Aus der

Klinik und Poliklinik für Dermatologie und Allergologie

Klinikum der Ludwig-Maximilians-Universität München

Unveiling Novel Features of HPV-related Cancers through In Silico and In Vitro Approaches

Dissertation

zum Erwerb des Doktorgrades der Medizin
an der Medizinischen Fakultät
der Ludwig-Maximilians-Universität München

vorgelegt von

Jiahua Li

aus

Shenzhen, China

Jahr

2025

Mit Genehmigung der Medizinischen Fakultät der
Ludwig-Maximilians-Universität München

Erstes Gutachten:	Prof. Dr. Markus Reinholz
Zweites Gutachten:	Prof. Dr. Lars E. French
Drittes Gutachten:	Prof. Dr. Sven Mahner

Dekan: Prof. Dr. med. Thomas Gudermann

Tag der mündlichen Prüfung: 26.02.2025

Table of content

List of abbreviations

AEs	Adverse events
ASCC	Anal squamous cell carcinoma
APCs	Antigen-presenting cells
ATM	Ataxia telangiectasia-mutated
ATR	Ataxia telangiectasia-mutated and Rad3-related
bp	Base pairs
CCL20	C-C motif chemokine ligand 20
CC	Cervical cancer
CDC42	Cell division control protein 42
CIN	Cervical intraepithelial neoplasia
CESC	Cervical squamous cell carcinoma and endocervical adenocarcinoma
CTLs	Cytotoxic T lymphocytes
DEGs	Differentially expressed genes
DMPs	Differentially methylated positions
EGFR	Epidermal growth factor receptor
GSH	Glutathione
HNC	Head and neck cancers
HNSCC	Head and neck squamous cell carcinomas
HR	High-risk
HLA-I	Human leukocyte antigen class I
HPV	Human papillomavirus
IL	Including interleukin
IFN-γ	Interferon-γ
LCR	Long control region
LR	Low-risk
MIP3α	Macrophage inflammatory protein 3 α
NADPH	Nicotinamide adenine dinucleotide phosphate hydrogen
NOX	Nicotinamide adenine dinucleotide phosphate hydrogen oxidase
NF-κB	Nuclear factor κ-light-chain-enhancer of activated B cells
ORR	Objective response rates
ORFs	Open reading frames
OS	Overall survival

PV	Papillomavirus
PD-1	Programmed cell death protein 1
PD-L1	Programmed cell death protein ligand 1
ROS	Reactive oxygen species
RIPK3	Receptor-interacting protein kinase 3
Rb	Retinoblastoma protein
scRNA-seq	Single-cell RNA sequencing
TERT	Telomerase reverse transcriptase
WGCNA	Weighted correlation network analysis
WHO	The World Health Organization
IFITM1	Transmembrane protein 1
TNF-α	Tumor necrosis factor α
UCHL1	Ubiquitin C-terminal hydrolase L1
VEGF	Vascular endothelial growth factor

List of publications

1. **Li J**, Li J, Zhang L, Liu X, Cao Y, Wang P, Wang X. Comparison of red light and blue light therapies for mild-to-moderate acne vulgaris: A randomized controlled clinical study. Photodermatol Photoimmunol Photomed. 2022 Sep;38(5):459-464. doi: 10.1111/phpp.12769. Epub 2022 Jan 19. PMID: 34981580.

2. Wei E, Reisinger A, **Li J**, French LE, Clanner-Engelshofen B, Reinholz M. Integration of scRNA-Seq and TCGA RNA-Seq to Analyze the Heterogeneity of HPV+ and HPV- Cervical Cancer Immune Cells and Establish Molecular Risk Models. Front Oncol. 2022 Jun 1;12:860900. doi: 10.3389/fonc.2022.860900. PMID: 35719936; PMCID: PMC9198569.

3. **Li J**, Wei E, Reisinger A, French LE, Clanner-Engelshofen BM, Reinholz M. Comparison of Different Anti-Demodex Strategies: A Systematic Review and Meta-Analysis. Dermatology. 2023;239(1):12-31. doi: 10.1159/000526296. Epub 2022 Oct 28. PMID: 36310014.

4. Wei E, **Li J**, Anand P, French LE, Wattad A, Clanner-Engelshofen B, Reinholz M. "From molecular to clinic": The pivotal role of CDC42 in pathophysiology of human papilloma virus related cancers and a correlated sensitivity of afatinib. Front Immunol. 2023 Mar 1;14:1118458. doi: 10.3389/fimmu.2023.1118458. PMID: 36936942; PMCID: PMC10014535.

Among the above publications,

Integration of scRNA-Seq and TCGA RNA-Seq to Analyze the Heterogeneity of HPV+ and HPV- Cervical Cancer Immune Cells and Establish Molecular Risk Models

and

"From molecular to clinic": The pivotal role of CDC42 in pathophysiology of human papilloma virus related cancers and a correlated sensitivity of afatinib

are used for my cumulative dissertation.

Your contribution to the publications

1.1 Contribution to paper I

Wei E, Reisinger A, Li J, French LE, Clanner-Engelshofen B, Reinholz M. Integration of scRNA-Seq and TCGA RNA-Seq to Analyze the Heterogeneity of HPV+ and HPV- Cervical Cancer Immune Cells and Establish Molecular Risk Models. Front Oncol. 2022 Jun 1;12:860900. doi: 10.3389/fonc.2022.860900. PMID: 35719936; PMCID: PMC9198569.

I am the co-author of this paper.

I participated in the proofreading and analysis of the results and provided suggestions for the manuscript.

1.2 Contribution to paper II

Wei E, Li J, Anand P, French LE, Wattad A, Clanner-Engelshofen B, Reinholz M. "From molecular to clinic": The pivotal role of CDC42 in pathophysiology of human papilloma virus related cancers and a correlated sensitivity of afatinib. Front Immunol. 2023 Mar 1;14:1118458. doi: 10.3389/fimmu.2023.1118458. PMID: 36936942; PMCID: PMC10014535.

I am the co-first author and corresponding author of this paper.

I completed the following work independently:

(i) Ensuring cell line stability for *in vitro* experiments in this study, conducting cell recovery, cell culture, and cell passaging for A431 and HaCaT cells.

(ii) Conducting ethanol precipitation to obtain plasmid DNA of HPV 16 E6/E7 from ampicillin-resistant colonies on LB agar plates.

(iii) Transfecting HPV 16 E6/E7 into A431 and HaCaT cells using X-tremeGENE 9 DNA transfection reagent, and treating HaCaT and A431 cells with different concentrations of afatinib.

(iv) Conducting cell viability and proliferation assays using WST-1, incubation and measurement of results. Statistically analyzing results in this step.

(v) Exploring potential research directions to explain the findings (molecular docking for a quaternary complex of CDC42-GTPase-effector interface-EGFR-afatinib).

(vi) Writing the majority of manuscript, including the title (except for some parts in the materials and methods I was not involved in).

(vii) Submitting the paper, communicating with editors and reviewers, and conducting supplementary experiments (for all *in vitro* experiments required by reviewers).

I also collaborated in the following part with other co-authors, especially Erdong Wei:

(i) Establishing the workflow and division of tasks.

(ii) Bioinformatics analyses and interpretation of the results.

(iii) Proofreading results (bioinformatic and experimental), producing figures in this study.

2. Introduction

2.1 Human papillomavirus (HPV):

Papillomavirus (PV) has undergone co-evolution with vertebrates for over 350 million years, enabling them an excellent adaptation to its host cells and establishing itself as a prevalent virus within the vertebrate kingdom. Within this category, our attention is specifically directed towards a type of PV virus, which is closely related to us (Homo sapiens, “Human”), known as human papillomavirus, the HPV, of which, more than 200 types have already been identified[1].

2.1.1 Prevalence of HPV infection

As a virus coexisting with humans since the very dawn of our species (approximately 400,000 to 700,000 years ago[2]), HPV has well adapted in the squamous epithelia of mucosal and skin covers of humans and has spread to every corner of the Earth with the migration of its host, exhibiting a remarkably high prevalence on a global scale.

Among males, the estimated global pooled prevalence stands at 31% for any HPV and 21% for high-risk HPV[3]. Among females with normal cervical cytology, the estimated global HPV infection rate in 2010 was 11.1%, later declined to 9.9% in 2019[4], a relatively pronounced decline, which is partly due to the worldwide implementation of HPV vaccines.

2.1.2 Types of HPV

Even though the prevalence of HPV infection among global population is very high, not all HPV types can cause tumors. HPV is thus classified into high-risk as well as low-risk types based on the potential of carcinogenicity. Among these, certain HPV types are of particular concern.

1. Low-risk (LR) HPV types:

These types of HPV are generally associated with benign conditions, such as genital warts, which rarely lead to the development of cancer and are considered less carcinogenetic. LR types include HPV 11, 6, 44, 42, and 43.

2. High-risk (HR) HPV types:

These types, including types 70, 68, 16, 33, 18, 31, 66, 34, 59, 35, 58, 39, 45, 56, 52, and 51 pose a greater risk of causing malignant transformations and are strongly associated with the development of various cancers. Among these, HPV 16 and HPV 18 are the most prevalent HR types. Approximately 50% of cervical cancer (CC) cases test positive for HPV 16, while 18% are attributed to HPV 18. Similarly, around 40% and 3–7% of high-grade cervical intraepithelial neoplasia (CIN) cases are associated with HPV 18 and 16, respectively[5, 6]. The rate of progression from HPV 18- or 16-infected cervical epithelium to CIN3 or worse within 10 years is approximately 15%, compared to other HR HPV types, significantly increased[7].

2.1.3 Structure and genome of HPV

The various characteristics of HPV we observed from a “macro” perspective, such as the different carcinogenicity rate of different types of HPV, are rooted in the difference of their microscopic features.

HPV has a diameter of approximately 60 nanometers, comprising a single circular double-stranded DNA and an icosahedral capsid, the DNA of which contains roughly 8,000 base pairs (bp). Only one strand of this DNA will be utilized as the template in transcription, and it encompasses three distinct genomic regions.[8].

The early region (E) of the genome encodes 6 open reading frames (ORFs): E1, E7, E2, E6, E4, and E5, which accounts for more than 50% of the viral genome and. Proteins with the same name as the corresponding ORF are encoded in this region, among which proteins E6 and E7 play the most important roles in the HPV-related carcinogenesis, whose function will be described in the next chapter. To avoid confusion, unless otherwise stated, E6 and E7 in the following text refer to the proteins rather than the ORFs in the genome.

The interaction capacity of E6 and E7 with intracellular sites, for examples, tumor-suppressor proteins, also serves as a crucial determinant for the different carcinogenicity between HR and LR HPV. Compared to HR HPV, E6 of LR HPV could not induced the degradation of p53 protein[9]. The affinity of E7 protein of LR HPV with retinoblastoma protein (Rb) is also lower than HR HPV[10]. An insert in the E6 gene of HR HPV, which creates an extra PDZ domain-binding motif in E6, may explain the greater ability of E6 of HR HPV to interact with proteins in host cells[11].

Located downstream of the early region, the late (L) region of the viral genome encodes L2 and L1, the minor and major capsid proteins, which are targeted by HPV vaccines, accounting for almost 40% of the whole genome.

At the last, LCR, the long control region is a fragment comprising approximately 850 base pairs, accounting for about 10% of the entire genome. It lacks any protein-coding function but encompasses sequences involved in transcriptional regulation as well as the origin of DNA replication.

2.2 HPV-mediated carcinogenesis

2.2.1 Initiation of HPV-mediated carcinogenesis

The most crucial factors contributing to the carcinogenic effects of HPV, as mentioned earlier, are the E6 and E7. Their interactions with tumor suppressor proteins are the most well-understood initial mechanism of HPV-mediated carcinogenesis[8]. Rb, together with other Rb pocket proteins, could be bound by E7 of HR HPV, which leads to the inhibition and proteasomal degradation of Rb [12]. Rb is an important factor in cell cycle regulation. The interaction of E7 with Rb prevents its binding with transcription factors of E2F-family, resulting in enhanced E2F-dependent transcription. This, in turn, promotes the transition into S-phase of the cell cycle, leading to an enhanced cell proliferation, concurrently with an escalation of viral gene transcription[13, 14].

In normal circumstances, p53 protein could counteract this dysregulated Rb/E2F pathway by inducing apoptosis and inhibiting cell growth[15]. However, this protective mechanism is blocked by the “accomplice” of E7, E6. E6 induces proteasomal degradation of p53, by firstly binding to cellular ubiquitin ligase E6AP, forming a E6/E6AP heterodimer, which subsequently recruits and degrade p53 in a E6/p53/E6AP complex structure[16]. In addition to its protein-level interaction, E6 of specific types of HPV (including 16, 31, 18, 11), may also directly bind to the DNA of p53, hindering the transcription of this tumor-suppressive gene[17].

2.2.2 Accumulation of carcinogenetic events intracellular

The initiation process mentioned above alone is not sufficient to complete the transformation of host cells into cancer cells. However, due to abnormalities in apoptosis and increased proliferation in the host cell caused in the initiation process, an accumulation of errors occurred in cell replication becomes much more possible[8]. The accumulation of these intracellular events that promote malignant transformation will ultimately manifest as instability of the host cell genome, which could be triggered by HPV through various following mechanisms[18]:

1. Integration of HPV DNA:

Performing whole-genome sequencing on HPV-integrated cancer cells unveils structural changes such as translocations, inversions, duplications and deletions. These alterations are frequently accompanied or connected by HPV integrants[19, 20]. Studies have found that E7 levels are associated with quantitative and structural abnormalities in chromosomes, which was observed only after the integration of HPV16 DNA into host cells[21]. In addition, the integration process requires the break of both strands of DNA in the host cell, a process that itself tends to destabilize the genome and introduce new errors.

2. Generating oxidative stress:

E6 of HPV 16 and 18 could lead to a reduction in expressional levels and enzymatic activity of glutathione (GSH) and catalase proteins, correlating with elevated oxidative stress and DNA damage, which were also observed through expressions of E7 and co-expression of E1 and E2 [22]. Furthermore, E6*I, a spliced isoform of E6, which exhibits the highest abundance of transcripts of RNAs in HPV-related cancers, could enhance the expression of genes associated with nicotinamide adenine dinucleotide phosphate hydrogen (NADPH) oxidase (NOX)-dependent reactive oxygen species (ROS) production and is considered to play a pivotal role in generating oxidative stress and subsequential DNA damage[23].

3. Modifying the length of telomeres:

The regulation of telomere length associated with the action of HPV oncoproteins is typically attributed to E6, while the effects of E7 manifest as either telomere shortening or lengthening in different contexts[18, 24]. Research findings indicate that in cells transfected with E6/E7, there is a correlation between low expression of Telomerase reverse transcriptase (TERT) and significant shortening of telomeres. This correlation is associated with an elevated frequency of anaphase bridges, serving as a marker for unsuccessful chromosome segregation[25]. This effect is particularly pronounced in the situation when E7 is the only one expressed, and a recover of the normal phenotype can be achieved by introducing TERT[25].

4. Impairing DNA repair mechanisms:

The DNA repair mechanism can correct some genomic errors caused by the above process, but unfortunately it is also targeted. E6 and E7 induce the activation of the ataxia telangiectasia-mutated (ATM) and consequently the Rad3-related (ATR) DNA damage repair pathways upon encountering single-strand or double-strand DNA breaks.[26]. While the conventional response of these pathways is to induce arrest in cell cycle, E7 disrupts this usual process by facilitating the degradation of a pivotal protein involved in the recovery in DNA damage, called claspin[27]. Consequently, host cells are misled into perceiving that DNA damage repair has taken place, allowing cells to proceed through mitosis despite the existence of DNA damage. This, in turn, exacerbates genome instability and contributes to the advancement of malignancy[8].

2.2.3 HPV-mediated immune evasion

After undergoing the transformation from innocent epithelial cells to cancer cells through the aforementioned steps, cancer cells are confronted with a critical challenge: how to successfully evade the host immune system? A fully functioning immune system possesses the capability to eliminate even some low-grade intraepithelial neoplasia lesions.

An intriguing phenomenon is observed: the epithelial neoplasms always develop from the "surface side" towards the "inner side" rather than the opposite, despite HPV only infecting the basal cells on the inner side of the epithelium. Part of the reason for this may be that only HPV-infected cells, which have migrated to the "surface side" of the epithelium and distanced themselves from the reach of host immune surveillance, can further develop into a cancerous lesion.

Another crucial factor contributing to this phenomenon is the nature of HPV itself. Due to the repressive effect of HPV E2 in the transcriptional level, the initial expression of HPV viral genes in host cells is low, minimizing the potential presentation to the host immune system. The initiation of carcinogenesis, marked by increased expressions of E6 and E7, occurs only after the integration of the viral genome into its host cells when the disruption of E2 takes place[8].

In addition to the strategy above, HPV can also interfere with host's immune system, through different approaches and at multiple phases.

HPV reduces the chance of the infected host cells being recognized by antigen-presenting cells (APCs). HPV modifies the expression of immunoproteasome subunits PSMB8 and PSMB9, proteins related to antigen processing[28]. HPV downregulate C-C motif chemokine ligand 20 (CCL20), a chemokine for Langerhans cells[29].

Even recognized by immune system, HPV minimizes the impact of immune response to host cells. E5 of HPV16 reduces the expression of human leukocyte antigen class I (HLA-I) on the cell surface, disrupting the recognition and killing effect of cytotoxic T lymphocytes (CTLs)[30]. HPV downregulates the expression of interferon-induced transmembrane protein 1 (IFITM1) andreceptor-interacting protein kinase 3 (RIPK3), evading antiproliferative effects and necroptosis mediated by tumor necrosis factor α (TNF-α) and Interferon-γ (IFN-γ)[31]. HPV also upregulates ubiquitin C-terminal hydrolase L1 (UCHL1), suppressing the innate immune response of host cells[32].

In addition, HPV reduces inflammatory activity, hindering both the recognition and the response intensity of immune system. HPV reduces the production of several proinflammatory cytokines, including macrophage inflammatory protein 3 α (MIP3α), TNF-α, interleukin (IL)-8 and IL-6 in host cells [32, 33]. E6 and E7 of high-risk HPV could inhibit pathway involved in inflammatory responses, such as nuclear factor κ-light-chain-enhancer of activated B cells (NF-κB) pathway, by binding to the coactivator of NF-kB in the nucleus[34].

2.3 HPV-related cancers

2.3.1 Global burden of HPV-related cancers

HPV infection is related with cancers at multiple sites, including CC, anogenital cancers and head and neck cancers (HNC). Annually, an estimated 625,600 women and 69,400 men develop HPV-related cancers, accounting for approximately 5% of all cancers worldwide[35], making it a significant public health challenge.

CC, the most common type among all HPV-related cancers, could be attributed to HPV infection in over 99% of cases[36]. In 2020 alone, they contributed to an estimated 604,100 new cancer cases and 341,800 cancer-related deaths worldwide. This positioned them as the third leading cause of both cancer incidence and mortality among all cancers in women. Particularly in Africa, Central America, and Southeast Asia, where screening test and HPV vaccine are less available, CC continue to be one of the primary causes of women cancer death[37].

Anogenital cancers (located in the vulva, vagina, penis, and anus) are another group of cancers closely related to HPV. According to estimates, a varying percentage from 15% to 48% of vulvar cancers, approximately 51% of penile cancers, 78% of vaginal cancers, and 88% of anal cancers are considered to be related to HPV infection[38]. In 2020, they collectively contributed to 150,100 new cancer cases and 57,900 cancer deaths globally, with the highest numbers of both incidence and mortality recorded in anal cancers [37]. Unlike CC, the routine screening for which has been introduced globally, anal cancers, including its most common type, anal squamous cell carcinoma (ASCC), are often first diagnosed at an advanced stage, which is quite challenging to manage, leaving the 5-year survival rates of locally advanced and metastatic ASCC cases only 65% and 32%, respectively[39].

Some types of HNC are considered to be caused by HPV, including those in oropharynx, oral cavity and larynx[40]. On a global scale, there are estimated 456,000 cases of HNC diagnosed annually, 37,200 cases from which could be attributed to HPV, specifically 3,800 in the larynx, 4,400 in the oral cavity and 29,000 in the oropharynx [41]. HNC primarily affect men, accounting for about 50% of all HPV-related cancers in men[41].

In addition to directly bringing pain and death, HPV-related cancers also cause a significant economic burden on healthcare systems. According to estimates, the total economic burden on the French National Health Insurance for potentially HPV-related cancers in the 3 years following cancer diagnosis is approximately €500 million (an annual cost of €14,629 per hospitalized patient) [42]. Among these, female genital and anal cancers account for one-third (€167 million), while HNC account for two-thirds (€343 million)[42]. In Sweden, the total cost of all HR HPV-attributable precancers and cancers is approximately €94 million each year[43]. Despite lacking direct data for citation, it is conceivable that, with lower fiscal revenue and higher incidence, the economic burden might be more pronounced in less developed countries.

2.3.2 Current challenges in treatment of HPV-related cancers

From a public health perspective, the optimal strategy to counter HPV-related cancers is reinforcing primary and secondary prevention, namely, promoting HPV vaccines and early screening. The World Health Organization (WHO) has been actively engaged in this endeavor for the past two decades. However, due to the relatively short duration of these efforts and the disparities in global development levels (including economic and health care system levels), the effectiveness of this strategy has been limited. As narrated in the previous section, there are still many newly diagnosed HPV-related cancer patients each year around the world.

In the early stages of the disease (*in situ* or precancerous), cure can be achieved through methods such as surgical resection. However, once the tumor breaches the basement membrane (local advanced stage), treatment becomes more complex, requiring more extensive surgical procedures and postoperative chemoradiotherapy. Especially for patients already in the metastatic stage, treatment becomes particularly challenging.

Advancements in targeted cancer therapy have brought new hope for patients with metastatic HPV-related cancers. The most significant advancement is the combination of traditional chemotherapy with angiogenesis inhibitors or/and immune checkpoint Inhibitors, such as vascular endothelial growth factor (VEGF) inhibitors, as well as programmed cell death protein 1 (PD-1) and PD-ligand 1 (PD-L1) inhibitors.

The combination of bevacizumab, a VEGF inhibitor, with cisplatin-paclitaxel resulted in higher overall survival (OS) (17.0 months vs. 13.3 months, $P = 0.004$) and increased objective response rates (ORR) (48% vs. 36%, $P = 0.008$)[44]. Pembrolizumab, a PD-1 monoclonal antibody, combined to platinum (cisplatin or carboplatin)-paclitaxel with or without bevacizumab, showed improved ORR (65.9% vs. 50.8%) and better OS at 24 months (50.4% vs. 40.4%, $P < 0.001$)[45]. Due to their outstanding therapeutic efficacy demonstrated in clinical studies, both bevacizumab and pembrolizumab have been recommended in European as the first-line management for distant recurrence or metastatic CC[46].

The therapeutic benefits of PD-L1/PD-1 inhibitors have also been investigated in patients with advanced ASCC and head and neck squamous cell carcinomas (HNSCC). However, compared to CC, these findings were less encouraging. Nivolumab and pembrolizumab improved OS in metastatic HNSCC, but durable responses to these treatments were achieved in less than 20% of patients[47]. Combined therapy of pembrolizumab, platinum, and 5-FU for HNSCC exhibited an increased OS and ORR, but also a 69.3% rate of grade 3–5 adverse events (AEs) [48], significantly higher than those observed in CC. Monotherapies for HNSCC had less AEs, but showed only low ORR of 13-18%[49]. In previously treated advanced ASCC patients, the ORR to pembrolizumab were only 3% and 15% in patients who have PD-L1 negative and positive tumors, respectively[50].

Effectiveness of angiogenesis inhibitors, apart from CC, also remains uncertain. Most clinical trials did not demonstrate the efficacy of angiogenesis inhibitors in the treatment of HNSCC, but embarrassingly, their association to toxicity[51]. Additionally, no relevant clinical trials have been found regarding angiogenesis inhibitors in ASCC.

In summary, for metastatic HPV-related cancer patients with poor response to first-line treatment, PD-L1 negative tumors, or those who cannot tolerate the AEs, there is still a lack of promising second-line treatment options. Although numerous clinical trials are attempting to validate the potential treatment options, including the administration of drugs with similar pharmacological mechanisms, the superiority of any treatment method has not yet been confirmed.

To navigate the complex clinical scenarios, we still need a more sophisticated arsenal with more advanced weaponry.

2.4 Methodological overview of integrating *in silico* and *in vitro* approaches

As a downstream discipline of application, the most significant developments in medicine since 20th century were based on the breakthrough in basic science. For example, the innovative diagnostic method based on X-ray and the discover of DNA in X-ray diffraction. However, since the theories of relativity and quantum mechanics, there have been no more disruptive breakthroughs in basic science in the past 100 years. Therefore, after translating the previously occurred breakthroughs in basic science into breakthroughs within its own field, progress in medical research has also become slow.

To gain a deeper understanding of medical topics, we must depend on advancements in other disciplines.

2.4.1 Development of bioinformatics technology

Over the past two decades, the most progressive scientific field is information and computational science, the breakthroughs in which have greatly advanced the development of various disciplines, including life sciences. Here are some key developments in bioinformatics technology over the last 20 years:

1. Revolution in high-throughput sequencing technologies: The emergence of second-generation sequencing technologies, such as Ion Torrent as well as Illumina, has made high-throughput sequencing more cost-effective. This development has propelled research in genomics, transcriptomics, and proteomics.

2. Single-cell sequencing technology: The emergence of single-cell sequencing technology empowers researchers to explore the gene expression and genomic information at the individual cell level, uncovering heterogeneity and diversity within cell populations, including cancers.

3. Advancements in proteomics: Continuous improvements in proteomic technologies, such as enhanced mass spectrometry and improved protein affinity purification methods, contribute to a more comprehensive understanding of protein structure and function.

4. Application of big data and artificial intelligence: The exponential growth of data in the field of bioinformatics underscores the importance of big data analytics and artificial intelligence applications. These technologies find applications in disease prediction, gene identification, drug development, and more.

2.4.2 The crucial role of integrating: why is *in vitro* validation decisive for *in silico* analysis?

Although bioinformatics technology has become an indispensable and powerful tool in the field of life sciences, it has some limitations.

In my opinion, the most significant limitation is: conclusions obtained from pure bioinformatic analysis may not align consistently with the real world. This phenomenon can be attributed to several factors.:

1. Interpretation biases in computational results: In bioinformatic analysis, the interpretation of computational results may be influenced by subjective biases of the researcher. This could include errors or biases in understanding gene expression, protein interactions, or other biological processes.

2. Data quality and accuracy: bioinformatic analysis relies on the quality and accuracy of input data. If the raw experimental data contains noise, bias, or other issues, the results of the analysis may be affected.

3. Limitations of computational models: Bioinformatic methods often use specific computational models to simulate biological processes. However, these models may simplify the complexity and diversity of real biological systems, leading to inconsistencies between computational results and real-world situations.

4. Limitations in data representation: Bioinformatic data often exists in a high-dimensional form, while our human understanding and representation are typically limited to three-dimensional space. When compressing high-dimensional data into understandable charts or graphs, information loss may result in a misunderstanding of the real-world situation.

To validate the findings in bioinformatic analysis, it is necessary to conduct experiments in the real world. Without this validation, even the most compelling results based on pure bioinformatic analysis could potentially result in disastrous outcomes in real-world scenarios.

Therefore, in our studies, we chose approaches integrating *in silico* (bioinformatic) analysis and *in vitro* experiments to investigate the HPV-related cancers from novel entry points, and expected to draw reliable research conclusions from the results.

In the following two sections, I will briefly introduce how the above methodology was applied to our research and outline the most significant findings we have attained.

2.5 Research design

2.5.1 Research design of paper I

This study was aimed to find the heterogenicity of immunological features between HPV- and HPV+ cancers and build prognostic gene risk models for both survival and relapse. As a preliminary investigation, this study was designed to be pure *in silico* analysis.

Data of single-cell RNA sequencing (scRNA-seq) and bulk RNA sequencing of cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC) were downloaded and analyzed, through which the differentially expressed genes (DEGs) and heterogenicity in immune infiltration between HPV+ and HPV- tumors were identified. Because they were inferred to play important roles in HPV-related cancers, DEGs and genes related to immune heterogenicity were recruited in the construction of prognostic risk models. Genes in the models were then annotated and investigated in other HPV-related cancers to further expand their validity.

2.5.2 Research design of paper II

In this study, we aimed to find the pivotal gene in the progression of HPV-related cancers, and identified potential promising drug targeted it.

We started from the entry point of DNA methylation. DNA methylation could be altered by HPV, but we still lack a detailed understanding of the role of this mechanism in carcinogenesis. Differentially methylated positions (DMPs) between HPV-related cancers and normal tissues were identified and the pathways they involved in were annotated. Genes in these pathways, together with genes identified through clustered visualized immune infiltration analysis based on the downstream scRNA-seq data, were pooled. Based on the genes obtained from the above steps and the data from The Cancer Genome Atlas (TCGA) database, we could construct a matrix including clinical data as well as expression profiling data. Then genes with the most significant correlation to the progression (advanced stage) of the HPV-related cancers could be screened. We could subsequently find a potential promising drug, which exhibited sensitivity to these genes. To validate the findings in the *in silico* analyses, *in vitro* experiments were conducted, administrating the identified drug in normal epithelial cells and squamous carcinoma cells transfected with or without HPV E6/E7 and the cell proliferation assay were performed.

2.6 Main findings

2.6.1 Main findings of paper I

Immune analysis showed a reduction of B cells and CD8+ T cells and an increase of Treg cells, CD4+ T cells, and epithelial cells in HPV+ CESC samples. In clinical aspect, patients with less CD8+ T cells and naive B cells have worse prognosis in terms of survival. Based on the genes related to CD8+ T cells, naive B cells, and DEGs in HPV+/- samples, a 9-genes prognostic risk model for survival in CESC patients and a 7-genes model for relapse, were established. IKZF3, FOXP3, and JAK3 were found to be included in both models and have a protective effect for CESC patients. Risk models based on them were further validated in other HPV-related cancers and similar protective effects were observed in HNSCC.

2.6.2 Main findings of paper II

Based on analysis of DNA methylation data from ASCC, CESC and their precancerous lesions, we initially demonstrated the similarity in the level of DNA methylation between HPV-related cancers. Subsequently, the pathways of these DMPs were found to be related to HPV infection, immune, oxidative stress, ferroptosis and necroptosis. Through weighted correlation network analysis (WGCNA), we found the top ten genes related to cancers in advanced stage. Following that, an examination was conducted to analyze the associations between the expressions of these genes and survival outcomes. Among them, cell division control protein 42 (CDC42) was the only statistically significant gene correlated to the survival of patients. In the next phase, we screened afatinib for its highest positive sensitivity correlation with CDC42 as a potential drug. The efficacy of afatinib was then validated in *in vitro* experiments, showing an increased inhibitory effect on A431 and HaCaT cells after transfection with HPV E6 and E7.

During our discussions on the findings, we encountered difficulty explaining the sensitivity of afatinib in cells with high CDC42 expression. Considering that CDC42 were only found intracellular, how could afatinib, an epidermal growth factor receptor (EGFR) inhibitor, interact with it and where is this sensitivity generated from? To explain this, through computational molecular docking method, we find a stable quaternary complex of CDC42-GTPase-effector interface-EGFR-afatinib, which may support our findings in the molecular level.

In conclusion, from demonstrating the molecular interaction of afatinib, the alteration in DNA methylation level and transcriptomic level, to the correlation of CDC42 with clinical features, we identified CDC42 as a pivotal gene in the progression of HPV-related cancers and highlighted afatinib as a potential promising drug.

3. Summary (in Englisch)

High-risk HPV has the potential to induce carcinogenesis in epithelial cells through diverse pathways. Although primary and secondary prevention has been promoted by the WHO over the past two decades, HPV-related cancers still demonstrate considerable global incidence rates and economic impacts and treatment options for patients with advanced HPV-related cancers still remain limited. To develop novel therapeutic drugs and more effectively address the challenges posed by HPV-related cancers, it is essential to explore these cancers from new perspectives. The advancements in bioinformatics technology provide a possibility for this, and with validations through *in vitro* experiments, its reliability limitations could be compensated.

In our studies, we utilized both bioinformatic analysis and *in vitro* experiments to investigate HPV-related cancers. Main findings in these studies involves the establishment of gene risk models for survival and relapse in patients with HPV-related cancers, the identification of the pivotal role played by CDC42 in cancer progression, and the screening of afatinib as a potential therapeutic drug, which unveiled novel features of these cancers and laid the foundation for the development of novel targeted therapy in the future.

4. Zusammenfassung (deutsch)

High-risk HPV hat das Potenzial, Karzinogenese in epithelialen Zellen durch vielfältige Wege zu induzieren. Obwohl die primäre und sekundäre Prävention in den letzten zwei Jahrzehnten von der WHO gefördert wurde, zeigen HPV-bezogene Krebserkrankungen weiterhin erhebliche globale Inzidenzraten und wirtschaftliche Auswirkungen, und die Behandlungsoptionen für Patienten mit fortgeschrittenen HPV-bezogenen Krebserkrankungen bleiben begrenzt. Um neuartige therapeutische Medikamente zu entwickeln und den Herausforderungen von HPV-bezogenen Krebserkrankungen wirksamer zu begegnen, ist es entscheidend, diese Krebserkrankungen aus neuen Perspektiven zu erforschen. Die Fortschritte in der Bioinformatik-Technologie bieten hierfür eine Möglichkeit, und mit Validierungen durch *in vitro*-Experimente könnten die Zuverlässigkeitsgrenzen ausgeglichen werden.

In unseren Studien haben wir sowohl bioinformatische Analysen als auch *in vitro*-Experimente genutzt, um HPV-bezogene Krebserkrankungen zu untersuchen. Die Hauptergebnisse dieser Studien umfassen die Entwicklung von genetischen Risikomodellen für das Überleben und das Wiederauftreten bei Patienten mit HPV-bezogenen Krebserkrankungen, die Identifizierung der entscheidenden Rolle von CDC42 im Fortschreiten von Krebs und das Screening von Afatinib als potenzielles therapeutisches Medikament. Diese Erkenntnisse enthüllen neue Merkmale auf diese Krebserkrankungen und legen den Grundstein für die Entwicklung neuartiger zielgerichteter Therapieansätze in der Zukunft.

5. Paper I

Wei E, Reisinger A, Li J, French LE, Clanner-Engelshofen B, Reinholz M. Integration of scRNA-Seq and TCGA RNA-Seq to Analyze the Heterogeneity of HPV+ and HPV- Cervical Cancer Immune Cells and Establish Molecular Risk Models. Front Oncol. 2022 Jun 1;12:860900. doi: 10.3389/fonc.2022.860900. PMID: 35719936; PMCID: PMC9198569.

6. Paper II

Wei E, Li J, Anand P, French LE, Wattad A, Clanner-Engelshofen B, Reinholz M. "From molecular to clinic": The pivotal role of CDC42 in pathophysiology of human papilloma virus related cancers and a correlated sensitivity of afatinib. Front Immunol. 2023 Mar 1;14:1118458. doi: 10.3389/fimmu.2023.1118458. PMID: 36936942; PMCID: PMC10014535.

References

1. Shimizu, A., R. Yamaguchi, and Y. Kuriyama, *Recent advances in cutaneous HPV infection.* J Dermatol, 2023. **50**(3): p. 290-298.
2. Stringer, C., *The origin and evolution of Homo sapiens.* Philos Trans R Soc Lond B Biol Sci, 2016. **371**(1698).
3. Bruni, L., et al., *Global and regional estimates of genital human papillomavirus prevalence among men: a systematic review and meta-analysis.* Lancet Glob Health, 2023. **11**(9): p. e1345-e1362.
4. Kombe Kombe, A.J., et al., *Epidemiology and Burden of Human Papillomavirus and Related Diseases, Molecular Pathogenesis, and Vaccine Evaluation.* Front Public Health, 2020. **8**: p. 552028.
5. Azuma, Y., et al., *Human papillomavirus genotype distribution in cervical intraepithelial neoplasia grade 2/3 and invasive cervical cancer in Japanese women.* Jpn J Clin Oncol, 2014. **44**(10): p. 910-7.
6. Bulk, S., et al., *The contribution of HPV18 to cervical cancer is underestimated using high-grade CIN as a measure of screening efficiency.* Br J Cancer, 2007. **96**(8): p. 1234-6.
7. Khan, M.J., et al., *The elevated 10-year risk of cervical precancer and cancer in women with human papillomavirus (HPV) type 16 or 18 and the possible utility of type-specific HPV testing in clinical practice.* J Natl Cancer Inst, 2005. **97**(14): p. 1072-9.
8. Cosper, P.F., et al., *Biology of HPV Mediated Carcinogenesis and Tumor Progression.* Semin Radiat Oncol, 2021. **31**(4): p. 265-273.
9. Li, X. and P. Coffino, *High-risk human papillomavirus E6 protein has two distinct binding sites within p53, of which only one determines degradation.* J Virol, 1996. **70**(7): p. 4509-16.
10. Munger, K., et al., *Complex formation of human papillomavirus E7 proteins with the retinoblastoma tumor suppressor gene product.* EMBO J, 1989. **8**(13): p. 4099-105.
11. Auslander, N., et al., *A unique insert in the genomes of high-risk human papillomaviruses with a predicted dual role in conferring oncogenic risk.* F1000Res, 2019. **8**: p. 1000.
12. Gheit, T., *Mucosal and Cutaneous Human Papillomavirus Infections and Cancer Biology.* Front Oncol, 2019. **9**: p. 355.
13. Helin, K., E. Harlow, and A. Fattaey, *Inhibition of E2F-1 transactivation by direct binding of the retinoblastoma protein.* Mol Cell Biol, 1993. **13**(10): p. 6501-8.
14. Chellappan, S., et al., *Adenovirus E1A, simian virus 40 tumor antigen, and human papillomavirus E7 protein share the capacity to disrupt the interaction between transcription factor E2F and the retinoblastoma gene product.* Proc Natl Acad Sci U S A, 1992. **89**(10): p. 4549-53.
15. Demers, G.W., C.L. Halbert, and D.A. Galloway, *Elevated wild-type p53 protein levels in human epithelial cell lines immortalized by the human papillomavirus type 16 E7 gene.* Virology, 1994. **198**(1): p. 169-74.
16. Martinez-Zapien, D., et al., *Structure of the E6/E6AP/p53 complex required for HPV-mediated degradation of p53.* Nature, 2016. **529**(7587): p. 541-5.
17. Lechner, M.S. and L.A. Laimins, *Inhibition of p53 DNA binding by human papillomavirus E6 proteins.* J Virol, 1994. **68**(7): p. 4262-73.
18. Porter, V.L. and M.A. Marra, *The Drivers, Mechanisms, and Consequences of Genome Instability in HPV-Driven Cancers.* Cancers (Basel), 2022. **14**(19).
19. Akagi, K., et al., *Genome-wide analysis of HPV integration in human cancers reveals recurrent, focal genomic instability.* Genome Res, 2014. **24**(2): p. 185-99.

20. Parfenov, M., et al., *Characterization of HPV and host genome interactions in primary head and neck cancers.* Proc Natl Acad Sci U S A, 2014. **111**(43): p. 15544-9.

21. Pett, M.R., et al., *Acquisition of high-level chromosomal instability is associated with integration of human papillomavirus type 16 in cervical keratinocytes.* Cancer Res, 2004. **64**(4): p. 1359-68.

22. Cruz-Gregorio, A., et al., *Human Papillomavirus Types 16 and 18 Early-expressed Proteins Differentially Modulate the Cellular Redox State and DNA Damage.* Int J Biol Sci, 2018. **14**(1): p. 21-35.

23. Paget-Bailly, P., et al., *Comparative RNA sequencing reveals that HPV16 E6 abrogates the effect of E6*I on ROS metabolism.* Sci Rep, 2019. **9**(1): p. 5938.

24. McMurray, H.R. and D.J. McCance, *Human papillomavirus type 16 E6 activates TERT gene transcription through induction of c-Myc and release of USF-mediated repression.* J Virol, 2003. **77**(18): p. 9852-61.

25. Plug-DeMaggio, A.W., et al., *Telomere erosion and chromosomal instability in cells expressing the HPV oncogene 16E6.* Oncogene, 2004. **23**(20): p. 3561-71.

26. Spriggs, C.C. and L.A. Laimins, *Human Papillomavirus and the DNA Damage Response: Exploiting Host Repair Pathways for Viral Replication.* Viruses, 2017. **9**(8).

27. Spardy, N., et al., *Human papillomavirus 16 E7 oncoprotein attenuates DNA damage checkpoint control by increasing the proteolytic turnover of claspin.* Cancer Res, 2009. **69**(17): p. 7022-9.

28. Evans, M., et al., *Antigen processing defects in cervical carcinomas limit the presentation of a CTL epitope from human papillomavirus 16 E6.* J Immunol, 2001. **167**(9): p. 5420-8.

29. Caberg, J.H., et al., *Increased migration of Langerhans cells in response to HPV16 E6 and E7 oncogene silencing: role of CCL20.* Cancer Immunol Immunother, 2009. **58**(1): p. 39-47.

30. Campo, M.S., et al., *HPV-16 E5 down-regulates expression of surface HLA class I and reduces recognition by CD8 T cells.* Virology, 2010. **407**(1): p. 137-42.

31. Ma, W., et al., *Human Papillomavirus Downregulates the Expression of IFITM1 and RIPK3 to Escape from IFNgamma- and TNFalpha-Mediated Antiproliferative Effects and Necroptosis.* Front Immunol, 2016. **7**: p. 496.

32. Karim, R., et al., *Human papillomavirus (HPV) upregulates the cellular deubiquitinase UCHL1 to suppress the keratinocyte's innate immune response.* PLoS Pathog, 2013. **9**(5): p. e1003384.

33. Richards, K.H., et al., *The human papillomavirus (HPV) E7 protein antagonises an Imiquimod-induced inflammatory pathway in primary human keratinocytes.* Sci Rep, 2015. **5**: p. 12922.

34. Tilborghs, S., et al., *The role of Nuclear Factor-kappa B signaling in human cervical cancer.* Crit Rev Oncol Hematol, 2017. **120**: p. 141-150.

35. Sung, H., et al., *Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries.* CA Cancer J Clin, 2021. **71**(3): p. 209-249.

36. Okunade, K.S., *Human papillomavirus and cervical cancer.* J Obstet Gynaecol, 2020. **40**(5): p. 602-608.

37. Ferlay, J., et al., *Cancer statistics for the year 2020: An overview.* Int J Cancer, 2021.

38. Plummer, M., et al., *Global burden of cancers attributable to infections in 2012: a synthetic analysis.* Lancet Glob Health, 2016. **4**(9): p. e609-16.

39. Astaras, C., A. Bornand, and T. Koessler, *Squamous rectal carcinoma: a rare malignancy, literature review and management recommendations.* ESMO Open, 2021. **6**(4): p. 100180.

40. Serrano, B., et al., *Epidemiology and burden of HPV-related disease.* Best Pract Res Clin Obstet Gynaecol, 2018. **47**: p. 14-26.

41. de Martel, C., et al., *Worldwide burden of cancer attributable to HPV by site, country and HPV type.* Int J Cancer, 2017. **141**(4): p. 664-670.

42. Abramowitz, L., et al., *Epidemiological and economic burden of potentially HPV-related cancers in France.* PLoS One, 2018. **13**(9): p. e0202564.

43. Ostensson, E., et al., *The economic burden of human papillomavirus-related pre-cancers and cancers in Sweden.* PLoS One, 2017. **12**(6): p. e0179520.

44. *Improved Survival with Bevacizumab in Advanced Cervical Cancer.* N Engl J Med, 2017. **377**(7): p. 702.

45. Colombo, N., et al., *Pembrolizumab for Persistent, Recurrent, or Metastatic Cervical Cancer.* N Engl J Med, 2021. **385**(20): p. 1856-1867.

46. Cibula, D., et al., *ESGO/ESTRO/ESP Guidelines for the management of patients with cervical cancer - Update 2023.* Radiother Oncol, 2023. **184**: p. 109682.

47. Ruffin, A.T., et al., *Improving head and neck cancer therapies by immunomodulation of the tumour microenvironment.* Nat Rev Cancer, 2023. **23**(3): p. 173-188.

48. Harrington, K.J., et al., *Pembrolizumab With or Without Chemotherapy in Recurrent or Metastatic Head and Neck Squamous Cell Carcinoma: Updated Results of the Phase III KEYNOTE-048 Study.* J Clin Oncol, 2023. **41**(4): p. 790-802.

49. Cramer, J.D., B. Burtness, and R.L. Ferris, *Immunotherapy for head and neck cancer: Recent advances and future directions.* Oral Oncol, 2019. **99**: p. 104460.

50. Marabelle, A., et al., *Pembrolizumab for previously treated advanced anal squamous cell carcinoma: results from the non-randomised, multicohort, multicentre, phase 2 KEYNOTE-158 study.* Lancet Gastroenterol Hepatol, 2022. **7**(5): p. 446-454.

51. Hyytiainen, A., et al., *Angiogenesis Inhibitors for Head and Neck Squamous Cell Carcinoma Treatment: Is There Still Hope?* Front Oncol, 2021. **11**: p. 683570.

Acknowledgements

I extend my heartfelt gratitude to all those who have provided valuable support and guidance throughout the completion of this dissertation.

First and foremost, I would like to express my gratitude to Professor Reinholz for being my supervisor. I am truly thankful for the invaluable opportunity he provided, enabling me to traverse great distances and undertake studies and work in Munich. Additionally, I extend my deep appreciation for the comfortable and warm laboratory atmosphere he has fostered for everyone.

Next, I would like to extend my gratitude to Dr. Clanner-Engelshofen. His ever-helpful nature and robust support in experimental techniques have been instrumental in helping me adapt to the laboratory environment. His calm and composed demeanor, coupled with a sense of humor in his approach, also earned my deep admiration.

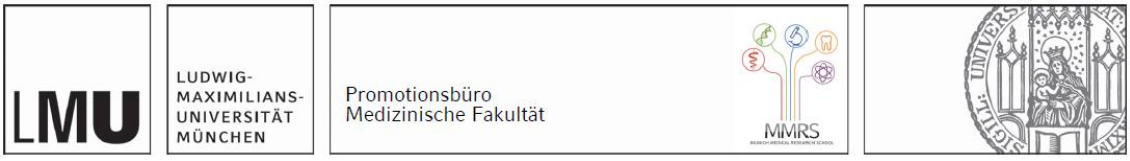
Furthermore, I am most thankful for my colleague and friend, Erdong Wei. His profound expertise in bioinformatics technology, rigorous logic, and remarkable proactiveness have left a lasting impression on me. Our collaboration and discussions have always been filled with enjoyment. The time spent together in Munich has been a delightful and memorable part of my life.

I am also grateful to all the laboratory colleagues who have provided assistance, especially Philipp, Amin, Takashi, and Claudia.

Finally, I wish to thank my wife, my parents, and my friends. They are my unwavering pillars of strength, providing constant support. Your companionship has made my journey in scientific research more steadfast and resilient.

Once again, thank you all for your support and companionship.

Affidavit



Affidavit

Li, Jiahua
_____
Surname, first name

_____
Street

_____
Zip code, town, country

I hereby declare, that the submitted thesis entitled:

Unveiling Novel Features of HPV-related Cancers through *In Silico* and *In Vitro* Approaches

……

is my own work. I have only used the sources indicated and have not made unauthorised use of services of a third party. Where the work of others has been quoted or reproduced, the source is always given.

I further declare that the dissertation presented here has not been submitted in the same or similar form to any other institution for the purpose of obtaining an academic degree.

Munich, 26.02.2025_____ _____Jiahua Li_____
place, date Signature doctoral candidate