors found by DNA sequencing than with other methods. A high DNA sequencing read count matched with HPV positivity as determined by all other methods, whereas a low read count matched with an HPV negative status (Supplementary Figure 3). However, there was one exception. This tumor was HPV positive by all methods, but had a low DNA sequencing read count. It was found to be HPV33 positive, for which no baits were used in this study. This patient had a local recurrence after 21 months, which was treated successfully. This tumor was on the cutoff point for a low and high DNA sequencing read count, based on Figure 1c, and had a read count of 16,911. All hypopharyngeal tumors had the same HPV status as determined by the different methods, except for DNA sequencing. Again, all tumors with a low HPV-DNA sequencing read count, were HPV negative as determined by other methods.

We found 40% of the oropharyngeal tumors to have transcriptionally active HPV, comparable to previously reported incidences in the Netherlands [54]. No transcriptionally active tumors were found among the hypopharyngeal tumors. Interestingly we found, next to the high number of transcriptionally active tumors, another 61% (42 out of the 69 DNA and RNA sequenced samples) to contain HPV-DNA, but transcriptionally inactive. This is a much higher number of tumors compared to the Wichmann study [48]. In our study HPV status was determined by next generation targeted sequencing and confirmed by classical methods, consisting of HPV PCR and/or p53 staining on p16 overexpressing tumors. HPV status in the study of Wichmann et al. was determined by INNO-LiPA HPV Genotyping Extra and RNA-HPV by RT-PCR of E6*I transcripts. Notably, we used baits in our DNA sequencing. This targeted sequencing enabled us to perform deep sequencing of the HPV-DNA resulting in the identification of many more silent HPV tumors. Moreover, we looked not only at HPV 16, but also at the other common HPV types, HPV 18 and 33. Therefore our results are a more comprehensive reflection of the presence of HPV-DNA in HNSCC than what has been previously published. Because we found such a high number of silent HPV tumors and equally as important, because the cohort was homogenously treated, we were able to look better into clinical behavior. We can therefore conclude from our study that transcriptionally inactive tumors have a similar behavior as HPV-DNA negative tumors.

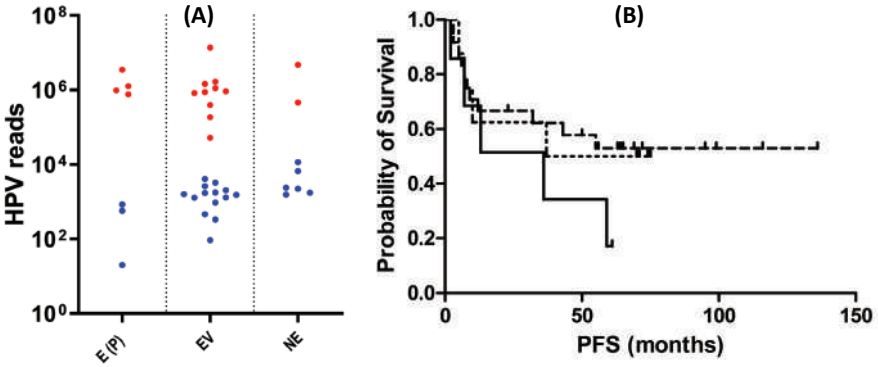
Locoregional control in the subgroups for oropharyngeal tumors is not significantly different, implying that there is no relation with treatment and HPV status in this series. Interestingly we find that locoregional control is better for the silent HPV hypopharyngeal tumors, than in tumors where no HPV was detected; however, the number of tumors are low in the group where no HPV was detected.

Since HPV transcriptionally active tumors have a favorable prognosis and treatment de-intensification trials are being conducted, It is of great importance to know how to deal with silent HPV tumors. Our study implies that detecting DNA only is not sufficient to determine the treatment strategy, since there is a high number of “silent HPV tumors” and treatment de-intensification for this group would be inappropriate. A future study, with a larger cohort, should confirm whether above notions hold ground and whether these silent HPV tumors can truly be grouped together with the HPV-DNA negative patient group.



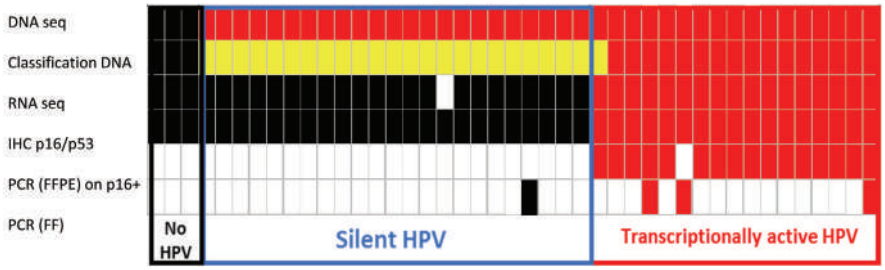
Supplemental Figure 1: Coverage distribution of HPV-DNA reads in HNSCC tumors (n=69)

Number of HPV-DNA reads per tumor sample per subsite. Black dots (no HPV), silent HPV tumors (blue dots) and HPV transcriptionally active tumors (red dots).



Supplemental Figure 2: Variants of HPV16 in oropharyngeal tumors (n=39)

(A) Coverage distribution of the amount of HPV 16 variant DNA reads in oropharyngeal tumors. No HPV containing tumors are shown as black dots, silent HPV as blue dots and HPV transcribed tumors red dots. **(B)** Progression-free survival for patients with a European (E(P); black line), a European variant (EV; broken line) or a Non-European (NE; dotted line) variant of HPV16. No significant difference in survival for these variants was found.



Supplemental Figure 3: Heatmap

Comprehensive summary of the different methods of HPV detection in oropharyngeal tumors. Tumors are ordered based on number of HPV-DNA sequencing reads. HPV positive samples as assessed by the specific method are shown in red, in black HPV negative samples. In yellow are the low number of HPV-DNA sequencing reads shown, as classified based on Figure 1c.

Supplemental Table 1: Patient and tumor characteristics of HNSCC cohort (n=71)

	Oropharynx (n=43)	Hypopharynx (n=28)	Total (n=71)
Variable			
Gender			
Female	13 (30)	7 (25)	20 (28)
Male	30 (70)	21 (75)	51 (72)
Age at diagnosis			
Median (range)	58 (37-75)	61 (48-76)	58 (37-76)
Mean (sd)	58 (8)	61 (7)	59 (8)
Disease stage			
III	4 (9)	8 (29)	12 (17)
IVA	29 (67)	18 (64)	47 (66)
IVB	10 (23)	2 (7)	12 (17)
cT classification			
T1	1 (2)	0 (0)	1 (1)
T2	6 (14)	9 (32)	15 (21)
T3	18 (42)	11 (39)	29 (41)
T4a	14 (33)	6 (21)	20 (28)
T4b	4 (9)	2 (7)	6 (8)
cN classification			
N0	5 (12)	5 (18)	10 (14)
N1	1 (2)	7 (25)	8 (11)
N2	31 (72)	16 (57)	47 (66)
N3	6 (14)	0 (0)	6 (8)
Smoker			
Former smoker	6 (14)	8 (29)	14 (20)
Never	2 (5)	1 (4)	3 (4)
Unknown	4 (9)	2 (7)	6 (8)
Yes	31 (72)	17 (61)	48 (68)
Alcohol consumption			
Former alcoholic	6 (14)	5 (18)	11 (15)
Never	1 (2)	1 (4)	2 (3)
Unknown	5 (12)	2 (7)	7 (10)
Yes	31 (72)	20 (71)	51 (72)
Cisplatin regimen			
daily (6mg/m2, 5 weeks)	5 (12)	9 (32)	14 (20)
3-weekly (100mg/m2, 3 times)	29 (67)	14 (50)	43 (61)
weekly (100-150mg/m2, 4 times)	9 (21)	5 (18)	14 (20)
Cumalitive cisplatin dose			
Low (<300mg/m2)	14 (33)	13 (46)	27 (38)
High (≥300 mg/m2)	29 (67)	15 (54)	44 (62)
Events			
DM	12 (28)	3 (11)	15 (21)
Local recurence	7 (16)	2 (7)	9 (13)
Dead at 5 years	22 (51)	10 (36)	32 (45)

The patient characteristics of 43 oropharyngeal and 28 hypopharyngeal tumors are shown. All patients received concurrent cisplatin-based chemoradiotherapy.

Supplemental Table 2: HPV16 and 18 baits

		Chromosome position	
Gene		**Start**	**End**
HPV16	E6	83	559
	E7	562	858
	E1	865	2813
	L1	5559	7151
HPV18	E6	105	581
	E7	590	907
	E1	914	2887
	L1	5430	7136

Regions of the HPV16 and 18 genome for which baits were designed.

References

[1] van Houten, V.M., et al., Biological evidence that human papillomaviruses are etiologically involved in a subgroup of head and neck squamous cell carcinomas. Int J Cancer, 2001. **93**(2): p. 232-5.

[2] Ragin, C.C. and E. Taioli, Survival of squamous cell carcinoma of the head and neck in relation to human papillomavirus infection: review and meta-analysis. Int J Cancer, 2007. **121**(8): p. 1813-20.

[3] Schlecht, N.F., Prognostic value of human papillomavirus in the survival of head and neck cancer patients: an overview of the evidence. Oncol Rep, 2005. **14**(5): p. 1239-47.

[4] Fakhry, C., et al., Improved survival of patients with human papillomavirus-positive head and neck squamous cell carcinoma in a prospective clinical trial. J Natl Cancer Inst, 2008. **100**(4): p. 261-9.

[5] Fallai, C., et al., Oropharyngeal squamous cell carcinoma treated with radiotherapy or radiochemotherapy: prognostic role of TP53 and HPV status. Int J Radiat Oncol Biol Phys, 2009. **75**(4): p. 1053-9.

[6] Lassen, P., et al., Prognostic impact of HPV-associated p16-expression and smoking status on outcomes following radiotherapy for oropharyngeal cancer: The MARCH-HPV project. Radiother Oncol, 2018. **126**(1): p. 107-115.

[7] Gabani, P., et al., Radiation therapy dose de-escalation compared to standard dose radiation therapy in definitive treatment of HPV-positive oropharyngeal squamous cell carcinoma. Radiother Oncol, 2019. **134**: p. 81-88.

[8] Ang, K.K., et al., Human papillomavirus and survival of patients with oropharyngeal cancer. N Engl J Med, 2010. **363**(1): p. 24-35.

[9] Mirghani, H. and P. Blanchard, Treatment de-escalation for HPV-driven oropharyngeal cancer: Where do we stand? Clin Transl Radiat Oncol, 2018. **8**: p. 4-11.

[10] Bhatia, A. and B. Burtness, Human Papillomavirus-Associated Oropharyngeal Cancer: Defining Risk Groups and Clinical Trials. J Clin Oncol, 2015. **33**(29): p. 3243-50.

[11] Masterson, L., et al., De-escalation treatment protocols for human papillomavirus-associated oropharyngeal squamous cell carcinoma: a systematic review and meta-analysis of current clinical trials. Eur J Cancer, 2014. **50**(15): p. 2636-48.

[12] Chen, A.M., et al., Reduced-dose radiotherapy for human papillomavirus-associated squamous-cell carcinoma of the oropharynx: a single-arm, phase 2 study. Lancet Oncol, 2017. **18**(6): p. 803-811.

[13] Amin, M.B., et al., The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. CA Cancer J Clin, 2017. **67**(2): p. 93-99.

[14] Westra, W.H., Detection of human papillomavirus (HPV) in clinical samples: evolving methods and strategies for the accurate determination of HPV status of head and neck carcinomas. Oral Oncol, 2014. **50**(9): p. 771-9.

[15] de Jong, M.C., et al., HPV and high-risk gene expression profiles predict response to chemoradiotherapy in head and neck cancer, independent of clinical factors. Radiother Oncol, 2010. **95**(3): p. 365-70.

[16] Baumann, M., et al., Radiation oncology in the era of precision medicine. Nat Rev Cancer, 2016. **16**(4): p. 234-49.

[17] Li, H., et al., Association of Human Papillomavirus Status at Head and Neck Carcinoma Subsites With Overall Survival. Jama Otolaryngology-Head & Neck Surgery, 2018. **144**(6): p. 519-525.

[18] Speel, E.J., HPV Integration in Head and Neck Squamous Cell Carcinomas: Cause and Consequence. Recent Results Cancer Res, 2017. **206**: p. 57-72.

[19] Olthof, N.C., et al., Next-generation treatment strategies for human papillomavirus-related head and neck squamous cell carcinoma: where do we go? Rev Med Virol, 2012. **22**(2): p. 88-105.

[20] Lace, M.J., et al., Human papillomavirus type 16 (HPV-16) genomes integrated in head and neck cancers and in HPV-16-immortalized human keratinocyte clones express chimeric virus-cell mRNAs similar to those found in cervical cancers. J Virol, 2011. **85**(4): p. 1645-54.

[21] Olthof, N.C., et al., Viral load, gene expression and mapping of viral integration sites in HPV16-associated HNSCC cell lines. Int J Cancer, 2015. **136**(5): p. E207-18.

[22] Olthof, N.C., et al., Comprehensive analysis of HPV16 integration in OSCC reveals no significant impact of physical status on viral oncogene and virally disrupted human gene expression. PLoS One, 2014. **9**(2): p. e88718.
[23] Brakenhoff, R.H., S. Wagner, and J.P. Klussmann, Molecular Patterns and Biology of HPV-Associated HNSCC. Recent Results Cancer Res, 2017. **206**: p. 37-56.
[24] 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, 2006. **25**(17): p. 2558-64.
[25] Cancer Genome Atlas, N., Comprehensive genomic characterization of head and neck squamous cell carcinomas. Nature, 2015. **517**(7536): p. 576-82.
[26] Leemans, C.R., P.J.F. Snijders, and R.H. Brakenhoff, The molecular landscape of head and neck cancer. Nat Rev Cancer, 2018. **18**(5): p. 269-282.
[27] 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, 2000. **92**(9): p. 709-20.
[28] Broglie, M.A., et al., Evaluation of type-specific antibodies to high risk-human papillomavirus (HPV) proteins in patients with oropharyngeal cancer. Oral Oncol, 2017. **70**: p. 43-50.
[29] Kreimer, A.R., et al., Human papillomavirus types in head and neck squamous cell carcinomas worldwide: a systematic review. Cancer Epidemiol Biomarkers Prev, 2005. **14**(2): p. 467-75.
[30] Nichols, A.C., et al., Does HPV type affect outcome in oropharyngeal cancer? J Otolaryngol Head Neck Surg, 2013. **42**: p. 9.
[31] Hobbs, C.G., et al., Human papillomavirus and head and neck cancer: a systematic review and meta-analysis. Clin Otolaryngol, 2006. **31**(4): p. 259-66.
[32] Randen-Brady, R., et al., In situ hybridization for high-risk HPV E6/E7 mRNA is a superior method for detecting transcriptionally active HPV in oropharyngeal cancer. Hum Pathol, 2019. **90**: p. 97-105.
[33] 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, 2013. **24**(11): p. 2740-5.
[34] Rietbergen, M.M., et al., Molecular characterization of p16-immunopositive but HPV DNA-negative oropharyngeal carcinomas. Int J Cancer, 2014. **134**(10): p. 2366-72.
[35] Begum, S., et al., Detection of human papillomavirus in cervical lymph nodes: a highly effective strategy for localizing site of tumor origin. Clin Cancer Res, 2003. **9**(17): p. 6469-75.
[36] Jordan, R.C., et al., Validation of methods for oropharyngeal cancer HPV status determination in US cooperative group trials. Am J Surg Pathol, 2012. **36**(7): p. 945-54.
[37] van Monsjou, H.S., et al., Human papillomavirus status in young patients with head and neck squamous cell carcinoma. Int J Cancer, 2012. **130**(8): p. 1806-12.
[38] 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, 2007. **121**(11): p. 2465-72.
[39] Lechner, M., et al., Targeted next-generation sequencing of head and neck squamous cell carcinoma identifies novel genetic alterations in HPV+ and HPV- tumors. Genome Med, 2013. **5**(5): p. 49.
[40] Chung, C.H., et al., Genomic alterations in head and neck squamous cell carcinoma determined by cancer gene-targeted sequencing. Ann Oncol, 2015. **26**(6): p. 1216-23.
[41] Balm, A.J., et al., High-dose superselective intra-arterial cisplatin and concomitant radiation (RADPLAT) for advanced head and neck cancer. Head Neck, 2004. **26**(6): p. 485-93.
[42] Essers, P.B.M., et al., Drug Sensitivity Prediction Models Reveal a Link between DNA Repair Defects and Poor Prognosis in HNSCC. Cancer Res, 2019. **79**(21): p. 5597-5611.
[43] Verhagen, C.V.M., et al., Fanconi anemia and homologous recombination gene variants are associated with functional DNA repair defects in vitro and poor outcome in patients with advanced head and neck squamous cell carcinoma. Oncotarget, 2018. **9**(26): p. 18198-18213.
[44] van der Heijden, M., et al., Biological Determinants of Chemo-Radiotherapy Response in HPV-Negative Head and Neck Cancer: A Multicentric External Validation. Front Oncol, 2019. **9**: p. 1470.

[45] Meulendijks, D., et al., HPV-negative squamous cell carcinoma of the anal canal is unresponsive to standard treatment and frequently carries disruptive mutations in TP53. Br J Cancer, 2015. **112**(8): p. 1358-66.
[46] Kluin, R.J.C., et al., XenofilteR: computational deconvolution of mouse and human reads in tumor xenograft sequence data. BMC Bioinformatics, 2018. **19**(1): p. 366.
[47] 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, 2015. **137**(12): p. 2846-57.
[48] Coordes, A., et al., Meta-analysis of survival in patients with HNSCC discriminates risk depending on combined HPV and p16 status. Eur Arch Otorhinolaryngol, 2016. **273**(8): p. 2157-69.
[49] Yamada, T., et al., Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, L2, and L1 coding segments. J Virol, 1995. **69**(12): p. 7743-53.
[50] Zuna, R.E., et al., HPV16 variant lineage, clinical stage, and survival in women with invasive cervical cancer. Infect Agent Cancer, 2011. **6**: p. 19.
[51] Snijders, P.J., A.J. van den Brule, and C.J. Meijer, The clinical relevance of human papillomavirus testing: relationship between analytical and clinical sensitivity. J Pathol, 2003. **201**(1): p. 1-6.
[52] Mirghani, H., et al., Human papilloma virus testing in oropharyngeal squamous cell carcinoma: what the clinician should know. Oral Oncol, 2014. **50**(1): p. 1-9.
[53] Braakhuis, B.J., et al., Genetic patterns in head and neck cancers that contain or lack transcriptionally active human papillomavirus. J Natl Cancer Inst, 2004. **96**(13): p. 998-1006.
[54] Henneman, R., et al., Incidence Changes of Human Papillomavirus in Oropharyngeal Squamous Cell Carcinoma and Effects on Survival in the Netherlands Cancer Institute, 1980-2009. Anticancer Res, 2015. **35**(7): p. 4015-22.

CHAPTER 6

Discussion

Massive sequencing of some cancer types (e.g. non-small-cell-lung cancer and melanoma) has revealed frequent, and often hotspot mutations in oncogenes that are therapeutically targetable. This led to dramatic changes in the treatment of these patients. Whole exome sequencing for loads of other tumors, including head and neck squamous cell carcinoma (HNSCC), has not proven to be as fruitful, given that the prevalence of targetable mutations in oncogenes is modest [1]. Most of precision medicine has focused on targeting somatic mutations affecting oncogenes. An alternative approach to find potential targets for precision medicine is to match specific chemotherapy and targeted therapy agents, to the deficiencies cancer cells harbor in terms of their mechanisms of DNA repair. Those cancer cells with defects in their DNA repair can be targeted by the concept of synthetic lethality. The concept is that two pathways must be inactivated to be lethal to a cell, whereas inactivation of one is not (see Figure 3, Chapter 1). This occurs when one pathway is a backup to the other, and accounts for tumors with defects in the DNA repair pathway of homologous recombination (HR). Drug inhibition of one pathway combined with an inactivating mutation of the other leads to tumor specific kill [2].

The research presented in this thesis describes studies looking into DNA repair deficits in HNSCC and how to exploit them. In this thesis we tried to find relevant variants and potentially (new) biomarkers in HR in vitro and in vivo, by functional outcome measurements and next generation sequencing. We next tried to exploit these defects. This research aims to improve the treatment of individual patients with HNSCC.

Discovery of (new) biomarkers in HNSCC

Driver genes in HNSCC: HR-related genes

Since the introduction of next generation sequencing major projects have been launched to create a comprehensive overview of all genes responsible for the initiation and progress of cancer. The Cancer Genome Atlas project (TCGA), which generated data across 33 cancer types, is an example of this. In the TCGA, the genomic data of 279 HNSCCs is included. Multiple advanced algorithms have been developed in order to identify important driver genes [3]. The general principle behind many of these methods is to select genes that are significantly more mutated than expected by chance, after correcting for factors such as gene length and mutational background processes [3]. The number of significantly mutated genes in HNSCC ranges between 10-200 [4-6]. Although these genes are more frequently mutated than expected by chance, most are infrequently mutated in absolute terms. Only a minority (see Table 1, Chapter 1) of all significantly mutated genes are established driver genes, mostly the ones which are mutated in a substantial proportion. The majority of the remaining should be considered putative driver genes, awaiting functional validation for involvement in carcinogenesis.

In a data-driven study there is a low *a priori* probability that a mutated gene has a strong predictive effect on treatment outcome. Most sequencing studies, however, try to correlate with outcome or treatment response and therefore the impact of the putative driver genes often remains unknown. Moreover, infrequent mutations require large series to prove that the mutation is prognostically important in a subset of patients. There are ways to overcome difficulties in establishing the role and importance of putative driver genes. One is to analyze infrequently mutated genes at the pathway level to show the importance of a specific pathway and another is to use a hypothesis driven approach in which a gene or pathway known to be important from literature data is specifically analyzed [7]. We integrated these approaches in our research as described in chapter 2 and 3. We hypothesized the presence of FA/HR deficiency in sporadic HNSCC. Firstly, because of the 500 to 700-fold increased risk on HNSCC in FA patients [8]. This suggests that inactivation of the FA pathway drives carcinogenesis, also in HNSCC. HNSCC cell lines derived from non-FA patients were indistinguishable from those derived from FA patients in terms of whole-arm translocations and numerous gains and losses [9, 10]. We therefore questioned whether somatic inactivation of FA genes is present in sporadic HNSCC, similar to inactivation of BRCA1 and BRCA2 in sporadic breast and ovarian cancer. Secondly FA cells are extremely sensitive to crosslinking agents and the addition of cisplatin to radiotherapy is associated with an approximately 8% decrease in cancer death of HNSCC in non-FA patients. This benefit suggests a subgroup of patients particularly sensitive to this crosslinking agent, possibly due to HR deficiency. The presence of FA/HR pathway aberrations and mutated genes in HNSCC has been shown in functional studies, but there was no link with sequencing [11, 12]. Although only a few sequence variants in FA/HR have been found in whole exome sequencing of large groups of head and neck tumors [5, 13, 14], this amounts to a substantial proportion of mutated tumors at the pathway level. We therefore tested FA/HR-pathway deficiency with multiple assessments in a panel of 29 HNSCC cell lines. We found a wide range of HR/FA pathway deficiencies. Next, we sequenced the DNA of the FA/HR genes from these cell lines in order to find DNA variants associated with the observed defects. Consequently, by using this hypothesis driven approach, we were able to detect an association of FA/HR pathway variants with functional outcome. We found that bi-allelic variants in FA/HR genes with an allele frequency below 2.5% were associated with MMC sensitivity in vitro. A homogenous group of patient tumor samples treated with a combination of radiotherapy and a crosslinking agent revealed an association with worse patient survival, but favorable response to cisplatin.

Human Papillomavirus (HPV)

HPV is an established biomarker. In the Radiation Therapy Oncology Group (RTOG) 0129 trial it was shown that patients with HPV-positive oropharyngeal squamous cell carcinomas (OPSCC) had a better clinical outcome compared with HPV-negative tumors when treated with radiation or a combination of radiotherapy and cisplatin [15]. This outcome led to a reevaluation of staging the HPV status in OPSCC and an adaptation

by the American Joint Committee on Cancer in the most recent TNM staging [16, 17]. Many patients previously categorized as TNM stage III or IV by the seventh edition TNM stage, are TNM stage I in the eighth edition staging, which reflects the better prognosis of HPV-positive tumors. The notably improved clinical outcome for patients with HPV-positive tumors, but high treatment related morbidity has led to an interest in the development of de-escalated therapeutic strategies with the aim of reducing short and long-term toxicity and maintaining or improving oncologic outcome [18].

From the above it is clear that reliable detection of HPV is essential. There is no standard approach for detecting HPV. Detection strategies differ both in design, as in their detection targets. These targets include HPV DNA, HPV RNA, viral oncoproteins, cellular proteins (e.g. p16 protein) and HPV-specific serum antibodies. In general, detection methods are based on HPV transcriptionally active tumors [19]. However, the role of tumors with transcriptionally inactive (silent) HPV remained unclear, mainly because of low numbers of such tumors in previous studies [20]. Their role was therefore studied in this thesis. Different methods for determining HPV status, including DNA and RNA next generation sequencing, immunohistochemical staining of $p16^{ink4a}$ and p53 and classic HPV specific PCR on p16 overexpressing samples were performed in a homogenously treated cohort of oropharyngeal and hypopharyngeal tumors. We found a considerable proportion of silent HPV in both oropharyngeal and hypopharyngeal tumors by DNA and RNA next generation sequencing of 52% and 74%, respectively. This is a much higher proportion than previously reported [21], probably as a result of the very sensitive targeted sequencing we performed, which included HPV baits. The silent HPV tumors found, turned out to be HPV-negative when assessed by other methods and their clinical behavior was as such too; they are associated with a poor prognosis. Our study indicates that transcriptionally inactive tumors are significantly different from transcriptionally active OPSCC and are frequently present. This is in line with the biological resemblance of transcriptionally inactive HPV tumors and HPV-negative tumors [20, 22]. According to our study, patients with silent HPV OPSCC should therefore be regarded and treated as HPV-negative. This study indicates that next generation sequencing of only DNA is not specific enough to determine HPV status.

Targeted therapy: exploiting driver genes

Large numbers of tumors have been sequenced and their data integrated and analyzed, so it is likely that the majority of (putative) driver genes have been discovered by now [23]. On the one hand research aims to establish the functional impact of these putative driver genes as we showed in chapter 2 and 3. On the other hand research moves from discovering potential driver genes to converting them into therapeutic targets or using them as biomarkers. Frequently mutated driver genes are interesting candidates for targeted therapy, an example for this are the three oncogenes EGFR, CCND1 and PIK3CA as mentioned in the Introduction. Targeted therapy by cetuximab against EGFR is used routinely in clinic. Interestingly the majority of frequently mutated

driver genes in HNSCC are tumor suppressor genes. This means that the majority of driver genes cannot be an option for targeted inhibitors.

However, loss of a tumor suppressor gene might be exploited via the synthetic lethality principle. The principle of synthetic lethality has been described in the Introduction. HR deficiency (HRD) is a prime example of an alteration that can be targeted with synthetic lethality. The FA/HR pathway is involved in the faithful repair if lesions that cause stalled replication forks and double strand breaks. When cells are unable to repair these lesion, accumulation of DNA damage occurs leading to genomic instability. Genomic instability can contribute to carcinogenesis but can also be exploited by drugs that require HR for repair. Two of these classes are PARPi and platinum-based drugs, such as cisplatin. We exploited the role of PARPi in HNSCC with FA/HR repair deficiencies as found in chapter 2 and describe this in chapter 4. The enzyme PARP1 binds at the site of single strand breaks of DNA and initiates DNA repair. PARP inhibitors prevent the release of PARP1 and therefore single strand breaks will not be repaired. The replication fork will stall and single strand breaks will be converted into double strand breaks upon which the cell will rely on faithful repair by the FA/HR pathway of the less reliable NHEJ and in the end cell death. In conclusion, HR-deficiency can be used as a target via the synthetic lethality principle.

Future perspectives

The research in this thesis helps making a step forward in precision medicine. We showed that bi-allelic variants in FA/HR genes are associated with genomic features of FA deficiency, both in vitro and in vivo. This provides novel opportunities for the delivery of precision medicine, as we showed by the exploitation of these defects with PARPi.

Our study of HRD started in the mid 2000s. Since then, HRD has been extensively studied. Traditionally, well known HR-mutated genes are BRCA1 and BRCA2, were mainly described in ovarian and breast cancer. It seems now that bi-allelic inactivating events occur across all tumor types, providing novel opportunities for the delivery of precision medicine approaches and expanding the population of patients who may benefit from targeting HR DNA repair defects [1]. The insights from these studies can be used to shed light on the important remaining questions after our study: are bi-allelic variants in FA/HR genes appropriate markers of HRD in HNSCC? Are there other markers for HRD? Are there other new treatment options?

Bi-allelic mutations in HR/FA genes

Germline BRCA1/2 mutation carriers typically lose the wild-type allele through loss of heterozygosity (LOH). The resulting genotype is known under various terms, including locus specific LOH and homozygous or bi-allelic mutation, alteration or inactivation. Bi-allelic FA/HR gene mutations are also found in tumors of patients that do not carry

germline HR/FA mutations, completely resulting from somatic events. Given the requirement of loss of tumor suppressor gene function for carcinogenesis, bi-allelic mutations are intuitive HRD markers. Experimental and clinical studies of (mainly) breast and ovarian tumors have shown that those with bi-allelic alterations in the FA/HR pathway are more sensitive to platinum-based drugs and PARP inhibitors and have a more favorable treatment response than those with mono-allelic or no alterations [24-27]. In breast and ovarian cancer, the pathogenicity of many BRCA1/2 variants is known. Therefore, a BRCA1/2 assay is already commercially available (BRCAnalysis, Myriad Genetics). In contrast, HNSCC has many variants of unknown significance (chapter 2 and 3). Moreover, when there are multiple alteration types, the cost of the analyses will rise. In vitro experiments can be used in the future to elucidate the significance of the variants, but these are labor-intensive and costly. Another option is to use HR-deficiency signatures to reclassify variants of unknown significance or simply rely on HR-deficiency signatures [28].

HRD signatures

HR deficiency results in genomic instability which results in a characteristic imprint of the genome. Unlike mutations in most oncogenes and tumor suppressor genes, mutations in DNA-repair related genes therefore provide a unique opportunity to assess phenotypic consequences by examining structural and mutational signatures obtained from the sequencing data. This does not only account for HR deficiency, but also for e.g. mismatch repair or nucleotide excision repair deficiency. There is increasing evidence for the presence of a characteristic genomic signature and that this can be targeted regardless of their exact mutation site [29].

Different signatures have been developed, among them are those based on copy number, expression or mutational sites. Of all the HRD copy number signatures which are developed, only one is commercially available: the HRDscore [30]. The HRDscore, which is derived from breast and ovarian cancer SNP array data, originally combines three independently developed HRD measures: LOH, telomeric imbalance and large-scale transitions. It is marketed as a predictive biomarker for platinum-based drugs and PARP inhibitors. Each of the three compounds of the HRDscore turns out to be increased in HNSCC that responds favorably to platinum based chemoradiotherapy [31]. Moreover, the HRDscore correlates well with other markers of HR deficiency and should therefore be tested and approved for use in HNSCC. Since copy number signatures represent the HR deficiency status in a tumor's genome, one could argue this is not necessarily the current HR deficiency status, since the transcriptome is more dynamic than the genome. Therefore, HR deficiency gene expression signatures are promising alternative HR deficiency markers. Since the commercialization of micro arrays various HR deficiency gene expression signatures have been developed [32-35]. We recently developed a HNSCC specific HR deficiency signature [36]. As the transcriptome is more dynamic than the genome, it reflects the current HRD status of the tumor. Tumors that are HR deficient based on their genome, could regain proficiency through secondary mutations. Signatures based on the transcriptome

would classify these tumors correctly as HR proficient. Another advantage are the low costs of RNA sequencing. The main disadvantage of these gene expression signatures however includes the inconsistency of identified gene sets. Furthermore, we do not know if the loss of different genes within the DNA repair pathways produces distinct gene expression patterns. Therefore, mutational signatures can be useful. Mutational signatures capture patterns of (small) mutations. They can be derived from any class of mutations, but base mutations are regularly used, because they are the most abundant in exome sequencing data. Signature 3 is a well-known HRD mutational signature, based on pan cancer exome sequencing data and associated with bi-allelic BRCA1/2 variants [37]. Another is HRDetect which combines several signatures and therefore has a higher sensitivity picking up variants [28]. In vitro studies have been conducted to establish a causal link between HRD and mutational signatures and show that refinement is possible. These signatures have not been studied in the clinic yet so further research is warranted.

The research discussed above shows that much progress has been made towards the detection of HR deficiency in cancer. However, most of this research pertains to breast and ovarian cancer, not HNSCC. For HNSCC the lessons learned from breast and ovarian cancer should be used to its advantage. We showed that bi-allelic genomic alterations cause HR deficiency in a limited number of HNSCC. A complete overview of causative alterations seems elusive and therefore a signature based on a phenotype caused by HR deficiency seems the way forward.

New treatment options

Therapeutic options consist, apart from surgery, of a combination of various agents, possibly in combination with various radiotherapy doses and fractionation schedules. Apart from agents targeting DNA repair deficiency, we have seen the rise of agents targeting aberrant pathways in HNSCC, such as EGFR, CCND and PIK3CA inhibitors. Moreover, much is to be expected from immune checkpoint inhibitors as mentioned in the introduction. Having all these options is a blessing and a pain, since we should be able to assess when to use which treatment and so improve survival and/or decrease toxicity. With the ability to asses HR deficiency in a tumor, we would be able to specifically treat patient and their tumors. On one hand by testing DNA repair targeting agents and on the other on the development of new agents. This thesis therefore is a step forward in precision medicine.

References

[1] Riaz, N., et al., Pan-cancer analysis of bi-allelic alterations in homologous recombination DNA repair genes. Nat Commun, 2017. **8**(1): p. 857.

[2] Begg, A.C., Predicting recurrence after radiotherapy in head and neck cancer. Semin Radiat Oncol, 2012. **22**(2): p. 108-18.

[3] Lawrence, M.S., et al., Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature, 2013. **499**(7457): p. 214-218.

[4] Leemans, C.R., P.J.F. Snijders, and R.H. Brakenhoff, The molecular landscape of head and neck cancer. Nat Rev Cancer, 2018. **18**(5): p. 269-282.

[5] Cancer Genome Atlas, N., Comprehensive genomic characterization of head and neck squamous cell carcinomas. Nature, 2015. **517**(7536): p. 576-82.

[6] Iorio, F., et al., A Landscape of Pharmacogenomic Interactions in Cancer. Cell, 2016. **166**(3): p. 740-754.

[7] Ignatiadis, N., et al., Data-driven hypothesis weighting increases detection power in genome-scale multiple testing. Nat Methods, 2016. **13**(7): p. 577-80.

[8] Kutler, D.I., et al., High incidence of head and neck squamous cell carcinoma in patients with Fanconi anemia. Arch Otolaryngol Head Neck Surg, 2003. **129**(1): p. 106-12.

[9] van Zeeburg, H.J., et al., Generation and molecular characterization of head and neck squamous cell lines of fanconi anemia patients. Cancer Res, 2005. **65**(4): p. 1271-6.

[10] Hermsen, M.A., et al., Cytogenetic characteristics of oral squamous cell carcinomas in Fanconi anemia. Fam Cancer, 2001. **1**(1): p. 39-43.

[11] Snyder, E.R., et al., Variation in cisplatinum sensitivity is not associated with Fanconi Anemia/BRCA pathway inactivation in head and neck squamous cell carcinoma cell lines. Cancer Lett, 2007. **245**(1-2): p. 75-80.

[12] Burkitt, K. and M. Ljungman, Compromised Fanconi anemia response due to BRCA1 deficiency in cisplatin-sensitive head and neck cancer cell lines. Cancer Lett, 2007. **253**(1): p. 131-7.

[13] Stransky, N., et al., The mutational landscape of head and neck squamous cell carcinoma. Science, 2011. **333**(6046): p. 1157-60.

[14] Agrawal, N., et al., Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1. Science, 2011. **333**(6046): p. 1154-7.

[15] Nguyen-Tan, P.F., et al., Randomized phase III trial to test accelerated versus standard fractionation in combination with concurrent cisplatin for head and neck carcinomas in the Radiation Therapy Oncology Group 0129 trial: long-term report of efficacy and toxicity. J Clin Oncol, 2014. **32**(34): p. 3858-66.

[16] 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, 2016. **17**(4): p. 440-451.

[17] Amin, M.B., et al., The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. CA Cancer J Clin, 2017. **67**(2): p. 93-99.

[18] Rosenberg, A.J. and E.E. Vokes, Optimizing Treatment De-Escalation in Head and Neck Cancer: Current and Future Perspectives. Oncologist, 2020.

[19] Bishop, J.A., et al., Detection of transcriptionally active high-risk HPV in patients with head and neck squamous cell carcinoma as visualized by a novel E6/E7 mRNA in situ hybridization method. Am J Surg Pathol, 2012. **36**(12): p. 1874-82.

[20] 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, 2015. **137**(12): p. 2846-57.

[21] Mirghani, H., et al., Human papilloma virus testing in oropharyngeal squamous cell carcinoma: what the clinician should know. Oral Oncol, 2014. **50**(1): p. 1-9.

[22] Braakhuis, B.J., et al., Genetic patterns in head and neck cancers that contain or lack transcriptionally active human papillomavirus. J Natl Cancer Inst, 2004. **96**(13): p. 998-1006.

[23] Bailey, M.H., et al., Comprehensive Characterization of Cancer Driver Genes and Mutations. Cell, 2018. **174**(4): p. 1034-1035.
[24] Mutter, R.W., et al., Bi-allelic alterations in DNA repair genes underpin homologous recombination DNA repair defects in breast cancer. J Pathol, 2017. **242**(2): p. 165-177.
[25] Sakai, W., et al., Functional restoration of BRCA2 protein by secondary BRCA2 mutations in BRCA2-mutated ovarian carcinoma. Cancer Res, 2009. **69**(16): p. 6381-6.
[26] Sakai, W., et al., Secondary mutations as a mechanism of cisplatin resistance in BRCA2-mutated cancers. Nature, 2008. **451**(7182): p. 1116-20.
[27] Drew, Y., et al., Therapeutic potential of poly(ADP-ribose) polymerase inhibitor AG014699 in human cancers with mutated or methylated BRCA1 or BRCA2. J Natl Cancer Inst, 2011. **103**(4): p. 334-46.
[28] Davies, H., et al., HRDetect is a predictor of BRCA1 and BRCA2 deficiency based on mutational signatures. Nat Med, 2017. **23**(4): p. 517-525.
[29] Le, D.T., et al., Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. Science, 2017. **357**(6349): p. 409-413.
[30] Telli, M.L., et al., Homologous Recombination Deficiency (HRD) Score Predicts Response to Platinum-Containing Neoadjuvant Chemotherapy in Patients with Triple-Negative Breast Cancer. Clin Cancer Res, 2016. **22**(15): p. 3764-73.
[31] Vossen, D.M., et al., Comparative genomic analysis of oral versus laryngeal and pharyngeal cancer. Oral Oncol, 2018. **81**: p. 35-44.
[32] Jazaeri, A.A., et al., Gene expression profiles of BRCA1-linked, BRCA2-linked, and sporadic ovarian cancers. J Natl Cancer Inst, 2002. **94**(13): p. 990-1000.
[33] Larsen, M.J., et al., Classifications within molecular subtypes enables identification of BRCA1/BRCA2 mutation carriers by RNA tumor profiling. PLoS One, 2013. **8**(5): p. e64268.
[34] Wang, Y., et al., Computational Investigation of Homologous Recombination DNA Repair Deficiency in Sporadic Breast Cancer. Sci Rep, 2017. **7**(1): p. 15742.
[35] Pitroda, S.P., et al., DNA repair pathway gene expression score correlates with repair proficiency and tumor sensitivity to chemotherapy. Sci Transl Med, 2014. **6**(229): p. 229ra42.
[36] Essers, P.B.M., et al., Drug sensitivity prediction models reveal a link between DNA repair defects and poor prognosis in HNSCC. Cancer Res, 2019.
[37] Alexandrov, L.B., et al., Signatures of mutational processes in human cancer. Nature, 2013. **500**(7463): p. 415-21.

CHAPTER 7

Summary

Head and neck squamous cell carcinoma (HNSCC) is an uncommon solid tumor. Patients often present at an advanced stage. The standard treatment for this patient group consists of a combination of radiotherapy and platinum-based chemotherapy. Unfortunately, not all patients benefit from this treatment and many patients suffer from severe acute and long term side effects. Knowing upfront who will benefit, or for whom it is possible to de-intensify treatment, would diminish these problems. The aim of this thesis was therefore to improve personalized treatment.

This thesis focusses on defects in the DNA repair pathway of HNSCC tumors. We chose to look into these particular defects, since it is known that aberrations in genes causing DNA repair defects, have impact on treatment response and patient outcome, for example for BRCA deficient tumors in breast cancer patients. Moreover, defects in the germline of DNA repair pathway genes, such as the Fanconi Anemia (FA) pathway, can predispose patients to cancer and cause hypersensitivity for platinum drugs in this group of patients.

Nowadays large genome studies are performed. These studies generate an enormous amount of novel alterations. A major challenge is to discriminate the alterations, or variants, with a clinical impact from the vast number of alterations. The way variant analysis is conducted, is therefore crucial for the nature and frequency of the variants that are found. It is of great importance to find functional alterations. We therefore first searched for functional defects in the DNA repair pathway by testing a large panel of HNSCC cell lines by multiple assays. We found evidence for DNA repair deficiency in a considerable proportion. Subsequently, we searched for the underlying DNA repair defect in these cell lines, by targeted DNA sequencing of a cohort of genes, including the FA and homologous recombination (HR) genes, whose products are well known players in DNA repair. We found that variant selection based on homozygosity and on SNP allele frequency are selection criteria to identify functional FA/HR repair defects.

The presence of these defects was confirmed in a separate patient cohort of HNSCC tumors all treated with chemoradiation. The incidence of these variants was 19%. Tumors with these variants were associated with a poor prognosis, but also with a better response to a high cumulative dose of the crosslinking chemotherapeutical agent cisplatin.

Another possibility in personalized medicine besides de-intensification, is the exploitation of these DNA repair defects by PARP inhibitors (PARPi) or other DNA repair inhibitors. PARPi is especially effective as a single agent in HR-deficient cells, which is explained by the concept of synthetic lethality. However, much lower doses of the PARPi olaparib are sufficient to induce radiosensitization in HR proficient cells and even lower PARPi doses in HR deficient cells. So olaparib could be considered in the development of new treatment strategies in combination with radiotherapy.

In addition to detection and exploration of DNA repair defects, we investigated the role of transcriptionally inactive HPV in HNSCC, since HPV is known to be an important biomarker. HPV-positive tumors are known to have a favorable prognosis over HPV-negative tumors. Most methods to detect HPV are based on transcriptionally active

HPV tumors. We find that transcriptionally inactive tumors should be regarded as HPV-negative tumors, since the survival of patients with these tumors is comparable and they show a biological resemblance with HPV-negative tumors.

This thesis concludes with a general discussion, on how our findings of DNA repair deficiency in HNSCC, both in vitro and in vivo, might lead to precision medicine in HNSCC by exploitation of these defects.

CHAPTER 8

Nederlandse samenvatting

Het plaveiselcelcarcinoom in het hoofd-halsgebied (HHPCC) is een niet veel voorkomende solide vorm van kanker. Patiënten presenteren zich vaak in een vergevorderd stadium van de ziekte. De standaardbehandeling voor deze patiëntengroep bestaat uit een combinatie van bestraling en op platinum gebaseerde chemotherapie. Helaas hebben niet alle patiënten baat bij deze behandeling en veel van hen hebben last van ernstige acute en late bijwerkingen. Als we van te voren zouden weten wie baat heeft bij de behandeling of bij wie het mogelijk is de behandeling minder intensief te maken, zouden deze problemen minder vaak voorkomen. Het doel van dit proefschrift was daarom om de gepersonaliseerde behandeling te verbeteren.

Het onderzoek beschreven in dit proefschrift richt zich op defecten in de zogenaamde DNA reparatie paden van HHPCC. We hebben ervoor gekozen juist deze defecten te onderzoeken, omdat bekend is dat afwijkingen in genen die DNA reparatie defecten tot gevolg hebben van belang zijn voor het effect van de behandeling en daarmee de prognose voor de patiënt mede bepalen, zoals bekend is bij BRCA deficiënte tumoren bij patiënten met borstkanker. Bovendien kunnen afwijkingen in de kiemcellen van DNA reparatie pad genen, zoals genen van het Fanconi Anemie (FA) pad, de kans op kanker verhogen, maar ook geassocieerd zijn met een betere respons op platinum bevattende medicatie bij deze patiënten.

Tegenwoordig worden grootschalige en gedetailleerde analyses van het genoom uitgevoerd. Deze studies laten een enorme hoeveelheid nieuwe afwijkingen zien. Het is een grote uitdaging om klinisch relevante afwijkingen, of varianten, te detecteren binnen het enorme aantal afwijkingen. De manier waarop de analyse van varianten wordt uitgevoerd is bepalend voor de aard en frequentie van de gevonden varianten. Het is van groot belang om functionele afwijkingen te vinden. We zochten daarom eerst naar functionele defecten in DNA reparatie door meerdere analyses uit te voeren in een grote groep HHPCC cellijnen. In een aanzienlijk aantal daarvan vonden we een stoornis in DNA reparatie. Vervolgens zochten we naar het onderliggende DNA reparatie defect in deze cellijnen door gerichte sequencing van een reeks genen, waaronder de FA en homologe recombinatie (HR) genen; de producten daarvan zijn bekende spelers bij DNA herstel. We toonden aan dat selectie van varianten gebaseerd op homozygotie en SNP allel frequentie, van groot belang is om functionele FA/HR reparatie defecten te identificeren.

De aanwezigheid van deze defecten werd bevestigd in een separaat cohort van patiënten met HHPCC tumoren, die allen behandeld werden met chemoradiatie. De incidentie van deze varianten was 19%. Tumoren met deze varianten hadden enerzijds een slechtere prognose, maar anderzijds een betere respons op een hoge cumulatieve dosis van het cross-linking chemotherapeuticum cisplatine, dan de groep zonder deze varianten.

Naast een minder intensieve behandeling, is het beïnvloeden van deze DNA reparatie defecten door poly (ADP-ribose) polymerase (PARP)-remmers of andere DNA schade reparatie remmers een mogelijkheid voor gepersonaliseerde behandeling. PARP-remmers zijn met name effectief als monotherapie in HR-deficiënte cellen, wat

verklaard wordt middels het concept van “synthetic lethality”. Opvallend is, dat een veel lagere dosis van de PARP-remmer olaparib voldoende is om radiosensibilisatie te induceren in cellen waar HR intact is, en zelfs een nog lagere dosis PARP-remmers in HR deficiënte cellen. Olaparib kan dus overwogen worden in de ontwikkeling van nieuwe behandelingsstrategieën in combinatie met bestraling.

Naast het detecteren en exploreren van DNA reparatie defecten, onderzochten we de rol van transcriptioneel inactief humaan papillomavirus (HPV) in HHPCC, aangezien HPV bekend staat als een belangrijke biomarker. HPV-positieve tumoren hebben een betere prognose dan HPV-negatieve tumoren. De meeste methodes om HPV aan te tonen zijn gebaseerd op transcriptioneel actieve HPV tumoren. Wij tonen aan dat transcriptioneel inactieve tumoren beschouwd moeten worden als HPV-negatieve tumoren, omdat enerzijds de overleving van patiënten met deze tumoren vergelijkbaar is met die van patiënten met HPV-negatieve tumoren en ze anderzijds biologische gelijkenissen vertonen.

Dit proefschrift eindigt met een algemene beschouwing over de vraag hoe onze bevindingen betreffende DNA reparatie deficiëntie in HHPCC, zowel in vitro als in vivo, zouden kunnen leiden tot een betere zorg op maat bij patiënten met HHPCC.

CHAPTER 9

Appendices

AUTHORS AND AFFILIATIONS

AC Annalisa L.E. Carli, Division of Cell Biology, The Netherlands Cancer Institute, Amsterdam, The Netherlands

AV Arno Velds, Genomics Core Facility, The Netherlands Cancer Institute, Amsterdam, The Netherlands

CV Caroline V.M. Verhagen, Division of Cell Biology, The Netherlands Cancer Institute, Amsterdam, The Netherlands; Department of Head and Neck Oncology and Surgery, The Netherlands Cancer Institute, Amsterdam, The Netherlands

CVens Conchita Vens, Division of Cell Biology, The Netherlands Cancer Institute, Amsterdam, The Netherlands; Department of Radiation Oncology, The Netherlands Cancer Institute, Amsterdam, The Netherlands

DV David M. Vossen, Division of Cell Biology, The Netherlands Cancer Institute, Amsterdam, The Netherlands; Department of Head and Neck Oncology and Surgery, The Netherlands Cancer Institute, Amsterdam, The Netherlands

FH Floor Hageman, Division of Cell Biology, The Netherlands Cancer Institute, Amsterdam, The Netherlands

JJ Jos Jonkers, Division of Molecular Pathology, The Netherlands Cancer Institute, Amsterdam, The Netherlands

KB Kerstin Borgmann, Laboratory of Radiobiology and Experimental Radiation Oncology, University Medical Centre Hamburg-Eppendorf, Hamburg, Germany

LM Lisanne Mout, Division of Cell Biology, The Netherlands Cancer Institute, Amsterdam, The Netherlands

LW Lodewijk F.A. Wessels, Molecular Carcinogenesis, The Netherlands Cancer Institute, Amsterdam, The Netherlands; Department of EEMCS, Delft University of Technology, Delft, The Netherlands

MB Michiel W.M. van den Brekel, Department of Head and Neck Oncology and Surgery, The Netherlands Cancer Institute, Amsterdam, The Netherlands; Institute of Phonetic Sciences, University of Amsterdam, Amsterdam, The Netherlands; Department of Oral and Maxillofacial Surgery, Academic Medical Center, University of Amsterdam, The Netherlands

MH Martijn van der Heijden, Division of Cell Biology, The Netherlands Cancer Institute, Amsterdam, The Netherlands; Department of Head and Neck Oncology and Surgery, The Netherlands Cancer Institute, Amsterdam, The Netherlands

MLV Marie-Louise van Velthuysen, Department of Pathology, The Netherlands Cancer Institute, Amsterdam, The Netherlands

MN Marja Nieuwland, Genomics Core Facility, The Netherlands Cancer Institute, Amsterdam, The Netherlands

MO Mark J. O'Connor, Oncology Innovative Medicines, AstraZeneca, Saffron Walden, United Kingdom

MV Marcel Verheij, Division of Cell Biology, The Netherlands Cancer Institute, Amsterdam, The Netherlands; Department of Radiation Oncology, The Netherlands Cancer Institute, Amsterdam, The Netherlands

MVJ Manon Verwijs-Janssen, Division of Cell Biology, The Netherlands Cancer Institute, Amsterdam, The Netherlands

RG Reidar Grénman, Department of Otorhinolaryngology, Turku University Hospital and University of Turku, Turku, Finland; Department of Medical Biochemistry and Genetics, Turku University, Turku, Finland

RH Rosemarie de Haan, Division of Cell Biology, The Netherlands Cancer Institute, Amsterdam, The Netherlands; Department of Radiation Oncology, The Netherlands Cancer Institute, Amsterdam, The Netherlands

RK Ron Kerkhoven, Genomics Core Facility, The Netherlands Cancer Institute, Amsterdam, The Netherlands

RKluin Roel J.C. Kluin, Genomics Core Facility, The Netherlands Cancer Institute, Amsterdam, The Netherlands

SW Stefan M. Willems, Department of Pathology, University Medical Center Utrecht, Utrecht, The Netherlands

TO Tim P.D. Oostendorp, Division of Cell Biology, The Netherlands Cancer Institute, Amsterdam, The Netherlands

TS Tesa M. Severson, Division of Molecular Pathology, The Netherlands Cancer Institute, Amsterdam, The Netherlands

VW Volkert B. Wreesmann, Department of Head and Neck Oncology and Surgery, The Netherlands Cancer Institute, Amsterdam, The Netherlands

AUTHOR CONTRIBUTIONS

Chapter 2

Fanconi anemia and homologous recombination gene variants are associated with functional DNA repair defects in vitro and poor outcome in patients with advanced head and neck squamous cell carcinoma

Caroline V.M. Verhagen, David M. Vossen, Kerstin Borgmann, Floor Hageman, Reidar Grénman, Manon Verwijs-Janssen, Lisanne Mout, Roel J.C. Kluin, Marja Nieuwland, Tesa M. Severson, Arno Velds, Ron Kerkhoven, Mark J. O'Connor, Martijn van der Heijden, Marie-Louise van Velthuysen, Marcel Verheij, Volkert B. Wreesmann, Lodewyk F.A. Wessels, Michiel W.M. van den Brekel, Conchita Vens

Study conceptualization and design:	CVens, MB, VW
Data acquisition and curation:	AV, CV, CVens, DV, FH, KB, LM, MB, MH, MLV, MN, MV, MVJ, RG, RK, Rkluin, TS
Data analysis and interpretation:	CV, CVens, DV, LW, MB
Resources:	AV, CV, CVens, LW, MB, MH, MV, RG, RK, RKluin, TS
Manuscript preparation:	CV, CVens, MB, MV
Manuscript editing and review:	AV, CV, CVens, DV, FH, KB, LM, LW, MB, MH, MLV, MN, MO, MV, MVJ, RG, RK, RKluin, TS, VW

Chapter 3

Role of variant allele fraction and rare SNP filtering to improve cellular DNA repair endpoint association

David M. Vossen, Caroline V.M. Verhagen, Reidar Grénman, Roel J.C. Kluin, Marcel Verheij, Michiel W.M. van den Brekel, Lodewyk F.A. Wessels, Conchita Vens

Study conceptualization and design:	CVens, LW
Data acquisition and curation:	CV, DV, LW
Data analysis and interpretation:	CVens, DV, LW
Resources:	CV, CVens, LW, MB, MV, RG
Manuscript preparation:	CVens, DV, LW
Manuscript editing and review:	CV, CVens, DV, LW, MB, MV, RG, RKluin

Chapter 4

Extent of radiosensitization by the PARP inhibitor olaparib depends on its dose, the radiation dose and the integrity of the homologous recombination pathway of tumor cells

Caroline V.M. Verhagen, Rosemarie de Haan, Floor Hageman, Tim P.D. Oostendorp, Annalisa L.E. Carli, Mark J. O'Connor, Jos Jonkers, Marcel Verheij, Michiel W.M. van den Brekel, Conchita Vens

Study conceptualization and design:	CVens
Data acquisition and curation:	AC, CV, CVens, FH, TO
Data analysis and interpretation:	CV, RH, CVens
Resources:	CV, CVens, JJ, MB, MV, RH
Manuscript preparation:	CV, RH, CVens
Manuscript editing and review:	AC, CV, CVens, FH, JJ, MB, MO, MV, RH, TO

Chapter 5

Transcriptionally inactive HPV is associated with poor prognosis in chemo-radiotherapy treated oropharyngeal squamous cell carcinomas patients

Caroline V.M. Verhagen, Roelof J.C. Kluin, Martijn vd Heijden, David M. Vossen, Marcel Verheij, Stefan M. Willems, Michiel W.M. van den Brekel, Conchita Vens

Study conceptualization and design:	CV, CVens
Data acquisition and curation:	CV, DV, MH, RKluin, SW
Data analysis and interpretation:	CV, CVens
Resources:	CV, CVens, MB, RKluin
Manuscript preparation:	CV, CVens
Manuscript editing and review:	CV, CVens, DV, MB, MH, MV, RKluin, SW

FUNDING

Het hier beschreven onderzoek werd mede mogelijk gemaakt door het Nederlands Kanker Instituut en het Verwelius fonds.

DANKWOORD

"The End Has No End", The Strokes.

Na ruim 12,5 jaar is het zover, ik mag mijn PhD verdedigen. In dit tijdsbestek zijn er vele mensen langsgekomen die op ieder hun eigen manier een bijdrage hebben geleverd.

Dr. Vens, Conchita, natuurlijk sta jij in dit dankwoord en niet voor niets als eerste. Jij hebt mij wegwijs gemaakt in de wetenschap. Van het leren kweken van cellen, gedetailleerde labverslagen maken en excel bestanden bijhouden, tot aan het op de juiste wijze omgaan met tumormateriaal ten behoeve van DNA en RNA isolatie, iets wat ik op een later moment zelfs anderen ging leren. Van begin tot eind was je nauw betrokken. Je kennis is indrukwekkend.

Dr. Wreesmann, Volkert, zonder jou was dit proefschrift er niet geweest. Als AIOS heb jij al het initiatief genomen dit project te schrijven naar aanleiding van je eigen promotie. We sparden in het AvL, maar ook in het Sint Lucas Andreas ziekenhuis of in een café. Je bewandelt zelden de meest gebruikelijke paden en vaart je eigen koers. Je bevlogenheid kan ik heel erg waarderen.

Prof. Dr. van de Brekel, Michiel, door de banden die jij met het Nijmeegse gelegd hebt ben ik in het AvL terecht gekomen. Aanvankelijk ook samen met Fons Balm, waren jullie betrokken vanuit de hoofd- en halschirurgie. Het is indrukwekkend hoeveel ballen jij in de lucht kan houden. Je hebt me de mogelijkheid gegeven me wetenschappelijk te ontwikkelen en deze kennis op verschillende mooie locaties wereldwijd ook te delen.

Prof. Dr. Verheij, Marcel, jij was al vanaf het begin betrokken en trouw aanwezig bij de vergaderingen. Je straalt naast je drukke werkzaamheden een enorme rust uit en weet altijd kritische vragen te stellen. Inmiddels heb je een stap gemaakt naar het mooie Radboudumc en ben je zo goed als mijn buurman. Dingen kunnen gek lopen.

Prof. Dr. Begg, Adrian, wat had ik je er graag bij gehad op deze dag. Jij stimuleerde de wetenschap door onder het genot van koffie en koekjes, een artikel te fileren. Je humor en warmte waardeerde ik enorm. Goede herinneringen heb ik aan onze squash wedstrijden, waarna je me altijd bedankte en complimenteerde over mijn spel, nadat je me volledig had ingemaakt. Het napraten aan de bar over het verloop van mijn promotie en wat verder opborrelde was een vanzelfsprekend onderdeel van deze activiteit.

De manuscriptcommissie. Dank voor het vrijmaken van tijd voor het lezen en beoordelen van dit manuscript.

Ik heb een fantastische tijd in het NKI-AvL gehad, waar de wetenschap op een hoog niveau bedreven wordt en dit ook gewaarborgd wordt. Naast het harde werken werd mijn tijd opgeleukt door de arts-onderzoekers, het NKI bedrijfshockeyteam, maar toch met name C2. Dank je wel Begg/Vens/Verheij, Stewart en Schellens-groep, met in het bijzonder Ingar, Geert, Monique, Ben, Hans, Fiona, Marion, Saske, Manon, Ingrid, Shuraila, Maarten, Rik en Artur. Jullie hebben mijn tijd gemaakt in het NKI.

Rosemarie, al enkele jaren terug hebben we ons gezamelijke artikel gepubliceerd, waarin jij met name het schrijfwerk op je genomen hebt. Wat fijn dat de start voor ons beiden was gemaakt.

Floor, jij kwam als geroepen! Je bent nauwkeurig en was een heel fijn persoon om erbij te hebben, zeker bij het uren en uren cellen tellen.

David, jij kwam wat later toen we met veel data zaten waar we heel hard een bio-informaticus voor nodig hadden. Vervolgens ben je zelfs vóór mij gepromoveerd. Wat ben jij belangrijk geweest bij het afronden van dit project en wat was het fijn met je te kunnen sparren.

Ook de Genomics Core Facility wil ik bedanken. Jullie hebben een wezenlijke bijdrage geleverd in het opzetten van de 'core' van deze studie.

In deze tijd is buiten de professionele werkzaamheden, het sociale minstens zo belangrijk geweest. De soms lange dagen in het lab, werden afgewisseld door met name uitstapjes met (leden van) voormalig Hurley Dames 6, inmiddels de Hurley oudjes. Wat heb ik naast het hockeyveld genoten van de "wanneer-waar moet ik zijn?"-TD-schema's, weekend tripjes en date diners. Ik had dit voor geen goud willen missen!

In het bijzonder wil ik hier Annelien noemen. Al snel werden we naast team-, ook huisgenoten. Samen kilometers naar hockey op en neer fietsen, struinen door de stad, avondjes op de bank, vakanties en nu delen we de aanwezigheid van allebei een klein mannetje in ons leven. Heel fijn dat je mijn paranimf wilt zijn!

Na mijn NKI-AvL tijd heb ik mijn studie gecombineerd met de KNO opleiding in het Radboudumc. De fijne werksfeer gecombineerd met de vele uitjes, bestaande uit weekendjes weg, dassendiners, borrels in Anneke en congressen, zijn oh zo waardevol. Mede AIOS, jullie zijn fantastisch geweest. Mijn dank ook aan de opleiders die zowel op professioneel als sociaal niveau een rol hebben gespeeld in dit project.

Al mijn andere lieve vrienden, die zo'n belangrijke rol spelen in mijn leven en altijd trouw bleven informeren hoe het ervoor stond met mijn proefschrift. Ook al werd de vraag in de loop van de jaren wat pijnlijker, jullie welgemeende interesse bleef. Wat was ik blij deze vraag eindelijk positief te kunnen beantwoorden. Ik zie uit naar nog vele fietstochtjes, dinertjes, weekendjes weg, borrels of wat dan ook.

Lieve Mariken, Arjan en Annet, wat een rijkdom om jullie om me heen te hebben. Jullie zijn er altijd voor mij en kennen me door en door. Een gezonde portie relativeringsvermogen is ons niet onbekend, en hiermee kunnen jullie alles altijd weer mooi in perspectief plaatsen. Jullie brengen een hele hoop extra familie mee. Lieve André, Elleke, Maik, Thomas, Emma, Soof, Jack en Kiki, wat fijn dat jullie er zijn.

Arjan, ik kon altijd op je terugvallen gedurende deze periode, onder andere met heel veel telefoongesprekken. Ik ben er trots op dat je straks naast me staat.

Papa en mama. Hoe kan ik met woorden goed beschrijven wat jullie voor me betekenen. Jullie hebben me laten groeien tot wie ik nu ben. Mij altijd vertrouwen gegeven en een veilige thuishaven. Meer kun je je als kind niet wensen. In sinterklaasgedichten werd jaarlijks gerefereerd naar de afronding van dit proefschrift en nu is het eindelijk zover, een afronding waar jullie nooit aan hebben getwijfeld.

Lieve Tom, de laatste woorden zijn voor jou. Ook al moet je dit plekje delen met Gijs. Het is onbeschrijfelijk fijn dat jullie in mijn leven zijn. We begonnen ons avontuur in Nieuw Zeeland en hebben er sindsdien vele beleefd. Na afronding van dit proefschrift is er (nog) meer tijd om van alle mooie dingen in ons leven te genieten. Er ligt zoveel moois voor ons in het verschiet. Ik kijk ernaar uit.