Aus dem
Pathologischen Institut

Institut der Universität München
Direktor: Prof. Dr. Frederick Klauschen

Reproducibility in cancer research: A study of the (epi)genomic, transcriptomic, and phenotypic stability in chromosomal translocation-driven pediatric sarcoma cell lines

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

vorgelegt von
Merve Kasan

aus
Seyhan

Jahr
2024

Mit Genehmigung der Medizinischen Fakultät

der Universität München

Berichterstatter:	Prof. Dr. Dr. Thomas Grünewald
Mitberichterstatter:	Prof. Dr. Lars Lindner Prof. Dr. Roland Dürr
Mitbetreuung durch den promovierten Mitarbeiter:	Dr. Florencia Cidre Aranaz
Dekan:	Prof. Dr. med. Thomas Gudermann
Tag der mündlichen Prüfung:	12.12.2024

Acknowledgements

I extend my deepest appreciation to my supervisor, Prof. Dr. med. Thomas Grünewald, PhD., and my mentor, Dr. Florencia Cidre Aranaz, for their unwavering support throughout this research endeavor. Their dedication to scientific inquiry and mentorship has been a constant source of inspiration for me. Without their initiatives and guidance, many of my accomplishments would not have been possible.

I would like to thank my colleagues and all collaborators for their contributions throughout this journey.

I am deeply indebted to my friends for accompanying me and providing invaluable support along the way. Their presence has been invaluable in maintaining my equilibrium during challenging times.

Foremost, I am profoundly grateful to my beloved family: my mother, Filiz, and my brother, Mert, for their boundless love, encouragement, and unwavering support throughout this demanding journey.

Lastly, my deepest gratitude lies with my remarkable father, Fikret, whose wisdom and guidance continue to resonate despite his early departure from this world.

Es ist nicht genug zu wissen
- man muss auch anwenden.

Es ist nicht genug zu wollen
- man muss auch tun.

Johann Wolfgang von Goethe

TABLE OF CONTENTS

1 ABBREVIATIONS

AI	Artificial intelligence
AP	Alkaline phosphatase
ATCC	American Type Culture Collection
ATP	Adenosine triphosphate
AUC	Area under curve
BH	Benjamini-Hochberg
bp	Base pair(s)
CCL2	CC-chemokine ligand 2
CC	Creative Commons
CC-BY	Creative Commons Attribution
CC-BY-NC	Creative Commons Attribution Non-Commercial
CDK	Cyclin-dependent kinase
cDNA	Complementary DNA
C. elegans	Caenorhabditis elegans
ChIP-seq	Chromatin immunoprecipitation and DNA sequencing
CGP	Cancer Genome Project
CHR	Cell cycle genes homology region
CIN	Cervical intraepithelial neoplasia
CNA	Copy-number alteration
COTF	chimeric oncogenic transcription factor
CT	Computed tomography
DEG	Differentially expressed genes
DGEA	Differentially gene expression analysis
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside triphosphate
dTTP	Deoxythymidine triphosphate
dUTP	Deoxyuridine triphosphate
DSMZ	German Collection of Microorganisms and Cell lines
ECACC	European Collection of Authenticated Cell Cultures
ED	Eucladian distance
EDTA	Ethylenediaminetetraacetic acid
EMBL	European Molecular Biology Laboratory
EMT	Epithelial-to-mesenchymal transition
ER	Estrogen receptor
ESC	Embryonic stem cell(s)
ESMO	European Society for Medical Oncology
ESR	Erythrocyte sedimentation rate
EwS	Ewing sarcoma
FBS	FOXM1 binding site
FC	Fold change

FCS	Fetal calf serum
FDR	False discovery rate
FIGO	International Federation of Gynecology and Obstetrics
FHIT	fragile histidine triad
GEO	Gene expression omnibus
GO	Gene ontology
GSA	Global screening array
GSE	Genetic suppressor element
GSEA	Gene set enrichment analysis
GWAS	Genome-wide association study
HER	Human Epidermal Growth Factor Receptor
hg19	Human (reference) genome 19
HPV	Human papillomavirus
hr	high-risk
HSC	Hematopoietic stem cell(s)
ICGC	International Cancer Genome Consortium
ICLAC	International Cell Line Authentication Committee
IC50	Half maximal inhibitory concentration
IGF	Insulin-Like Growth Factor
IRS	Immune Reactive Score
IL-2	Interleukin 2
kb	Kilobase(s)
m	month
MRI	Magnet resonance imaging
mRNA	Messenger RNA
mSat	Microsatellite
MSC	Mesenchymal stem cell(s)
NA	Not available/applicable
NES	Normalized Enrichment Score
NGS	Next-generation sequencing
NIH	National Institutes of Health
QC	Quality control
Pap-smear	Papanicolaou test
PBS	Phosphate-buffered saline
PCA	Pirincipal component analysis
PCR	Polymerase chain reaction
PDX	Patient-derived xenograft
PET	Positron-emission tomography
pPNET	Peripheral primitive neuroectodermal tumor
PR	Progesterone receptor
qRT-PCR	Quantitative real-time polymerase chain reaction REB
RNA	Ribonucleic acid
RNA-Seq	RNA sequencing

rpm	Rotations per minute
RPMI (medium)	Roswell Park Memorial Institute (medium)
scRNA-seq	single-cell RNA sequencing
SEM	Standard error of the mean
shRNA	Short hairpin RNA
shCtr	shControl
SNA	Single nucleotide alteration
SNP	Single nucleotide polymorphism(s)
STR	Short tandem repeat
TARGET	Therapeutically Applicable Research to Generate Effective Treatments
TCGA	The Cancer Genome Atlas
TF	Transcription factor
TNF	Tumor necrosis factor
TNM	tumor-node-metastasis
TRIS	Tris(hydroxymethyl)aminomethane
TSO	Template switch oligo
VCF	variant call format
UV	Ultraviolet
WGS	Whole genome sequencing
WHO	World Health Organization
WT	wild type
2D	two-dimensional
3D	three-dimensional

2 ABSTRACT

2.1. German

Seit Jahrzehnten dienen humane Zelllinien als unverzichtbare Werkzeuge in der Krebsforschung (1). Obwohl diese Forschungsmodelle aufgrund ihrer einfachen Handhabung und Fähigkeit wertvolle Einblicke in die Krebsbiologie zu liefern, weithin anerkannt und bevorzugt sind, haben jüngste Studien Fragen nach ihrer Zuverlässigkeit aufgeworfen (2–5). Im Laufe des letzten Jahrzehnts haben Forscher wichtige Merkmale hinsichtlich der genomischen Stabilität von Krebszelllinien in künstlichen Kulturbedingungen und der Reproduzierbarkeit von Ergebnissen, die mit diesen Modellen erzielt wurden, identifiziert (2,3,6). Ende der 2010er Jahre lieferten zwei wegweisende Studien neue Einblicke in die genomische und phänotypische Instabilität der hochmutierten adulten Karzinomzelllinien HeLa und MCF-7 (2,3). Diese Studien enthüllten eine bemerkenswerte Vielfalt derselben Krebszelllinie in verschiedenen Laboren sowie die Instabilität der jeweiligen Zelllinie in langandauernder Zellkultur, die zur genomischen und transkriptomischen Heterogenität und daraus resultierenden phänotypischen Variationen führte (2,3). Diese Erkenntnisse unterstreichen die Einschränkungen der Reproduzierbarkeit, die mit der Verwendung von menschlichen Zelllinienmodellen in der Krebsforschung verbunden sind. Dennoch bleibt die Verallgemeinerbarkeit solcher Schlussfolgerungen auf andere Krebszelllinien weitgehend unerforscht. Diese Doktorarbeit hatte zum Ziel, das Maß an Variationen in Krebszelllinien mit einer geringeren Mutationslast (oligomutiert) zu untersuchen. Um dieses Ziel zu erreichen, wurde das Ewing-Sarkom (EwS), ein pädiatrischer Knochenkrebs, als repräsentatives Modell ausgewählt, da es für seine minimale somatische Mutation bekannt ist (7–9).

Das EwS wird durch eine chromosomale Umordnung verursacht, die zu einem chimären onkogenen Transkriptionsfaktor (COTF) führt, der eine entscheidende Rolle bei der Regulation der Transkription und Biologie des Tumors spielt (8,10). Die Hypothese lautete, dass EwS-Zelllinien eine größere genetische und phänotypische Stabilität als adulten Karzinomzelllinien aufweisen würden. Die Analyse der molekularen und phänotypischen Merkmale von 11 Stämmen der EwS-Zelllinie A-673 aus verschiedenen Laboren zeigte eine bemerkenswerte genomische und phänotypische Einheitlichkeit, im Gegensatz zu Kontrollgruppen von adulten Karzinom-HeLa- und MCF-7-Stämmen. Darüber hinaus wurden neu erworbene A-673-, HeLa- und MCF-7-Zelllinien einer 12-monatigen kontinuierlichen Zellkultur einer longitudinalen Analyse unterzogen. Bemerkenswerterweise zeigte die EwS-Zelllinie A-673 eine außergewöhnliche Stabilität in Bezug auf das Genom und Transkriptom sowie im Hinblick auf Arzneimittelempfindlichkeit im Vergleich zu Zelllinien von Karzinomen des Erwachsenenalters. Zusätzlich wurden vier weitere EwS-Zelllinien in die longitudinale Analyse aufgenommen, um die beobachtete Stabilität der A-673-Zelllinie zu validieren. Die Analyse ergab, dass zwar alle fünf EwS-Zelllinien eine höhere Stabilität als adulten Krebszelllinien des Erwachsenenalters aufwiesen, jedoch nach 12 Monaten kontinuierlicher Passage unterschiedliche Grade von (epi)genomischen, transkriptomischen und phänotypischen Veränderungen beobachtet wurden. Dies suggeriert, dass die Stabilität von Zelllinien ein Spektrum darstellt, selbst innerhalb derselben Krebsentität. Die bemerkenswerte Stabilität, die bei COTF-getriebenen pädiatrischen Sarkomzelllinien beobachtet wurde, unterstreicht ihr Potenzial als zuverlässige Modelle für die Krebsforschung und therapeutische Tests. Diese Entdeckung hat tiefgreifende Auswirkungen auf präklinische Studien und deutet darauf hin, dass solche Zelllinienmodelle konsistentere und reproduzierbarere Ergebnisse liefern könnten, was die Übertragung von Forschungsergebnissen in die klinische

Anwendung erleichtert. Darüber hinaus beleuchtet sie das Spektrum der Reproduzierbarkeit von wissenschaftlichen in vitro Ergebnissen und betont die Wichtigkeit einer sorgfältigen Berücksichtigung der Zelllinienecharakteristika bei der experimentellen Gestaltung und Ergebnisinterpretation, selbst innerhalb derselben Tumorentität.

Essenzielle Abschnitte dieser Dissertation wurden im Preprint aufgenommen:

Merve Kasan, Jana Siebenlist, Martin Sill, Rupert Öllinger, Enrique de Álava, Didier Surdez, Uta Dirksen, Ina Oehme, Katia Scotlandi, Olivier Delattre, Martina Müller-Nurasyid, Roland Rad, Konstantin Strauch, Thomas G. P. Grünewald, Florencia Cidre-Aranaz, Genomic and phenotypic stability of fusion-driven pediatric Ewing sarcoma cell lines, bioRxiv 2023.11.20.567802

2.2. English

Since decades, human cell lines serve as indispensable tools in cancer research (1). While these research models are widely recognized and favorable due to their relatively easy usage and capability to provide invaluable insight into cancer biology, recent studies raised questions about their reliability (2–5). Over the past decade, researchers have identified significant challenges with the genomic stability of cancer cell lines in unnatural culture environments and the reproducibility of results obtained using these models (2,3,6,11). In the late 2010s, two pioneering studies have shed light on the genomic and phenotypic instability of highly mutated adult carcinoma cell lines: HeLa and MCF-7 (2,3). These studies revealed a remarkable diversity of the respective same cancer cell line across different laboratories, along with the instability of a cell line in prolonged cell culture leading to genomic and transcriptomic heterogeneity and consequential phenotypic variations (2,3).

These findings underscored the limitations in reproducibility associated with using human cell line models in cancer research. Nevertheless, while the implications of these

observations on adult carcinoma cell lines are becoming increasingly apparent, the generalizability of such conclusions to other cancer cell lines remains largely unexplored.

This thesis aimed to investigate the level of variation in cancer cell lines with a lower mutational burden (oligomutated). To achieve this goal, Ewing sarcoma (EwS), a malignant pediatric bone cancer, was selected as a representative model due to its reputation for harboring minimal somatic mutation (7–9). EwS is caused by a chromosomal rearrangement that results in a chimeric oncogenic transcription factor (COTF), which plays a crucial role in regulating the transcription and biology of the tumor (8,10). The hypothesis was that EwS cell lines would exhibit greater genetic and phenotypic stability than adult carcinoma cell lines. Analyzing the molecular and phenotypic traits of 11 EwS cell line A-673 strains from various laboratories showed remarkable genomic and phenotypic uniformity, contrasting control groups of adult carcinoma HeLa and MCF-7 strains. Additionally, newly purchased A-673, HeLa and MCF-7 cell lines were subjected to a 12-months continuous cell culture for a longitudinal analysis. Notably, EwS cell line A-673 exhibited exceptional stability in terms of genomic and transcriptomic levels as well as drug sensitivity, when compared to adult carcinoma cell lines. Further, four additional EwS cell lines were included in the longitudinal analysis to investigate the observed stability in the A-673 cell line. The analysis revealed that although all five EwS cell lines showed higher stability than their adult carcinoma counterparts, varying degrees of (epi)genomic, transcriptomic, and phenotypic alterations were observed after 12 months of continuous passaging. This indicated that cell line stability is a spectrum even within the same cancer entity.

The observed remarkable stability of COTF-driven pediatric sarcoma cells underscores their potential as faithful models for cancer research and therapeutic drug testing. Since these cell line models could offer more consistent and reproducible outcomes, this

discovery may have profound implications for preclinical studies, which may aid in the translation of research findings into clinical applications. Moreover, this thesis sheds light on the spectrum of reproducibility in in vitro scientific results, emphasizing the importance of carefully considering cell line characteristics in experimental design and result interpretation, even within the same tumor entity.

Key sections of this thesis were included in the preprint:

Merve Kasan, Jana Siebenlist, Martin Sill, Rupert Öllinger, Enrique de Álava, Didier Surdez, Uta Dirksen, Ina Oehme, Katia Scotlandi, Olivier Delattre, Martina Müller-Nurasyid, Roland Rad, Konstantin Strauch, Thomas G. P. Grünewald, Florencia Cidre-Aranaz, Genomic and phenotypic stability of fusion-driven pediatric Ewing sarcoma cell lines, bioRxiv 2023.11.20.567802

3 INTRODUCTION

Cancer is one of the most complicated diseases that humankind has been facing for centuries (12,13). According to the World Health Organization (WHO), cancer ranks second in mortality rates following cardiovascular diseases, affecting individuals across all age groups (14–16). With the escalating number of cancer cases, unraveling the disease's biological mechanisms and identifying novel vulnerabilities for preventive measures have become imperative (13,14,17).

Historically, cancer posed an enigmatic puzzle for early physicians (18). However, recent technological advancements, including sophisticated research models and high-throughput sequencing, hold promise in improving the understanding of this disease (19–21). In this context, the stability and reliability of research models in cancer field are pivotal factors, influencing the accuracy and translatability of findings from bench to bedside (22,21,23,24). Addressing these aspects is fundamental for advancing the comprehension of cancer biology and ultimately developing more effective therapeutic strategies (25).

3.1. Essence of Cancer Research: Tumor Models

3.1.1 The need for cancer modelling

Understanding cancer, a complex phenomenon since the 18th century, requires reliable research models that accurately mimic human biology (26,21). Establishing such models is crucial for comprehensive examinations of cancer biology and its interconnected structures (13,27). Research models serve as fundamental tools in deciphering the mechanisms of tumor initiation, growth, and therapy response (11,21). They offer a controlled setting to explore genetic mutations, cellular interactions, and the tumor microenvironment—critical elements in cancer development (26). Tumor models,

whether cell cultures, animal models, or advanced technologies like organoids and patient-derived xenografts, are designed to mimic the multifaceted nature of human tumors (11,23,28). The comphrensive review by Thomas et al. on the history of cancer models condenses the cycle of modeling into five steps: a) system perception and question establishment; b) model design and setup, whereby the hypothesis could be tested with experimental observations; c) model testing, whose output will serve for the clinical data comparison and evaluation; d) result interpretation and consistency analysis between the result and the original system; and e) model optimization (28). Considering these elements, the first successful animal tumor model was reported in 1918 on rabbits exploring the potential link between coal tar exposure and carcinogenesis in mammalian skin cells (29). After that, numerous other animal tumor models emerged, propelling advancements in understanding cancer biology (30–32).

A pivotal development unfolded several decades later with a groundbreaking report by Gey et al. in 1952, which described establishing the first human-derived cell line derived from a 31-year-old African-American woman diagnosed with cervix adenocarcinoma (33,34). This monumental achievement signified a remarkable leap forward in cancer research, fundamentally transforming the landscape by enabling the study of human cancer cells within controlled laboratory settings (33,35). This cell line, named HeLa, after the patients' name, was derived from and expeditiously provided to cancer research laboratories all over the world, as well as to pharmaceutical companies (34–36). As a consequence, it became the most outstanding model in the community (37).

3.1.2 Cancer Models

3.1.2.1 in Vivo *Models*

In in vivo studies, which translates to ‘in the living body of a plant or animal’ in English, various biological mechanisms are investigated within whole living organisms (38). In

cancer research, in vivo studies are performed using diverse tumor models, encompassing not only those derived from animals but also human material such as xenografts (23,39). Animal tumor models involving species like nematode *C. elegans*, zebrafish, mice, and rabbits play a pivotal role in investigating genetics and cancer biology, as well as in preclinical research for new anti-tumor therapies (40). Important historical milestones underlining the evolution of these models include Yamagiwa et al.'s first animal tumor model, which marked the beginning of extensive research in this area (28,29). Additionally, research of Skipper et al. on chemotherapeutic agents in murine leukemia and Hart & Fidler's demonstrations on metastatic tumor growth mechanisms were highly influential (41–44). Moreover, the introduction of *C. elegans* as a model for cancer-specific gene mutations and the discussions on zebrafish as a cancer model by White et al. opened new avenues in cancer genomics (23,39,40,45–47).

Technological advancements in previous decades have significantly expanded the possibilities and broadened the horizons of genetic engineering in the realm of in vivo research models (48). Today, in addition to traditional syngeneic mouse models and murine tumor transfers to genetically matched immune-competent mice, various advanced models are available for the cancer research community (31,32,49). These include orthotopic tumor models, cell line-derived xenografts (CDX), patient-derived xenografts (PDX), and humanized mouse models, each offering unique advantages for different research needs (23,39,47). These models are crucial in advancing our understanding of cancer and developing novel therapeutic strategies (23,50).

3.1.2.2 in Vitro *Models*

The Latin phrase ‘in vitro’ translating to ’ outside the living body and in an artificial environment’ refers to studies involving isolated cells, tissues, microorganisms, or other biological molecules isolated from their original environment (51). In the early 1900s, a biologist-anatomist Ross Graham Harrison successfully observed a notable increase in

nerve fiber growth in artificial tissue cultures for several weeks (52). Harrison's innovative method preserved cells outside the body under controlled laboratory conditions, facilitating prolonged examination, commonly referred to as monolayer or 2D cell culture—a practice still one of the essential techniques in modern cancer research (53–55). Initially, mammalian cell culture relied on four primary techniques: organ culture, primary explant culture, organotypic culture and cell culture (**Fig. 1**) (56,57). Each method offered distinct advantages and paved the way for studying cell behavior, responses to stimuli, and exploring disease mechanisms outside the complex biological environment of living organisms (55,57).

Modern cancer researchers benefit from the groundwork laid by earlier scientists in advancing cell culture methods. Important advancements in cell culture techniques include demonstrating medium changes and sub-culture, developing trypsinization methods, and exploring endless cell passaging possibilities (57–63). Additionally, Gey et al. established the first (cancer) cell line, HeLa, and Moore et al. demonstrated tumor cell proliferation in suspension culture, among many others (33,63,64).

Shortly, two-dimensional cell culture became a routinely used model thanks to its affordable cost and ease of use (55).

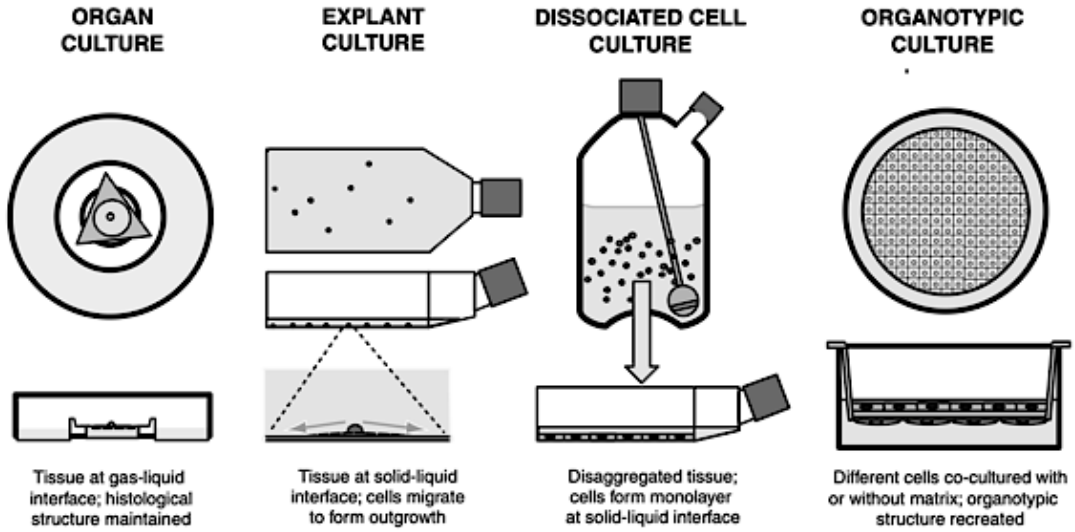


Figure 1:Different culture methods. The diagram illustrates various culture techniques from left to right. It depicts organ culture on a filter disk, explant cultures in a flask, a stirred vessel facilitating enzymatic disaggregation, and a filter well showcasing an array of cells. This depiction offers insight into diverse approaches for cell cultivation. Figure from the book section "Basic Principles of Cell Culture" by Freshney et *al*. 2006 (57) with permission under license number 5926630066999.

The tumor microenvironment's (TME) profound impact on tumorigenesis, metastasis, and drug sensitivity in malignant cells is well-established and recognized as a hallmark of cancer (65–68).

For the last few decades, three-dimensional (3D) tumor model technologies, such as spheroids, organoids, scaffolds, and hydrogels, have been providing a deeper insight into tumor biology by mimicing the TME (69,70). These advancements in 3D models provide a better understanding of tumor-stroma interactions, metastatic processes, and drug responses (24,71,72).

Cell lines, immortalized cell populations derived from cancerous tissues, and two-dimensional (2D) cell culture are fundamental models in cancer research, offering advantages in scalability, ease of manipulation, and cost-effectiveness (73,74). Furthermore, the extensive array of cancer cell lines derived from diverse cancer types is invaluable for examining tumor characteristics (73,74). Despite the more physiologically relevant nature of 3D models, 2D cell models and cancer cell lines remain indispensable in foundational cancer research and early-stage drug discovery (55).

3.2. Era of multiomic data

The six hallmarks of cancer, including maintaining proliferative signaling, avoiding growth suppression, refusing apoptosis, authorizing limitless replicative potential, stimulating invasion, and spread to other organs, and encouraging angiogenesis are

well-established (66,75,76). Cancer cells' uncontrollable proliferation and development ability result from a myriad of unique genetic abnormalities, which disrupt various cellular pathways, the complexity of tracking the consequences of these abnormalities adds another layer of challenge in cancer research, underscoring the need for sophisticated analytical tools and methodologies (66,77). Fortunately, the scientific and technological improvements over the last decades made the applicability of multiple high throughput sequencing possible in cancer research (78).

3.2.1 Rise of the sequencing technology

The introduction of the double-stranded DNA model by James Watson and Francis Crick, and the concept of genetic self-replication, marked a pivotal milestone in the dawn of a new era in molecular biology (79). This DNA model provided a framework for understanding biological inheritance and paved the way for many more scientific discoveries (79).

Nucleic acid sequencing refers to determining the order of nucleotides in DNA or RNA. Frederick Sanger's pioneering work in sequencing insulin's amino acids in 1955 led to a shift in focus towards these sequences, enabling the decoding of genetic information (80). Subsequently, Fiers et *al*. completed the sequencing of RNA in 1976, and Wu and Padmanabhan et al. developed early DNA sequencing strategies in the 1970s. These milestones marked significant progress in understanding the genetic code (81–85).

In the late 1970s, Fredrick Sanger and his team developed a technique called the ‘dideoxy’ chain-termination method, which is now known as the ‘Sanger Method’ (86,87). This method enabled rapid DNA sequencing, which was a significant breakthrough in the field of genetics. The first complete genome sequencing of an organism was carried out using the Sanger method, and the organism in question was the bacterium *Haemophilus influenzae* (88). Furthermore, innovations in labeling techniques,

enhanced polymerases, and capillary gel electrophoresis have pushed genome projects forward, such as the Human Genome Project (89–91). The Human Genome Project was an extensive international effort that aimed to unlock the mysteries of the human genome and involved support from scientists across various disciplines (89,92). By the early 2000s, several draft versions of the human genome were available, topping in the announcement of the completion of the draft phase in 2003 (93). Approximately 99% of the human reference genome was made available in its final form (94).

The advanced sequencing methods and devices marked the onset of high-throughput sequencing, also known as 'Next-generation Sequencing' (NGS) (95,96). The acceleration of genomic innovations over the last decade is a testament to the impact of the NGS era (97). Today, rapid whole-genome sequencing and resequencing, transcriptional and translational profiling, epigenome sequencing, and many other resourceful techniques are accessible, affordable, and widely employed by the scientific community (**Fig. 2**) (98). Sequencing technology's versatility and accessibility have transformed molecular biology, genetics, disease diagnosis, and targeted therapy development (80,91,97,99).

3.2.2 High throughput sequencing in cancer research

High-throughput sequencing has revolutionized cancer research in the 21st century. This technology allows for the quick and efficient decoding of large quantities of cancer DNA and RNA, enabling researchers to perform in-depth (epi)genomic and transcriptomic investigations (100). To gain a better understanding of cancer biology several large-scale cancer genome projects were initiated including The Cancer Genome Atlas (TCGA), Therapeutically Applicable Research to Generate Effective Treatments (TARGET) by the National Cancer Institute, Cancer Genome Project (CGP) by the Wellcome Trust Sanger Institute, and Pan-cancer Analysis of Whole Genomes

(PCAWG) by the International Cancer Genome Consortium (98,101–103) (**Fig. 2**). These initiatives delivered crucial insights, identifying genetic alterations driving diverse cancer types and uncovering pathways of tumor progression, revealing potential treatment targets (104). They drafted genomic landscapes across cancers, enabling classification into molecular subtypes and fostering tailored treatment approaches (105,106).

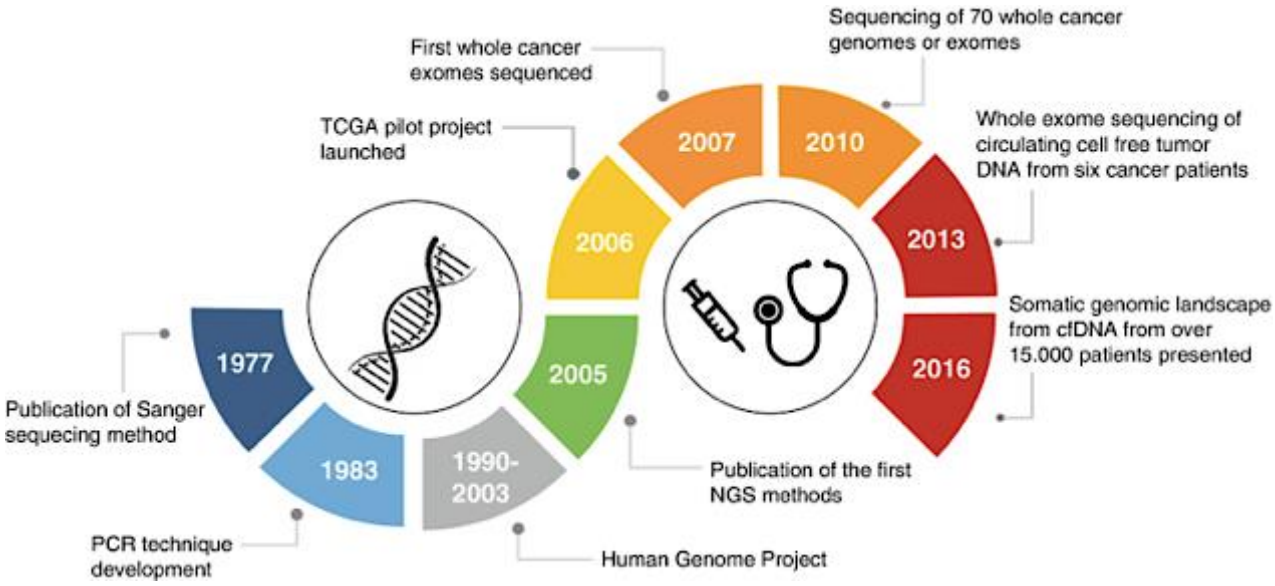


Figure 2: Key milestones in sequencing technologies. Figure from Morganti et al. (98) with permission under license number 5926970016917.

3.2.3 Genome sequencing

NGS technology encompasses two essential approaches: Exome Sequencing, and Whole Genome Sequencing (WGS), introduced in the early 21st century. Both methods have been proven invaluable in cancer research (107). On the one hand, Exome Sequencing is a cost-effective strategy that selectively targets the genome's protein-coding regions (108). This approach efficiently identifies mutations within cancer-associated genes, revealing potential drivers of the disease (107,108). On the other hand, WGS enables researchers to examine an organism's entire genome comprehensively, providing a broad view of genetic alterations and structural variations

that play a role in cancer (108,109). By analyzing different regions of a single tumor, researchers can gain insights into diverse genetic profiles and develop tailored treatments (110,111) (**Fig. 3**).

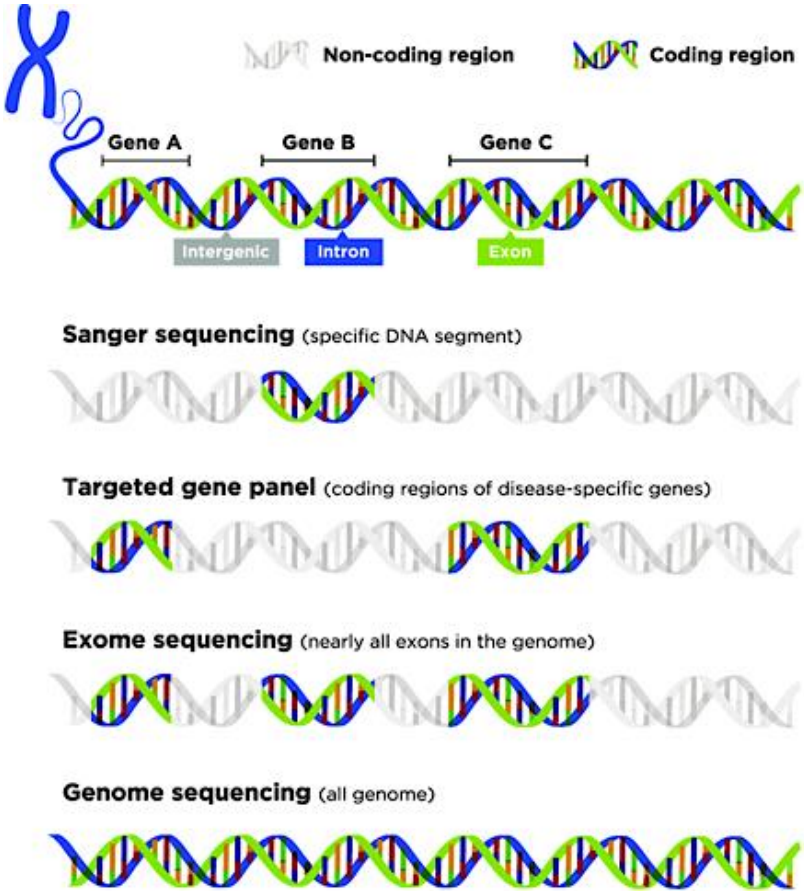


Figure 3: Various genetic testing methods aim at distinct genome segments. While Sanger sequencing focuses on a limited portion, targeted gene panels analyze coding regions of specific genes, whole exome sequencing captures nearly all coding sequences, and whole genome sequencing spans almost all regions of the genome. Figure from Devarajan et al. (111) published under CC BY-NC-ND license.

3.2.4 Transcriptome sequencing

The evolution of NGS brought a transformation in transcriptomic research, enabling precise quantification of gene expression and exploration of non-coding RNA (112). While traditional methods like Northern blotting and microarrays existed since the 1970s, their limitations, such as high background noise and limited dynamic range, sparked the need for innovation (101,113–115). The bulk RNA sequencing (RNA-seq) technology was introduced in 2008, inducing a dramatic shift in the field of transcriptomics

(116,117). RNA-seq offered the unprecedented ability to quantify gene expression levels with exceptional accuracy, sensitivity, and reproducibility (118).

Furthermore, it enabled the detection of alternative splicing, post-transcriptional modifications, and the identification of novel transcripts (119,120). As a result, RNA-seq became instrumental in studying various biological processes, including disease mechanisms and developmental biology (112).

In the 2010s, single-cell RNA-seq (scRNA-seq) technology was introduced (121), which allowed transcriptome profiling at the level of individual cells, shedding light on cell-to-cell variability and heterogeneity (117,122). This breakthrough technology provided unprecedented insights into the intricate tapestry of gene expression within complex biological systems (123).

3.2.5 Epigenome sequencing

NGS has expanded its applications beyond genomics and transcriptomics to encompass epigenetics, particularly DNA methylation and histone modifications, crucial for gene expression and cellular differentiation (124,125). Histone modifications refer to a range of chemical changes to histone proteins, which are important regulators of chromatin structure and gene expression (126).

These epigenetic marks, such as acetylation, methylation, phosphorylation, ubiquitination, and others, finely control the accessibility of DNA to transcriptional machinery, thus fine-tuning gene regulation (127,128). In cancer research, advanced sequencing techniques like Chromatin Immunoprecipitation followed by sequencing (ChIP-seq) have revolutionized the understanding of how histone modifications modulate gene expression in cancer cells (129–131). By systematically profiling histone modification patterns across cancer genomes with NGS, researchers can gain unprecedented insights into the epigenetic landscape, critical for identifying promising

therapeutic targets and furthering precision medicine approaches in cancer (132,133). DNA methylation, another key epigenetic process, involves the covalent addition of a methyl group to cytosine residues, mainly within CpG dinucleotides, particularly in gene promoter regions (134). This modification can suppress gene expression by interfering with transcription factor binding or recruiting chromatin-modifying enzymes (125). Furthermore, changes in DNA methylation patterns have been intricately linked to various diseases, including cancer and developmental disorders (135). Recent advances in DNA methylation analysis techniques, such as bisulfite sequencing, have revolutionized our understanding of the epigenetic landscape at single-nucleotide resolution (136,137). However, bisulfite sequencing has limitations, such as high cost, low throughput, and DNA quality requirements, leading to the development of alternative methods (137).

3.3. Questioning Reproducibility in Preclinical Research

To foster scientific progress, reliability and reproducibility of previous scientific results lay the groundwork for transformative discoveries (138,139). Unfortunately, the field of cancer research grapples with well-documented challenges in this regard, which may be a significant consequence considering the complexity of cancer. The Reproducibility Project: Cancer Biology, initiated in 2013, is a testament to these concerns, underscoring the need to assess the replicability of esteemed reports from 2010 to 2012 (4,22,140). In cancer research, in vitro cell culture is fundamental in advancing our understanding of cancer biology and drug development (1,73). Nevertheless, it is accompanied by the potential risk for cell line misidentification and/or contamination, genetic drift, clonal selection, and adaptation, as well as phenotypic alterations due to long-term exposure to an artificial culturing environment (2,3,141,142). These issues have gained substantial attention over the last decade due to their profound impact on

research accuracy and applicability (4). Cell line misidentification (i.e., labeling a cell line incorrectly), and contamination (i.e., inadvertent introduction of foreign cells or microorganisms) are the some of the most common and fundamental challenges of in vitro experiments (141–144). Several prominent examples underscore the gravity of these challenges. For instance, studies revealed the frequent misidentification of MCF-7 breast cancer cell line, mislabeled 37 out of 122 head and neck cancer cell lines, high authentication and cross-contamination issues in thyroid cancer cell lines, and cross-contamination of the widely employed HeLa cervix cancer cell line, all of which raised questions on the reliability and robustness of the findings derived from their use (34,144–148).

Addressing these issues requires rigorous quality control measures, including the authentication of cell lines through DNA profiling and regular screenings for contamination, as well as a notification system for the readers to evaluate the results with care (142,146,147). Initiatives like the International Cell Line Authentication Committee (ICLAC), European Collection of Authenticated Cell Cultures (ECACC), and the ATCC offer authenticated cell lines as reference materials, bolstering consistency across laboratories (142,144). Additionally, open-access databases and repositories containing comprehensive genomic and phenotypic data can facilitate cross-study comparisons, which may give the researchers insight into the reliability and reproducibility of cancer research (149,150). A substantial portion of the discrepancies observed in cancer cell line research can be ascribed to widespread mislabeling and contamination issues (142).

Additionally, cancer cells are renowned for their adaptability to shifting environmental conditions, which causes instability within continuous cell culture systems (2,151). Prolonged cell culture has been shown to induce genetic instability and clonal selection dynamics in cancer cells, potentially leading to inconsistencies between initial and

subsequent experimental outcomes (2,143). In 2018, Ben-David et al. examined 27 strains of MCF-7 breast carcinoma cell line, uncovering their substantial genetic and phenotypic heterogeneity (2). This diversity was evident not only at the single nucleotide level but also in shared copy number variations (CNV) across the MCF-7 strains (2). The observed genetic heterogeneity correlated with distinct gene expression patterns, morphological variations, and differences in drug sensitivity profiles among the strains. Specifically, when exposed to over 300 anti-cancer compounds, the 27 MCF7 strains displayed highly varied drug responses, with some exhibiting strong sensitivity to certain compounds while others showed resistance (2).

These findings were further validated in an additional 13 cell lines included in the study and emphasized the influence of specific culture conditions on the evolution of cell lines, driven by positive clonal selection (2).

In 2019, Liu et al. conducted a comprehensive cross-laboratory study on the HeLa cervix carcinoma cell line, revealing extensive diversity among different HeLa variants and those sourced from various laboratories (3). HeLa, being the first immortal cancer cell line, has played a significant role in biological research over the years (37). However, due to its extensive passaging and transfer between labs, multiple variants have emerged, including CLL2 (considered the original), HeLa S3 (also known as CLL2.2), and HeLa Kyoto, isolated from an early culture in Japan (3,37,152–155). Liu et al. gathered 14 HeLa cell lines from 13 laboratories and maintained them under uniform culture conditions (3). Analysis of their copy numbers, mRNA, protein levels, and turnover rates unveiled substantial genetic heterogeneity, particularly in the CLL2 and Kyoto variants (3). This heterogeneity was evident in variations observed through whole-genome sequencing (WGS) as well as in transcriptomic and phenotypic characteristics (3). Furthermore, their longitudinal analysis revealed the progressive evolution of the HeLa cell line, with notable differences observed between low-passage and high-

passage derivatives (3). These findings highlight the inherent challenges associated with utilizing established cell lines for cancer research and stress the importance of thorough characterization and validation of cell line models (3).

Furthermore, while considerable attention has been directed towards extensively mutated adult carcinoma cell lines such as HeLa and MCF-7, there has been relatively less focus on understanding the genetic stability and evolutionary trajectories of other cancer types (2,3,6,146).

Oligomutated pediatric sarcomas, characterized by a single COTF driving tumorigenesis, represent a distinct yet understudied subset of malignancies. One example is EwS, defined by a solitary genetic alteration, may be termed ‘oligomutated’, indicating a lower mutation burden compared to adult carcinoma cell lines (7–9,156). However, the degree to which these pediatric sarcomas undergo evolutionary changes and accumulate genetic alterations over time remains largely unexplored. A comprehensive understanding of the evolutionary dynamics of oligomutated pediatric sarcomas is imperative for elucidating the reproducibility and applicability of research findings within this specific context.

4 CANCER ENTITIES INCLUDED IN STUDY

4.1. Breast carcinoma

4.1.1 Clinical aspects

Breast carcinoma, a prevalent and life-threatening malignancy, predominantly affects individuals of the female gender and it is ranked as the most frequently diagnosed cancer among women globally (157–159). Age, with a peak incidence between 50 and 69 years, and hormonal factors, such as human epidermal growth factor receptor

(*HER2*), estrogen receptor (ER), and progesterone receptor (PR) status, play essential roles in breast carcinoma's epidemiology (159,160).

Breast carcinoma originates from epithelial cells in the breast ducts or lobules (161,162). Genetic and molecular alterations drive uncontrolled cell growth, resulting in tumor formation (160,161). Symptoms often include breast lumps, changes in size or shape, nipple abnormalities (retraction, inversion, discharge), and skin changes varying by subtype, size, stage, and patient characteristics (162).

The diagnosis of breast carcinoma typically involves clinical evaluation, precise imaging, and biopsy, procedures that provide crucial histopathological insights for confirming the diagnosis and ascertaining hormone receptor status (162–164). Imaging techniques such as ultrasound, magnetic resonance imaging (MRI), and positron emission tomography (PET) aid in tumor assessment (165). Staging, accomplished through the tumor-node-metastasis (TNM) system, guides therapeutic planning and prognosis assessment by stratifying breast carcinoma (166,167). Treatment is typically multimodal, determined by factors like subtype, biomarker & mutation status (Ki-67, *BRCA2*, *TP53*) tumor size, and metastasis, with primary interventions including tumor removal surgery, followed by adjuvant and chemotherapy (162,168).

4.1.2 Genetic and transcriptomic aspects

Breast carcinomas exhibit diverse clinical and histopathological characteristics, accompanied by distinct genomic and transcriptomic features (**Fig. 4**) (169–171). This heterogeneity extends beyond inter-patient variation to encompass intra-tumor heterogeneity (172,173). Notably, susceptibility genes have emerged over time. Inherited mutations in *BRCA1* and *BRCA2* have been linked to a significantly elevated risk of hereditary breast carcinoma; further mutations in *ATM*, *CHEK2*, *NBS1*, *PALB2*, *RAD50*, and *BRIP1* have shown a strong association, doubling the risk of this disease

(165,174–177). Significant associations have also been observed for *LZTR1*, *ATR*, and *BARD1* (178).

TP53 and *PIK3CA* represent frequently mutated genes in breast carcinoma, with mutations identified in roughly 30% of cases (162,177,179,180). *TP53* encodes a pivotal tumor suppressor protein that regulates the cell cycle and DNA repair (181). At the same time, *PIK3CA* plays a critical role in governing cell growth and survival through the p110α subunit of the phosphatidylinositol-3-kinase (PI3K) pathway (179,182). Among the notable chromosomal rearrangements in breast carcinoma, *ERBB2* (*HER2*) gene amplification or overexpression is a prominent example, occurring in 20-25% of breast cancer cases (183–185). This amplification leads to the overactivation of the *HER2* receptor, promoting cell proliferation and survival, and is associated with aggressive disease and therapy resistance (183,186).

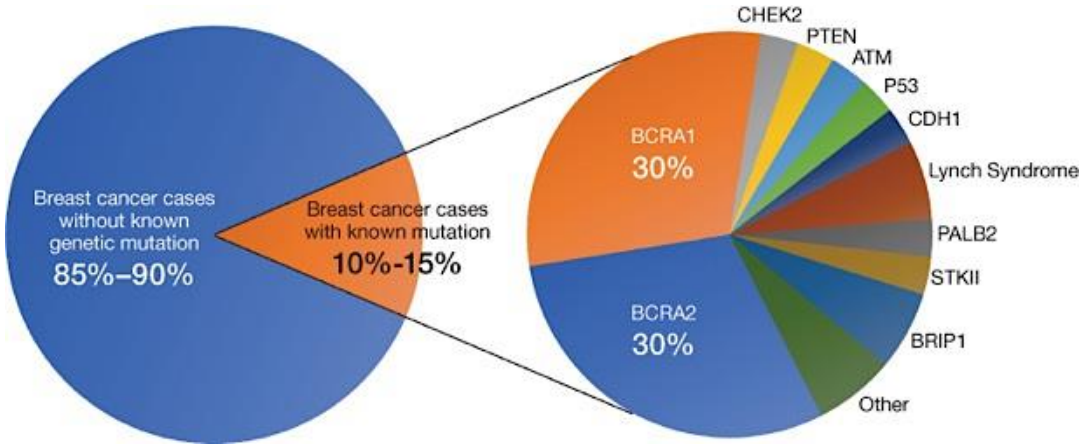


Figure 4: Diverse genetic mutations seen in breast carcinoma. Figure from Hasson, Menes, and Sonnenblick (171), published under CC BY-NC license.

Genomic instability in breast carcinoma is further substantiated by frequent copy number alterations (CNAs), such as *MYC* proto-oncogene amplification, observed in up to 15% of breast carcinomas (187).

Conversely, deletion of the *CDKN2A* gene, found in about 4% of breast tumors, results in the loss of tumor suppressor function, promoting uncontrolled cell division (188,189).

4.2. Cervix carcinoma

4.2.1 Clinical aspects

Cervical carcinoma, a highly prevalent malignancy worldwide, arises from the cellular transformation in cervix (190). Its incidence peaks between ages 30 and 50, primarily affecting adults (157,190–192). Persistent high-risk *human papillomavirus* (HPV) infection, notably HPV-16 and HPV-18, is one of the major risk factors in cervical carcinoma development (**Fig. 5**) (193–195). Socio-economic disparities, limited healthcare access, smoking, weakened immunity, and history of STIs further contribute to its widespread occurrence and global burden (196–198). Histopathologically, cervix carcinoma presents two primary subtypes based on the cell type of origin: squamous cell carcinoma and adenocarcinoma (190). In addition, rare subgroups include adenosquamous carcinoma, small cell carcinoma, lymphoma, and sarcomas (199).

Common clinical symptoms include abnormal vaginal bleeding, post-coital bleeding, and pelvic pain are the most common symptoms (196,200). Notably, its presentation varies based on tumor stage, subtype, and patient characteristics. In advanced stages, urinary and rectal symptoms due to local invasion and abdominal and back pain can occur (190). Considering its silent progression and unspecific symptoms, early detection through cervical screening (through the Papanicolaou (Pap) smear for cytology and high-risk (hr)-HPV testing) is crucially important (201–203).

Accurate diagnosis relies on clinical evaluation, imaging studies, and histopathological scrutiny through biopsy (colposcopy-guided cervical punch or cone biopsy) (190,196,199). Imaging techniques like MRI and CT scans help evaluate tumor size, extent, and lymph node involvement (190,199).

Staging, using the International Federation of Gynecology and Obstetrics (FIGO) system, guides therapy and prognosis assessment based on tumor size, lymph node involvement, and distant metastasis (190,200,204). Cervical carcinoma therapy is

multifaceted, considering factors like immunohistochemical subtyping, biomarkers (e.g., p16, HPV status), and tumor staging to tailor treatments (190,196,199,205). Surgery is often the primary intervention for tumor removal in early stages, while concurrent chemoradiotherapy is crucial for advanced cases (190,199,205). Targeted therapies such as bevacizumab as well as immunotherapies like pembrolizumab show promise in recurrent or metastatic disease (206,207).

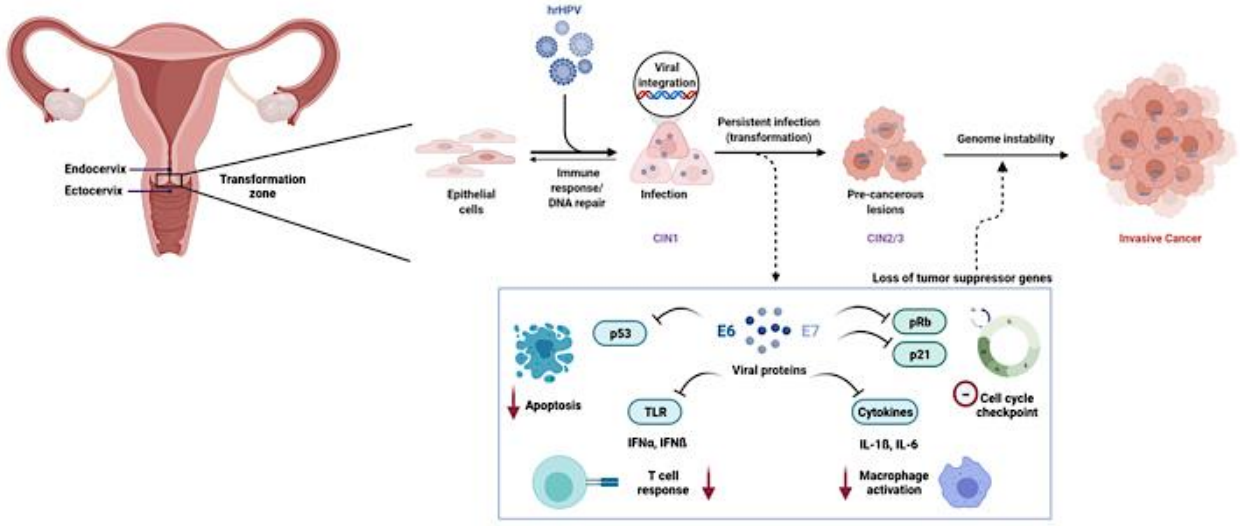


Figure 5: Development of cervical carcinoma. Epithelial cells in the cervix transformation zone develop lesions due to persistent high-risk HPV (hr-HPV) infection. Lesions can either resolve or progress from CIN1 to CIN2 and CIN3 upon viral integration. Viral proteins E6 and E7 hinder apoptosis via TP53, cell cycle regulation via p21, T-cell response via toll-like receptors (TLR), and macrophage activation via cytokines. This results in inadequate immune response, viral replication, uncontrolled cell proliferation, genome instability, leading to CIS or CC. Figure from review by Ramachandran and Dörk (195) published under CC BY license.

4.2.2 Genomic and Transcriptomic Aspects

Cervix carcinoma is characterized by a multifaceted genomic landscape with diverse genetic alterations (208). Notably, infection with hr-HPV types is a pivotal event in this malignancy's development, leading to the incorporation of viral DNA into the host genome and subsequent genomic instability (208–210). Apart from HPV, it has been thoroughly studied that several genetic alterations contribute to cervix carcinoma's pathogenesis. These include somatic mutations in tumor suppressor genes like *TP53* and those involved in the PI3K-AKT-mTOR pathway promoting cell survival and

proliferation as well as in the critical oncogenes from RAS/RAF/MAPK signaling, which leads to unrestricted cell growth (208,211).

Chromosomal rearrangements, such as the amplification of the long arm of chromosome 3 (3q), have been associated with cervix carcinoma (212). This region harbors several oncogenes, including *TP63* and *PIK3CA*, whose amplification results in increased expression, driving cellular proliferation and survival in cervical carcinoma (212,213). Moreover, recurring E322K substitutions within the *MAPK1* gene, alongside mutations in a subset of genes, specifically *HLA-B*, *EP300*, *FBXW7*, *NFE2L2*, *ERBB*, *CBFB*, and *ELF3* have been observed in this tumor entity (208). Copy number alterations are frequent genomic events in cervical carcinoma (208,213,214). For instance, loss of heterozygosity in the chromosomal region containing the fragile histidine triad (*FHIT*) gene on chromosome 3p is a recurring event (215). This alteration impacts DNA repair processes and contributes to genomic instability, further complicating the molecular landscape of cervical carcinoma (213,215).

4.3. Ewing sarcoma

4.3.1 Clinical aspects

Ewing sarcoma (EwS) is as the second most common malignant bone or soft tissue tumor in adolescents and young adults, peaking in incidence around the age of 15 and exhibiting a slight male predilection (9,216). It can manifest at various anatomical sites, predominantly within the pelvic region, femur, tibia, ribs, or, in the case of the soft-tissue variant, in the chest and pleural space, or musculature (9). The occurrence of EwS is attributed to genetic factors, as evidenced by familial clustering of EwS cases (217–219). Notably, a demographic pattern emerges, with higher incidence rates among Caucasians compared to Asian/Native American or African-American populations

(9,220). Morphologically, EwS shares characteristics with the small round blue cell tumor group, making diagnosis challenging (221). Despite rigorous genetic scrutiny, the precise cellular origin of EwS still needs to be discovered (7–9,156). Putative candidates include mesenchymal stem cells (MSCs) and neural crest-originated stem cells, both implicated in the potential genesis of EwS (9,222–224).

The most common symptom of this malignancy is localized, nonspecific pain, which can be misinterpreted as ‘growing pains’, potentially leading to delayed diagnosis (9,225). In localized stages, swelling in the affected area or a palpable tumor mass may occur (9,225). In more advanced or metastatic stages, additional symptoms, such as fever, night sweats, or significant weight loss (greater than 10% over the last six months), may manifest (9). The diagnosis of EwS relies on clinical evaluation, imaging of the primary lesion using an X-ray, MRI, and, if necessary, a CT scan, and molecular and histopathological assessment (226,227).

Staging EwS, dependent on tumor mass and metastasis status, significantly impacts treatment procedures. Currently, EwS is managed based on its stage with a combination of surgical resection, (neoadjuvant) chemotherapy, and radiotherapy (226,228).

4.3.2 Genomic and Transcriptomic Aspects

EwS is characterized by a unique genetic structure that revolves around the oncogenic fusion of the *EWSR1* gene with an ETS family transcription factor, predominantly *FLI1* (9). This fusion, often initiated by the chromosomal translocation t(11;22)(q24;q12), stands as the hallmark of EwS pathogenesis, giving rise to chimeric fusion proteins like *EWSR1::FLI1* or *EWSR1::ERG*, which function as aberrant transcription factors (9,222,229–231). Despite this characteristic fusion, EwS exhibits a genomically silent nature, with minimal inherited predisposition sites beyond the primary fusion event. Although, some studies have reported the recurrence of somatic mutations in genes like

STAG2 and *TP53*, which can impact disease prognosis and aggressiveness, the genetic landscape of EwS remains relatively simple, with secondary alterations being rare (7–9,156). When secondary alterations do arise in this oligo-mutated pediatric sarcoma, they typically involve genes associated with essential cellular processes like cell cycle regulation, DNA repair, or signaling pathways (7,8,156,222). Acting as the primary transcriptional regulator, the EWS*::*ETS fusion protein leverages GGAA-microsatellites to modulate chromatin structure, thereby inducing changes in gene expression (**Fig. 6**) (232,233). This highlights the pivotal role of the EWSR1*::*FLI1 fusion in molding the genomic architecture of EwS and propelling oncogenesis in this pediatric sarcoma, contributing to its genomically silent nature.

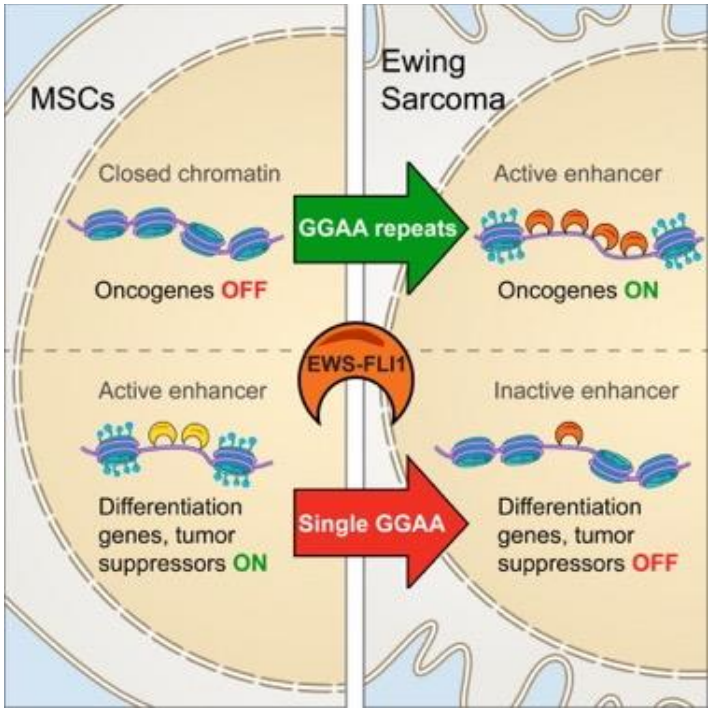


Figure 6: Regulatory role of EWSR1*::*FLI1. Functional role of EWSR1*::*FLI1. EWSR1*::*FLI1 has a dual regulatory function: it can enhance the expression of oncogenes by activating GGAA-microsatellites (mSats) and suppress the activity of conserved enhancers. MSC refers mesenchymal stem cell. Figure from Riggi *et al.* (232) with permission under license number 5926640930964.

5 STUDY CONCEPT, AIMS, AND OBJECTIVES

The primary objective of this project is to address the broader issue of reproducibility in research models, specifically human cell lines, and investigate their in vitro evolution and cross-laboratory diversity within pediatric oligomutated sarcomas. This investigation is motivated by the current knowledge gap, especially when compared to the vast amount of research on the effects of continuous cell culture on adult carcinoma cells. This research aims to provide insights into the reliability of cell line models in the context of pediatric oligomutated sarcomas, particularly those driven by COTF, presented by EwS, addressing a crucial aspect of research methodology and contributing to the broader goal of enhancing the reproducibility of findings in cell-based studies.

The overarching goal is to unveil the extent of cell line evolution in pediatric oligomutated sarcomas and assess the genetic and phenotypic stability in contrast to observations made in highly mutated adult carcinoma cell lines. This thesis aims to contribute critical insights into the reproducibility of scientific findings in diverse cancer cell lines, hypothesizing that COTF-driven pediatric sarcoma cell lines exhibit greater genetic and phenotypic stability than previously studied adult carcinoma cell lines. Through rigorous analysis and comparison, this study endeavors to advance the understanding of in vitro cancer cell evolution, potentially shedding light on the reliability and reproducibility of scientific results in the context of pediatric oligomutated sarcomas. This study includes the widely used EwS cell line A-673, the extensively studied adult breast carcinoma cell line MCF-7, and the seminal immortal cervix carcinoma cell line HeLa, renowned for their genetic instability in vitro.

Additionally other EwS cell lines MHH-ES-1, SK-ES-1, SK-N-MC, and TC-71 were included for a transversal and in-depth analysis of the in vitro cell line evolution within the same cancer entity.

AIMS AND OBJECTIVES

1- Assess Genomic Stability of Fusion-Driven EwS

- Investigate and compare the genomic evolution concerning single nucleotide polymorphism (SNP) alterations in the fusion-driven EwS cell line A-673 versus the previously studied adult carcinoma cell lines HeLa and MCF-7.
- Explore and compare genomic variability among strains from different laboratories.
- Analyze the longitudinal genomic evolution over time.

2- Analyze Transcriptomic Stability of Fusion-Driven EwS

- Explore and compare the transcriptomic stability and homogeneity in the fusion-driven EwS cell line (A-673) and the adult carcinoma cell lines.
- Investigate and compare transcriptomic homogeneity among strains across laboratories.
- Study the longitudinal transcriptomic evolution over time.

3- Evaluate Drug Response Uniformity of Fusion-Driven EwS

- Assess and compare the uniformity and stability in drug response in fusion-driven EwS A-673 cell lines versus adult carcinoma strains.
- Investigate and compare phenotypic stability among strains across laboratories.
- Examine the longitudinal phenotypic stability over time.

4- In-depth Analysis of Stability in the Same Sarcoma Entity

- Investigate the evolution of different cell lines within the same sarcoma entity.
- Explore and compare the degree of (epi)genomic, transcriptomic, and phenotypic evolution over time.

6 MATERIALS AND METHODOLOGY

6.1. MATERIALS

6.1.1 Cell lines

Table 1: List of cell lines

Article	Specification	Supplier
A-673	EwS, wild type (WT)	ATCC, Manassas, VA, USA
A-673	EwS, WT, strain	Delattre, O., Paris, France
A-673	EwS, WT, strain	Dirksen, U., Essen, Germany
A-673	EwS, WT, strain	De Álava, E., Seville, Spain
A-673	EwS, WT, strain	Grünewald, T., Munich, Germany
A-673	EwS, WT, strain	Kovar, H., Vienna, Austria
A-673	EwS, WT, strain	Scotlandi, K., Bologna, Italy
A-673	EwS shcontrol (234)	Grünewald, T., Munich, Germany
A-673	EwS TR/shEF1(234)	Grünewald, T., Munich, Germany
HeLa	cervix carcinoma, WT	DSMZ, Braunschweig, Germany
HeLa	cervix carcinoma, WT, strain	Öhme, I., Heidelberg, Germany
HeLa	cervix carcinoma, WT, strain	Jung, A., Munich, Germany
MCF-7	breast carcinoma, WT	DSMZ, Braunschweig, Germany
MCF-7	breast carcinoma, WT, strain	Jung, A., Munich, Germany
MCF-7	breast carcinoma, WT, strain	E. Butt, Würzburg, Germany
MHH-ES-1	EwS, WT	DSMZ, Braunschweig, Germany
TC-71	EwS, WT	DSMZ, Braunschweig, Germany
SK-ES-1	EwS, WT	DSMZ, Braunschweig, Germany
SK-N-MC	EwS, WT	DSMZ, Braunschweig, Germany

6.1.2 Chemicals and reagents

Table 2: List of chemicals and reagents

Chemical / reagent	Specification	Manufacturer
Agarose	1kg	Carl Roth
Aqua bidestillata	NA	H. Kerndl
Crystal violet	NA	Sigma-Aldrich
CutSmart buffer	10x	New England Bio Labs
Dimethyl sulfoxide (DMSO)	Sterile-filtered	Sigma-Aldrich

DNA ladders	GeneRuler 1kb Plus GeneRuler 100bp Plus	Thermo Fisher Scientific
dNTPs	10mM	Sigma-Aldrich
Ethanol	≥99.8%, denatured	Carl Roth
Fetal calf serum (FCS)	NA	Sigma-Aldrich
Formaldehyde	Pierce, 16% methanol-free	Thermo Fisher Scientific
Hydrochloric acid (HCl)	0.1N solution, endotoxin-free	Merck
Hematoxylin counterstain	Based on Gill‘s formulation	Vector Laboratories
NU6140	50mg 10mg 5mg	Biotechne Tocris Merck
Nuclease-free H2O	NA	Carl Roth
Penicillin/Streptomycin	Penicillin: 10,000U/ml; Streptomycin: 10,000µg/ml	Biochrom
Phosphate-buffered saline (PBS)	Dulbecco, 500ml	Biochrom
Plasmocure	100mg/ml	Invivogen
Radioimmuno-precipitation assay (RIPA) buffer	Pierce	Thermo Fisher Scientific
RNAse A	Invitrogen PureLink RNase A 20mg/ml	Thermo Fisher Scientific
Roswell Park Memorial Institute (RPMI) 1640 medium	Supplemented with stable glutamine	Biochrom
SYBR safe	400µl solution	Thermo Fisher Scientific
SYBR Select Mastermix	Contains SYBR GreenER dye, AmpliTaq DNA polymerase UP, dNTPs with dUTP/dTTP-mixture, heatlabile UDG and optimized buffer components	Applied Biosystems
Trypan blue	0.4%	Sigma-Aldrich, Thermo Fisher Scientific
Trypsin	(10x) Trypsin (1:250)/EDTA-Solution (0,5%/0,2 %)	Biochrom

6.1.3 Compounds

Table 3: List of compounds

Compound	Mechanism of action	Manufacturer
BI 2536	PLK1 inhibition (mitotic block, apoptosis)	Selleckchem
CCT245737	CHK1 Inhibition	Selleckchem
Doxorubicin	DNA Topoisomerase II Inhibition	Selleckchem
Etoposide	DNA Topoisomerase II Inhibition	Selleckchem
JQ1	BET bromodomain protein inhibition	Adooq Bioscience
OSI-906 Linsitinib)	IGF-1R Inhibition	Adooq Bioscience
NU6140	CDK2 inhibition	Merck/Millipore
Rapamycin	Specific mTOR inhibition	Selleckchem
Vincristine sulfate	Interaction with tubulin resulting in mitosis inhibition at metaphase	Selleckchem
17-AGG (Tanespimycin)	Heat shock protein 90 (HSP90) inhibition	Adooq Bioscience
Adavosertib MK-1775	Selective inhibition of WEE1 kinase	Selleckchem
CFI-402257	Selective inhibition of Mps1/TTK	MedChemExpress
Clofarabine	Inhibition of ribonucleotide reductase	Selleckchem
Elesclomol	Induction of ROS-mediated apoptosis	Selleckchem
Gemcitabine	Prevention of chain elongation resulting in DNA synthesis inhibition	Selleckchem
Mercaptopurine	Purine metabolism inhibition through interference with nucleic acid synthesis	Selleckchem
Metformin	AMPK pathway activation & mTOR signaling inhibition	Selleckchem
Mithramycin	DNA/RNA polymerase inhibitor, DNA binding transcriptional inhibitor, TNF-α and Fas ligand-induced apoptosis mediator	Cayman chemical
Olaparib	PARP enzymes (PARP1 and PARP2) inhibition	Selleckchem
Triapine	Ribonucleotide reductase inhibition	Selleckchem

6.1.4 Commercial kits

Table 4: List of commercial kits

Usage	Kit	Components	Manufacturer

Genomic DNA extraction	NucleoSpin	Lysis Buffer T1, Lysis Buffer B3, Wash Buffer BW, Wash Buffer B5 (concentrate), Elution Buffer BE, Proteinase K (lyophilized), Proteinase Buffer PB, NucleoSpin Tissue Columns (light green rings), Collection Tubes (2ml)	Machery-Nagel
RNA extraction	NucleoSpin RNA	Lysis Buffer RA1, Wash Buffer RAW2, Wash Buffer RA3 (concentrate), Membrane Desalting Buffer MDB, Reaction Buffer for rDNase, rDNase (RNase-free, lyophilized), RNase-free H2O, NucleoSpin Filters (violet rings), NucleoSpin RNA Columns (light blue rings - plus Collection Tubes), Collection Tubes (2ml), Collection Tubes (1.5ml)	Machery-Nagel

6.1.5 Consumables

Table 5: List of consumables

Consumable	Specification	Manufacturer
Cell culture flasks	T150 (150cm2), T75 (75cm2), T25 (25cm2) T175 (175cm2), T75 (75cm2), T25 (25cm2)	TPP Corning
Cell culture plates	96-well, 12-well, 6-well	TPP
Filters	Rotilabo, sterile, 0.45 µm pore size	Carl Roth
Freezing containers	Mr. Frosty	Thermo Fisher Scientific
Gloves	Kimtech, Purple Nitrile Nitril NextGen	Kimberly-Clark Meditrade
Hemocytometers (single use)	C-Chip Neubauer Improved	NanoEnTek
Laboratory film	Parafilm	Bemis
Microscope slides	TOMO adhesive glass slide	Matsunami Glass
PCR tube strips	4titude	Brooks Life Sciences
Petri dishes	10cm GBO 10cm Nunclon	Greiner Thermo Fisher Scientific
Pipet tips	1250µl, 200µl, 100µl, 20µl, 10µl SurPhob SafeSeal	Biozym
qRT-PCR plate seals	4ti-0560	Brooks Life Sciences
qRT-PCR plates	Framestar, 96-well, semiskirted	Brooks Life Sciences

Reaction tubes	50ml, 15ml 2ml, 1.5ml TPX Polymethylpentene (PMP) tubes for DNA sonication	Greiner, Falcon Diagenode

6.1.6 Technical equipment and instruments

Table 6: Technical equipment and instruments

Device/Equipment	Model specification	Manufacturer
Aspirators	Vacusafe FTA-1	Integra Biosan
Automatic ice maker	SPR-80	Nordcap
Cell counter	Countess II	Invitrogen
Centrifuges	Heraeus Megafuge 40R	Thermo Fisher Scientific
	Heraeus Megafuge 8R	Thermo Fisher Scientific
	5415R	Eppendorf
	Universal 320	Hettich
	Rotina 320R	Hettich
	4K15C	Sigma
Electrophoresis gel chambers	Sub-cell GT	Bio-Rad
	40-0911, 40-1410, 40-0708	Peqlab
Electrophoresis gel imager	Multiimage Light Cabinet	Alpha Innotech
Electrophoresis power suppliers	PowerPac 300 Model 200	Bio-Rad Bio-Rad
Flasks and bottles	Erlenmeyer flask (Duran 500ml) Laboratory flask (Duran 1000ml, 500ml, 250ml, 100ml)	Schott
Photo (plate) scanner	Epson Perfection V370 Photo	Epson
Fridges and freezers	4°C, -20°C	Bosch, Siemens
	-80°C	Thermo Fisher Scientific
Hemocytometers	Neubauer Improved	Hartenstein
Incubators	HERAcell 240i, Forma 3111	Thermo Fisher Scientific
	CB-170	Binder
Laminar flow cabinets	Safe 2020, Maxisafe 2020, Herasafe	Thermo Fisher Scientific
Liquid nitrogen tank	Arpege 70	Air Liquide Medical
Manual counter	Analog	Hartenstein
Microscopes	Axiovert 200	Zeiss
	Axiovert 25	Zeiss
	Axioplan 2 imaging	Zeiss
	Primovert	Zeiss

Multistep pipet	Handy Step	Brand
PCR cyclers	T100 Thermal Cycler Mastercycler pro	Bio-Rad Eppendorf
Pipets	Pipetman 1000µl, 200µl, 100µl, 20µl, 10µl, 2µl	Gilson
Pipetting assistants	Pipetboy 2 Accu-jet pro	Integra Brand
Plate readers	Orion II microplate luminometer Varioskan	Titertek-Berthold Thermo Fisher Scientific
qRT-PCR cycler	CFX Connect	Bio-Rad
Racks	For 15ml and 50ml tubes, For 1.5ml and 2ml tubes For cryotubes, For PCR tubes	Hartenstein
Scissors	Surgical, 160mm	Hartenstein
Sequencing system		Illumina
Spectrophotometers	DS-11 Nanodrop ND-1000	DeNovix Peqlab
Table centrifuges	PerfectSpin mini Sprout LSE Spectrafuge 3-180 Qualitron DW-41	Peqlab Biozym Corning Neolab Thermo Fisher Scientific
Thermoblocks and Thermoshakers	Thermomixer comfort, Thermomixer compact, ThermoStat plus TS-100	Eppendorf Biosan
Vortexers	Vortex Genie 2 LSE 7-2020	Scientific_Industries Corning Neolab

6.2. METHODOLOGY

6.2.1 Provenience of cell lines and cell culture conditions

For long-term culture assays the following early passage (<5 passages) human cancer cell lines were acquired: the cervix carcinoma cell line HeLa, the human breast carcinoma cell line MCF-7, the EwS MHH-ES-1, SK-ES-1, SK-N-MC, and TC-71 cell lines were purchased from the German Collection of Microorganism and Cell Cultures

(DSMZ). The A-673 EwS cell line was purchased from the American Type Culture Collection (ATCC). For cross-laboratory assays the various wild type strains of A-673, HeLa and MCF-7 cell lines with an undefined number of passages were assembled: A-673 wild type strains were collected from O. Delattre from INSERM U830 'Genetics and Biology of Cancers' Institut Curie Research Center, Paris France, U. Dirksen from International Ewing Sarcoma Research Group, Westdeutsches Tumorzentrum Essen (WTZ), Universität Duisburg-Essen (Essen, Germany), E. de Álava from Molecular Pathology of Sarcomas Laboratory, Hospital Universitario Virgen del Rocio (Sevilla Spain), K. Scotlandi from Laboratory of Experimental Oncology, Rizzoli Orthopedic Institute, University of Bologna (Bologna, Italy), and H. Kovar from Molecular Biology of Solid Tumors, St. Anna Kinderkrebsforschung, Children´s Cancer Research Institute (Vienna, Austria). Wild type MCF7 strains were kindly provided by E. Butt from University Clinic of Würzburg, Germany. HeLa strains were gifted by A. Jung from Pathology Institute of LMU (Munich, Germany), and I. Öhme from German Cancer Research Center (DKFZ, Germany). Single cell clones derived from A-673 cell lines with either a neutral manipulation (A-673/TR/shcontrol) or an inducible shRNA construct against its *EWSR1::FLI1* fusion transcript (A-673/TR/shEF1) were previously described by the Grünewald laboratory (234) in Munich, Germany.

Cultured at 37°C with 5% CO_2, the cell lines were maintained in RPMI 1640 supplemented with 10% fetal bovine serum and 1% penicillin-streptomycin. Daily monitoring and biweekly passaging using Trypsin-EDTA were conducted when cells reached approximately 70% confluency. All cell lines underwent routine mycoplasma contamination testing by nested PCR, and their purity and authenticity were confirmed by short tandem repeat (STR)-profiling. All primers for mycoplasma purity testing were retrieved from Eurofins Genomics, Ebersberg, Germany (**Table 1**). And mycoplasma

test was performed as according to the laboratory protocol designed by Dr. Dr. Martin F. Orth (235).

Table 7: Primers used for Mycoplasma purity testing of the cell lines.

Name/target	Sequence	Purpose
Myco-F1	5'-ACACCATGGGAGCTGGTAAT-3'	Mycoplasma PCR
Myco-F1t	5'-ACACCATGGGAGTTGGTAAT-3'	Mycoplasma PCR
Myco-F2	5'-GTTCTTTGAAAACTGAAT-3'	Mycoplasma PCR
Myco-F2a	5'-ATTCTTTGAAAACTGAAT-3'	Mycoplasma PCR
Myco-F2cc	5'-GCTCTTTCAAAACTGAAT-3'	Mycoplasma PCR
Myco-R1	5'-CTTCATCGACTTTCAGACCCAAGGCAT-3'	Mycoplasma PCR
Myco-R1ac	5'-CTTCATCGACTTCCAGACCCAAGGCAT-3'	Mycoplasma PCR
Myco-R1cat	5'-CCTCATCGACTTTCAGACCCAAGGCAT-3'	Mycoplasma PCR
Myco-R1tt	5'-CTTCTTCGACTTTCAGACCCAAGGCAT-3'	Mycoplasma PCR
Myco-R2	5'-GCATCCACCAAAAACTCT-3'	Mycoplasma PCR
Myco-R2at	5'-GCATCCACCAAATACTCT-3'	Mycoplasma PCR
Myco-R2ca	5'-GCATCCACCACAAACTCT-3'	Mycoplasma PCR

Table 8: STR profiling of cell lines included in cross-laboratory analysis.

		D5S818	D13S317	D7S820	D16S539	vWA	TH01	AMG	TPOX	CSF1PO
A-673_1 (m0)	A1	11	8	10	11	15	9,3	X	8	11
	A2	12	13	12		18				12
A-673_2 (m6)	A1	11	8	10	11	15	?	X	8	11
	A2		13	12		18				
A-673_3 (m12)	A1	11	8	10	11	15	9,3	X	8	11
	A2		13	12		18				
A-673_4	A1	11	8	11	7	15	10	?	9	8
	A2	12	13	13	11	18				10
A-673_5	A1	11	8	6	11	15	9,3	X	8	6
	A2			10		18				11
	A1	11	8	6	11	15	9,3	X	8	11

A-673_6	A2			10		18				12
	A1	11	8	6	11	15	9,3	X	8	6
A-673_7	A2	12	13	10		18				11
	A1	11	8	6	11	15	9,3	X	8	?
A-673_8	A2		13	10		18				11
	A1	11	8	6	11	15	9,3	X	8	6
A-673_9	A2			10		18				11
	A1	11	8	6	11	15	9,3	X	8	11
A-673_10	A2	12		10		18				12
	A1	11	8	6	11	15	9,3	X	8	11
A-673_11	A2	12		10		18				12
	A1	11	12	6	9	16	7	X	8	9
HeLa_1	A2	12	14	8	10	18			12	10
	A1	11	12	6	9	16	7	X	8	9
HeLa_2	A2	12	14	8	?	18				
	A1	11	12	6	9	16	7	X	8	9
HeLa_3	A2	12	14	8	?	18				
	A1	11	12	6	9	16	7	X	8	9
Hela_4	A2	12	14	8	10	18			12	10
	A1	11	12	6	9	16	7	X	8	9
Hela_5	A2	12	14	8	10	18			12	10
	A1	11	11	6	11	14	6	X	9	10
MCF_1	A2	12		8	12	15			12	
	A1	11	11	6	11	14	6	X	9	9,3
MCF_2	A2	12		8	12	15			12	
	A1	11	11	6	11	14	6	X	9	9,3
MCF_3	A2	12		8	12	15			12	
	A1	11	11	6	11	14	6	X	9	9,3
MCF_4	A2	12		8	12	15			12	
	A1	11	11	6	11	14	6	X	9	?
MCF_5	A2	12		8	12	15			12	

Table 9: STR profiling of additional EwS cell lines.

		D5S818	**D13S317**	**D7S820**	**D16S539**	**vWA**	**TH01**	**AMG**	**TPOX**	**CSF1PO**
MHH-ES-1_1	A1	13	8	6	11	16	8	X	8	11
	A2			9		17	9			
MHH-ES-1_2	A1	13	8	6	11	16	8	X	8	11
	A2		13	9		17	9			
MHH-ES-1_3	A1	13	8	6	11	16	8	X	8	11
	A2			9		17	9			
SK-ES_1	A1	12	8	10	11	14	6	X	8	11
	A2		9	11		17	9,3	Y		
SK-ES_2	A1	12	8	10	11	14	6	X	5	11
	A2		9	11		17	9,3	Y	8	
	A1	12	8	10	11	14	6	X	8	11

SK-ES_3	A2		9	11		17	9,3	Y		
SK-N-MC_1	A1	11	11	8	12	17	9,3	X	9	9,3
	A2					18			11	
SK-N-MC_2	A1	11	11	6	12	17	10	X	9	10
	A2			8		18			11	
SK-N-MC_3	A1	11	11	8	12	17	9,3	X	9	10
	A2					18			11	
TC-71_1	A1	9,3	11	10	11	17	9,3	X	8	?
	A2		12		11				9	9,3
TC-71_2	A1	9,3	11	10	11	17	?	X	5	10
	A2		12		14			Y	8	11
TC-71_3	A1	9,3	11	10	11	17	9,3	X	8	10
	A2				14			Y	9	11

6.2.2 DNA extraction, methylation, and global screening arrays

When the cell culture flasks reached about 70% confluency, cell lysates were prepared, and total DNA extraction was carried out using the NucleoSpin Tissue kit from Macherey-Nagel. This kit utilizes spin-column-based technology for efficient purification of DNA from various tissue samples, following a standard protocol provided by the manufacturer. For each sample, 900 ng of extracted DNA was used for subsequent analysis. Genomic DNA was profiled using the Illumina Infinium Global Screening array, a widely used genotyping platform that utilizes bead array technology to assay genetic variations across the genome.

Additionally, methylation analysis was performed using the Illumina MethylationEPIC array, specifically designed to interrogate DNA methylation patterns at high resolution. Both genotyping and methylation profiling were conducted at the Molecular Epidemiology Unit of the German Research Center for Environmental Health (Helmholtz Center, Munich, Germany).

6.2.3 DNA methylation data analysis

The initial pre-processing of the raw methylation was performed in R version 3.3.1 by Martin Sill (236). Raw signal intensities were obtained from IDAT-files using the minfi

Bioconductor package version 1.21.4 in R version 3.3.1 (236,237). Each sample underwent individual normalization, involving background correction (shifting of the 5% percentile of negative control probe intensities to 0) and dye-bias correction (scaling the mean of normalization control probe intensities to 10,000) for both color channels. The methylated and unmethylated signals were corrected individually. Further methylation data management was kindly performed by Jana Siebenlist. Beta values were calculated from the retransformed intensities using an offset of 100 (as recommended by Illumina). Out of 865,859 probes on the EPIC array, 105,454 probes were masked according to Zhou *et al.* (238) as well as 16,944 probes on the X and Y chromosomes. In total, 743,461 probes were kept for downstream analysis. The beta values were transformed to M-values with the logit2 function of the *minfi* package version 1.42.0, R version 4.2.0 (236,237). A probe-wise differential methylation analysis (239) was performed using the *limma* package version 3.52.4 in R version 4.2.0 by comparing six and twelve months of culturing with the initial time point (m0) as reference (236,240). Significant differentially methylated CpG probes were extracted with the *decideTests* function of the *limma* package with an FDR<0.05 (Benjamini-Hochberg) (240).

All significantly differentially methylated (total hypo- and hyper-methylated) CpG sites were visualized using PRISM 9 (GraphPad Software Inc. CA, USA) (241).

6.2.4 Global screening array (GSA) data analysis

The initial processing and quality control (QC) of the raw genotyping data was performed using PLINK version 1.9 (SNP call rate >95%, Hardy-Weinberg exact test $<1e^{-6}$, and variants on Y chromosome were excluded) (242). In total 526,610 variants out of 696,726 passed the QC filters. Infinium GSA v3.0 annotation file was used to filter for in-exon or non-synonymous variants. To determine single nucleotide alterations (SNA) after six and 12 months in cell culture (m6 and m12), each cell line was compared to its

m0 version (number of consistent alleles and changes from homozygous to heterozygous) using the Variant Call Format (VCF) file generated by PLINK 1.9. Further data analysis was performed in R version 4.2.1, using the vcfR package (243) version 1.14.0 for parsing and analysing variant call format (VCF) files generated from NGS data., among other data processing packages described below. The distance between two time points for each cell line was computed in R version 4.2.1 using the proxy package version 0.4-27 (244). The eigenvectors, generated for dimension reduction in PLINK version 1.9, were used as input (242). The heatmap was generated in R version 4.2.1 using pheatmap package version 1.0.12 (236,245).

6.2.5 RNA extraction, RNA sequencing and analysis

RNA extraction was conducted when cell culture flasks reached approximately 70% confluency, utilizing the NucleoSpin RNA kit from Macherey-Nagel, Germany, in accordance with the manufacturer's instructions. Following extraction, RNA quality was assessed using a Nanodrop Spectrophotometer ND-1000 (Thermo Fisher) and quantified on a Qubit instrument (Life Technologies). Subsequently, for each sample, 50–100 ng of RNA in three biological and two technical replicates were utilized as input material and profiled on an Illumina NextSeq 500 system at the Institute of Molecular Oncology and Functional Genomics in Rechts der Isar University Hospital (TranslaTUM Cancer Center, Munich, Germany). For library preparation for bulk 3'-sequencing of poly(A)-RNA, a protocol described by Parekh *et al.* was followed (246). Briefly, barcoded cDNA of each sample was generated with a Maxima RT polymerase (Thermo Fisher) using oligo-dT primer containing barcodes, unique molecular identifiers (UMIs) and an adapter. 5' ends of the cDNAs were extended by a template switch oligo (TSO) and after pooling of all samples full-length cDNA was amplified with primers binding to the TSO-site and the adapter. cDNA was fragmented and TruSeq-Adapters ligated with the

NEBNext® Ultra™ II FS DNA Library Prep Kit for Illumina® (NEB) and 3'-end-fragments were finally amplified using primers with Illumina P5 and P7 overhangs. P5 and P7 sites were exchanged to allow sequencing of the cDNA in read1 and barcodes and UMIs in read2 to achieve better cluster recognition. The library was sequenced with 75 cycles for the cDNA in read1 and 16 cycles for the barcodes and UMIs in read2. Initial data processing was conducted using the published Drop-seq pipeline (v1.0) to generate sample- and gene-wise UMI tables by Dr. Rupert Öllinger (247). After eliminating the transcripts with very low counts (sums of all samples <10), RNA-seq data in count matrix format was batch corrected using ComBat-Seq function of R package sva version 3.44.0 (248), and differential gene expression analysis (DGEA) was performed using DESeq2 version 1.36.0 on R version 4.2.1 (236,249). Combat-Seq adjusted data was used as count input for DESeqDataSet. To prevent potential false discoveries resulting from the detection of minimally expressed genes, the 40% lowest expressed genes across samples were excluded (remaining expressed genes *N*=10,257).

For the examination of long-term cultured EwS cell lines a DGEA was performed using the top 60% expressed genes included in the raw count matrix (*N*=27,143, all EwS cell line samples were analyzed in one batch). For DGEA between two samples, genes with an adjusted *p* value less than 0.01 and an absolute log2 fold change greater than one were classified as DEG. Principal component analysis was employed to keep the overarching characteristics of the dataset with the *plotPCA* function. To comprehensively display the degree of variability between strains in each tumor type, gene-specific CV of the transcriptomic data was calculated. In the long-term culture experiments, the log2 fold change (log2FC) of gene expression for each cell line at 6 and 12 months (m6 and m12) was evaluated relative to the baseline values at the initial time point (m0).

6.2.6 Gene set enrichment analysis (GSEA)

To identify alterations in gene-sets over time following 6 months of continuous cell culture within each cancer entity, a gene expression dataset comprising raw transcript counts for each cell line (A-673, HeLa, and MCF-7) was utilized for gene set enrichment analysis (GSEA) (see section RNA sequencing (RNA-seq). GSEA was performed on GSEA software version 4.3 developed by the Broad Institute and the Molecular Signatures Database (MSigDB) library was employed (250,251). Enrichment statistics were weighted, and the metric for ranking genes was determined as signal-to-noise ratio. Each GSEA analysis was performed individually for every cell line, and comprehensive GSEA reports were generated. Following the completion of GSEA results for each cell line, the results were combined for each cancer entity using the tidyverse package in R and the combined dataset was then visualized using the ggplot2 package in R version 4 (236,252).

6.2.7 Drug screening and data analysis

All A-673, HeLa and MCF-7 strains, as well as MHH-ES-1, SK-ES-1, SK-N-MC, and TC-71 EwS cell lines, were tested against a core drug library consisting of 10 cytotoxic or cytostatic agents, or an extended drug library consisting of 20 agents (**Materials section, Compounds**). The experimental procedures involved the seeding of cells into 96-well plates at a standardized density of 5×10^3 cells per well in triplicates. Following cell attachment, which typically occurred approximately 4 h post-seeding, serially diluted concentrations of each compound were meticulously added to the wells, with concentrations ranging from 1×10^{-5} µM to 10 µM. Throughout the experimental course, DMSO was utilized as the vehicle control to ensure the fidelity of the results. Following approximately 72 h incubation period maintained at 37 °C with 5% CO_2 in a humidified atmosphere, the assessment of cell viability was carried out utilizing a meticulously

prepared solution of 25 µg/ml of resazurin salt (Sigma-Aldrich) as previously described (253). It is noteworthy that each compound-cell line combination was subjected to rigorous testing, with all assays performed in four biological replicates to ensure the robustness and reliability of the experimental findings. After acquiring the raw viability measurements, normalization of the cell viability data was conducted by aligning them with the corresponding DMSO control values. This normalization step aimed to account for any variations attributed to experimental conditions or technical factors. Subsequently, the area under the curve (AUC) for each cell line was computed using the PharmacoGx package version 3.0.2 within the R statistical environment, specifically version 4.2.1 (254). To evaluate the similarity or dissimilarity between drug sensitivity profiles across different cell lines, Euclidean distances (ED) were calculated. This involved determining the distance between the mean AUC value of all strains and the AUC value of each individual cell line using the formula:

$function(x1, x2)\ sqrt(sum((x1 - x2)^2)) = ED$,

where x1 is the mean value of AUC of all strains and x2 the AUC of individual cell lines. This analytical approach facilitated the quantification of the degree of variation in drug response observed among the diverse cancer entities under investigation. Moreover, to provide a comprehensive illustration of the variability in drug response across different cancer types, the standard error of ED was visually represented. This visualization method was designed to accommodate potential differences in sample size, ensuring an accurate depiction of the variability in drug response profiles across the various cancer cell lines examined. Additionally, any changes in drug sensitivity profiles during the extended culture periods of six and twelve months (m6 and m12) were thoroughly examined. These analyses were conducted in comparison to the baseline measurements taken at the initial time point (m0). This comparative approach allowed for a comprehensive assessment of how drug sensitivity evolved over time, shedding

light on any potential adaptations or alterations in the response of the cell lines to the tested compounds.

6.2.8 Other bioinformatic and statistical analyses

Unless specified otherwise, all data analyses concerning to genomics, methylation, transcriptomics, and drug sensitivity were carried out in R version 4.2.1 (236). A range of R packages was utilized for various aspects of data processing and visualization. For data processing, the readxl package version 1.4.3 was employed for efficient handling of Excel files, while the tidyverse package version 2.0, reshape2 package version 1.4.4, cowplot package version 1.1.1, Rfast package version 2.0.8, and data.table package version 1.14.8 were utilized for diverse data manipulation tasks (252,255,256).

Data visualization was facilitated by the ggplot2 package version 3.4.1, PupillometryR package version 0.0.4, and circlize package version 0.4.15 for circle plot (252,257,258). Furthermore, for specialized analyses such as principal component analysis (PCA) and volcano plots, the ggplot2 package version 3.4.1 was used (252). Spearman's correlation analyses of quantitative data concerning both mRNA and drug response were carried out using the Hmisc package version 4.7-2 (259). Specific figures, including Figures 9b, 11, 12, 14b,14c,15,16 and Supplementary Figure 1b and e, were generated using PRISM 9 (GraphPad Software Inc., Ca, USA) (241). To enhance the robustness and comprehensiveness of transcriptomic datasets, data from this study were integrated with those from Liu *et al*. (3). This combined dataset underwent batch using ComBat-Seq function of package sva version 3.44.0 (260), ensuring consistency and accuracy across different datasets and experimental batches.

7 RESULTS

7.1. EwS cell line is genomically more stable than the adult carcinoma cell lines

To investigate the clonal or genetic heterogeneity of COTF-driven sarcoma cell lines across laboratories and compare them with their adult carcinoma derivates, 11 distinct A-673 cell line strains were included from seven research groups. Additionally, five strains of HeLa, human cervix carcinoma cell line, and MCF-7, human breast carcinoma cell line, from three and two different laboratories, respectively were included (**Fig. 7a**), All cell line strains included in the study had an unknown passage number, yet they were deemed eligible for cell biology research purposes. Among the 11 A-673 strains, nine were wild type strains acquired from seven different laboratories, while two underwent genetic modifications. These two genetically modified cell lines, provided by Dr. Dr. M.F. Orth, were engineered using the dox-inducible pLKO-Tet-On all-in-one system containing a fusion transcript-specific shRNA for *EWSR1::FLI1* (234). One cell line served as the shRNA control, representing a neutral manipulation, while the other featured shRNA targeting *EWSR1::FLI1* (234). This experimental setup allowed assessing whether these genetic alterations cause significant differences at the genomic, transcriptomic, and phenotypic levels. To analyse the possible impact of continuous cell culture on the cell lines, a newly purchased strain (passage number < 5) was obtained for each cell line from international cell repositories (see Methods section). These new strains were cultured continuously for 12 months and analysed at three different time points (corresponding to month 0, 6, and 12, hereafter referred to as m0, m6, m12; (**Fig. 7a**) (261).

All strains were passaged in the same cell culture conditions to minimize experimental bias before the genomic, transcriptomic, and phenotypical comparison and underwent

routine authenticity control using STR profiling. Additionally, routine Mycoplasma detection tests ensured the absence of microbial contamination (see Methods section).

Initially, a comparative genomic analysis across laboratories was performed involving A-673 cell line strains and genotyped each newly acquired cell lines (m0) and their corresponding offspring from m6 on Illumina Global Screening Arrays (GSA). This approach enabled the tracking of genetic alterations over time, explicitly pinpointing point mutations at reference SNPs represented on this array platform. The analysis primarily focused on comparing the status—homozygous for reference allele, alternate allele, or heterozygous—of non-synonymous SNPs across these cell lines (**Table 4**). After six months of cell culture, a matching rate of over 98% of the genotyped SNPs was noted in each cell line (261). Upon closer examination of the in-exon SNP counts, it was observed that HeLa and MCF-7 cells exhibited 2.33-fold and 1.15-fold higher rates of genetic evolution, respectively, when compared to A-673 EwS cells (**Fig. 7b**) (261). This suggests that HeLa and MCF-7 cells are more prone to genetic changes than A-673 EwS cells when kept in cell culture conditions for extended periods.

Notably, SNP status alterations were consolidated differently among chromosomes within these cell lines. In A-673, most SNP status alterations were observed on chromosome 5, whereas MCF-7 exhibited predominant alterations on chromosome 19. Conversely, and consistent with the previous multi-omic comparison, HeLa demonstrated pronounced changes after six months, particularly on chromosomes 2 and 9 (3). In 2019, Liu et al. revealed prevalent copy number variation (CNV) differences across various chromosomes, notably including chromosomes 2 and 9 (3).

The convergence of the findings of this thesis with theirs suggests a consistent pattern of genomic instability in these specific chromosomal regions within the HeLa cell line, with this thesis highlighting their heightened instability over time.

SNP_ID	Gene	A-673_1	A-673_2
rs1042718	*ADRB2*	0/1	0/0
rs11167756	*ARAP3*	0/1	0/0
rs11962165	*HCG11*	0/0	0/1
rs12658464	*ZNF474*	0/1	0/0
rs2229882	*MAP3K1*	0/1	0/0
rs2547	*CXCL14*	0/1	1/1
rs33409	*SYNPO*	0/1	0/0
rs3733695	*PCDHB7*	0/1	1/1
rs3822625	*MAP3K1*	0/1	0/0
rs79790545	*LVRN*	0/1	0/0

Table 10: Exemplary illustration of Non-synonymous SNP status comparison in A-673_1 (m0) and A-673_2 (m6)

Subsequently, the comparative analysis of non-synonymous SNPs impacting coding sequences and splicing regions across 11 distinct A-673 strains, including two genetically modified strains, unveiled an exceptional 98.9% consistency in shared SNPs among all strains (**Fig. 7c**). This striking finding stands in stark contrast to observations in MCF-7 cells, where Ben-David et al. reported a mere 35% of SNPs shared across all screened 27 MCF-7 strains (2). Further data analysis revealed that the SNP alterations were mainly within silent SNPs, and no specific functionality correlations were identified.

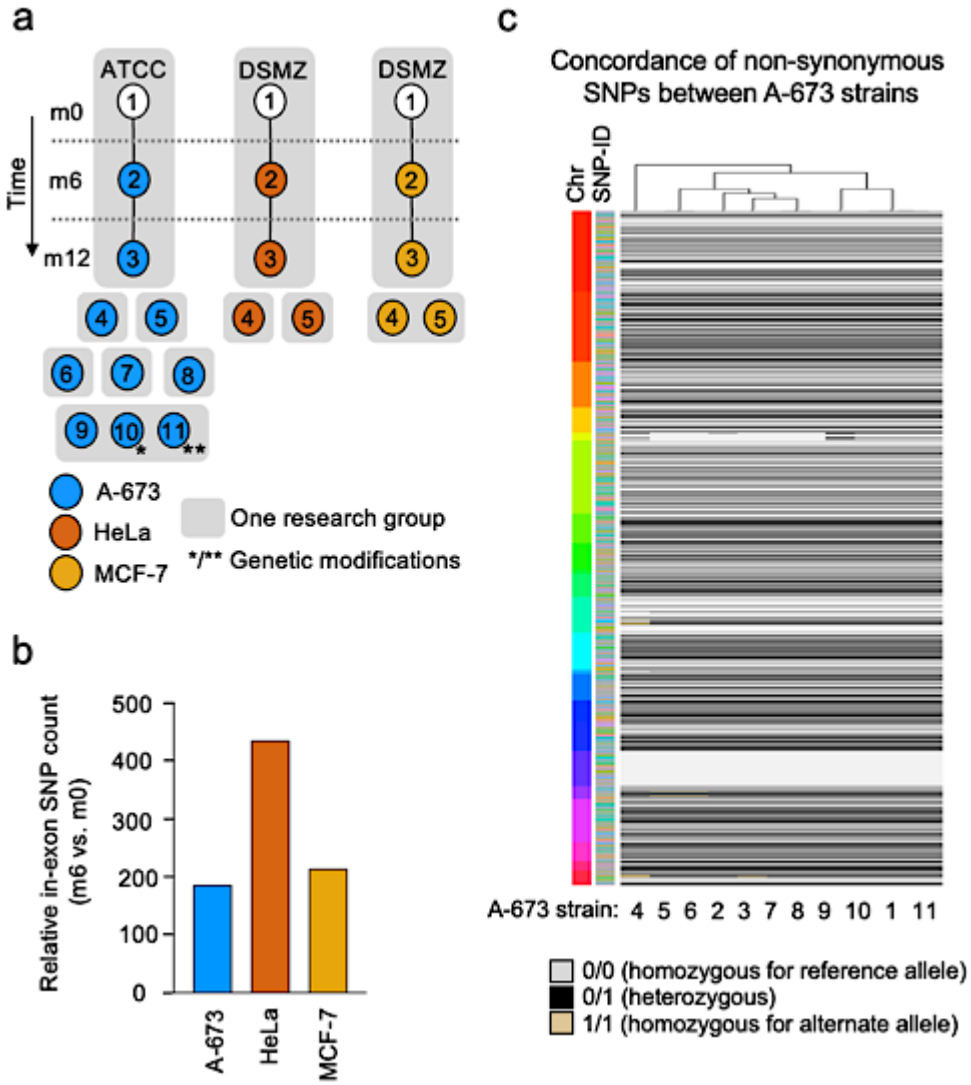


Figure 7: Longitudinal and cross-laboratory genomic stability in COTF-driven pediatric sarcoma cell line strains in contrast to adult carcinoma strains. a. Newly acquired wild type A-673 EwS, HeLa cervix carcinoma and MCF-7 breast carcinoma cell lines (A-673_1, HeLa_1 and MCF-7_1) at the initial time point (m0) were kept in culture over six months (m6; A-673_2, HeLa_2 and MCF-7_2) and twelve months (m12; A-673_3, HeLa_3 and MCF-7_3). Additional cell line strains were gathered from seven, three and two laboratories, respectively, and labeled A-673_4 to A-673_9, HeLa_4 to HeLa_5, and MCF-7_4 to MCF-7_5. Single cell clones with either a neutral manipulation (*, A-673_10) or an inducible shRNA construct targeting its *EWSR1::FLI1* translocation (**, A-673_11) were included in the cross-laboratory comparison. ATCC, American Type Culture Collection, DSMZ (German Collection of Microorganism and Cell Cultures). **B.** The relative counts of in-exon SNPs in A-673, HeLa, and MCF-7 were assessed after six months of continuous culture, with initial time point (m0) values used as reference (m6 vs. m0). **c.** A heatmap depicting the status (homozygous for reference allele, alternate allele, or heterozygous) of non-synonymous SNPs across 11 A-673 strains. Chromosomes are represented on the left color bar, while different SNP-IDs are shown on the right color bar (*N*=1,599). Figure and figure legend from Kasan et *al.*, 2023 preprint available on bioRxiv, (261), also in press in Nature Communications under CC-BY license.

Despite the rigorous efforts in genomic data analysis, a significant limitation was encountered due to batch discrepancy, which hindered a complete longitudinal comparison (month 0 versus month 12) of carcinoma and EwS cell lines. Although all samples from months 0 and 6 underwent processing together, sequencing of MCF-7 and HeLa m12 derivatives occurred separately. These batch discrepancies resulted in reduced SNP identification and compromised the accuracy of comparative assessment.

7.2. EwS cell line displays remarkably stable and homogenous transcriptome

To compare cross-laboratory heterogeneity in transcriptomic level between adult carcinoma cell lines and oligo-mutated pediatric sarcomas, RNA sequencing (RNA-Seq) was performed on 11 A-673, five HeLa, and five MCF-7 strains. Principal component analysis (PCA) of transcriptomic profiles showed striking differences between both carcinoma-derived cells and EwS cell lines: The A-673 EwS strains exhibited a notably tighter clustering pattern compared to the extensive diversity observed within the HeLa and MCF-7 carcinoma strains (**Fig. 8a**). Even despite the inclusion of two genetically modified A-673 strains, the EwS cluster maintained its remarkable transcriptomic homogeneity (**Fig. 8a**) (261). This is particularly notable given the smaller sample size of the carcinoma cell lines. The higher transcriptomic homogeneity of A-673 strains was further confirmed by comparing the coefficient of variation (CV) of gene expression of A-673 strains to adult carcinoma strains (**Fig. 8b**) (261). The differential impact on transcriptomic stability observed in this study could explain the tighter clustering and lower transcriptomic variability observed in A-673 EwS cells compared to the adult carcinoma cells.

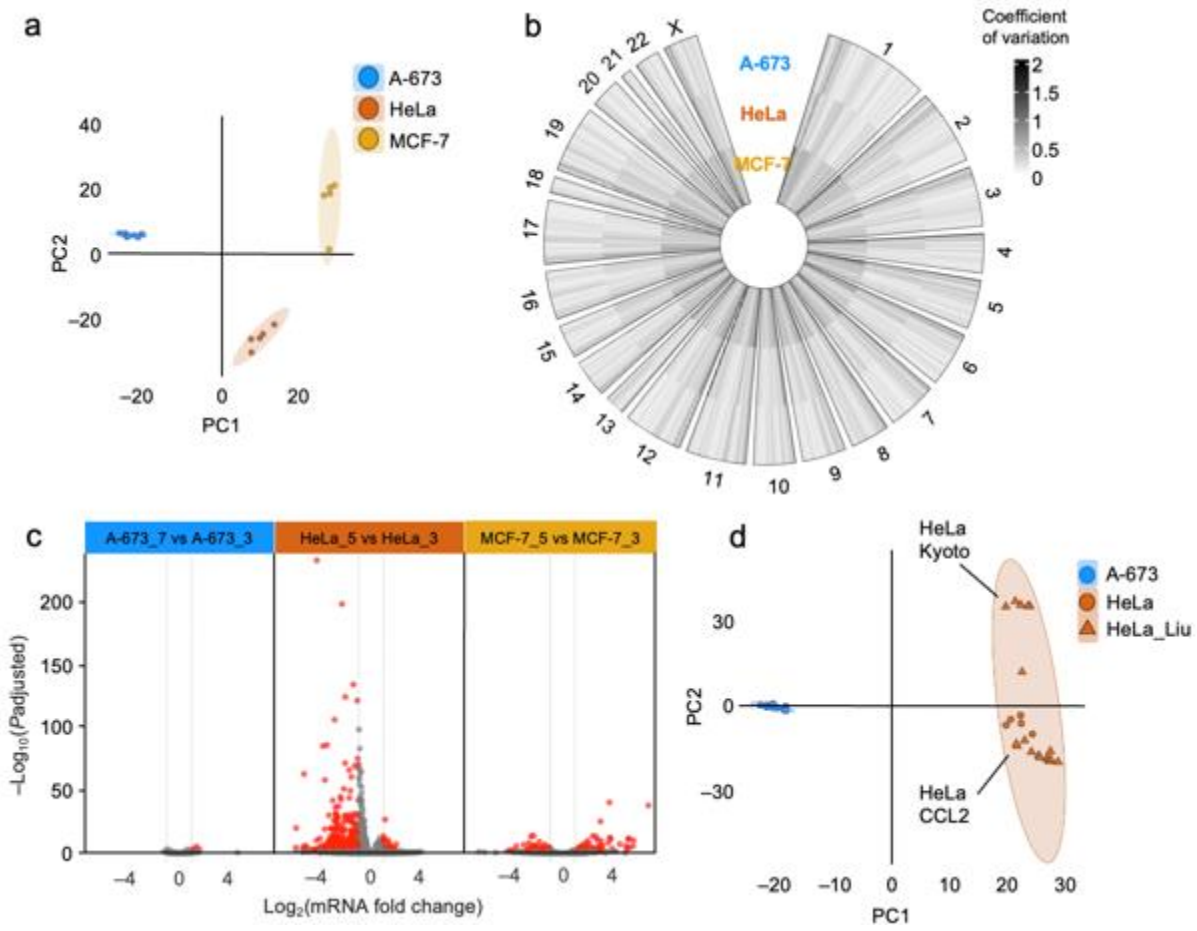


Figure 8: Remarkable transcriptomic homogeneity in COTF-driven pediatric sarcoma strains compared to adult carcinoma strains. Transcriptomic PCA portrays the distribution of 11 A-673, five HeLa and five MCF-7 strains, covering a total of 10,256 transcripts. **B.** Circle plot displays the coefficient of variation (CV) of expressed genes per chromosome (top 60% quantile) for all A-673, HeLa and MCF-7 cell line strains. **c.** Volcano plot visualizes the DEG obtained by comparison of two A-673, HeLa and MCF-7 strains with the highest variance (A-673_7 vs A-673_3, HeLa_5 vs HeLa_3 and MCF-7_5 vs MCF-7_3). The red dots represent significantly differentially expressed genes (BH adjusted P<0.01; |FC|>1). **D.** Combined transcriptomic PCA integrates A-673 and HeLa strains in the dataset and that of Liu *et al.* (HeLa_Liu) (N=13,569 transcripts). Figure and figure legend from Kasan *et al.,* 2023 preprint available on bioRxiv (261) also in press in Nature Communications under CC-BY license.

For an in-depth investigation, the two strains displaying the highest transcriptomic variability were compared in each cancer type. These strains were A-673_7 and A-673_3 for EwS, HeLa_5 and HeLa_3 for cervical cancer, and MCF-7_5 and MCF-7_3 for breast cancer. Analysis of differentially expressed genes (DEGs) notably showed that the HeLa strains displayed over 60 times more DEGs (380 transcripts; 39 up-

regulated, 341 down-regulated) meeting the criteria of |fold change (FC)|>1, Benjamini-Hochberg (BH) adjusted *P*<0.01, in comparison to A-673 EwS strains which exhibited only five transcripts, all of which were up-regulated. Similarly, the MCF-7 strains showed a 20-fold increase in DEGs (108 transcripts; 57 up-regulated, 51 down-regulated) compared to A-673 EwS strains (**Fig. 8c**, **Supp. Fig. 1a**) (261). Furthermore, RNA-seq data derived from this thesis was combined with that of Liu *et al.*, which included 14 different HeLa strains from various laboratories (3). PCA clearly showed a specific clustering of HeLa strains of this thesis with their HeLa-CCL2 strains, suggesting a probable shared ancestry (**Fig. 8d**) (261). As expected, the HeLa strains exhibited greater heterogeneity than the A-673 strains (**Fig. 8d**, **Supp. Fig. 1b**) (261). In addition, the heterogeneity observed between the HeLa-CCL2 and Kyoto strains by Liu *et al.* suggests that integrating HeLa-Kyoto within the cell panel studied in this thesis would have even further increased the observed variability of HeLa cell strains when compared to the COTF-driven A-673 strains (3).

Next, the global gene expression of the newly purchased cell lines (m0) of each cancer type and their respective m12 offspring was compared. Corresponding with the outcomes from the cross-laboratory analysis, a significant difference in global gene expression was evident in HeLa and MCF-7 cells in contrast to A-673 (A-673 cell line had a median $\log_2$FC zero, (ranging from –4.25 to 4.22); for HeLa median $\log_2$FC was 0.47 (ranging from –3.39 to 16.89), and for MCF-7 median $\log_2$FC was also 0.47 (ranging from –3.46 to 15.81); *P*<0.0001, two-sided Wilcoxon signed-rank test, **Supp. Fig. 1c,d**) (261). Finally, a Gene Set Enrichment Analysis (GSEA) was performed, which provided valuable insights into the molecular changes in cell lines over six months.

Here, the A-673 cell line exhibited notable downregulation within the mitotic spindle pathways compared to its status at month 0**.** Conversely, cervix carcinoma cell lines

displayed significant enrichment scores not only in the mitotic spindle pathway but also in pathways related to protein secretion, TNF Beta signaling, epithelial-mesenchymal transition, MYC targets, and many others (**Supp. Fig. 2**).

7.3. COTF-driven EwS cell line exhibits a uniform and stable phenotype

To explore the intricate relationship between genomic and transcriptomic alterations and their impact on phenotype, the thesis delved into drug sensitivity, comparing COTF-driven EwS cells with their highly mutated adult carcinoma counterparts. A panel of cell lines consisting of 11 A-673 EwS strains (including two genetically modified A-673 strains), five HeLa cervical cancer strains, and five MCF-7 breast cancer strains were tested with ten different compounds targeting various cellular pathways. The methodological approach involved resazurin assays, wherein cells were incubated with each respective drug in 96-well plates under standardized conditions for 72 h. Viability data was calculated as a percentage relative to control wells incubated with vehicle (DMSO). Dose-response curves and the measurement of the AUC and IC50 values was subsequently performed for each compound. First, an exploratory data analysis using PCA with AUC values as input to identify the fundamental diversity within the dataset was performed. Leveraging PCA's dimensionality reduction properties, this approach allowed to explore the data complexity while capturing essential patterns and relationships in a condensed form. Like in the transcriptomic and genomic layers, notable diversity was observed in adult carcinoma strains from various laboratories, contrasting with the homogeneous COTF-driven EwS cell strains (**Fig. 9a**).

Despite the inclusion of two cell lines with genetic modifications, remarkable uniformity in drug response was displayed across the 11 A-673 cells (261).

To illustrate drug response of the panel of cell lines, a compound also screened in the study by Ben-David *et al.* on MCF-7 cells, 17-AGG, was selected and drug response curves were generated using the viability data (2). 17-AGG demonstrated a great dissimilarity in sensitivity variation across HeLa and MCF-7 carcinoma strains, and a remarkable stability in response in A-673 strains (**Fig. 9b-d**).

Subsequently, the ED between drug sensitivity profiles of a specific cell line were calculated, allowing the measurement of the overall AUC to mean across all cell lines. Using ED was instrumental in acknowledging the multidimensional nature of drug response data, considering the magnitude and directionality of differences in response patterns among individual cell lines. This methodological choice was pivotal in establishing an unbiased and standardized metric for assessing drug sensitivity, ensuring an unbiased comparison across diverse cancer types. Building upon the prior findings, a stark contrast in drug response was observed between adult carcinoma strains and EwS strains across all screened compounds ($P<0.005$, one-sided Wilcoxon signed-rank test) (**Fig. 10a**).

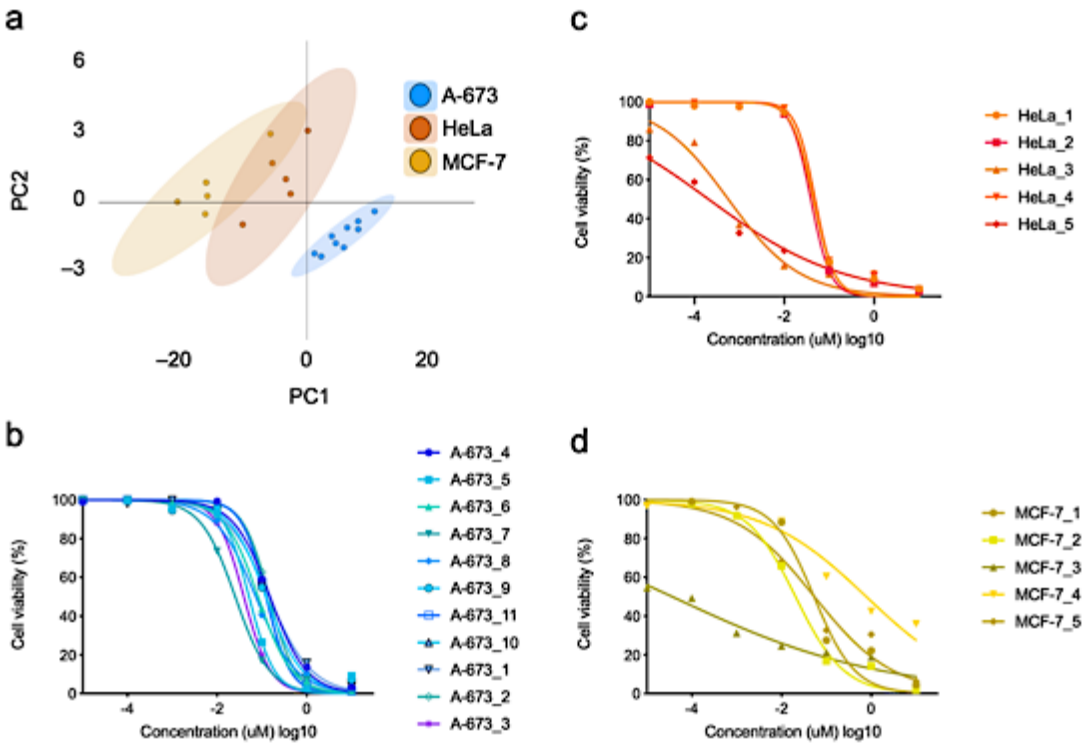


Figure 9: The exceptional phenotypic uniformity in A-673 in contrast to HeLa and MCF-7 cell lines. A. Drug sensitivity PCA and **b-d.** drug response curves for 17-AGG of 11 A-673, five HeLa and five MCF-7 cell line strains.

To validate the uniformity in drug response among COTF-driven EwS cells and their disparity from carcinoma cell lines, the data was subjected to a Spearman's correlation test. Once again, these results underscored the higher similarity among EwS strains compared to adult carcinoma strains Mean Spearman's ρ of A-673 strains was 0.94 (ranging from 0.95 to 0.93), HeLa strains 0.87 (ranging from 0.91 to 0.83), and for MCF-7 0.88 (ranging from 0.92 to 0.85) (**Fig. 10b**) (261). This robust correlation analysis reinforces the notion that COTF-driven EwS cells exhibit a significantly more uniform drug response compared to adult carcinoma cell lines.

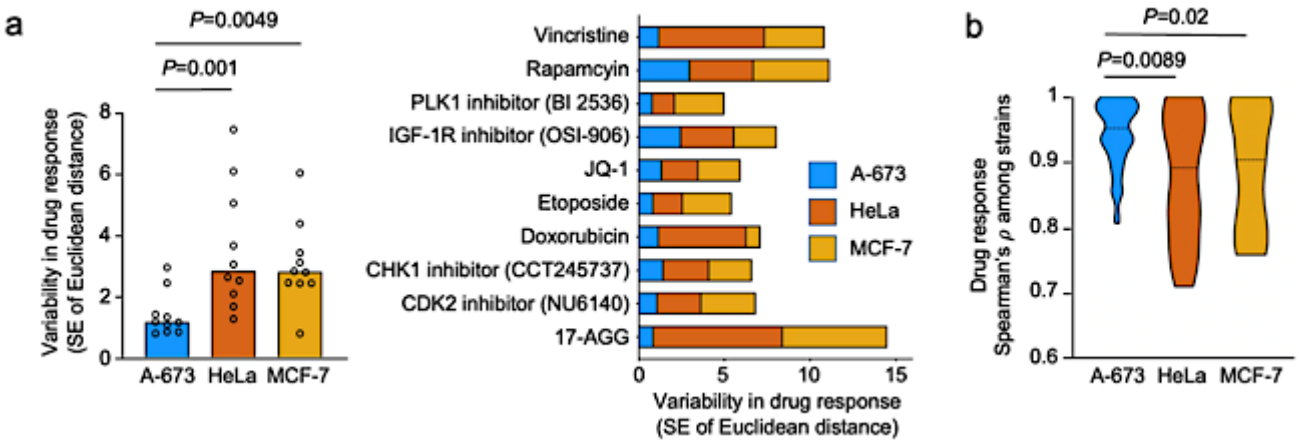


Figure 10: Uniform drug-response in COTF-driven A-673 strains. **a.** Left, jointed variability in drug sensitivity in all A-673, HeLa, and MCF-7 strains depicted as standard error of ED, each compound is shown as a black circle, one-sided Wilcoxon signed-rank test. Right, standard error of ED for each specific screened compound. **B.** Violin plot depicting the distribution of Spearman's *ρ* for drug response across 11 A-673, five HeLa, and five MCF-7 cell line strains. Dotted black line shows the median (one-sided Wilcoxon rank-sum test). Figure and figure legend from Kasan *et al.* 2023 preprint available on bioRxiv (261), also in press in Nature Communications under CC-BY license.

Moreover, to investigate the impact of prolonged culture on potential changes in drug sensitivity, a drug screening was performed on newly purchased A-673, MCF-7, and HeLa cells (m0). This comprehensive analysis was repeated at two additional predetermined intervals after continuous culturing (m6 and m12). Subsequently, the raw cell viability measurements at a standardized concentration of 1 µM were compared with their respective parental cells for each compound. Aligning with standards in compound screening assays (262), this approach allowed the consideration of diverse cancer types and mechanisms while presenting nuanced details in drug response. In parallel with earlier observations, a remarkably uniform phenotype in A-673 after 6 and 12 months contrasted the sensitivity instabilities observed in HeLa and MCF-7 cell lines (**Fig. 11**) (261)**.** Displaying raw viability data for A-673, HeLa, and MCF-7 cell lines over 12 months distinctly highlights the relative stability of A-673 cells compared to the pronounced viability changes observed in carcinoma cell lines.

Notable shifts in drug sensitivity were observed in carcinoma cells over time with seven out of 10 compounds, compared to only slight changes in A-673 cells with only 2 out of the ten compounds tested (**Fig. 11**) (261). The varying degree of drug sensitivity shifts among these cell lines underscore the remarkable stability in drug response observed in A-673 cells compared to the more pronounced changes seen in carcinoma cells over the 12 months.

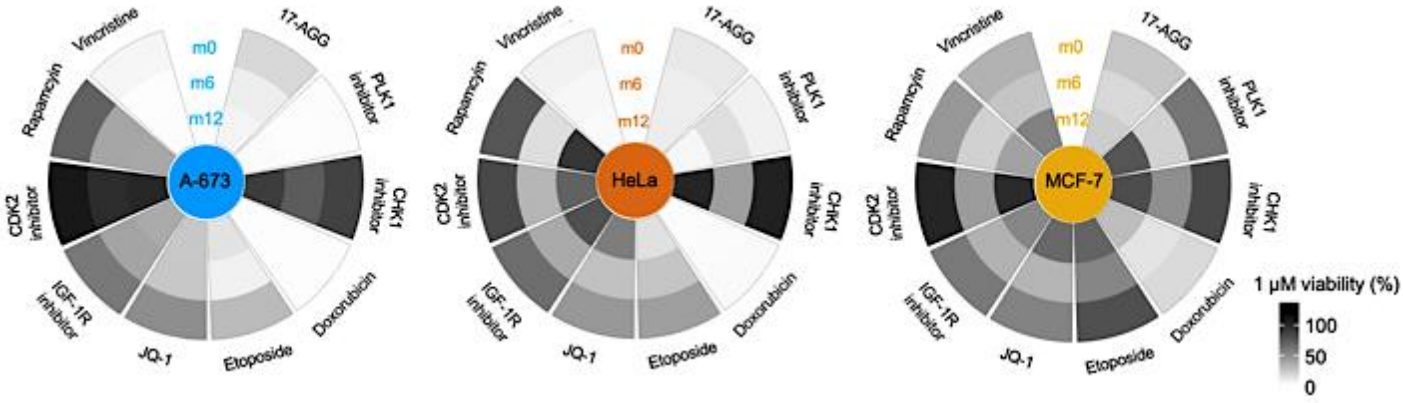


Figure 11: Steady drug sensitivity in A-673 cell line after long term cell culture. Circle plots display the raw viability of A-673, HeLa and MCF-7 cell lines for each compound (at 1 µM) after 0, 6, and 12 months of continuous long-term culture (m0, m6, and m12). Figure and figure legend from Kasan *et al.* 2023 preprint available on bioRxiv (261), also in press in Nature Communications under CC-BY license.

Furthermore, images of each cell line at each predetermined time point were captured, keeping consistent confluency levels, to visually assess cell morphology over time. Interestingly, all cells maintained a predominantly stable morphology throughout the 12 months. Despite the observed changes in drug sensitivity, visual assessment of cell morphology did not reveal significant diversity (**Supp. Fig. 3**).

7.4. Cell line stability is a spectrum even within the same tumor entity

Finally, to further investigate if the absence of genomic and phenotypic evolution in EwS cell line A-673 was a general feature of EwS cell lines, four additional EwS cell lines, namely MHH-ES1_m0, SK-ES-1_m0, SK-N-MC_m0, and TC-71_m0 were purchased,

and subjected to a 12-month culture period (**Fig 12a**). Hereafter comprehensive genomic and epigenomic analyses of these cell lines were performed using Illumina GSA and MethylationEPIC BeadChip arrays.

The analyses of different EwS cell lines revealed that, while all EwS cell lines exhibit relative stability in genomic, transcriptomic, and phenotypic level, compared to the A-673 cell line, there is a spectrum of variability among these cell lines. This spectrum is evident in both non-synonymous SNP alterations and differentially methylated CpG sites over time among these EwS cell lines (**Fig. 12b, c**). Notably, while the majority of in-exon SNPs (median of 99.6%, ranging from 99.3% to 99.8%) remained unchanged after 12 months of continuous culture across these cell lines, differences in stability were noticeable (261). For example, A-673 and MHH-ES-1 cell lines showed less stability while TC-71 stood out for its remarkable genomic immutability (**Fig. 12b**).

Further, an in-depth exploration of the transcriptional landscape, provided a nuanced comprehension of the dynamic evolution observed in EwS cell lines over an extended period. The analysis of DEGs within each EwS cell line at the 12-month mark reveals a spectrum of transcriptional alterations (**Fig. 12d**) (261). Notably, TC-71 stood out with minimal changes in gene expression, indicating remarkable transcriptional stability over 12 months. In contrast, SK-ES-1 exhibited a distinctive profile, revealing the highest number of DEGs after 12 months of continuous culture— (219 transcripts; 99 up-regulated, 120 down-regulated, marking a 50% increase compared to A-673) (**Fig. 12d**). This comprehensive exploration traverses the complex and varied dynamics within the genomic, epigenomic, and transcriptional layers, portraying the diverse spectrum of evolutionary behaviors intrinsic to EwS cell lines.

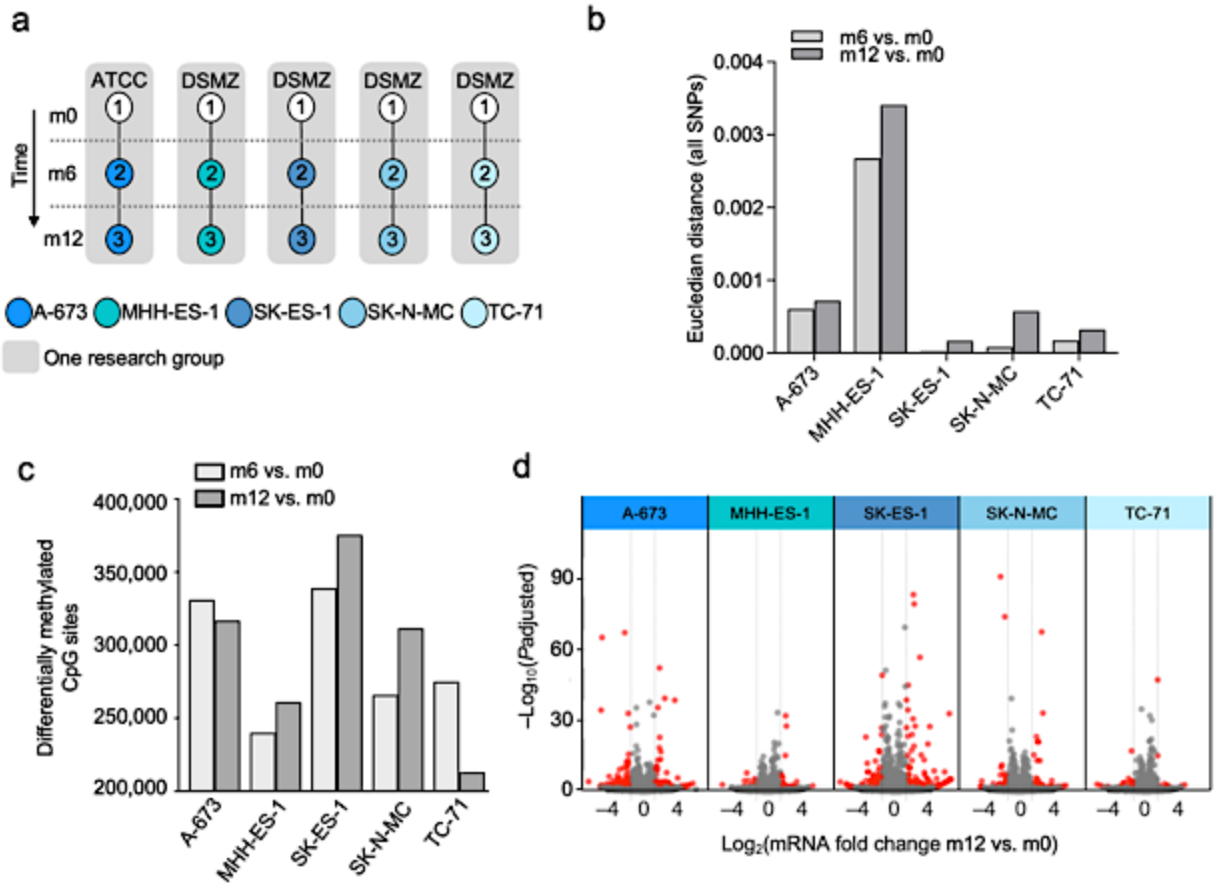


Figure 12: In-depth analysis of stability on individual cell lines from the same COTF-driven sarcoma entity. a. Newly acquired wild-type EwS cell lines A-673, MHH-ES-1, SK-ES-1, SK-N-MC, and TC-71 EwS (_1) were kept in culture for six months (m6; _2), and 12 months (m12; _3). ATCC, American Type Culture Collection, DSMZ (German Collection of Microorganism and Cell Cultures). **b.** Bar plot displays the Euclidean distance between the reference version of each cell line (_1) and its offspring at 6 months (_2) and 12 months (_3) after continuous culture (m6 vs. m0; m12 vs. m0), considering all SNPs. **c.** Bar plot depicting the number of differentially methylated CpG sites (including differentially hypo- and hyper-methylated) for A-673, MHH-ES-1, SK-ES-1, SK-N-MC, and TC-71 after six (m6) and 12 months (m12) of continuous culture, referencing the initial time point (m0) values.
d. Volcano plot of DEG after 12 months (m12) of continuous culture, using the respective initial time point (m0) values as reference for A-673, MHH-ES-1, SK-ES-1, SK-N-MC, and TC-71. The red dots denote significantly differentially expressed genes (BH adjusted $P<0.01$; $|FC|>1$). Figure and figure legend from Kasan *et al*. 2023 preprint available on bioRxiv (261), also in press in Nature Communications under CC-BY license.

Additionally, the study involved subjecting newly acquired EwS cell lines to an extended drug library comprising ten additional compounds (**Fig. 13**). This included drugs that have recently been implicated in EwS field, such as elesclomol, olaparib, and gemcitabine (263–265).

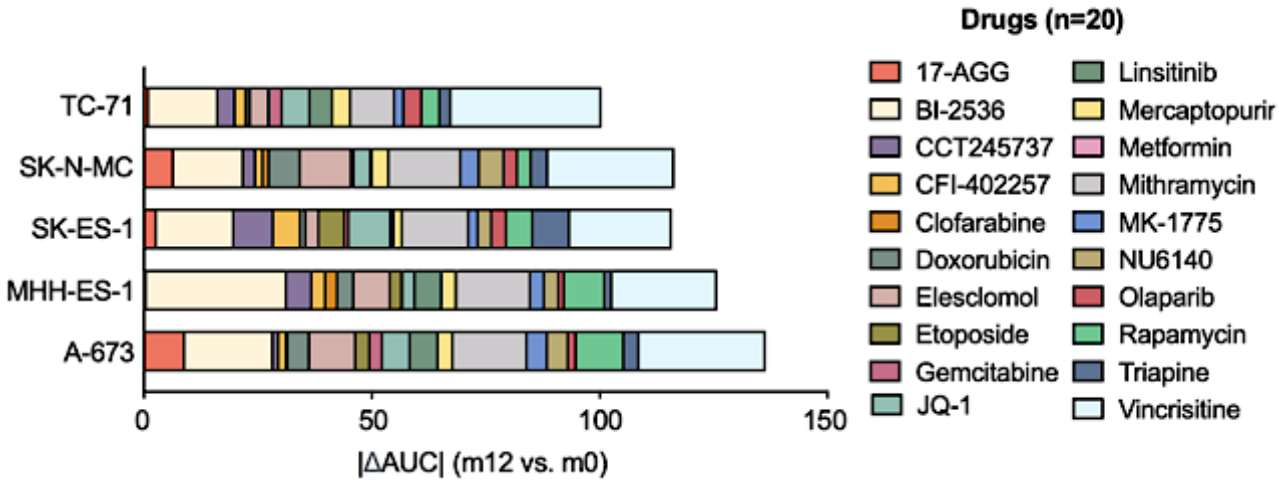


Figure 13: Drug sensitivity change in each EwS cell line after 12 months. The relative variation in drug sensitivity for each EwS cell line measured as the absolute delta of mean AUC values from four biological replicates, comparing the m12 and m0 versions for the extended drug library (20 compounds) on each cell line. Figure and figure legend from Kasan *et al*. 2023 preprint available on bioRxiv (261), also in press in Nature Communications under CC-BY license.

In this analysis, a spectrum of inter-cell line variability in the collective drug response over time was depicted, ranging from the less stable A-673 to the notably stable TC-71 EwS cell line (**Fig. 13**) (261). To provide a detailed illustration of drug sensitivity alters in these cell lines over time, three compounds were selected: previously mentioned 17-AGG as well as Etoposide, and Doxorubicin, which are topoisomerase II inhibitors and used in the chemotherapy regimen for EwS (266).

The drug response curves were generated for A-673 and TC-71 cell lines using viability data collected at months 0, 6, and 12. Notably, TC-71 cells showed a remarkably consistent and robust drug response curve for all three compounds compared to A-673 cells (**Fig 14a-c**), which highlights the chemical response stability of TC-71 cells even after 12 months of culture, suggesting potential therapeutic advantages in drug screening assays and preclinical studies.

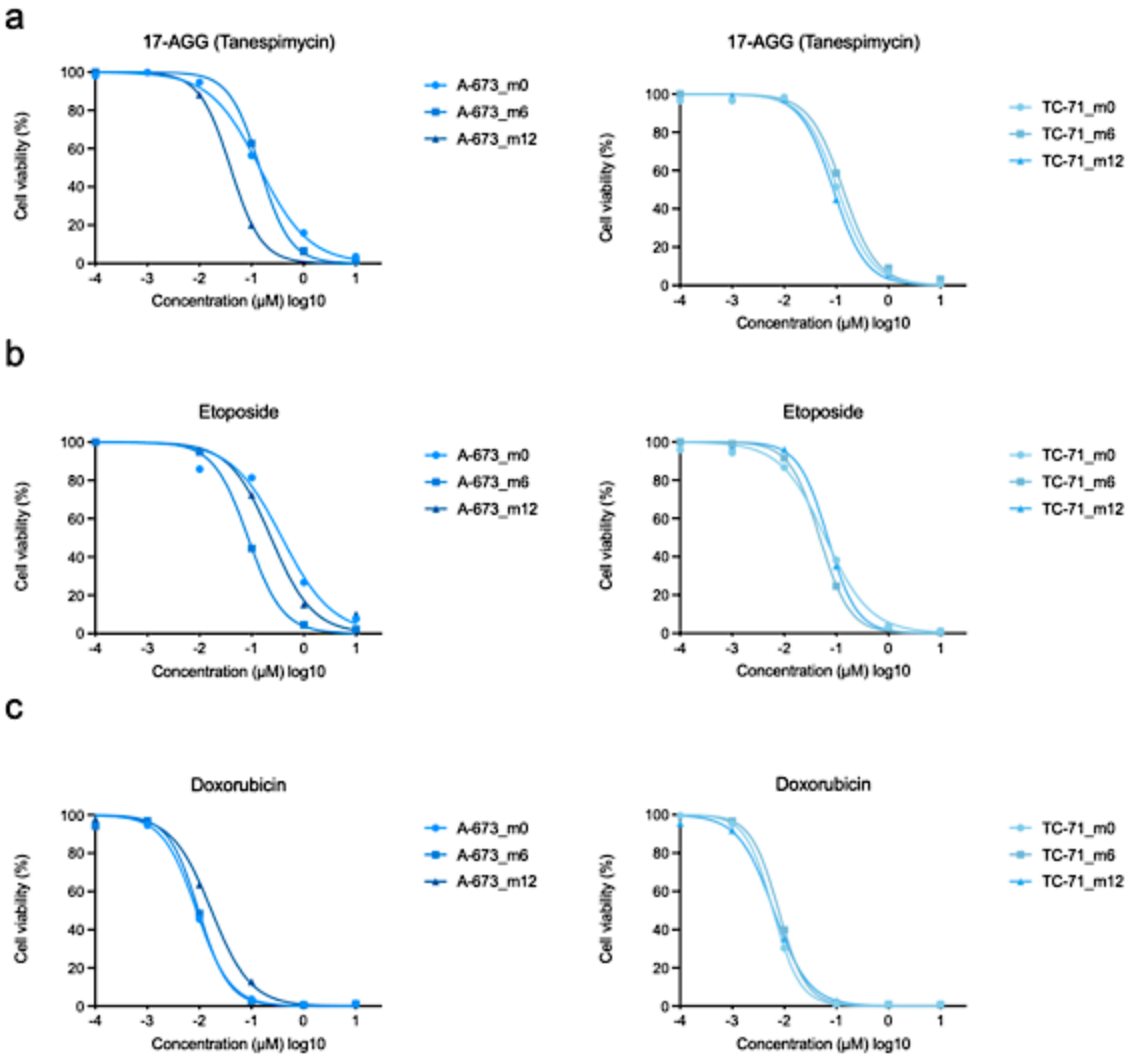


Figure 14: Drug response curves for A-673 and TC-71 cell lines depicting the sensitivity shift over 12 months for three compounds: a.17-AGG, **b.** Etoposide c. Doxorubicin.

Finally, ranking plots were generated to allow for a detailed exploration of the stability spectrum across different data layers, as observed in EwS cell lines cultivated for 12 months (**Fig. 15**) (261). By integrating diverse genomic, transcriptomic, and phenotypic metrics, these plots comprehensively depict how each cell line evolves under continuous culture conditions. These visualizations unveil the underlying dynamics within the EwS cell line landscape, from the intricate interplay of genetic alterations to the dynamic shifts in gene expression patterns and cellular phenotypes.

Such comprehensive evaluations serve as indispensable tools for researchers aiming to optimize experimental designs and enhance the translational relevance of their findings in EwS research.

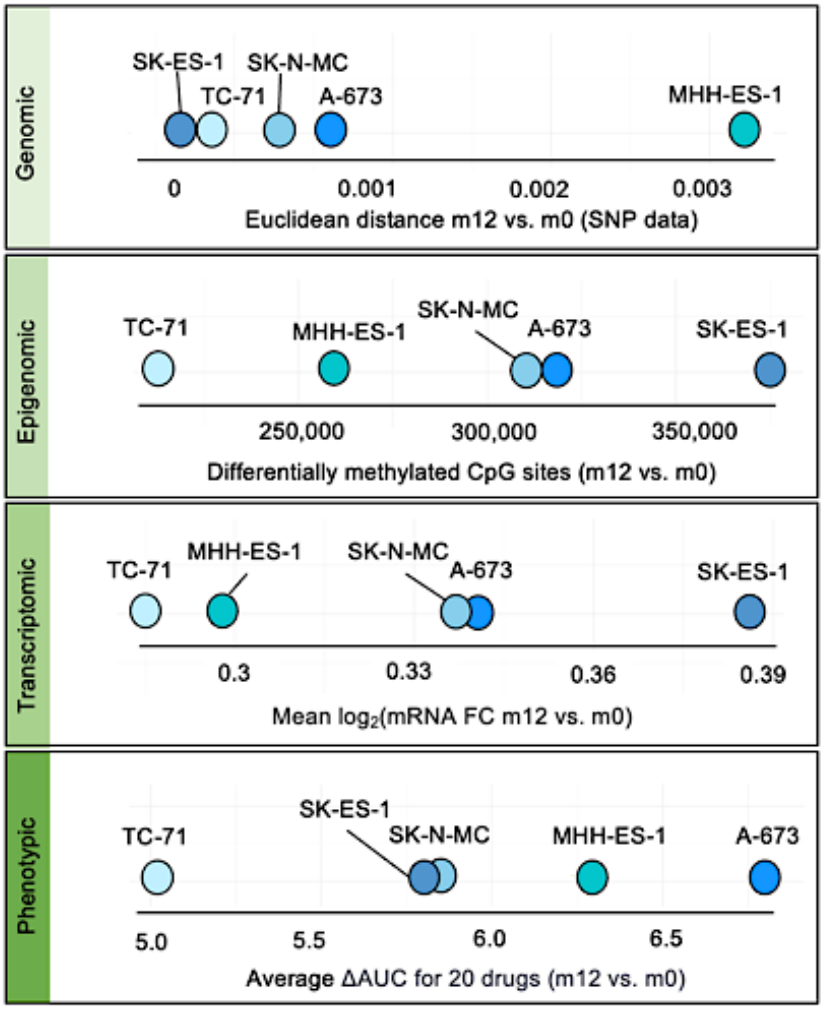


Figure 15: Cell line stability spectrum in EwS. Ranking plots illustrate the positioning of EwS cell lines based on their evolution degree due to long-term culturing for 12 months across various datasets. Figure and figure legend from Kasan et al. 2023 preprint available on bioRxiv (261), also in press in Nature Communications under CC-BY license.

In conclusion, this thesis sheds light on the translational implications of findings from previous research, particularly regarding the stability and heterogeneity observed in adult carcinoma cell lines and the challenges surrounding reproducibility in cancer research using cell lines as research models.

The results suggest that such observations may not necessarily generalize to other cancer cell lines, particularly those harboring a stable genetic background and a defined driver mutation, such as the COTF found in EwS (**Fig. 16**) (261).

Significantly, these findings suggest that research using COTF-driven cell line models like EwS, even following genetic modifications and extended periods of continuous culture should be reproducible. Furthermore, the results of this study unveil the variable degree of evolution among individual cell lines within the same cancer entity, emphasizing the importance of choice of cell line for the experimental outcome in question. These insights deepen the understanding of the molecular and phenotypic dynamics within cancer cell lines, with profound implications for designing and interpreting preclinical studies to advance therapeutic strategies for COTF-driven pediatric cancers like EwS.

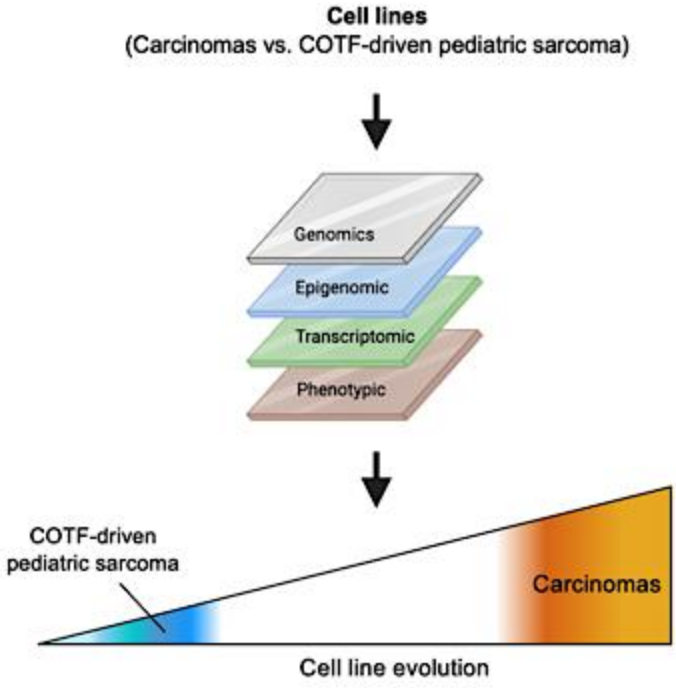


Figure 16: Summary illustration of the main results of this thesis. Figure and figure legend from Kasan *et al*. 2023 preprint available on bioRxiv (261), also in press in Nature Communications under CC-BY license.

8 DISCUSSION

In this thesis, the primary objective was to investigate the validity of generalizing previous findings from studies focusing on the heterogeneity and instability observed in adult carcinoma cell lines, such as MCF-7 and HeLa, to a different paradigm: single COTF-driven pediatric sarcoma cell lines renowned for their genomic and transcriptomic stability (2,3). The reproducibility and reliability of cell line models in preclinical research are crucial for ensuring the accuracy and applicability of experimental results, particularly in the context of advancing therapeutic strategies in cancer treatment (2–4,6). Given the significant attention placed on adult carcinoma cell lines in previous studies, this inquiry aimed to address a critical gap in knowledge by probing whether the same degree of heterogeneity and evolution is mirrored in long-term cell cultures of pediatric oligomutated sarcomas, particularly those propelled by COTFs like EwS. Shedding light on the clonal and genetic heterogeneity within pediatric sarcomas underscores the importance of utilizing a reliable research model to enhance the precision and relevance of experimental findings, ultimately shaping the implications for therapeutic development and clinical translation. This thesis focused on three distinct cancer entities, each represented by well-established cell line models: cervix carcinoma (HeLa), breast carcinoma (MCF-7), and EwS, specifically using the A-673 cell line. These cell lines serve as invaluable tools in cancer research, each embodying unique characteristics, and molecular profiles reflective of their respective cancer types (1,37,234,267). Examining these cell lines collectively aimed to explore the broader implications of findings across different cancer types and genetic backgrounds, with a particular focus on elucidating the stability and evolution dynamics within the context of COTF-driven pediatric sarcomas.

The findings of this thesis align with previous studies, confirming the observed instability of adult carcinoma cell lines during prolonged culture and the significant diversity

observed across different research settings (2,3,6). Despite the relatively small sample size in the adult carcinoma "control" group, thesis results emphasize the potential limitations of generalizing these conclusions to other cell lines. Specifically, it suggests that such instability may not universally apply to all cell lines, especially to ones characterized by a stable genetic background and a distinct driver mutation, as exemplified by the COTF present in EwS (8,10,156). By focusing on COTF-driven cell line models like EwS, the investigation unveils a promising approach for achieving reproducible research outcomes, even in the presence of genetic alterations and prolonged culture durations. Furthermore, this study reveals the varying degrees of (epi)genomic and phenotypic evolution among cell lines of the same cancer type. This underscores the necessity of considering specific cell line attributes when interpreting and replicating research findings, highlighting the complex interplay between cellular characteristics and result reproducibility in cancer research.

The analysis yielded valuable insights into the genomic and transcriptomic stability of the examined cell lines. Specifically, A-673 EwS cells exhibited remarkable stability at both the genetic and transcriptional levels compared to the adult carcinoma cell lines HeLa and MCF-7. This stability suggests a robustness in the genomic and transcriptomic architecture of A-673 cells, which could enhance their reliability as a model system for studying EwS and related conditions. In contrast, HeLa and MCF-7 cells showed more pronounced genetic alterations, particularly in the form of non-synonymous SNPs, over the 6-month culture period. This shared observation may imply various biological consequences.

For example, it could be hypothesized that this genomic instability may lead to the dysregulation of critical genes involved in cell cycle control or tumor suppression pathways, such as *PTEN* and *CDKN2A* located on chromosomes 2 and 9, respectively, and their roles in the PI3K-AKT-mTOR and DNA repair pathways (268,269). The

comparative analysis of non-synonymous SNPs impacting coding sequences and splicing regions across 11 distinct A-673 strains from different laboratories revealed a remarkable 98.9% consistency in shared SNPs among all strains. This finding stands in contrast to observations in 27 MCF-7 strains, where Ben-David *et al*. reported a mere 35% of SNPs shared across all screened strains (2). The stark disparity in SNP consistency between the two cell lines is likely due to the distinct genomic landscapes of the A-673 and MCF-7 cell lines. A-673 cells, originating from a pediatric patient with EwS, may exhibit genomic stability attributed to the regulatory role of the main driver, *EWSR1::FLI1* fusion oncogene. Conversely, MCF-7 cells, derived from a metastatic breast cancer site of a 69-year-old patient, likely manifest a heightened mutation load due to their metastatic nature (160,170,270).

Further the analyses revealed notable distinctions in transcriptomic and phenotypic characteristics between A-673 EwS cell lines and adult carcinoma cell lines. Specifically, a remarkable transcriptomic homogeneity among A-673 EwS strains was observed, including genetically modified strains, in contrast to the substantial diversity observed in HeLa and MCF-7 strains. Furthermore, GSEA uncovered a sparse number of significantly enriched pathways in EwS, contrasting with the abundance observed in cervix carcinoma.

The unique genetic landscape of A-673 EwS cells, characterized by a single fusion oncogene presented a distinct transcriptomic stability compared to adult carcinoma cells (8,9,222). The dominant regulatory role of this single fusion oncogene in maintaining transcriptomic potentially reflects on the cellular behavior, disease progression, and therapeutic response observed clinically in EwS (9,222,271,272).

For instance, the stability of A-673 EwS cells' transcriptome might be influenced by the regulatory impact of the fusion oncogene on crucial pathways like cell cycle control (e.g., CDK4) or signaling pathways (like the IGF-1 pathway) that are commonly implicated in

EwS progression and therapeutic responsiveness (273–275). In contrast, the multifactorial genetic makeup with high mutation burden in HeLa and MCF-7 cells may contribute to diverse transcriptomic profiles and volatility, potentially involving genes related to hormone receptor pathways (e.g., estrogen receptor genes like ESR1 in breast cancer) and cellular proliferation pathways (such as the p53 pathway in cervical cancer) that play crucial roles in disease behavior and therapy response (37,168,170,276).

Further, drug sensitivity assays revealed a consistent response in the A-673 cell line across both cross-laboratory and longitudinal assays for all tested compounds, contrasting with the varied responses observed in adult carcinoma strains. Moreover, the outcomes suggested a potential correlation between transcriptomic analysis and drug assays, indicating enrichment in mitotic spindle pathways and sensitivity alterations to specific drugs like vincristine and PLK-1 inhibitors, which present intriguing insights. Notably, a captivating phenomenon unfolded in drug sensitivity assays, wherein carcinoma cells exhibit an initial increase in sensitivity followed by the development of resistance, sparking speculation about underlying biological mechanisms.

Rigorously validated for robustness, this study confirms the reliability of these observations regarding potential adaptations in signaling pathways. The phenomenon of cell lines in long-term cell culture gaining sensitivity after the first six months and developing resistance in the following six months raises intriguing questions about the underlying mechanisms.

One plausible explanation could be the activation of compensatory signaling pathways or the emergence of adaptive mutations during prolonged culture. Studies suggest that prolonged exposure to specific compounds can trigger compensatory responses within cancer cells, leading to the development of resistance mechanisms as a survival strategy (277,278).

All cell images captured over time revealed stable morphology, likely due to cellular plasticity and adaptability, a phenomenon observed in various cancer cell lines under prolonged culture conditions (279–281). This stability may indicate preserved cellular homeostasis, highlighting the resilience of these cell lines to environmental changes and extended culturing (282). Notably, a subjective observation revealed that as time progressed, all three cell lines had an increased need for more frequent passaging and medium changes, leading to questions about their metabolic demands and growth characteristics. Although specific proliferation or metabolomic assays were not conducted to delve into this instance, this subjective observation underscores the dynamic nature of cell behaviour over extended culture periods. This trend aligns with existing literature suggesting that prolonged culture conditions can induce changes in cellular metabolism and growth patterns (2,3,6). Understanding these dynamics is crucial for comprehending the factors influencing cell behaviour during extended culture, offering valuable insights for optimizing cell culture conditions in research and potential clinical applications.

Moving forward, further exploration into the biological mechanisms driving these differences, including factors influencing drug response and cell cycle regulation, will be imperative for advancing the understanding of pediatric sarcomas and guiding targeted therapeutic strategies. While this thesis primarily focused on genomic and transcriptomic analyses, understanding the correlation of these findings with biological pathways, and conducting functional validation studies will be essential to elucidate the underlying mechanisms behind the observed phenotypic differences.

The analysis of different EwS cell lines revealed a spectrum of stability across various molecular profiling layers, including genomic, transcriptomic, and epigenomic analyses. This variability underscores the complex nature of tumor biology and highlights the dynamic response to environmental conditions. The observed differences in stability or

evolution among cell lines could stem from factors such as inherent genetic variability, variations in cell culture techniques, and interactions with the tumor microenvironment. The genomic stability of each cell line may be affected by its unique genetic landscape, which includes not only known *FET::ETS* rearrangements but also additional cell line-specific mutations and alterations in copy numbers (8,283,284). For instance, the A-673 cell line is known in EwS research due to an additional activating *BRAF* mutation, which may influence its stability and evolution dynamics (285,286). These findings have significant implications for cancer research and therapeutic development. Understanding the spectrum of stability within a cancer entity can inform the selection of appropriate cell line models for studying specific aspects of tumor biology or testing therapeutic interventions. For example, identifying highly stable cell lines, such as TC-71, may offer advantages in longitudinal studies or preclinical drug screening assays, where consistent molecular profiles and drug response patterns are desirable. Conversely, cell lines exhibiting greater variability, such as SK-ES-1, provide valuable insights into the plasticity of cancer cells and potential mechanisms underlying drug resistance or sensitivity shifts.

In conclusion, this thesis contributes to the ongoing discourse on the reproducibility of scientific findings, particularly in cancer research. By thoroughly examining the stability and heterogeneity of commonly used cell line models in preclinical studies, light is shed on the intricate dynamics driving genomic, transcriptomic, and phenotypic variations across different cancer entities. The robust stability demonstrated in COTF-driven pediatric sarcoma cells, as evidenced by the A-673 and TC-71 cell lines, underscores their potential as dependable models for investigating cancer biology and assessing therapeutic interventions. This discovery carries profound implications for the design and interpretation of preclinical studies, suggesting that certain cell line models may

offer more consistent and reproducible results, thus enhancing the translatability of research findings to clinical applications. This thesis also underscores the intrinsic diversity and instability inherent in adult carcinoma cell lines like HeLa and MCF-7. Despite their extensive utilization in cancer research, the findings suggest that their genomic and phenotypic variability could present challenges in reproducibility and result interpretation. This emphasizes the necessity of comprehending the specific attributes of cell line models and taking their genetic background into account when planning experiments and analyzing outcomes.

9 CONCLUSION, LIMITATIONS AND PERSPECTIVES

This thesis delves into the stability and heterogeneity of commonly used cancer cell line models, shedding light on the reproducibility and reliability of preclinical studies. Through comprehensive analyses encompassing genomic, epigenomic, transcriptomic, and phenotypic aspects of A-673 EwS cells and adult carcinoma cell lines HeLa and MCF-7, insights into the dynamics of cellular behavior in prolonged culture are provided. The primary aim was to assess the transferability of previous findings, primarily focusing on adult carcinoma cell lines, to other cancer cell types, particularly those with a stable genetic background driven by specific oncogenic drivers like the COTF found in EwS. By comparing these distinct cell line models, this thesis addresses the critical need for reliable and reproducible preclinical research tools in cancer studies.

The findings suggest that the instability and heterogeneity observed in adult carcinoma cell lines may not necessarily apply to other cancer cell types. COTF-driven pediatric sarcoma cells, represented by the A-673 and TC-71 cell lines, demonstrate remarkable stability in genomic, transcriptomic, and phenotypic profiles, even following extended culture periods and genetic modifications. This underscores the potential of these cell line models as reliable tools for cancer research, offering more consistent and reproducible results compared to their adult carcinoma counterparts.

However, this thesis also underscores the challenges associated with utilizing adult carcinoma cell lines in preclinical research. The substantial variability and instability observed in HeLa and MCF-7 cells emphasize the importance of understanding specific cell line characteristics and considering genetic background when interpreting experimental results.

Caution should be exercised when extrapolating findings from these cell lines to other cancer types or clinical settings due to their genomic and phenotypic heterogeneity potentially introducing confounding factors that impact result reproducibility.

Additionally, this study did not correlate genomic findings with specific biological pathways or functional outcomes, which may yield additional information to elucidate the biological significance of identified alterations. Furthermore, employing single-cell sequencing techniques and functional validation studies using CRISPR-based gene editing and RNA interference approaches could provide a more nuanced understanding of cell line heterogeneity and validate the biological significance of identified alterations. Integrating multi-omics data from genomic, transcriptomic, epigenomic, and proteomic analyses, as well as metabolomics profiling, could offer a comprehensive view of the molecular landscape of pediatric sarcomas and provide deeper insights into the biological mechanisms underlying these tumors. Addressing these limitations should be a priority for future investigations to enhance the precision and reliability of genomic analyses in cancer cell line research.

However, interpreting the functional relevance of molecular findings and correlating them with phenotypic outcomes presents a significant challenge. Future research should aim to elucidate the functional consequences of molecular alterations observed across different layers and their implications for cancer biology and therapeutic response.

Furthermore, the research underscores the imperative for enhanced transparency and standardization in data analysis methods to bolster data reproducibility. This objective can be achieved through the development of user-friendly omics analysis tools and platforms that adhere strictly to best practices and established guidelines.

Conversely, collaboration and knowledge sharing within the scientific community can help overcome challenges associated with cell line variability, advancing the understanding of cancer biology, and leading to enhanced therapeutic strategies for cancer patients.

Moving forward, addressing these limitations, and continuing to explore the underlying mechanisms driving cell line stability and heterogeneity will be essential. Standardizing protocols, improving bioinformatic expertise, and integrating multi-omics data from genomic, transcriptomic, epigenomic, and proteomic analyses could enhance the understanding of cancer biology and inform the development of effective therapeutic strategies.

10 REFERENCES

1. Sharma SV, Haber DA, Settleman J. Cell line-based platforms to evaluate the therapeutic efficacy of candidate anticancer agents. Nat Rev Cancer. 2010 Apr;10(4):241–53.

2. Ben-David U, Siranosian B, Ha G, Tang H, Oren Y, Hinohara K, et al. Genetic and transcriptional evolution alters cancer cell line drug response. Nature [Internet]. 2018 Aug [cited 2024 Feb 28];560(7718):325–30. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6522222/

3. Liu Y, Mi Y, Mueller T, Kreibich S, Williams EG, Van Drogen A, et al. Multi-omic measurements of heterogeneity in HeLa cells across laboratories. Nat Biotechnol. 2019 Mar;37(3):314–22.

4. Errington TM, Denis A, Perfito N, Iorns E, Nosek BA. Challenges for assessing replicability in preclinical cancer biology. Rodgers P, Franco E, editors. eLife [Internet]. 2021 [cited 2024 Feb 28];10:e67995. Available from: https://doi.org/10.7554/eLife.67995

5. Souren NY, Fusenig NE, Heck S, Dirks WG, Capes-Davis A, Bianchini F, et al. Cell line authentication: a necessity for reproducible biomedical research. EMBO J. 2022 Jul 18;41(14):e111307.

6. Cooper JR, Abdullatif MB, Burnett EC, Kempsell KE, Conforti F, Tolley H, et al. Long Term Culture of the A549 Cancer Cell Line Promotes Multilamellar Body Formation and Differentiation towards an Alveolar Type II Pneumocyte Phenotype. PLOS ONE [Internet]. 2016 Oct 28 [cited 2024 Mar 2];11(10):e0164438. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0164438

7. Brohl AS, Solomon DA, Chang W, Wang J, Song Y, Sindiri S, et al. The genomic landscape of the Ewing Sarcoma family of tumors reveals recurrent STAG2 mutation. PLoS Genet. 2014 Jul;10(7):e1004475.

8. Crompton BD, Stewart C, Taylor-Weiner A, Alexe G, Kurek KC, Calicchio ML, et al. The genomic landscape of pediatric Ewing sarcoma. Cancer Discov. 2014 Nov;4(11):1326–41.

9. Grünewald TGP, Cidre-Aranaz F, Surdez D, Tomazou EM, de Álava E, Kovar H, et al. Ewing sarcoma. Nat Rev Dis Primer. 2018 Jul 5;4(1):5.

10. Vibert J, Saulnier O, Collin C, Petit F, Borgman KJE, Vigneau J, et al. Oncogenic chimeric transcription factors drive tumor-specific transcription, processing, and translation of silent genomic regions. Mol Cell. 2022 Jul 7;82(13):2458-2471.e9.

11. Sajjad H, Imtiaz S, Noor T, Siddiqui YH, Sajjad A, Zia M. Cancer models in preclinical research: A chronicle review of advancement in effective cancer research. Anim Models Exp Med [Internet]. 2021 Mar 29 [cited 2024 Feb 27];4(2):87–103. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8212826/

12. Schwab ED, Pienta KJ. Cancer as a complex adaptive system. Med Hypotheses. 1996 Sep;47(3):235–41.

13. Koutsogiannouli E, Papavassiliou AG, Papanikolaou NA. Complexity in cancer biology: is systems biology the answer? Cancer Med. 2013 Apr;2(2):164–77.

14. Health (US) NI of, Study BSC. Understanding Cancer. In: NIH Curriculum Supplement Series [Internet] [Internet]. National Institutes of Health (US); 2007 [cited 2024 Feb 27]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK20362/

15. de Martel C, Georges D, Bray F, Ferlay J, Clifford GM. Global burden of cancer attributable to infections in 2018: a worldwide incidence analysis. Lancet Glob Health. 2020 Feb;8(2):e180–90.

16. WHO. Global cancer burden growing, amidst mounting need for services [Internet]. 2024 [cited 2024 Mar 20]. Available from: https://www.who.int/news/item/01-02-2024-global-cancer-burden-growing--amidst-mounting-need-for-services

17. Yazdanpanah N, Dochy F, Darmstadt GL, Peters GJ, Tsitlakidis A, Aifantis EC, et al. Cancer: A Complex Problem Requiring Interdisciplinary Research. In: Rezaei N, editor. Cancer Treatment: An Interdisciplinary Approach [Internet]. Cham: Springer Nature Switzerland; 2023 [cited 2024 Mar 26]. p. 1–45. Available from: https://doi.org/10.1007/16833_2022_116

18. Di Lonardo A, Nasi S, Pulciani S. Cancer: we should not forget the past. J Cancer. 2015;6(1):29–39.

19. Gagan J, Van Allen EM. Next-generation sequencing to guide cancer therapy. Genome Med [Internet]. 2015 Jul 29 [cited 2024 Mar 26];7(1):80. Available from: https://doi.org/10.1186/s13073-015-0203-x

20. Colomer R, Miranda J, Romero-Laorden N, Hornedo J, González-Cortijo L, Mouron S, et al. Usefulness and real-world outcomes of next generation sequencing testing in patients with cancer: an observational study on the impact of selection based on clinical judgement. eClinicalMedicine [Internet]. 2023 Jun 1 [cited 2024 Mar 26];60. Available from: https://www.thelancet.com/journals/eclinm/article/PIIS2589-5370(23)00206-7/fulltext

21. Golebiewska A, Fields RC. Advancing preclinical cancer models to assess clinically relevant outcomes. BMC Cancer [Internet]. 2023 Mar 10 [cited 2024 Feb 27];23(1):230. Available from: https://doi.org/10.1186/s12885-023-10715-7

22. Errington TM, Iorns E, Gunn W, Tan FE, Lomax J, Nosek BA. An open investigation of the reproducibility of cancer biology research. eLife [Internet]. 2014 [cited 2024 Feb 28];3:e04333. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4270077/

23. Amalinei C, Grigoraș A, Pricope DL, Pricop BI. Cancer Study: Cell to the Animal Models. In: Pathak S, Banerjee A, Bisgin A, editors. Handbook of Animal Models and its Uses in Cancer Research [Internet]. Singapore: Springer Nature; 2022 [cited 2024 Mar 20]. p. 1–23. Available from: https://doi.org/10.1007/978-981-19-1282-5_2-1

24. Weigelt B, Ghajar CM, Bissell MJ. The need for complex 3D culture models to unravel novel pathways and identify accurate biomarkers in breast cancer. Adv Drug Deliv Rev. 2014 Apr;69–70:42–51.

25. Begley CG, Ellis LM. Raise standards for preclinical cancer research. Nature [Internet]. 2012 Mar [cited 2024 Mar 26];483(7391):531–3. Available from: https://www.nature.com/articles/483531a

26. Reyes-Aldasoro CC. Modelling the Tumour Microenvironment, but What Exactly Do We Mean by “Model”? Cancers [Internet]. 2023 Jul 26 [cited 2024 Mar 20];15(15):3796. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10416922/

27. Costanza R, Wainger L, Folke C, Mäler KG. Modeling Complex Ecological Economic Systems: Toward an evolutionary, dynamic understanding of people and nature. BioScience [Internet]. 1993 Sep 1 [cited 2024 Feb 27];43(8):545–55. Available from: https://doi.org/10.2307/1311949

28. Thomas RM, Van Dyke T, Merlino G, Day CP. Concepts in Cancer Modeling: A Brief History. Cancer Res [Internet]. 2016 Oct 15 [cited 2024 Feb 27];76(20):5921–5. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5117812/

29. Yamagiwa K, Murayama K. Summary of the Results of Experiments on the Pathogenesis of Epithelial Growths1: I. The Experimental Production of Mammary Carcinoma on Rabbits. J Cancer Res [Internet]. 1924 Apr 1 [cited 2024 Feb 27];8(1):119–36. Available from: https://doi.org/10.1158/jcr.1924.119

30. Li L, Wazir J, Huang Z, Wang Y, Wang H. A Comprehensive Review of Animal Models for Cancer Cachexia: Implications for Translational Research. Genes Dis [Internet]. 2023 Sep 13 [cited 2024 Feb 27];101080. Available from: https://www.sciencedirect.com/science/article/pii/S235230422300363X

31. Tan JHL, Hwang YY, Chin HX, Liu M, Tan SY, Chen Q. Towards a better preclinical cancer model – human immune aging in humanized mice. Immun Ageing [Internet]. 2023 Sep 27 [cited 2024 Feb 27];20(1):49. Available from: https://doi.org/10.1186/s12979-023-00374-4

32. Zhou Y, Xia J, Xu S, She T, Zhang Y, Sun Y, et al. Experimental mouse models for translational human cancer research. Front Immunol. 2023;14:1095388.

33. Gey GO, Coffman WD, Kubicek MT. Tissue culture studies of the proliferative capacity of cervical carcinoma and normal epithelium. Gey G.O., Coffman W.D. and Kubicek M.T. Cancer Res. 1952;12.

34. Lucey BP, Nelson-Rees WA, Hutchins GM. Henrietta Lacks, HeLa cells, and cell culture contamination. Arch Pathol Lab Med. 2009 Sep;133(9):1463–7.

35. Scherer WF, Syverton JT, Gey GO. Studies on the propagation in vitro of poliomyelitis viruses. IV. Viral multiplication in a stable strain of human malignant epithelial cells (strain HeLa) derived from an epidermoid carcinoma of the cervix. J Exp Med. 1953 May;97(5):695–710.

36. Masters JR. HeLa cells 50 years on: the good, the bad and the ugly. Nat Rev Cancer. 2002 Apr;2(4):315–9.

37. Landry JJM, Pyl PT, Rausch T, Zichner T, Tekkedil MM, Stütz AM, et al. The Genomic and Transcriptomic Landscape of a HeLa Cell Line. G3 GenesGenomesGenetics [Internet]. 2013 Aug 1 [cited 2024 Mar 6];3(8):1213–24. Available from: https://doi.org/10.1534/g3.113.005777

38. Merriam-Webster. Merriam-Webster.com Dictionary. 2024 [cited 2024 Mar 26]. in vivo. Available from: https://www.merriam-webster.com/dictionary/in%20vivo

39. Kandir S. Animal Models for Cancer Research: The Choice of the Right Model System. In: Pathak S, Banerjee A, Bisgin A, editors. Handbook of Animal Models and its Uses in Cancer Research [Internet]. Singapore: Springer Nature; 2022 [cited 2024 Mar 20]. p. 1–16. Available from: https://doi.org/10.1007/978-981-19-1282-5_3-1

40. Yee NS, Ignatenko N, Finnberg N, Lee N, Stairs D. ANIMAL MODELS OF CANCER BIOLOGY. Cancer Growth Metastasis [Internet]. 2015 Dec 10 [cited 2024 Feb 27];8(Suppl 1):115–8. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4676433/

41. Skipper HE, Schabel FM, Wilcox WS. EXPERIMENTAL EVALUATION OF POTENTIAL ANTICANCER AGENTS. XIII. ON THE CRITERIA AND KINETICS ASSOCIATED WITH “CURABILITY” OF EXPERIMENTAL LEUKEMIA. Cancer Chemother Rep. 1964 Feb;35:1–111.

42. Skipper HE, Schabel FM, Wilcox WS. EXPERIMENTAL EVALUATION OF POTENTIAL ANTICANCER AGENTS. XIV. FURTHER STUDY OF CERTAIN BASIC CONCEPTS UNDERLYING CHEMOTHERAPY OF LEUKEMIA. Cancer Chemother Rep. 1965 Apr;45:5–28.

43. Skipper HE, Perry S. Kinetics of normal and leukemic leukocyte populations and relevance to chemotherapy. Cancer Res. 1970 Jun;30(6):1883–97.

44. Hart IR, Fidler IJ. Cancer invasion and metastasis. Q Rev Biol. 1980 Jun;55(2):121–42.

45. Siddiqui SS, Loganathan S, Krishnaswamy S, Faoro L, Jagadeeswaran R, Salgia R. C. elegans as a model organism for in vivo screening in cancer: effects of human c-Met in lung cancer affect C. elegans vulva phenotypes. Cancer Biol Ther [Internet]. 2008 Jun 1 [cited 2024 Feb 27];7(6):856–63. Available from: https://doi.org/10.4161/cbt.7.6.5842

46. White R, Rose K, Zon L. Zebrafish cancer: the state of the art and the path forward. Nat Rev Cancer [Internet]. 2013 Sep [cited 2024 Feb 27];13(9):624–36. Available from: https://www.nature.com/articles/nrc3589).

47. Yusuf K, Umar S, Ahmed I. Animal Models in Cancer Research. In: Pathak S, Banerjee A, Bisgin A, editors. Handbook of Animal Models and its Uses in Cancer Research [Internet]. Singapore: Springer Nature; 2022 [cited 2024 Mar 20]. p. 1–20. Available from: https://doi.org/10.1007/978-981-19-1282-5_17-1

48. Cheon DJ, Orsulic S. Mouse models of cancer. Annu Rev Pathol. 2011;6:95–119.

49. van der Mijn JC, Laursen KB, Fu L, Khani F, Dow LE, Nowak DG, et al. Novel genetically engineered mouse models for clear cell renal cell carcinoma. Sci Rep [Internet]. 2023 May 22 [cited 2024 Mar 26];13(1):8246. Available from: https://www.nature.com/articles/s41598-023-35106-7

50. Jin J, Yoshimura K, Sewastjanow-Silva M, Song S, Ajani JA. Challenges and Prospects of Patient-Derived Xenografts for Cancer Research. Cancers. 2023 Aug 31;15(17):4352.

51. Merriam-Webster. “In vitro.” [Internet]. 2024 [cited 2024 Mar 26]. Available from: https://www.merriam-webster.com/dictionary/in%20vitro. Accessed 26 Mar. 2024.

52. Harrison RG, Greenman MJ, Mall FP, Jackson CM. Observations of the living developing nerve fiber. Anat Rec [Internet]. 1907 [cited 2024 Feb 27];1(5):116–28. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/ar.1090010503

53. Breslin S, O’Driscoll L. Three-dimensional cell culture: the missing link in drug discovery. Drug Discov Today. 2013 Mar;18(5–6):240–9.

54. Souza AG, Marangoni K, Fujimura PT, Alves PT, Silva MJ, Bastos VAF, et al. 3D Cell-SELEX: Development of RNA aptamers as molecular probes for PC-3 tumor cell line. Exp Cell Res. 2016 Feb 15;341(2):147–56.

55. Kapałczyńska M, Kolenda T, Przybyła W, Zajączkowska M, Teresiak A, Filas V, et al. 2D and 3D cell cultures – a comparison of different types of cancer cell cultures. Arch Med Sci AMS [Internet]. 2018 Jun [cited 2024 Feb 27];14(4):910–9. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6040128/

56. Schaeffer WI, Simkins SG, Lamore BJ. A system employing a minimal defined medium for the selection of tumorigenic cells. Vitro Cell Dev Biol J Tissue Cult Assoc. 1990 Jul;26(7):737–40.

57. Freshney RI. Basic Principles of Cell culture. In: Culture of Cell for Tissue Engineering. John Wiley & Sons, Inc; 2006.

58. Carrel A, Burrows MT. Cultivation of Adult Tissues and Organs Outside of the Body: The First Note [Internet]. American Medical Association; 1910. Available from: https://books.google.de/books?id=96wXnQEACAAJ

59. Rous P, Jones FS. A METHOD FOR OBTAINING SUSPENSIONS OF LIVING CELLS FROM THE FIXED TISSUES, AND FOR THE PLATING OUT OF INDIVIDUAL CELLS. J Exp Med. 1916 Apr 1;23(4):549–55.

60. Dulbecco R. Production of Plaques in Monolayer Tissue Cultures by Single Particles of an Animal Virus. Proc Natl Acad Sci U S A. 1952 Aug;38(8):747–52.

61. Kellner G, Broda E, Suschny O, Rücker W. Effects of trypsin treatment on tissue in culture. Exp Cell Res [Internet]. 1959 Aug 1 [cited 2024 Feb 27];18(1):168–71. Available from: https://www.sciencedirect.com/science/article/pii/0014482759903003

62. Furger C. Live Cell Assays: From Research to Regulatory Applications. John Wiley & Sons; 2016. 286 p.

63. Marquis CP. Mammalian Cell Culture. In: Biotechnology. Encyclopedia of Life Support Systems (EOLSS); 2022.

64. Moore GE, Mount D, Tara G, Schwartz N. GROWTH OF HUMAN TUMOR CELLS IN SUSPENSION CULTURES. Cancer Res. 1963 Dec;23:1735–41.

65. Weinberg RA. Coevolution in the tumor microenvironment. Nat Genet. 2008 May;40(5):494–5.

66. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011 Mar 4;144(5):646–74.

67. Hanahan D, Coussens LM. Accessories to the crime: functions of cells recruited to the tumor microenvironment. Cancer Cell. 2012 Mar 20;21(3):309–22.

68. Spill F, Reynolds DS, Kamm RD, Zaman MH. Impact of the physical microenvironment on tumor progression and metastasis. Curr Opin Biotechnol. 2016 Aug;40:41–8.

69. Fatehullah A, Tan SH, Barker N. Organoids as an in vitro model of human development and disease. Nat Cell Biol. 2016 Mar;18(3):246–54.

70. Zanoni M, Piccinini F, Arienti C, Zamagni A, Santi S, Polico R, et al. 3D tumor spheroid models for in vitro therapeutic screening: a systematic approach to enhance the biological relevance of data obtained. Sci Rep. 2016 Jan 11;6:19103.

71. Di Modugno F, Colosi C, Trono P, Antonacci G, Ruocco G, Nisticò P. 3D models in the new era of immune oncology: focus on T cells, CAF and ECM. J Exp Clin Cancer Res [Internet]. 2019 Mar 22 [cited 2024 Feb 27];38(1):117. Available from: https://doi.org/10.1186/s13046-019-1086-2

72. Sutherland S, Egger B, Tenenbaum J. Building 3D Generative Models from Minimal Data. Int J Comput Vis [Internet]. 2024 Feb 1 [cited 2024 Feb 27];132(2):555–80. Available from: https://doi.org/10.1007/s11263-023-01870-2

73. Gillet JP, Varma S, Gottesman MM. The clinical relevance of cancer cell lines. J Natl Cancer Inst. 2013 Apr 3;105(7):452–8.

74. Dai X, Cheng H, Bai Z, Li J. Breast Cancer Cell Line Classification and Its Relevance with Breast Tumor Subtyping. J Cancer. 2017;8(16):3131–41.

75. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell. 2000 Jan 7;100(1):57–70.

76. Hanahan D. Hallmarks of Cancer: New Dimensions. Cancer Discov. 2022 Jan;12(1):31–46.

77. Garraway LA, Lander ES. Lessons from the cancer genome. Cell. 2013 Mar 28;153(1):17–37.

78. Mooney SD. Progress Towards the Integration of Pharmacogenomics in Practice. Hum Genet [Internet]. 2015 May [cited 2024 Feb 27];134(5):459–65. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4362928/

79. Watson JD, Crick FHC. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. Nature [Internet]. 1953 Apr [cited 2024 Feb 27];171(4356):737–8. Available from: https://www.nature.com/articles/171737a0

80. Stretton AOW. The first sequence. Fred Sanger and insulin. Genetics [Internet]. 2002 Oct [cited 2024 Feb 27];162(2):527–32. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1462286/

81. Min Jou W, Haegeman G, Ysebaert M, Fiers W. Nucleotide sequence of the gene coding for the bacteriophage MS2 coat protein. Nature. 1972 May 12;237(5350):82–8.

82. Padmanabhan R, Wu R. Nucleotide sequence analysis of DNA. IV. Complete nucleotide sequence of the left-hand cohesive end of coliphage 186 DNA. J Mol Biol. 1972 Apr 14;65(3):447–67.

83. Padmanabhan R, Wu R. Nucleotide sequence analysis of DNA. IX. Use of oligonucleotides of defined sequence as primers in DNA sequence analysis. Biochem Biophys Res Commun. 1972 Sep 5;48(5):1295–302.

84. Padmanabhan R, Wu R, Bode VC. Arrangement of DNA in lambda bacteriophage heads. 3. Location and number of nucleotides cleaved from lambda-DNA by micrococcal nuclease attack on heads. J Mol Biol. 1972 Aug 21;69(2):201–7.

85. Fiers W, Contreras R, Haegemann G, Rogiers R, Van de Voorde A, Van Heuverswyn H, et al. Complete nucleotide sequence of SV40 DNA. Nature. 1978 May 11;273(5658):113–20.

86. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A. 1977 Dec;74(12):5463–7.

87. Sanger F. Determination of nucleotide sequences in DNA. Science. 1981 Dec 11;214(4526):1205–10.

88. Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, et al. Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. Science. 1995 Jul 28;269(5223):496–512.

89. Collins FS, Fink L. The Human Genome Project. Alcohol Health Res World [Internet]. 1995 [cited 2024 Mar 2];19(3):190–5. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6875757/

90. Karger BL, Guttman A. DNA Sequencing by Capillary Electrophoresis. Electrophoresis [Internet]. 2009 Jun [cited 2024 Mar 2];30(Suppl 1):S196–202. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2782523/

91. Shendure J, Lieberman Aiden E. The expanding scope of DNA sequencing. Nat Biotechnol. 2012 Nov;30(11):1084–94.

92. Hood L, Rowen L. The Human Genome Project: big science transforms biology and medicine. Genome Med [Internet]. 2013 Sep 13 [cited 2024 Mar 2];5(9):79. Available from: https://doi.org/10.1186/gm483

93. Human Genome Project Fact Sheet [Internet]. [cited 2024 Mar 2]. Available from: https://www.genome.gov/about-genomics/educational-resources/fact-sheets/human-genome-project

94. Human Genome Project: Sequencing the Human Genome | Learn Science at Scitable [Internet]. [cited 2024 Mar 2]. Available from: http://www.nature.com/scitable/topicpage/dna-sequencing-technologies-key-to-the-human-828

95. Ronaghi M, Karamohamed S, Pettersson B, Uhlén M, Nyrén P. Real-time DNA sequencing using detection of pyrophosphate release. Anal Biochem. 1996 Nov 1;242(1):84–9.

96. Brenner S, Johnson M, Bridgham J, Golda G, Lloyd DH, Johnson D, et al. Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. Nat Biotechnol. 2000 Jun;18(6):630–4.

97. Goodwin S, McPherson JD, McCombie WR. Coming of age: ten years of next-generation sequencing technologies. Nat Rev Genet [Internet]. 2016 Jun [cited 2024 Feb 27];17(6):333–51. Available from: https://www.nature.com/articles/nrg.2016.49

98. Morganti S, Tarantino P, Ferraro E, D'Amico P, Viale G, Trapani D, et al. Role of Next-Generation Sequencing Technologies in Personalized Medicine. In: Pravettoni G, Triberti S, editors. P5 eHealth: An Agenda for the Health Technologies of the Future [Internet]. Cham: Springer International Publishing; 2019 [cited 2024 Feb 27]. p. 125–54. Available from: https://doi.org/10.1007/978-3-030-27994-3_8

99. Green RC, Berg JS, Grody WW, Kalia SS, Korf BR, Martin CL, et al. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. Genet Med [Internet]. 2013 Jul [cited 2024 Feb 27];15(7):565–74. Available from: https://www.nature.com/articles/gim201373

100. Arnedos M, Vicier C, Loi S, Lefebvre C, Michiels S, Bonnefoi H, et al. Precision medicine for metastatic breast cancer--limitations and solutions. Nat Rev Clin Oncol. 2015 Dec;12(12):693–704.

101. Aaltonen LA, Abascal F, Abeshouse A, Aburatani H, Adams DJ, Agrawal N, et al. Pan-cancer analysis of whole genomes. Nature [Internet]. 2020 Feb [cited 2024 Feb 27];578(7793):82–93. Available from: https://www.nature.com/articles/s41586-020-1969-6

102. The Cancer Genome Atlas Program (TCGA) - NCI [Internet]. 2022 [cited 2024 Feb 27]. Available from: https://www.cancer.gov/ccg/research/genome-sequencing/tcga

103. Cancer Genome Project [Internet]. [cited 2024 Feb 27]. Available from: https://www.sanger.ac.uk/group/cancer-genome-project/

104. Sosinsky A, Ambrose J, Cross W, Turnbull C, Henderson S, Jones L, et al. Insights for precision oncology from the integration of genomic and clinical data of 13,880 tumors from the 100,000 Genomes Cancer Programme. Nat Med [Internet]. 2024 Jan [cited 2024 Feb 27];30(1):279–89. Available from: https://www.nature.com/articles/s41591-023-02682-0

105. Hulsen T. Sharing Is Caring—Data Sharing Initiatives in Healthcare. Int J Environ Res Public Health [Internet]. 2020 Jan [cited 2024 Feb 27];17(9):3046. Available from: https://www.mdpi.com/1660-4601/17/9/3046

106. ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium. Pan-cancer analysis of whole genomes. Nature. 2020 Feb;578(7793):82–93.

107. Mardis ER. Next-generation DNA sequencing methods. Annu Rev Genomics Hum Genet. 2008;9:387–402.

108. Meyerson M, Gabriel S, Getz G. Advances in understanding cancer genomes through second-generation sequencing. Nat Rev Genet. 2010 Oct;11(10):685–96.

109. Ng PC, Kirkness EF. Whole genome sequencing. Methods Mol Biol Clifton NJ. 2010;628:215–26.

110. Gerlinger M, Horswell S, Larkin J, Rowan AJ, Salm MP, Varela I, et al. Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. Nat Genet. 2014 Mar;46(3):225–33.

111. Devarajan P, Chertow GM, Susztak K, Levin A, Agarwal R, Stenvinkel P, et al. Emerging Role of Clinical Genetics in CKD. Kidney Med. 2022 Apr;4(4):100435.

112. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet. 2009 Jan;10(1):57–63.

113. Alwine JC, Kemp DJ, Stark GR. Method for detection of specific RNAs in agarose gels by transfer to diazobenzyloxymethyl-paper and hybridization with DNA probes. Proc Natl Acad Sci U S A. 1977 Dec;74(12):5350–4.

114. Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science. 1995 Oct 20;270(5235):467–70.

115. Quackenbush J. Microarray data normalization and transformation. Nat Genet. 2002 Dec;32 Suppl:496–501.

116. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods. 2008 Jul;5(7):621–8.

117. Velmeshev D. Genomics Research from Technology Networks. 2019 [cited 2024 Feb 28]. Recent Advances in Single-Cell Genomics Techniques. Available from: http://www.technologynetworks.com/genomics/articles/recent-advances-in-single-cell-genomics-techniques-324695

118. Zhao S, Fung-Leung WP, Bittner A, Ngo K, Liu X. Comparison of RNA-Seq and microarray in transcriptome profiling of activated T cells. PloS One. 2014;9(1):e78644.

119. Wang ET, Sandberg R, Luo S, Khrebtukova I, Zhang L, Mayr C, et al. Alternative isoform regulation in human tissue transcriptomes. Nature. 2008 Nov 27;456(7221):470–6.

120. Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat Genet [Internet]. 2008 Dec [cited 2024 Mar 26];40(12):1413–5. Available from: https://www.nature.com/articles/ng.259

121. Tang F, Lao K, Surani MA. Development and applications of single-cell transcriptome analysis. Nat Methods. 2011 Apr;8(4 Suppl):S6-11.

122. Picelli S, Björklund ÅK, Faridani OR, Sagasser S, Winberg G, Sandberg R. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nat Methods [Internet]. 2013 Nov [cited 2024 Feb 28];10(11):1096–8. Available from: https://www.nature.com/articles/nmeth.2639

123. Cao J, Packer JS, Ramani V, Cusanovich DA, Huynh C, Daza R, et al. Comprehensive single-cell transcriptional profiling of a multicellular organism. Science. 2017 Aug 18;357(6352):661–7.

124. Takai D, Gonzales FA, Tsai YC, Thayer MJ, Jones PA. Large scale mapping of methylcytosines in CTCF-binding sites in the human H19 promoter and aberrant hypomethylation in human bladder cancer. Hum Mol Genet. 2001 Nov 1;10(23):2619–26.

125. Cedar H, Bergman Y. Linking DNA methylation and histone modification: patterns and paradigms. Nat Rev Genet. 2009 May;10(5):295–304.

126. Kouzarides T. Chromatin modifications and their function. Cell. 2007 Feb 23;128(4):693–705.

127. Allis CD, Berger SL, Cote J, Dent S, Jenuwien T, Kouzarides T, et al. New nomenclature for chromatin-modifying enzymes. Cell. 2007 Nov 16;131(4):633–6.

128. Goldberg AD, Allis CD, Bernstein E. Epigenetics: a landscape takes shape. Cell. 2007 Feb 23;128(4):635–8.

129. Dawson MA, Kouzarides T. Cancer epigenetics: from mechanism to therapy. Cell. 2012 Jul 6;150(1):12–27.

130. Grosselin K, Durand A, Marsolier J, Poitou A, Marangoni E, Nemati F, et al. High-throughput single-cell ChIP-seq identifies heterogeneity of chromatin states in breast cancer. Nat Genet [Internet]. 2019 Jun [cited 2024 Feb 28];51(6):1060–6. Available from: https://www.nature.com/articles/s41588-019-0424-9

131. Gopi LK, Kidder BL. Integrative pan cancer analysis reveals epigenomic variation in cancer type and cell specific chromatin domains. Nat Commun [Internet]. 2021

Mar 3 [cited 2024 Feb 28];12(1):1419. Available from: https://www.nature.com/articles/s41467-021-21707-1

132. Beltrán-García J, Osca-Verdegal R, Mena-Mollá S, García-Giménez JL. Epigenetic IVD Tests for Personalized Precision Medicine in Cancer. Front Genet. 2019;10:621.

133. Garcia-Martinez L, Zhang Y, Nakata Y, Chan HL, Morey L. Epigenetic mechanisms in breast cancer therapy and resistance. Nat Commun [Internet]. 2021 Mar 19 [cited 2024 Feb 28];12(1):1786. Available from: https://www.nature.com/articles/s41467-021-22024-3

134. Jones PA, Takai D. The role of DNA methylation in mammalian epigenetics. Science. 2001 Aug 10;293(5532):1068–70.

135. Esteller M. Epigenetics in cancer. N Engl J Med. 2008 Mar 13;358(11):1148–59.

136. Bock C. Analysing and interpreting DNA methylation data. Nat Rev Genet [Internet]. 2012 Oct [cited 2024 Feb 28];13(10):705–19. Available from: https://www.nature.com/articles/nrg3273

137. Laird PW. Principles and challenges of genomewide DNA methylation analysis. Nat Rev Genet. 2010 Mar;11(3):191–203.

138. Ioannidis JPA, Allison DB, Ball CA, Coulibaly I, Cui X, Culhane AC, et al. Repeatability of published microarray gene expression analyses. Nat Genet. 2009 Feb;41(2):149–55.

139. Freedman LP, Cockburn IM, Simcoe TS. The Economics of Reproducibility in Preclinical Research. PLoS Biol. 2015 Jun;13(6):e1002165.

140. Errington TM, Mathur M, Soderberg CK, Denis A, Perfito N, Iorns E, et al. Investigating the replicability of preclinical cancer biology. Pasqualini R, Franco E, editors. eLife [Internet]. 2021 Dec 7 [cited 2024 Feb 28];10:e71601. Available from: https://doi.org/10.7554/eLife.71601

141. American Type Culture Collection Standards Development Organization Workgroup ASN-0002. Cell line misidentification: the beginning of the end. Nat Rev Cancer [Internet]. 2010 Jun [cited 2024 Feb 28];10(6):441–8. Available from: https://www.nature.com/articles/nrc2852

142. Capes-Davis A, Theodosopoulos G, Atkin I, Drexler HG, Kohara A, MacLeod RAF, et al. Check your cultures! A list of cross-contaminated or misidentified cell lines. Int J Cancer. 2010 Jul 1;127(1):1–8.

143. Hughes P, Marshall D, Reid Y, Parkes H, Gelber C. The costs of using unauthenticated, over-passaged cell lines: how much more data do we need? BioTechniques. 2007 Nov;43(5):575, 577–8, 581-582 passim.

144. Lorsch JR, Collins FS, Lippincott-Schwartz J. Fixing problems with cell lines. Science [Internet]. 2014 Dec 19 [cited 2024 Feb 28];346(6216):1452–3. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5101941/

145. Schweppe RE, Klopper JP, Korch C, Pugazhenthi U, Benezra M, Knauf JA, et al. Deoxyribonucleic acid profiling analysis of 40 human thyroid cancer cell lines reveals cross-contamination resulting in cell line redundancy and misidentification. J Clin Endocrinol Metab. 2008 Nov;93(11):4331–41.

146. Zhao M, Sano D, Pickering CR, Jasser SA, Henderson YC, Clayman GL, et al. Assembly and Initial Characterization of a Panel of 85 Genomically Validated Cell Lines from Diverse Head and Neck Tumor Sites. Clin Cancer Res [Internet]. 2011 Nov 30 [cited 2024 Feb 28];17(23):7248–64. Available from: https://doi.org/10.1158/1078-0432.CCR-11-0690

147. Horbach SPJM, Halffman W. The ghosts of HeLa: How cell line misidentification contaminates the scientific literature. PloS One. 2017;12(10):e0186281.

148. Korch C, Hall EM, Dirks WG, Ewing M, Faries M, Varella-Garcia M, et al. Authentication of M14 melanoma cell line proves misidentification of MDA-MB-435 breast cancer cell line. Int J Cancer. 2018 Feb 1;142(3):561–72.

149. Holliday DL, Speirs V. Choosing the right cell line for breast cancer research. Breast Cancer Res BCR. 2011 Aug 12;13(4):215.

150. Domcke S, Sinha R, Levine DA, Sander C, Schultz N. Evaluating cell lines as tumour models by comparison of genomic profiles. Nat Commun. 2013;4:2126.

151. Wass AV, Butler G, Taylor TB, Dash PR, Johnson LJ. Cancer cell lines show high heritability for motility but not generation time. R Soc Open Sci [Internet]. 2020 Apr 15 [cited 2024 Feb 28];7(4):191645. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7211847/

152. Bottomley RH, Trainer AL, Griffin MJ. ENZYMATIC AND CHROMOSOMAL CHARACTERIZATION OF HELA VARIANTS. J Cell Biol [Internet]. 1969 Jun 1 [cited 2024 Mar 6];41(3):806–15. Available from: https://doi.org/10.1083/jcb.41.3.806

153. Rutledge S. What HeLa Cells Are You Using? Authorea. 2015 Aug 7;

154. Frattini A, Fabbri M, Valli R, De Paoli E, Montalbano G, Gribaldo L, et al. High variability of genomic instability and gene expression profiling in different HeLa clones. Sci Rep. 2015 Oct 20;5:15377.

155. Adey A, Burton JN, Kitzman JO, Hiatt JB, Lewis AP, Martin BK, et al. The haplotype-resolved genome and epigenome of the aneuploid HeLa cancer cell line. Nature [Internet]. 2013 Aug [cited 2024 Mar 6];500(7461):207–11. Available from: https://www.nature.com/articles/nature12064

156. Tirode F, Surdez D, Ma X, Parker M, Le Deley MC, Bahrami A, et al. Genomic landscape of Ewing sarcoma defines an aggressive subtype with co-association of STAG2 and TP53 mutations. Cancer Discov. 2014 Nov;4(11):1342–53.

157. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018 Nov;68(6):394–424.

158. Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. CA Cancer J Clin. 2023 Jan;73(1):17–48.

159. DeSantis CE, Ma J, Gaudet MM, Newman LA, Miller KD, Goding Sauer A, et al. Breast cancer statistics, 2019. CA Cancer J Clin. 2019 Nov;69(6):438–51.

160. Perou CM, Sørlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. Nature. 2000 Aug 17;406(6797):747–52.

161. Polyak K. Breast cancer: origins and evolution. J Clin Invest. 2007 Nov;117(11):3155–63.

162. Henry NL, Shah PD, Haider I, et al. Chapter 88: Cancer of the Breast. In: Abeloff's Clinical Oncology. 6th edition. Philadelphia, Pa: Elsevier; 2020.

163. Schwartz GF, Cantor RI, Biermann WA. Neoadjuvant Chemotherapy Before Definitive Treatment for Stage III Carcinoma of the Breast. Arch Surg [Internet]. 1987 Dec 1 [cited 2024 Feb 28];122(12):1430–4. Available from: https://doi.org/10.1001/archsurg.1987.01400240078014

164. Wolff AC, Hammond MEH, Hicks DG, Dowsett M, McShane LM, Allison KH, et al. Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. J Clin Oncol Off J Am Soc Clin Oncol. 2013 Nov 1;31(31):3997–4013.

165. Breast cancer [Internet]. [cited 2024 Feb 28]. Available from: https://www.who.int/news-room/fact-sheets/detail/breast-cancer

166. Giuliano AE, Edge SB, Hortobagyi GN. Eighth Edition of the AJCC Cancer Staging Manual: Breast Cancer. Ann Surg Oncol. 2018 Jul;25(7):1783–5.

167. Gradishar WJ, Moran MS, Abraham J, Aft R, Agnese D, Allison KH, et al. Breast Cancer, Version 3.2022, NCCN Clinical Practice Guidelines in Oncology. J Natl Compr Cancer Netw JNCCN. 2022 Jun;20(6):691–722.

168. Smith I, Robertson J, Kilburn L, Wilcox M, Evans A, Holcombe C, et al. Long-term outcome and prognostic value of Ki67 after perioperative endocrine therapy in postmenopausal women with hormone-sensitive early breast cancer (POETIC): an open-label, multicentre, parallel-group, randomised, phase 3 trial. Lancet Oncol. 2020 Nov;21(11):1443–54.

169. Nik-Zainal S, Davies H, Staaf J, Ramakrishna M, Glodzik D, Zou X, et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. Nature. 2016 Jun 2;534(7605):47–54.

170. Angus L, Smid M, Wilting SM, van Riet J, Van Hoeck A, Nguyen L, et al. The genomic landscape of metastatic breast cancer highlights changes in mutation and signature frequencies. Nat Genet. 2019 Oct;51(10):1450–8.

171. Hasson SP, Menes T, Sonnenblick A. Comparison of Patient Susceptibility Genes Across Breast Cancer: Implications for Prognosis and Therapeutic Outcomes . Pharmacogenomics Pers Med [Internet]. 2020 Jul 27 [cited 2024

Mar 2];13:227–38. Available from: https://www.dovepress.com/comparison-of-patient-susceptibility-genes-across-breast-cancer-implic-peer-reviewed-fulltext-article-PGPM

172. Martelotto LG, Ng CK, Piscuoglio S, Weigelt B, Reis-Filho JS. Breast cancer intra-tumor heterogeneity. Breast Cancer Res [Internet]. 2014 May 20 [cited 2024 Feb 28];16(3):210. Available from: https://doi.org/10.1186/bcr3658

173. Lüönd F, Tiede S, Christofori G. Breast cancer as an example of tumour heterogeneity and tumour cell plasticity during malignant progression. Br J Cancer [Internet]. 2021 Jul [cited 2024 Feb 28];125(2):164–75. Available from: https://www.nature.com/articles/s41416-021-01328-7

174. Walsh T, King MC. Ten genes for inherited breast cancer. Cancer Cell. 2007 Feb;11(2):103–5.

175. Zheng W, Long J, Gao YT, Li C, Zheng Y, Xiang YB, et al. Genome-wide association study identifies a novel breast cancer susceptibility locus at 6q25.1. Nat Genet [Internet]. 2009 Mar [cited 2024 Feb 28];41(3):324–8. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2754845/

176. Han MR, Zheng W, Cai Q, Gao YT, Zheng Y, Bolla MK, et al. Evaluating genetic variants associated with breast cancer risk in high and moderate-penetrance genes in Asians. Carcinogenesis [Internet]. 2017 May 1 [cited 2024 Feb 28];38(5):511–8. Available from: https://doi.org/10.1093/carcin/bgx010

177. Breast Cancer Association Consortium, Dorling L, Carvalho S, Allen J, González-Neira A, Luccarini C, et al. Breast Cancer Risk Genes - Association Analysis in More than 113,000 Women. N Engl J Med. 2021 Feb 4;384(5):428–39.

178. Wilcox N, Dumont M, González-Neira A, Carvalho S, Joly Beauparlant C, Crotti M, et al. Exome sequencing identifies breast cancer susceptibility genes and defines the contribution of coding variants to breast cancer risk. Nat Genet [Internet]. 2023 Sep [cited 2024 Feb 28];55(9):1435–9. Available from: https://www.nature.com/articles/s41588-023-01466-z

179. Kalinsky K, Jacks LM, Heguy A, Patil S, Drobnjak M, Bhanot UK, et al. PIK3CA mutation associates with improved outcome in breast cancer. Clin Cancer Res Off J Am Assoc Cancer Res. 2009 Aug 15;15(16):5049–59.

180. Freitag CE, Mei P, Wei L, Parwani AV, Li Z. Genetic alterations and their association with clinicopathologic characteristics in advanced breast carcinomas: focusing on clinically actionable genetic alterations. Hum Pathol. 2020 Aug;102:94–103.

181. Vousden KH, Lane DP. p53 in health and disease - PubMed. Nat Rev Mol Cell Biol [Internet]. 2007 [cited 2024 Mar 26];8(4)(275–283). Available from: https://pubmed.ncbi.nlm.nih.gov/17380161/

182. Yuan T, Cantley L. PI3K pathway alterations in cancer: variations on a theme. Oncogene [Internet]. 2008 Sep 18 [cited 2024 Mar 26];27(41):5497–510. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3398461/

183. Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. Science. 1987 Jan 9;235(4785):177–82.

184. Sircoulomb F, Bekhouche I, Finetti P, Adélaïde J, Hamida AB, Bonansea J, et al. Genome profiling of ERBB2-amplified breast cancers. BMC Cancer [Internet]. 2010 Oct 8 [cited 2024 Feb 28];10(1):539. Available from: https://doi.org/10.1186/1471-2407-10-539

185. Hsu PY, Hsu HK, Lan X, Juan L, Yan PS, Labanowska J, et al. Amplification of distant estrogen response elements deregulates target genes associated with tamoxifen resistance in breast cancer. Cancer Cell. 2013 Aug 12;24(2):197–212.

186. Korkola J, Gray JW. Breast cancer genomes--form and function. Curr Opin Genet Dev. 2010 Feb;20(1):4–14.

187. Green AR, Aleskandarany MA, Agarwal D, Elsheikh S, Nolan CC, Diez-Rodriguez M, et al. MYC functions are specific in biological subtypes of breast cancer and confers resistance to endocrine therapy in luminal tumours. Br J Cancer [Internet]. 2016 Apr [cited 2024 Feb 28];114(8):917–28. Available from: https://www.nature.com/articles/bjc201646

188. ShahidSales S, Mehramiz M, Ghasemi F, Aledavood A, Shamsi M, Hassanian SM, et al. A genetic variant in CDKN2A/B gene is associated with the increased risk of breast cancer. J Clin Lab Anal [Internet]. 2018 [cited 2024 Feb 28];32(1):e22190. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/jcla.22190

189. Aftab A, Shahzad S, Hussain HMJ, Khan R, Irum S, Tabassum S. CDKN2A/P16INK4A variants association with breast cancer and their in-silico analysis. Breast Cancer Tokyo Jpn. 2019 Jan;26(1):11–28.

190. Jhungran A, Russell AH, Seiden MV, Duska LR, Goodman A, Lee S, et al. Chapter 84: Cancers of the Cervix, Vulva, and Vagina. In: Abeloff's Clinical Oncology. 6th edition. Philadelphia, Pa: Elsevier; 2020.

191. Hildesheim A, Herrero R, Castle PE, Wacholder S, Bratti MC, Sherman ME, et al. HPV co-factors related to the development of cervical cancer: results from a population-based study in Costa Rica. Br J Cancer. 2001 May 4;84(9):1219–26.

192. Schiffman M, Castle PE, Jeronimo J, Rodriguez AC, Wacholder S. Human papillomavirus and cervical cancer. Lancet Lond Engl. 2007 Sep 8;370(9590):890–907.

193. de Sanjosé S, Diaz M, Castellsagué X, Clifford G, Bruni L, Muñoz N, et al. Worldwide prevalence and genotype distribution of cervical human papillomavirus DNA in women with normal cytology: a meta-analysis. Lancet Infect Dis. 2007 Jul;7(7):453–9.

194. Bosch FX, Burchell AN, Schiffman M, Giuliano AR, de Sanjose S, Bruni L, et al. Epidemiology and natural history of human papillomavirus infections and type-specific implications in cervical neoplasia. Vaccine. 2008 Aug 19;26 Suppl 10:K1-16.

195. Ramachandran D, Dörk T. Genomic Risk Factors for Cervical Cancer. Cancers [Internet]. 2021 Jan [cited 2024 Mar 2];13(20):5137. Available from: https://www.mdpi.com/2072-6694/13/20/5137

196. Small W, Bacon MA, Bajaj A, Chuang LT, Fisher BJ, Harkenrider MM, et al. Cervical cancer: A global health crisis. Cancer. 2017 Jul 1;123(13):2404–12.

197. Arbyn M, Weiderpass E, Bruni L, de Sanjosé S, Saraiya M, Ferlay J, et al. Estimates of incidence and mortality of cervical cancer in 2018: a worldwide analysis. Lancet Glob Health. 2020 Feb;8(2):e191–203.

198. Tsu BV, Beierschmitt C, Ryan AP, Agarwal R, Mitchell PS, Daugherty MD. Diverse viral proteases activate the NLRP1 inflammasome. eLife. 2021 Jan 7;10:e60609.

199. DeVita (Jr.) VT, Rosenberg SA, Lawrence TS. Devita, Hellman, and Rosenberg's Cancer: Principles and Practice of Oncology. Wolters Kluwer; 2023. 1968 p.

200. Bhatla N, Aoki D, Sharma DN, Sankaranarayanan R. Cancer of the cervix uteri. Int J Gynaecol Obstet Off Organ Int Fed Gynaecol Obstet. 2018 Oct;143 Suppl 2:22–36.

201. Saslow D, Solomon D, Lawson HW, Killackey M, Kulasingam SL, Cain J, et al. American Cancer Society, American Society for Colposcopy and Cervical Pathology, and American Society for Clinical Pathology screening guidelines for the prevention and early detection of cervical cancer. CA Cancer J Clin. 2012;62(3):147–72.

202. Arbyn M, Simon M, Peeters E, Xu L, Meijer CJLM, Berkhof J, et al. 2020 list of human papillomavirus assays suitable for primary cervical cancer screening. Clin Microbiol Infect Off Publ Eur Soc Clin Microbiol Infect Dis. 2021 Aug;27(8):1083–95.

203. Perkins RB, Wentzensen N, Guido RS, Schiffman M. Cervical Cancer Screening: A Review. JAMA. 2023 Aug 8;330(6):547–58.

204. Bhatla N, Berek JS, Cuello Fredes M, Denny LA, Grenman S, Karunaratne K, et al. Revised FIGO staging for carcinoma of the cervix uteri. Int J Gynaecol Obstet Off Organ Int Fed Gynaecol Obstet. 2019 Apr;145(1):129–35.

205. Cervical Cancer Treatment (PDQ®) - NCI [Internet]. 2024 [cited 2024 Mar 26]. Available from: https://www.cancer.gov/types/cervical/hp/cervical-treatment-pdq

206. Chung HC, Ros W, Delord JP, Perets R, Italiano A, Shapira-Frommer R, et al. Efficacy and Safety of Pembrolizumab in Previously Treated Advanced Cervical Cancer: Results From the Phase II KEYNOTE-158 Study. J Clin Oncol Off J Am Soc Clin Oncol. 2019 Jun 10;37(17):1470–8.

207. Tewari KS, Sill MW, Penson RT, Huang H, Ramondetta LM, Landrum LM, et al. Bevacizumab for advanced cervical cancer: final overall survival and adverse event analysis of a randomised, controlled, open-label, phase 3 trial (Gynecologic Oncology Group 240). Lancet Lond Engl. 2017 Oct 7;390(10103):1654–63.

208. Ojesina AI, Lichtenstein L, Freeman SS, Pedamallu CS, Imaz-Rosshandler I, Pugh TJ, et al. Landscape of genomic alterations in cervical carcinomas. Nature [Internet]. 2014 Feb [cited 2024 Mar 2];506(7488):371–5. Available from: https://www.nature.com/articles/nature12881

209. zur Hausen H. Papillomaviruses in the causation of human cancers - a brief historical account. Virology. 2009 Feb 20;384(2):260–5.

210. Doorbar J, Quint W, Banks L, Bravo IG, Stoler M, Broker TR, et al. The biology and life-cycle of human papillomaviruses. Vaccine. 2012 Nov 20;30 Suppl 5:F55-70.

211. Zhang L, Wu J, Ling MT, Zhao L, Zhao KN. The role of the PI3K/Akt/mTOR signalling pathway in human cancers induced by infection with human papillomaviruses. Mol Cancer [Internet]. 2015 Apr 17 [cited 2024 Mar 2];14:87. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4498560/

212. Voutsadakis IA. 3q26 Amplifications in Cervical Squamous Carcinomas. Curr Oncol Tor Ont. 2021 Jul 29;28(4):2868–80.

213. Burk RD, Chen Z, Saller C, Tarvin K, Carvalho AL, Scapulatempo-Neto C, et al. Integrated genomic and molecular characterization of cervical cancer. Nature [Internet]. 2017 Mar [cited 2024 Mar 2];543(7645):378–84. Available from: https://www.nature.com/articles/nature21386

214. Liu W, Jiang Q, Sun C, Liu S, Zhao Z, Wu D. Developing a 5-gene prognostic signature for cervical cancer by integrating mRNA and copy number variations. BMC Cancer [Internet]. 2022 Feb 21 [cited 2024 Mar 2];22(1):192. Available from: https://doi.org/10.1186/s12885-022-09291-z

215. Hidalgo A, Baudis M, Petersen I, Arreola H, Piña P, Vázquez-Ortiz G, et al. Microarray comparative genomic hybridization detection of chromosomal imbalances in uterine cervix carcinoma. BMC Cancer [Internet]. 2005 Jul 9 [cited 2024 Mar 2];5(1):77. Available from: https://doi.org/10.1186/1471-2407-5-77

216. Gaspar N, Hawkins DS, Dirksen U, Lewis IJ, Ferrari S, Le Deley MC, et al. Ewing Sarcoma: Current Management and Future Approaches Through Collaboration. J Clin Oncol Off J Am Soc Clin Oncol. 2015 Sep 20;33(27):3036–46.

217. Postel-Vinay S, Véron AS, Tirode F, Pierron G, Reynaud S, Kovar H, et al. Common variants near TARDBP and EGR2 are associated with susceptibility to Ewing sarcoma. Nat Genet. 2012 Feb 12;44(3):323–7.

218. Grünewald TGP, Bernard V, Gilardi-Hebenstreit P, Raynal V, Surdez D, Aynaud MM, et al. Chimeric EWSR1-FLI1 regulates the Ewing sarcoma susceptibility gene EGR2 via a GGAA microsatellite. Nat Genet. 2015 Sep;47(9):1073–8.

219. Machiela MJ, Grünewald TGP, Surdez D, Reynaud S, Mirabeau O, Karlins E, et al. Genome-wide association study identifies multiple new loci associated with Ewing sarcoma susceptibility. Nat Commun [Internet]. 2018 Aug 9 [cited 2024 Apr 6];9(1):3184. Available from: https://www.nature.com/articles/s41467-018-05537-2

220. Jawad MU, Cheung MC, Min ES, Schneiderbauer MM, Koniaris LG, Scully SP. Ewing sarcoma demonstrates racial disparities in incidence-related and sex-related

differences in outcome: an analysis of 1631 cases from the SEER database, 1973-2005. Cancer. 2009 Aug 1;115(15):3526–36.

221. Cidre-Aranaz F, Watson S, Amatruda JF, Nakamura T, Delattre O, de Alava E, et al. Small round cell sarcomas. Nat Rev Dis Primer [Internet]. 2022 Oct 6 [cited 2024 Mar 2];8(1):1–22. Available from: https://www.nature.com/articles/s41572-022-00393-3

222. Riggi N, Cironi L, Provero P, Suvà ML, Kaloulis K, Garcia-Echeverria C, et al. Development of Ewing's sarcoma from primary bone marrow-derived mesenchymal progenitor cells. Cancer Res. 2005 Dec 15;65(24):11459–68.

223. von Levetzow C, Jiang X, Gwye Y, von Levetzow G, Hung L, Cooper A, et al. Modeling initiation of Ewing sarcoma in human neural crest cells. PloS One. 2011 Apr 29;6(4):e19305.

224. Tanaka T, Narazaki M, Kishimoto T. IL-6 in inflammation, immunity, and disease. Cold Spring Harb Perspect Biol. 2014 Sep 4;6(10):a016295.

225. Widhe B, Widhe T. Initial symptoms and clinical features in osteosarcoma and Ewing sarcoma. J Bone Joint Surg Am. 2000 May;82(5):667–74.

226. Casali PG, Bielack S, Abecassis N, Aro HT, Bauer S, Biagini R, et al. Bone sarcomas: ESMO-PaedCan-EURACAN Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol Off J Eur Soc Med Oncol. 2018 Oct 1;29(Suppl 4):iv79–95.

227. Meyer JS, Nadel HR, Marina N, Womer RB, Brown KLB, Eary JF, et al. Imaging guidelines for children with Ewing sarcoma and osteosarcoma: a report from the Children's Oncology Group Bone Tumor Committee. Pediatr Blood Cancer. 2008 Aug;51(2):163–70.

228. Ewing Sarcoma Treatment (PDQ®) - NCI [Internet]. 2024 [cited 2024 Mar 2]. Available from: https://www.cancer.gov/types/bone/hp/ewing-treatment-pdq

229. Delattre O, Zucman J, Plougastel B, Desmaze C, Melot T, Peter M, et al. Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. Nature. 1992 Sep 10;359(6391):162–5.

230. Delattre O, Zucman J, Melot T, Garau XS, Zucker JM, Lenoir GM, et al. The Ewing family of tumors--a subgroup of small-round-cell tumors defined by specific chimeric transcripts. N Engl J Med. 1994 Aug 4;331(5):294–9.

231. Sorensen PH, Lessnick SL, Lopez-Terrada D, Liu XF, Triche TJ, Denny CT. A second Ewing's sarcoma translocation, t(21;22), fuses the EWS gene to another ETS-family transcription factor, ERG. Nat Genet. 1994 Feb;6(2):146–51.

232. Riggi N, Knoechel B, Gillespie SM, Rheinbay E, Boulay G, Suvà ML, et al. EWS-FLI1 Utilizes Divergent Chromatin Remodeling Mechanisms to Directly Activate or Repress Enhancer Elements in Ewing Sarcoma. Cancer Cell [Internet]. 2014 Nov 10 [cited 2024 Mar 10];26(5):668–81. Available from: https://www.cell.com/cancer-cell/abstract/S1535-6108(14)00408-5

233. Grünewald TGP, Delattre O. Cooperation between somatic mutations and germline susceptibility variants in tumorigenesis - a dangerous liaison. Mol Cell Oncol. 2016 May;3(3):e1086853.

234. Orth MF, Surdez D, Faehling T, Ehlers AC, Marchetto A, Grossetête S, et al. Systematic multi-omics cell line profiling uncovers principles of Ewing sarcoma fusion oncogene-mediated gene regulation. Cell Rep. 2022 Dec 6;41(10):111761.

235. Orth MF. Systematic multi-omics profiling of Ewing sarcoma cell lines [Internet] [Text.PhDThesis]. Ludwig-Maximilians-Universität München; 2021 [cited 2024 Mar 20]. Available from: https://edoc.ub.uni-muenchen.de/27750/

236. R Core Team. R: A Language and Environment for Statistical Computing. 2021.

237. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics [Internet]. 2014 May 15 [cited 2024 Apr 6];30(10):1363–9. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4016708/

238. Zhou W, Laird PW, Shen H. Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes. Nucleic Acids Res. 2017 Feb 28;45(4):e22.

239. Maksimovic J, Phipson B, Oshlack A. A cross-package Bioconductor workflow for analysing methylation array data. F1000Research. 2016;5:1281.

240. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015 Apr 20;43(7):e47.

241. PRISM 9 [Internet]. USA: GraphPad Software Inc.; Available from: www.graphpad.com

242. Shaun Purcell. PLINK 1.9 [Internet]. 2009. Available from: http://pngu.mgh.harvard.edu/purcell/plink/

243. Knaus BJ, Grünwald NJ. vcfr: a package to manipulate and visualize variant call format data in R. Mol Ecol Resour. 2017 Jan;17(1):44–53.

244. David Meyer. proxy: Distance and Similarity Measures [Internet]. 2022. Available from: https://cran.r-project.org/web/packages/proxy/index.html

245. Raivo Kolde. pheatmap: Pretty Heatmaps. 2019.

246. Parekh S, Ziegenhain C, Vieth B, Enard W, Hellmann I. The impact of amplification on differential expression analyses by RNA-seq. Sci Rep. 2016 May 9;6(1):25533.

247. Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, et al. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell. 2015 May 21;161(5):1202–14.

248. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinforma Oxf Engl. 2012 Mar 15;28(6):882–3.

249. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.

250. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005 Oct 25;102(43):15545–50.

251. Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J, et al. PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet [Internet]. 2003 Jul [cited 2024 Apr 6];34(3):267–73. Available from: https://www.nature.com/articles/ng1180

252. Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, et al. Welcome to the Tidyverse. J Open Source Softw [Internet]. 2019 Nov 21 [cited 2024 Apr 6];4(43):1686. Available from: https://joss.theoj.org/papers/10.21105/joss.01686

253. Musa J, Cidre-Aranaz F. Drug Screening by Resazurin Colorimetry in Ewing Sarcoma. Methods Mol Biol Clifton NJ. 2021;2226:159–66.

254. Smirnov P, Safikhani Z, El-Hachem N, Wang D, She A, Olsen C, et al. PharmacoGx: an R package for analysis of large pharmacogenomic datasets. Bioinforma Oxf Engl. 2016 Apr 15;32(8):1244–6.

255. Wickham H. Reshaping Data with the reshape Package. J Stat Softw. 2007 Nov 13;21:1–20.

256. Dowle M, Srinivasan A, Gorecki J, Chirico M, Stetsenko P, Short T, et al. data.table: Extension of “data.frame” [Internet]. 2023 [cited 2023 Nov 8]. Available from: https://cran.r-project.org/web/packages/data.table/index.html

257. Gu Z, Gu L, Eils R, Schlesner M, Brors B. circlize Implements and enhances circular visualization in R. Bioinforma Oxf Engl. 2014 Oct;30(19):2811–2.

258. Forbes SH. PupillometryR: An R package for preparing and analysing pupillometry data. J Open Source Softw [Internet]. 2020 Jun 20 [cited 2024 Apr 6];5(50):2285. Available from: https://joss.theoj.org/papers/10.21105/joss.02285

259. Jr FEH, functions) CD (contributed several functions and maintains latex. Hmisc: Harrell Miscellaneous [Internet]. 2023 [cited 2023 Nov 8]. Available from: https://cran.r-project.org/web/packages/Hmisc/index.html

260. Stein CK, Qu P, Epstein J, Buros A, Rosenthal A, Crowley J, et al. Removing batch effects from purified plasma cell gene expression microarrays with modified ComBat. BMC Bioinformatics. 2015 Feb 25;16:63.

261. Kasan M, Siebenlist J, Sill M, Öllinger R, Álava E de, Surdez D, et al. Genomic and phenotypic stability of fusion-driven pediatric Ewing sarcoma cell lines [Internet].

bioRxiv; 2023 [cited 2024 Mar 2]. p. 2023.11.20.567802. Available from: https://www.biorxiv.org/content/10.1101/2023.11.20.567802v1

262. Hughes. Principles of early drug discovery. Br J Pharmacol [Internet]. 2011 [cited 2024 Mar 10]; Available from: https://bpspubs.onlinelibrary.wiley.com/doi/10.1111/j.1476-5381.2010.01127.x

263. Mora J, Castañeda A, Perez-Jaume S, Lopez-Pousa A, Maradiegue E, Valverde C, et al. GEIS-21: a multicentric phase II study of intensive chemotherapy including gemcitabine and docetaxel for the treatment of Ewing sarcoma of children and adults: a report from the Spanish sarcoma group (GEIS). Br J Cancer. 2017 Sep 5;117(6):767–74.

264. Marchetto A, Ohmura S, Orth MF, Knott MML, Colombo MV, Arrigoni C, et al. Oncogenic hijacking of a developmental transcription factor evokes vulnerability toward oxidative stress in Ewing sarcoma. Nat Commun [Internet]. 2020 May 15 [cited 2024 Mar 2];11(1):2423. Available from: https://www.nature.com/articles/s41467-020-16244-2

265. Takagi M, Ogawa C, Iehara T, Aoki-Nogami Y, Ishibashi E, Imai M, et al. First phase 1 clinical study of olaparib in pediatric patients with refractory solid tumors. Cancer. 2022 Aug 1;128(15):2949–57.

266. Gupta A, Riedel RF, Shah C, Borinstein SC, Isakoff MS, Chugh R, et al. Consensus recommendations in the management of Ewing sarcoma from the National Ewing Sarcoma Tumor Board. Cancer. 2023 Nov 1;129(21):3363–71.

267. Soule HD, Vazguez J, Long A, Albert S, Brennan M. A human cell line from a pleural effusion derived from a breast carcinoma. J Natl Cancer Inst. 1973 Nov;51(5):1409–16.

268. Feinberg AP, Vogelstein B. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. Nature [Internet]. 1983 Jan [cited 2024 Mar 2];301(5895):89–92. Available from: https://www.nature.com/articles/301089a0

269. Duesberg P, Fabarius A, Hehlmann R. Aneuploidy, the primary cause of the multilateral genomic instability of neoplastic and preneoplastic cells. IUBMB Life. 2004 Feb;56(2):65–81.

270. Cellosaurus - Cell line encyclopedia [Internet]. [cited 2024 Mar 2]. Available from: https://www.cellosaurus.org/index.html

271. Tageja N. Prognostic indicators for Ewing's sarcoma. The Lancet [Internet]. 2010 Jul 24 [cited 2024 Mar 2];376(9737):232. Available from: https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(10)61141-5/fulltext

272. Stahl D, Gentles AJ, Thiele R, Gütgemann I. Prognostic profiling of the immune cell microenvironment in Ewing´s Sarcoma Family of Tumors. Oncoimmunology. 2019;8(12):e1674113.

273. Grohar PJ, Kim S, Rangel Rivera GO, Sen N, Haddock S, Harlow ML, et al. Functional Genomic Screening Reveals Splicing of the *EWS-FLI1* Fusion

Transcript as a Vulnerability in Ewing Sarcoma. Cell Rep [Internet]. 2016 Jan 26 [cited 2024 Mar 2];14(3):598–610. Available from: https://www.sciencedirect.com/science/article/pii/S2211124715015041

274. Franzetti GA, Laud-Duval K, van der Ent W, Brisac A, Irondelle M, Aubert S, et al. Cell-to-cell heterogeneity of EWSR1-FLI1 activity determines proliferation/migration choices in Ewing sarcoma cells. Oncogene [Internet]. 2017 Jun [cited 2024 Mar 2];36(25):3505–14. Available from: https://www.nature.com/articles/onc2016498

275. Barghi F, Shannon HE, Saadatzadeh MR, Bailey BJ, Riyahi N, Bijangi-Vishehsaraei K, et al. Precision Medicine Highlights Dysregulation of the CDK4/6 Cell Cycle Regulatory Pathway in Pediatric, Adolescents and Young Adult Sarcomas. Cancers [Internet]. 2022 Jan [cited 2024 Mar 2];14(15):3611. Available from: https://www.mdpi.com/2072-6694/14/15/3611

276. Koboldt DC, Fulton RS, McLellan MD, Schmidt H, Kalicki-Veizer J, McMichael JF, et al. Comprehensive molecular portraits of human breast tumours. Nature [Internet]. 2012 Oct [cited 2024 Feb 28];490(7418):61–70. Available from: https://www.nature.com/articles/nature11412

277. Sharma SV, Lee DY, Li B, Quinlan MP, Takahashi F, Maheswaran S, et al. A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. Cell. 2010 Apr 2;141(1):69–80.

278. Burrell RA, McGranahan N, Bartek J, Swanton C. The causes and consequences of genetic heterogeneity in cancer evolution. Nature [Internet]. 2013 Sep [cited 2024 Mar 2];501(7467):338–45. Available from: https://www.nature.com/articles/nature12625

279. Shen S, Clairambault J. Cell plasticity in cancer cell populations. F1000Research [Internet]. 2020 Jun 22 [cited 2024 Mar 2];9:F1000 Faculty Rev-635. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7309415/

280. Campiglio CE, Figliuzzi M, Silvani S, Tironi M, Conti S, Boschetti F, et al. Influence of Culture Substrates on Morphology and Function of Pulmonary Alveolar Cells In Vitro. Biomolecules [Internet]. 2021 Apr 30 [cited 2024 Mar 2];11(5):675. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8147120/

281. Gnocchi D, Nikolic D, Paparella RR, Sabbà C, Mazzocca A. Cellular Adaptation Takes Advantage of Atavistic Regression Programs during Carcinogenesis. Cancers [Internet]. 2023 Jan [cited 2024 Mar 2];15(15):3942. Available from: https://www.mdpi.com/2072-6694/15/15/3942

282. Gomes AP, Blenis J. A nexus for cellular homeostasis: the interplay between metabolic and signal transduction pathways. Curr Opin Biotechnol [Internet]. 2015 Aug [cited 2024 Mar 2];34:110–7. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4490161/

283. Cole SW, Sood AK. Molecular pathways: beta-adrenergic signaling in cancer. Clin Cancer Res Off J Am Assoc Cancer Res. 2012 Mar 1;18(5):1201–6.

284. Zack TI, Schumacher SE, Carter SL, Cherniack AD, Saksena G, Tabak B, et al. Pan-cancer patterns of somatic copy number alteration. Nat Genet [Internet]. 2013 Oct [cited 2024 Mar 2];45(10):1134–40. Available from: https://www.nature.com/articles/ng.2760

285. Lochhead P, Kuchiba A, Imamura Y, Liao X, Yamauchi M, Nishihara R, et al. Microsatellite Instability and BRAF Mutation Testing in Colorectal Cancer Prognostication. JNCI J Natl Cancer Inst [Internet]. 2013 Aug 7 [cited 2024 Mar 2];105(15):1151–6. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3735463/

286. Seppälä TT, Böhm JP, Friman M, Lahtinen L, Väyrynen VMJ, Liipo TKE, et al. Combination of microsatellite instability and BRAF mutation status for subtyping colorectal cancer. Br J Cancer [Internet]. 2015 Jun [cited 2024 Mar 2];112(12):1966–75. Available from: https://www.nature.com/articles/bjc2015160

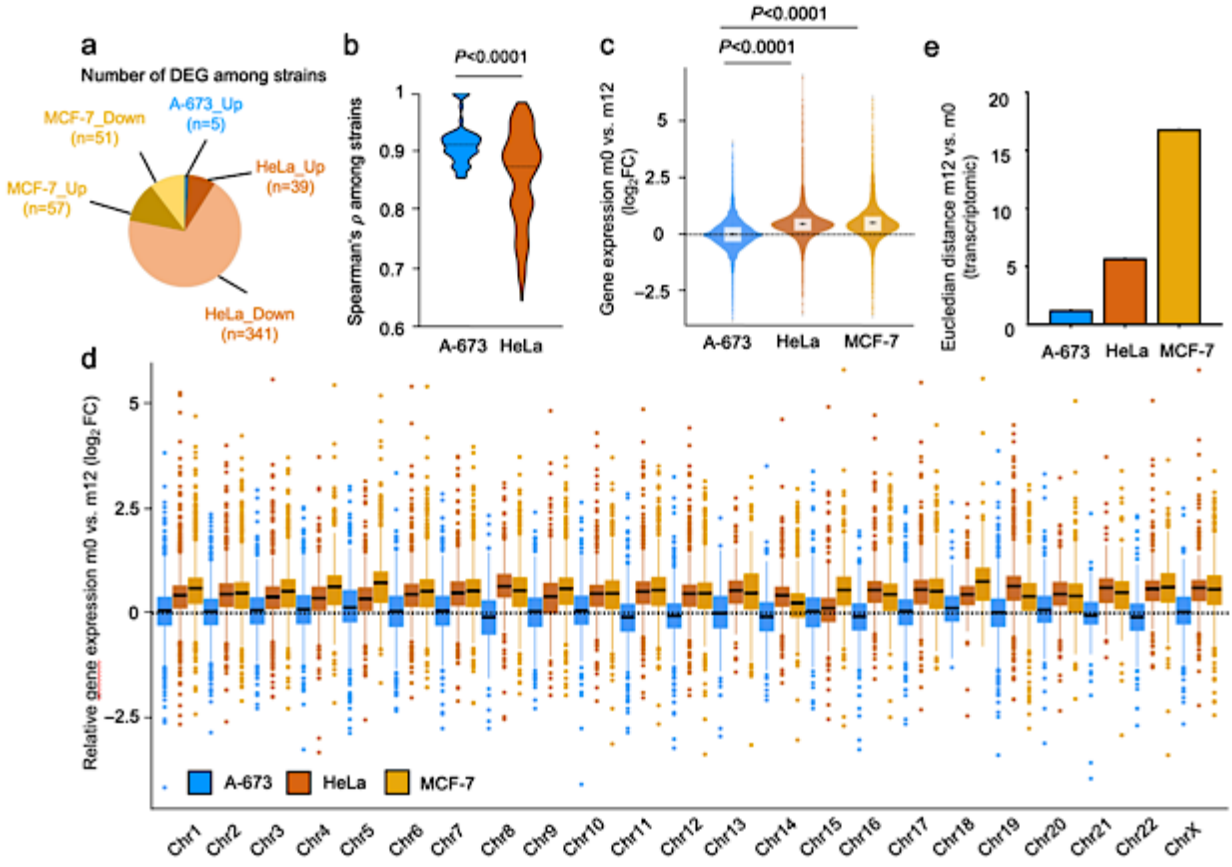
11 APPENDIX

11.1. Supplementary Tables and Figures

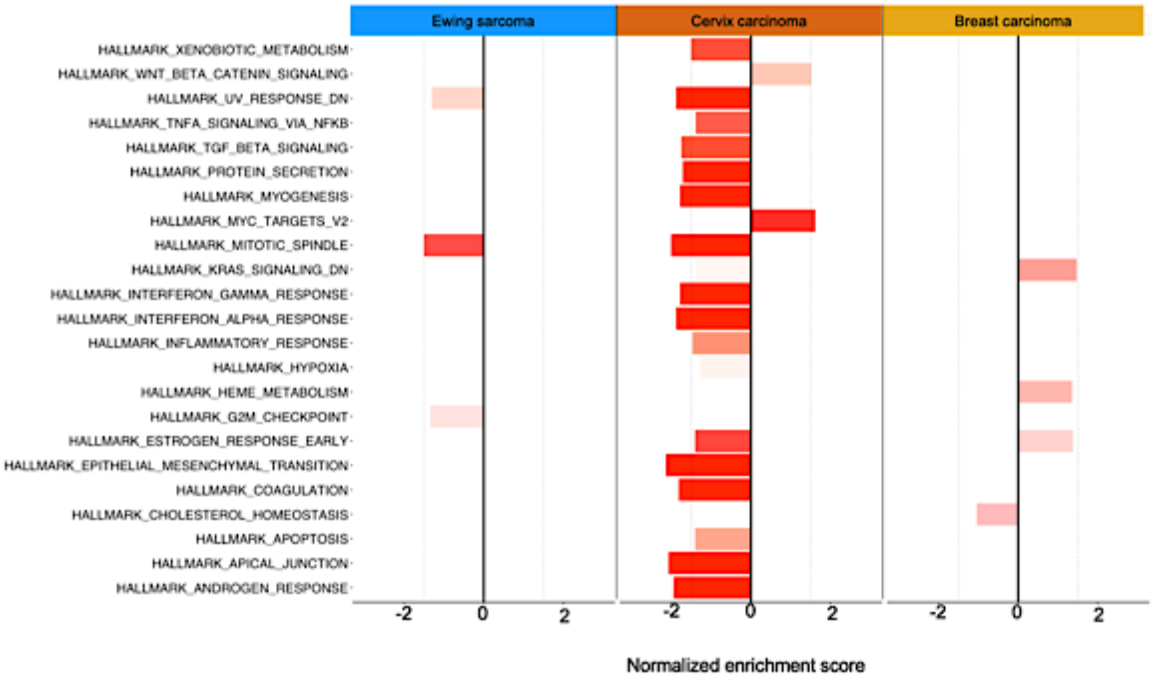
Table 11: DEGs across comparison of two strains with the highest variability.

Chr	Gene	Base mean	Log_2FC	*p* value	Adj. *p* value	Comparison
chr11	*CEND1*	475,777	-13,693	3,34E-11	1,59758E-08	MCF-7_5 vs. MCF-7_3
chr10	*ACTA2*	202,300	-4,420	9,50E-239	1,5427E-234	HeLa_5 vs. HeLa_3
chr6	*CGA*	832,906	-3,912	1,39E-89	2,50532E-86	HeLa_5 vs. HeLa_3
chr12	*BCAT1*	135,218	-3,597	1,84E-90	3,74272E-87	HeLa_5 vs. HeLa_3
chr16	*CDH3*	23,364	-3,419	1,37E-08	6,97556E-07	HeLa_5 vs. HeLa_3
chr5	*EGR1*	672,497	-2,970	8,51E-111	2,3053E-107	HeLa_5 vs. HeLa_3
chr15	*LINC00052*	105,828	-2,776	2,26E-16	2,60257E-14	HeLa_5 vs. HeLa_3
chr2	*MLPH*	309,524	-2,749	1,45E-29	3,51301E-27	HeLa_5 vs. HeLa_3
chr9	*TPM2*	415,444	-2,744	1,16E-11	5,81656E-09	MCF-7_5 vs. MCF-7_3
chr6	*MARCKS*	812,942	-2,588	2,14E-17	3,35515E-14	MCF-7_5 vs. MCF-7_3
chr8	*CLU*	3631,970	-2,383	3,66E-18	9,20434E-15	MCF-7_5 vs. MCF-7_3
chr9	*ASS1*	393,923	-2,106	8,42E-76	1,14009E-72	HeLa_5 vs. HeLa_3
chr11	*GSTP1*	1611,288	-1,924	8,05E-13	5,06081E-10	MCF-7_5 vs. MCF-7_3
chr8	*CLU*	9821,501	-1,436	3,58E-139	1,9403E-135	HeLa_5 vs. HeLa_3
chr17	*LLGL2*	378,518	-1,377	3,04E-16	3,40432E-14	HeLa_5 vs. HeLa_3
chrX	*SMC1A*	722,158	-1,208	6,95E-04	0,045806346	A-673_7 vs. A-673_3
chr19	*HNRNPM*	1159,023	-1,166	2,58E-04	0,020405921	A-673_7 vs. A-673_3
chr9	*RPL35*	18299,226	-1,118	4,17E-126	1,3537E-122	HeLa_5 vs. HeLa_3
chr1	*HMGN2*	677,700	-1,088	1,72E-03	0,088914254	A-673_7 vs. A-673_3
chr12	*GAPDH*	11455,426	-1,082	1,49E-79	2,20437E-76	HeLa_5 vs. HeLa_3
chr19	*FTL*	11229,056	-1,061	1,89E-74	2,1944E-71	HeLa_5 vs. HeLa_3
chr4	*OCIAD2*	1011,670	-1,048	2,47E-04	0,020405921	A-673_7 vs. A-673_3
chr1	*PARP1*	925,654	-1,016	3,18E-03	0,119098891	A-673_7 vs. A-673_3
chr20	*ID1*	680,658	1,018	1,78E-16	2,06246E-14	HeLa_5 vs. HeLa_3
chr3	*HES1*	2143,985	1,129	1,03E-06	0,000407906	A-673_7 vs. A-673_3
chr10	*AKR1C2*	904,782	1,156	3,90E-30	9,75337E-28	HeLa_5 vs. HeLa_3
chr9	*HSPA5*	1209,115	1,187	9,02E-06	0,001528763	A-673_7 vs. A-673_3
chr16	*NQO1*	1375,265	1,211	3,85E-05	0,005716676	A-673_7 vs. A-673_3
chr11	*MALAT1*	33768,499	1,214	3,67E-03	0,131856375	A-673_7 vs. A-673_3
chr20	*ID1*	724,977	1,316	2,57E-04	0,020405921	A-673_7 vs. A-673_3
chr10	*DDIT4*	468,315	1,346	1,92E-03	0,088914254	A-673_7 vs. A-673_3
chr1	*ID3*	1620,048	1,412	9,12E-09	5,41401E-06	A-673_7 vs. A-673_3
chr11	*TAF1D*	558,767	1,418	2,37E-04	0,020405921	A-673_7 vs. A-673_3
chr5	*BDP1*	746,663	1,624	1,81E-06	0,000494611	A-673_7 vs. A-673_3
chr12	*DDIT3*	543,712	1,637	2,85E-03	0,112770345	A-673_7 vs. A-673_3
chr5	*RHOBTB3*	797,004	1,999	1,16E-13	7,76211E-11	MCF-7_5 vs. MCF-7_3

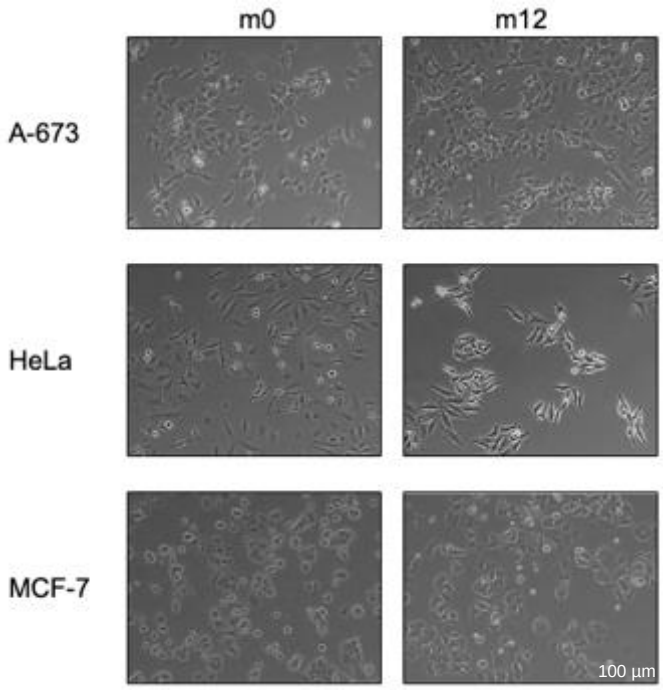
chr14	*FOS*	747,916	2,434	1,00E-17	2,01102E-14	MCF-7_5 vs. MCF-7_3
chr8	*TMEM64*	539,388	2,709	2,34E-17	3,35515E-14	MCF-7_5 vs. MCF-7_3
chr6	*PKIB*	303,461	2,833	3,90E-14	3,27012E-11	MCF-7_5 vs. MCF-7_3
chr10	*CALML5*	777,387	3,083	1,20E-29	4,0139E-26	MCF-7_5 vs. MCF-7_3
chr10	*MSMB*	173,787	3,839	2,06E-16	2,59071E-13	MCF-7_5 vs. MCF-7_3
chr19	*KLK11*	208,829	3,921	2,48E-15	2,49057E-12	MCF-7_5 vs. MCF-7_3
chr7	*SHH*	73,459	4,660	1,03E-10	4,48346E-08	MCF-7_5 vs. MCF-7_3
chr6	*HLA-DRB1*	84,860	5,333	5,51E-15	5,03637E-12	MCF-7_5 vs. MCF-7_3
chr4	*AREG*	246,140	6,995	1,31E-42	6,56276E-39	MCF-7_5 vs. MCF-7_3



Supplementary Figure 1: a. Pie chart illustrating the count of DEGs categorized as up-regulated and down-regulated in each comparison. **b.** Gene-specific Spearman's ρ analysis conducted across A-673 and HeLa cell line strains using the combined transcriptomic datasets depicted in Fig 10d. The median is indicated by dotted black lines, assessed using a one-sided Wilcoxon rank-sum test. **c, d.** Assessment of relative gene expression changes in A-673, HeLa, and MCF-7 cell lines after twelve months of continuous culture (m12 versus m0). Boxplots represent the interquartile range and mean (two-sided Wilcoxon signed-rank test). **d.** Distribution of genes per chromosome. **e.** Bar plot visualizing the Euclidean distance between the initial (m0) of A-673, HeLa, and MCF-7 cell lines and their progeny after twelve months of continuous cell culture. Figure and figure legend from Kasan et al. 2023 preprint available on bioRxiv (261), also in press in Nature Communications under CC-BY license.



Supplementary Figure 2: Bar plot depicting enriched pathways after 6 months of continuous cell culture for each cancer type. Darker colors indicate lower *p*-values. Positive scores signify upregulation, while negative scores denote downregulation.



Supplementary Figure 3: Micrographs illustrating cell morphology at the initial time point (m0) and after 12 months of continuous cell culture (m12). The scale bar represents 100 µm.

11.2. List of tables

11.3. List of figures

Eidesstattliche Versicherung

Merve Kasan

Ich erkläre hiermit an Eides statt,

dass ich die vorliegende Dissertation mit dem Titel

” *Reproducibility in cancer research: A study of the (epi)genomic, transcriptomic, and phenotypic stability in chromosomal translocation-driven pediatric sarcoma cell lines*”

selbstständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

Heidelberg, 17.12.2024 Merve Kasan

Erklärung zur Übereinstimmung der gebundenen Ausgabe der Dissertation mit der elektronischen Fassung

Hiermit erkläre ich, dass die elektronische Version der eingereichten Dissertation mit dem Titel

"*Reproducibility in cancer research: A study of the (epi)genomic, transcriptomic, and phenotypic stability in chromosomal translocation-driven pediatric sarcoma cell lines*"

in Inhalt und Formatierung mit den gedruckten und gebundenen Exemplaren übereinstimmt.

Heidelberg, 17.12.2024 Merve Kasan