T.C.
DOKUZ EYLÜL ÜNIVERSITESI IZMIR
INTERNATIONAL
BIOMEDICINE AND GENOME
INSTITUTE

A SYSTEMS BIOLOGY APPROACH FOR THE IDENTIFICATION OF GENE EXPRESSION SIGNATURES FOR CANCER DIAGNOSIS, PROGNOSIS, AND RESPONSE TO RADIOTHERAPY

HALİL IBRAHİM TOY

MOLECULAR BIOLOGY AND GENETICS

PhD THESIS

IZMIR-2022

TEZ KODU: DEU.HSI.PhD-2017850017

T.C.

DOKUZ EYLÜL ÜNIVERSITESI IZMIR
INTERNATIONAL
BIOMEDICINE AND GENOME
INSTITUTE

A SYSTEMS BIOLOGY APPROACH FOR THE IDENTIFICATION OF GENE EXPRESSION SIGNATURES FOR CANCER DIAGNOSIS, PROGNOSIS, AND RESPONSE TO RADIOTHERAPY

MOLECULAR BIOLOGY AND GENETICS

PhD THESIS

HALİL İBRAHİM TOY

Supervisor: Asst. Prof. ATHANASIA PAVLOPOULOU

Co-supervisor: Assoc. Prof. GÖKHAN KARAKÜLAH

THESIS CODE: DEU.HSI.PhD-2017850017

Dokuz Eylül University Izmir International Biomedicine and Genome Institute Department of Genomics and Molecular Biotechnology, Molecular Biology and Genetics PhD program student Halil İbrahim TOY has successfully completed the PhD thesis entitled "**A systems biology approach for the identification of gene expression signatures for cancer diagnosis, prognosis, and response to radiotherapy**" on the date of 01/07/2022.

HEAD OF THE COMMITTEE
Asst. Prof. Dr. Athanasia PAVLOPOULOU

MEMBER
Asst. Prof. Dr. Yavuz OKTAY

MEMBER
Assoc. Prof. Dr. Şerif ŞENTÜRK

MEMBER
Prof. Dr. Ayse NALBANTSOY

MEMBER
Prof. Dr. Ralph MEUWISSEN

DOKTORA TEZ SAVUNMA SINAVI TUTANAĞI

Dokuz Eylül Üniversitesi İzmir Uluslararası Biyotıp ve Genom Enstitüsü Genom Bilimleri ve Moleküler Biyoteknoloji Anabilim Dalı Moleküler Biyoloji ve Genetik Doktora Programı 2017850017 numaralı öğrencisi Halil İbrahim TOY'nın, "**Sistem biyolojisi yaklaşımlarıyla kanserin tanısı, prognozu ve radyoterapiye yanıt için gen ekspresyon imzalarının tanımlanması**" konulu Doktora tezini 01/07/2022 tarihinde yapılan savunma sınavı sonucunda başarılı olmuştur.

BAŞKAN
Dr. Öğr. Üyesi Athanasia PAVLOPOULOU

JÜRİ ÜYESİ
Dr. Öğr. Üyesi Yavuz OKTAY

JÜRİ ÜYESİ
Doç. Dr. Şerif ŞENTÜRK

JÜRİ ÜYESİ
Prof. Dr. Ayse NALBANTSOY

JÜRİ ÜYESİ
Prof. Dr. Ralph MEUWISSEN

INDEXES

INDEXES i
INDEX OF TABLES iv
INDEX OF FIGURES v
LIST OF ABBREVIATIONS vi
ACKNOWLEDGEMENTS vii
ABSTRACT 1
ÖZET 2
1. INTRODUCTION AND OBJECTIVES 2
1.1. Statement and Importance of the Problem 3
1.2. Aims of the Study 3
1.3. Hypothesis of the Study 4
2. GENERAL INFORMATION 4
2.1. HOX transcript antisense RNA (HOTAIR) 4
2.2. Resistance to Radiation Therapy 7
2.3. Acute Myeloid Leukemia 10
3. MATERIALS AND METHODOLOGY 13
3.1. Type of the Study 13
3.2. Time and Place of the Study 13
3.3. Materials of the Study 13
3.4. Variables of the Study 13
3.5. Programming Languages and Tools for Data Processing 14
3.6. Data Analysis 14
3.6.1. Meta-analysis of the prognostic significance of HOTAIR in cancer 15
3.6.1.1. Search Strategy and Study Eligibility Criteria 15
3.6.1.2. Study Selection, Data Extraction, and Quality 15
3.6.1.3. Statistical Analyses 16
3.6.2. Identification of Molecular Determinants of Cancer Cell Resistance to Ionizing Radiation 16
3.6.2.1. Data acquisition 16
3.6.2.2. Microarray-Based Transcriptomic Data Analysis 18
3.6.2.3. RNA-Seq-Based Transcriptomic Data Analysis 19
3.6.2.4. Pathway Enrichment Analysis 20

3.6.2.5. Functional Interactions Networks 21
3.6.2.6. Survival Analysis 21
3.6.2.7. Tissue-Wise Differential Gene Expression Analysis 21
3.6.3. Identification of DEGs Correlated with Long-Term Survival in AML patients 21
4. RESULTS 22
4.1. HOTAIR as an Unfavorable Prognostic Biomarker for Cancer 22
4.1.1. Study Selection and Characteristics of Eligible Studies 22
4.1.2. Association between High HOTAIR Expression and Overall Survival in Diverse Cancers 29
4.1.3. HOTAIR Overexpression is Associated with Cancer Recurrence and Progression 31
4.1.4. Publication Bias 32
4.1.5. Sensitivity Analysis 33
4.2. Identification of Molecular Determinants of Cancer Cell Resistance to Ionizing Radiation 34
4.2.1. Identification of Differential Expression Patterns in Radioresistance versus Radiosensitive Cancer Cells 35
4.2.2. Identification of Cancer Cell Radioresistance Related Genes 39
4.2.3. Differentially Expressed Radiogenes in Cancer vs. Normal Tissue 44
4.2.4. Radiogenes are Potential Cancer Prognostic Markers 46
4.3. Identification of DEGs Correlated with Long-Term Survival in AML patients 47
4.3.1. Inferring Favorable Prognostic Markers from the Transcriptomic Profiles of Long-Term Survivors of AML 47
5. DISCUSSION 53
5.1. Meta Analysis of the Prognostic Value of HOTAIR in Cancer 53
5.2. Identification of Molecular Determinants of Cancer Cell Resistance to Ionizing Radiation 56
5.3. Identification of Favorable Prognostic Markers from the Transcriptomic Profiles of Long-Term Survivors of AML 61
6. CONCLUSIONS AND FUTURE ASPECTS 65
7. REFERENCES 67
8. ADDENDA 100
8.1. References (Table 1) 100

8.2. Ethics Committee Report.....107
8.3. Curriculum Vitae.....108
8.4. Publications from the Thesis.....109

INDEX OF TABLES

Table 1. Samples from each transcriptomic dataset analyzed in this study.. 19

Table 2. Main characteristics of the studies included in the meta-analysis. 26

Table 3. Distribution of the over-represented cancer-related WikiPathways in the DEGs of each transcriptome dataset. 38

Table 4. Gene symbol and expression status of the 36 differentially expressed radiogenes (radioresistant vs. radiosensitive cancer cells). 43

INDEX OF FIGURES

Figure 1. Associations of human lncRNsA with biological molecules (Kazimierczyk et al., 2020). 6

Figure 2. HOTAIR major biological functions (McFadden & Hargrove, 2016). 7

Figure 3. Illustration of the molecular mechanisms of ionizing radiation to cells. 9

Figure 4. Molecular mechanisms underlying radioresistance (Buckley et al., 2020). 10

Figure 5. Flow chart of the process for study selection. 25

Figure 6. Forest plots of combined analyses on the association of survival with HOTAIR expression 31

Figure 7. Forest plots of combined analyses for overall survival (OS) associated with HOTAIR expression in different groups 32

Figure 8. Begg's funnel plots of publication bias. 34

Figure 9. Sensitivity analysis of each eligible study. 35

Figure 10. Flowchart diagram of the overall methodology employed in this study. .36

Figure 11. Heatmap representing color-coded expression levels of DEGs. 37

Figure 12. Schematic workflow outlining the selection of radiogenes. 42

Figure 13. Donut chart depicting the over-represented cancer-relevant pathways across the 88 common genes. 43

Figure 14. STRING interaction network of the products of the 36 radiogenes. 45

Figure 15. Differential expression patterns of radiogenes in the corresponding cancer tissue; BRCA, breast invasive carcinoma; COAD, colon adenocarcinoma; LAML, acute myeloid leukemia. 46

Figure 16. Kaplan-Meier curves depicting the prognostic value of radiogenes for overall survival in invasive breast carcinoma (a) CHEK1, (b) PLK1, (c) MAP2K1 and (d) NFKBIA. 47

Figure 17. Identification of favorable prognostic factors based on transcriptomic

profiles of LTSs. 51

Figure 18. Radiation-induced DNA damage and its repair mechanisms. 58

LIST OF ABBREVIATIONS

AML	Acute Myeloid Leukemia
bp	base pair
CI	Confidence Interval
DEG	Differentially Expressed Genes
DFS	Disease-Free Survival
FDR	False Discovery Rate
GEO	Gene Expression Omnibus (GEO)
GSEA	Gene Set Enrichment Analysis
GTEx	Genotype-Tissue Expression
HOTAIR	HOX transcript antisense RNA
HR	Hazard Ratio
IR	ionizing radiation
lncRNA	long non-coding RNA
LTS	Long-Term Survivors
MFS	Metastasis-Free Survival
MRD	Minimal Residual Disease
NCBI	National Center for Biotechnology Informatıon
OS	Overall Survival
PFS	Progression-Free Survival
RFS	Recurrence-Free Survival
RNA-Seq	RNA sequencing
RR	Radioresistant
RS	Radiosensitivity
RT	Radiation Therapy
STRING	Search Tool for the Retrieval of Interacting Genes

TCGA The Cancer Genome Atlas

ACKNOWLEDGEMENTS

Firstly, I would like to express my sincere thanks and gratitude to my advisor Asst. Prof. Dr. Athanasia PAVLOPOULOU for all of her support and guidance during my doctoral studies. Her scholarly and timely advice, meticulous scrutiny, and scientific approach have motivated me to carry out substantial research work. I shall always keep your advice in mind and thank you so much. I believe that we could contribute more in the field of science by continuing our excellent collaboration.

I owe a great debt of gratitude and appreciation to my co-supervisor Assoc. Prof. Dr. Gökhan KARAKÜLAH especially for his support and expert advice in the computational approaches at every stage of my research.

I would also like to extend my thanks to all Pavlopoulou lab members, especially Hande YILMAZ, Didem ÖKMEN, Işıl TAKAN and Melih ÖZBEK for their kind help and friendship, Halil İbrahim PAZARBAŞI, Egemen Erdem GÜLER and Deniz DOĞAN for their valuable friendship during PhD studies.

Moreover, I would like to acknowledge the scholarships I received from the Council of Higher Education (CoHE) (YÖK 100/2000), as well as the Scientific and Technological Research Institution of Turkey (TUBITAK BIDEB 2211-C) during my PhD studies.

Special thanks to the members of my PhD committee Asst. Prof. Dr. Yavuz OKTAY, Assoc. Prof. Dr. Şerif ŞENTÜRK, Prof. Ayşe NALBANTSOY, and Prof. Ralph MEUWISSEN for their valuable time and suggestions.

Finally and most importantly, I would like to thank wholeheartedly my family for their continuous and unrequited support, love, and encouragement. In addition, I dedicate this research to my daughter, Sophia TOY.

A systems biology approaches for the identification of gene expression signatures for cancer diagnosis, prognosis, and response to radiotherapy

Halil İbrahim Toy, Izmir International Biomedicine and Genome Institute, Dokuz Eylul University, Health Campus, Balcova, 35340, Izmir / Turkey

ABSTRACT

Cancer represents one of the leading causes of death worldwide. Despite the availability of anticancer therapeutic regimens, the incidence, prevalence, and mortality of cancer are increasing exponentially. The accurate diagnosis and prognosis of cancer, as well as assessment of patients' response to anticancer treatments, are of critical importance in clinical decision-making. Therefore, there is an emerging need for the identification of powerful molecular markers for the early detection and outcomes of cancer. Despite the exemplary efforts made to date, a comprehensive effort to determine potential diagnostic, prognostic, and radioresistance molecular signatures across human cancers is still lacking. The scope of our proposed studies was to identify diagnostic, prognostic, and radioresistance gene signatures for different types of human cancers, based on publicly available bibliographic and high-throughput data. To this end, systematic reviews and meta-analyses, gene expression profiling, and transcriptome-wide association analyses were performed in diverse cancers through integrative in silico approaches. The findings of those studies could be potentially taken into consideration in the clinical setting for the accurate and timely diagnosis of cancer, monitoring cancer patients, and the administration of appropriate antineoplastic therapy.

Keywords: cancer diagnosis, cancer prognosis, radioresistance, transcriptomics, meta-analysis, systems biology

Sistem biyolojisi yaklaşımlarıyla kanserin tanısı, prognozu ve radyoterapiye yanıt için gen ekspresyon imzalarının tanımlanması

Halil İbrahim TOY, İzmir Uluslararası Biyotıp ve Genom Enstitüsü, Uluslararası Biyotıp ve Genom Enstitüsü, Dokuz Eylül Üniversitesi, Sağlık Yerleşkesi, Balçova, 35340, İzmir / Türkiye

ÖZET

Kanser dünya çapında önde gelen ölüm nedenlerinden biridir. Antikanser tedavi yöntemlerinin mevcudiyetine rağmen, hem kanser insidansı ve prevalansı hem de mortalitesi halen yüksektir. Kanserin doğru teşhisi ve prognozu yanısıra hastaların antikanser tedavilerine yanıtlarının değerlendirilmesi klinik karar vermede kritik öneme sahiptir. Bu nedenle, kanserin erken teşhisi ve neticesi için güçlü moleküler belirteçlerin tanımlanmasına yönelik artan bir ihtiyaç bulunmaktadır. Bugüne kadar gösterilen örnek çabalara rağmen, insan kanserlerinde potansiyel tanısal, prognostik ve kanser tedavisine karşı dirençlilik sağlayan moleküler biyobelirteçleri belirlemek için kapsamlı çalışmalar halen yetersiz kalmaktadır. Önerilen çalışmalarımızın kapsamı, kamuya açık farklı kanser türlerine ait bibliyografik ve yüksek verimli verilerden faydalanılarak tanısal, prognostik ve radyo-direnç gen imzalarını belirlemektir. Bu amaçla, çeşitli kanser türlerinde sistematik incelemeler ve meta-analizler, gen ekspresyon profili oluşturma ve transkriptom çapında ilişkilendirme analizleri gibi bütünleştirici insiliko yaklaşımları uygulandı. Bu bulgular, kanserin doğru ve zamanında teşhisi, kanser hastalarının izlenmesi ve uygun antineoplastik tedavinin uygulanması için klinik ortamda potansiyel argümanlar olarak değerlendirilebilir.

Anahtar kelimeler: kanser tanısı, kanser prognozu, radyorezistan, transkriptomik, meta-analiz, sistem biyolojisi

1. INTRODUCTION AND OBJECTIVES

1.1. Statement and Importance of the Problem

Cancer represents one of the leading causes of death worldwide. Despite the availability of anticancer therapeutics, the incidence, prevalence, and mortality of cancer are increasing exponentially. The diagnosis and prognosis of cancer are of critical importance in clinical decision-making. Therefore, there is an emerging need for the identification of powerful molecular diagnostic and prognostic markers, like genes and proteins, for the early detection and outcomes of cancer (Brennan et al., 2010a; Hartwell et al., 2006a; van 't Veer & Bernards, 2008a) https://www.cancer.gov/news-events/press-releases/2020 (accessed on 7 June 2021). Radiation therapy or radiotherapy (RT) represents one of the optimal, most widely used modalities in the treatment of multiple cancers, either alone or combined with other curative anti-cancer modalities like chemotherapy or immunotherapy (A. C. Begg et al., 2011a; Delaney et al., 2005a; Schoenhals et al., 2016a; C. Tang et al., 2014a). Although advances in radiotherapy contribute greatly to cancer patients' improvement of overall survival and quality of life, radioresistance represents a major limiting factor in the effective treatment of different types of cancers (Baskar et al., 2012a; Pavlopoulou et al., 2016a). The advent of high throughput sequencing technologies during the last decade enabled the generation of an enormous amount of cancer-relevant gene expression data (Pavlopoulou et al., 2015). Those data are stored in international public repositories including The Cancer Genome Atlas (TCGA) (Weinstein et al., 2013a), and Gene Expression Omnibus (GEO) of National Center for Biotechnology Informatıon (NCBI) (Edgar et al., 2002a). TCGA provides a vast amount of data and metadata from the major types and subtypes of cancer (Weinstein et al., 2013b), while NCBI's GEO includes heterogeneous data from high-throughput gene expression and genomic hybridization experiments (Edgar et al., 2002a).

1.2. Aims of the Study

The scope of our proposed study is to identify diagnostic, prognostic, and radioresistance gene signatures for diverse types of human cancers, as well as different stages of cancer, based on gene expression data currently available in TCGA, and NCBI's GEO. To this end, the expression patterns of all differentially expressed genes in cancers were investigated through an integrative in silico approach.

1.3. Hypothesis of the Study

Despite the exemplary efforts made to date, a comprehensive effort to determine potential diagnostic, prognostic, and radioresistance molecular signatures across human cancers is lacking. The identification of an optimal diagnostic, prognostic, and radioresistance gene signature comprising the minimum number of genes for each type of cancer, as well as different stages of cancer, is of particular clinical importance (Brennan et al., 2010b; Hartwell et al., 2006b; van 't Veer & Bernards, 2008b) (https://www.cancer.gov/news-events/press-releases/2020 accessed on 7 June 2021). Therefore, there is an emerging need for the identification of powerful biomarkers for cancer diagnosis, prognosis and response to radiotherapy given the increase in cancer cases and deaths. Specific gene signatures would enable proper decision-making for accurate diagnosis, prognosis, and treatment plan (Saville et al., 2019).

2. GENERAL INFORMATION

2.1. HOX transcript antisense RNA (HOTAIR)

The long non-coding RNAs (lncRNAs) are non-protein-coding RNAs ≥ 200 bp in length, transcribed by the RNA polymerase II. LncRNAs can be capped, polyadenylated and spliced, but they lack a functional open reading frame. It is estimated that approximately 27% (i.e., up to 60,000) of the annotated genes in the human genome encode lncRNAs, while the number of protein-coding genes ranges from 20,000 to 25,000 (Nagano & Fraser, 2011; Rinn & Chang, 2012). They largely interact with cellular biomolecules to regulate gene expression at the transcriptional, posttranscriptional, and epigenetic level for a myriad of cellular functions (**Figure 1**) (Geisler & Coller, 2013; Nagano & Fraser, 2011). LncRNAs have emerged as critical components of cancer pathophysiology, being involved in one or more hallmarks of cancer, such as proliferation and metastasis (Prensner & Chinnaiyan, 2011a; Schmitt & Chang, 2016). They can act either as oncogenes or tumor suppressors, or indirectly, through interaction with oncogenes and tumor suppressors, such as MYC

protooncogene (MYC) and tumor protein p53 (TP53), respectively(Prensner & Chinnaiyan, 2011b, 2011a).

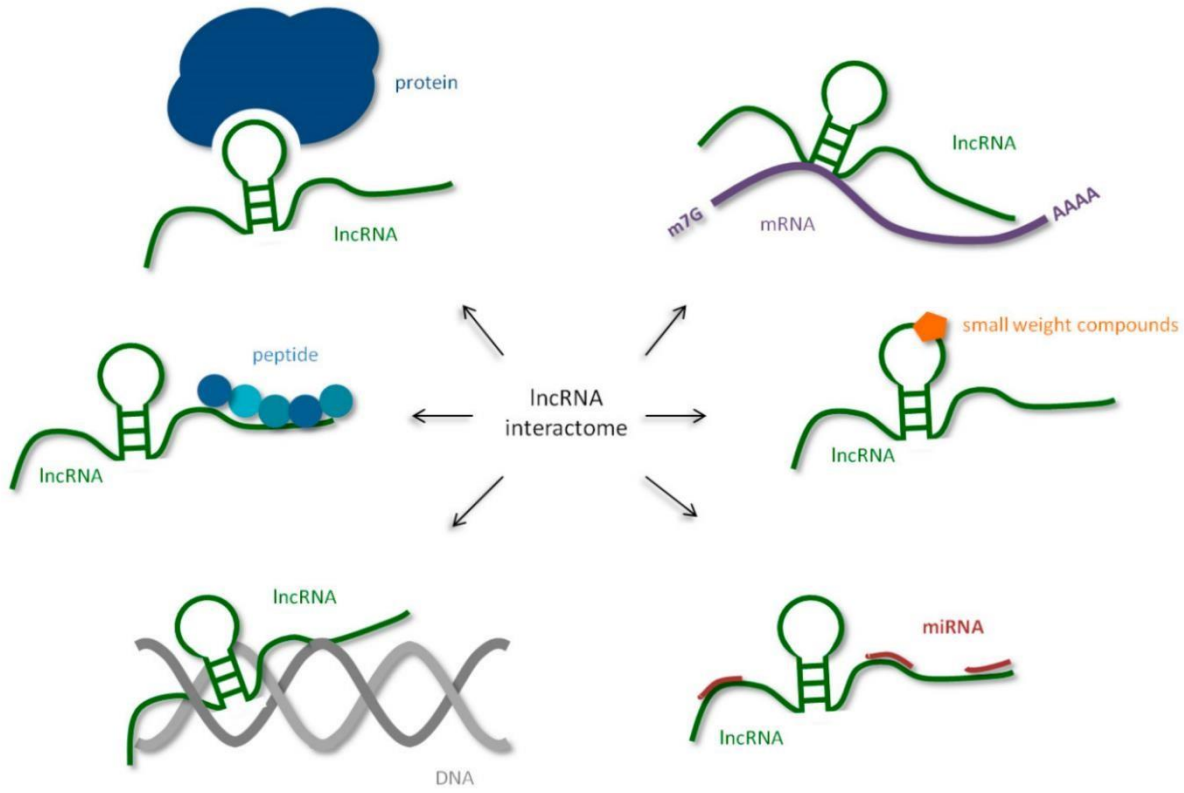


Figure 1. Associations of human lncRNsA with biological molecules (Kazimierczyk et al., 2020).

One of the most well-studied lncRNAs is HOX transcript antisense RNA (HOTAIR) that is located within the HOMEOBOX C (HOXC) gene cluster on chromosome 12q13.13 (Rinn et al., 2007). HOTAIR is 2158 bp long and consists of six exons. HOTAIR orthologues are restricted to eutherian mammals (Cunningham et al., 2019). HOTAIR is known to bind to the Polycomb Repressive Complex 2 (PRC2) and the histone H3K4 demethylase LSD1, and serve as a scaffold to assemble these regulators at the HOXD gene cluster, where it establishes a transcriptionally repressive chromatin structure, thereby resulting in epigenetic repression of the HOXD gene locus (**Figure 2**) (Tsai et al., 2010). HOTAIR has been shown to function as an oncogene, since its expression is dysregulated in multiple types of cancers, including breast, lung,

liver, renal, hepatocellular, gastric, nasopharyngeal, cervical, colorectal, bladder, pancreatic cancer, as well as melanoma, leukemia, etc. (Bhan & Mandal, 2015a; Cai et al., 2014; HOTAIR Is Upregulated in Acute Myeloid Leukemia and That Indicates a Poor Prognosis - PubMed, n.d.; Qu et al., 2019; Woo & Kingston, 2007). Furthermore, HOTAIR is suggested to promote cancer progression and contribute largely to cancer cell invasion and metastasis (Endo et al., 2013; Gupta et al., 2010a; Wu et al., 2015a; P. Zheng et al., 2018). The multifunctional HOTAIR is implicated in different aspects of cancer pathophysiology by regulating gene expression at the transcriptional, posttranscriptional, and epigenetic level (Gupta et al., 2010b; Hajjari & Salavaty, 2015a; Liu et al., 2014; Tang & Hann, 2018a). Of note, several studies suggest that HOTAIR expression is highly predictive of cancer patient survival rates in diverse cancer types (Chen et al., 2013; Q. Deng et al., 2014a; J. Li et al., 2015; Nakagawa et al., 2013; Tan et al., 2018; Troiano, Caponio, et al., 2017; Xiong et al., 2018; S. Zhang et al., 2014a; Y. Zhang, Zhou, et al., 2018)

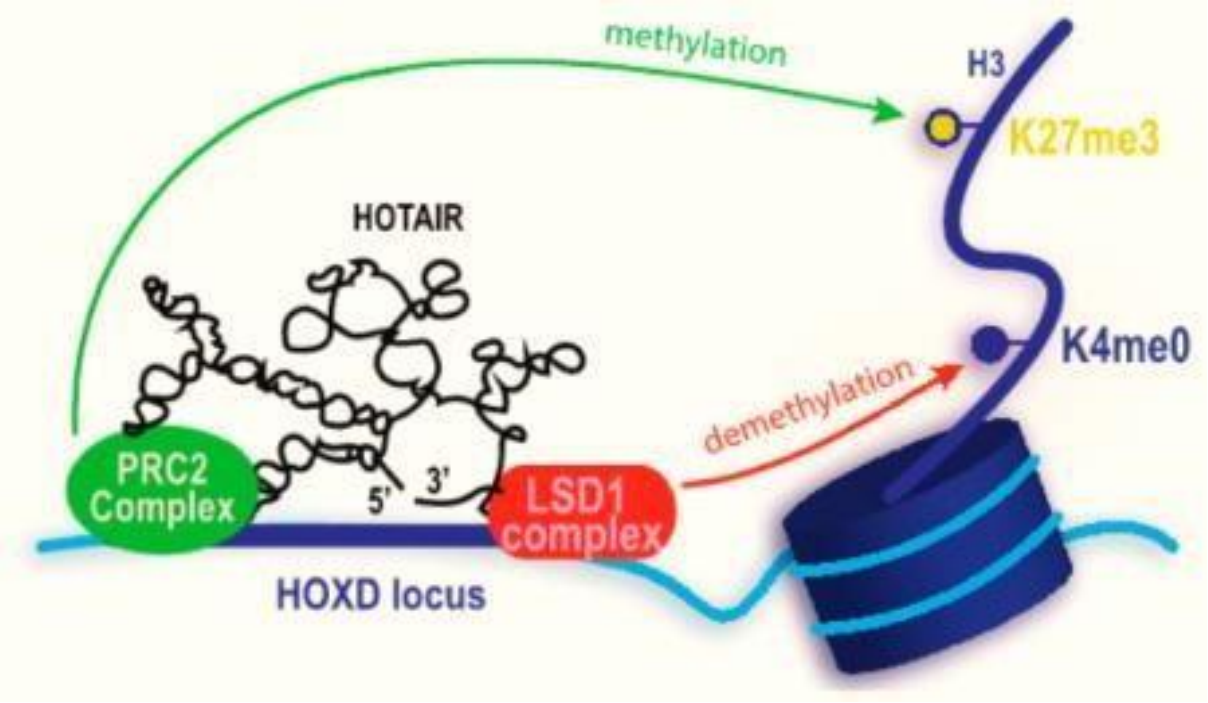


Figure 2. HOTAIR major biological functions (McFadden & Hargrove, 2016).

In this study, we conducted a comprehensive and updated meta-analysis to further investigate the prognostic value of HOTAIR expression for cancer patients. The potential clinical applications of our findings are also discussed towards the prognostic application of HOTAIR to multiple and different types of cancers.

2.2. Resistance to Radiation Therapy

Radiation therapy or radiotherapy (RT) represents one of the optimal, most widely used modalities in the treatment of multiple cancers, either alone or combined with other curative anti-cancer modalities like chemotherapy (A. C. Begg et al., 2011b; Delaney et al., 2005b) or immunotherapy (Schoenhals et al., 2016b; C. Tang et al., 2014b). It is estimated that approximately 50% of all cancer patients worldwide undergo radiotherapy throughout their illness trajectory (Baskar et al., 2012b).

Advances in radiotherapy contribute greatly to cancer patients' improvement of overall survival and quality of life (Baskar et al., 2012b). The aim of radiotherapeutic regimens is to sensitize cancer cells to IR, specifically and efficiently, in order to eliminate them and prevent cancer recurrence and relapse, minimizing at the same time the adverse effects of radiation on healthy tissue. RT affects cancer cells either directly, by inducing genomic (DNA) lesions, or indirectly, through the generation of DNA damaging intermediates through the interaction with water, like reactive oxygen/nitrogen species (ROS/RNS) and free radicals (e.g., hydrogen ion, hydroxide, etc.) (**Figure 3**) (Gong et al., 2021; Mikkelsen & Wardman, 2003; Yamamori et al., 2012).

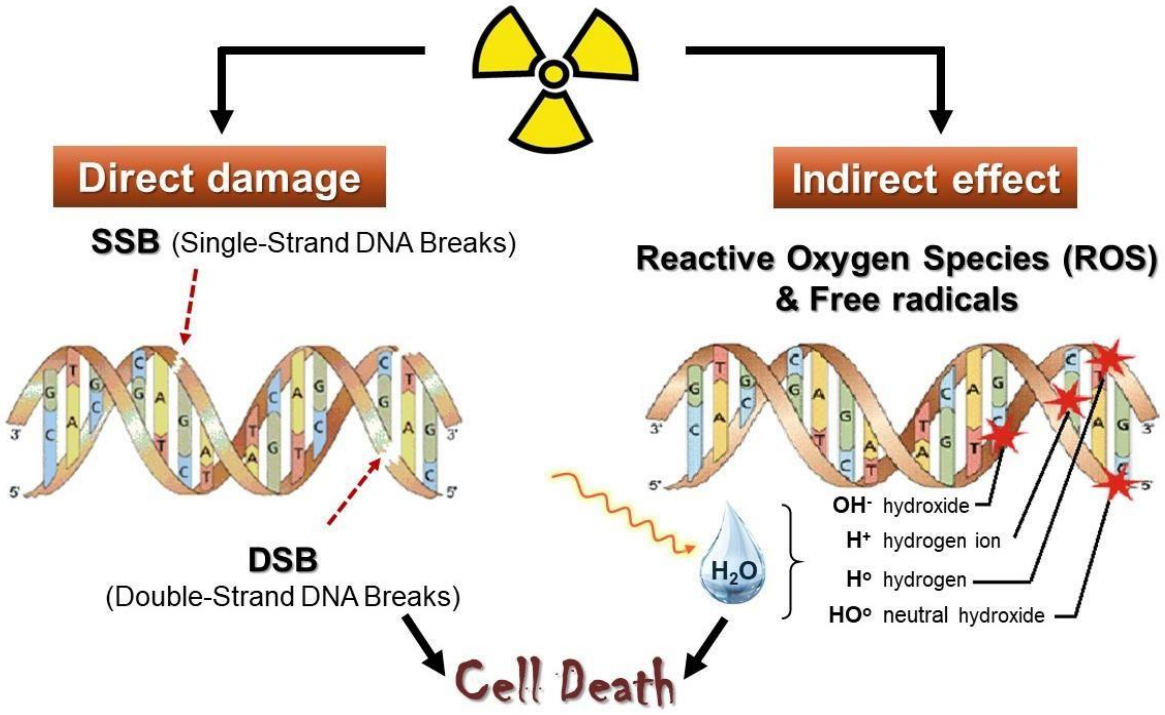


Figure 3. Illustration of the molecular mechanisms of ionizing radiation to cells.

However, cancer cells have the capacity to develop incredible tolerability and resistance to RT, thereby evading death. Radioresistance represents a major limiting factor in the effective treatment of different types of cancers. The response of tumor cells to radiation depends both on the resistance mechanisms of the cells and, also, on the accelerated repopulation of the tumor bulk by cells that have developed further radioresistance (Pavlopoulou et al., 2016b, 2017a). As noted in previous studies, the genes that are differentially expressed (either up- or down-regulated) between radioresistant (RR) and radiosensitive (RR) cancer cells are generally implicated in DNA damage response and repair (DDR/R) pathways, apoptosis, hypoxia, or response to oxidative stress, etc. (Pavlopoulou et al., 2016b, 2017a) (**Figure 4**). The complexity of radiation resistance mechanisms suggests the involvement of different and diverse biological mechanisms.

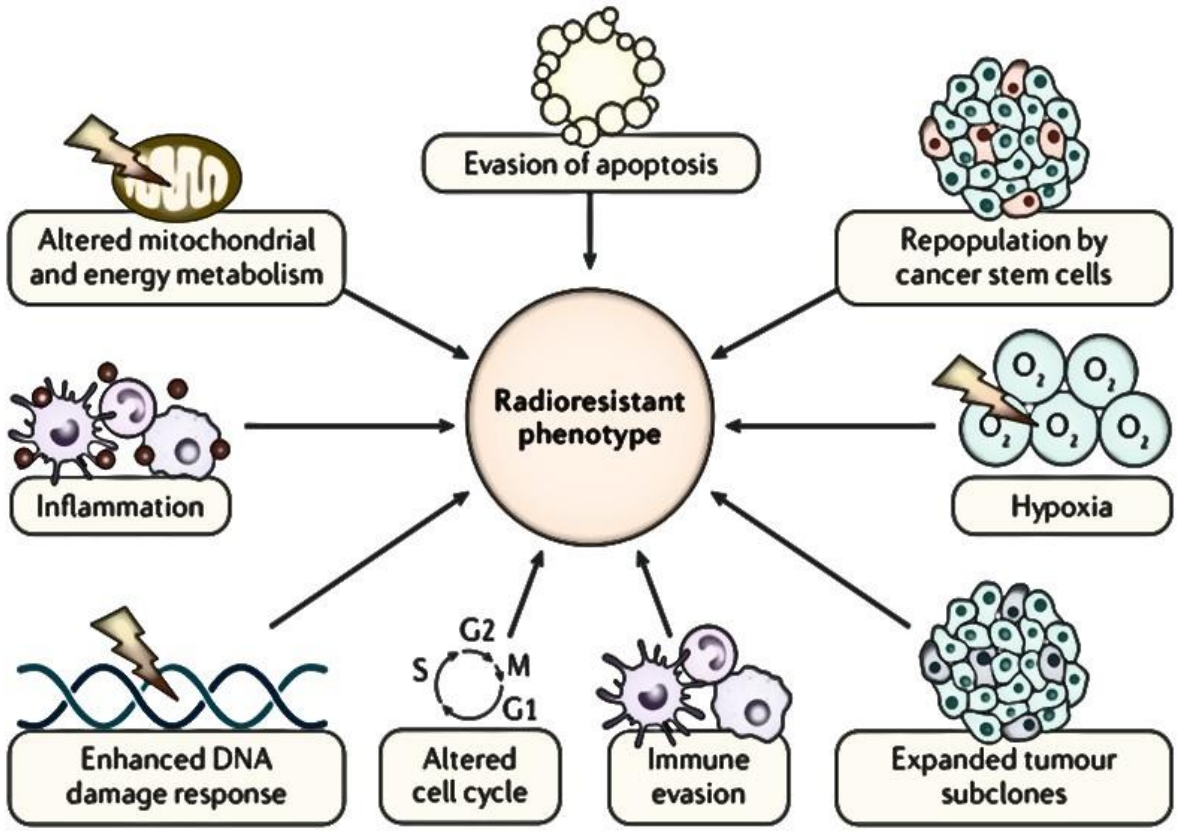


Figure 4. Molecular mechanisms underlying radioresistance (Buckley et al., 2020).

During the last decade, the advances in high-throughput (HTP) "omics" technologies (e.g., RNA-Seq and microarrays) enabled the generation of an enormous amount of gene expression data. Data produced with HTP technologies are stored in international public repositories such as NCBI's GEO (Gene Expression Omnibus) (https://www.ncbi.nlm.nih.gov/gds/); GEO DataSets contains both original records and curated datasets (Barrett et al., 2013; Clough & Barrett, 2016).

Accumulated knowledge over years of research on the biological effects of radiation points toward the development of holistic approaches to "big data" analysis by employing systems biology methodologies (Beheshti et al., 2019; Kanakoglou et al., 2020; Spratt & Speers, 2019; Unger, 2014). In this study, we employed a rigorous systems biology approach to unravel the molecular determinants of resistance of cancer cells to IR, based solely on HTP data. To this end, publicly available transcriptome datasets relevant to cancer cell response to radiation were retrieved

from NCBI's GEO, and specifically, cancer cell lines that displayed enhanced resistance to radiation. Statistical analyses were carried out to identify the differentially expressed genes (DEGs) between radioresistant and radiosensitive tumor cells. Furthermore, functional annotation of genes allowed us to identify specific biological pathways implicated in cancer cell resistance to radiation. Our findings could be applied in the laboratory and clinical setting as biomarkers for the design of targeted and personalized radiotherapy regimens in order to effectively sensitize cancer cells to radiation, enhance tumor control and thereby minimize tumor recurrence and metastasis.

2.3. Acute Myeloid Leukemia

Acute myeloid leukemia (AML) is a complex and heterogeneous blood condition, and the most prevalent form of acute leukemia in adults, representing 32% of all adult leukemia cases (https://www.cancer.net/cancer-types/leukemia-acute-myeloidaml/statistics (accessed on 7 June 2021)). The average age of AML diagnosis is 67 years (https://seer.cancer.gov/statfacts/html/amyl.html (accessed on 7 June 2021)). AML is caused by the uncontrolled clonal proliferation of abnormal immature white blood cells or leukemic blasts. Since blasts accumulate in the bone marrow, peripheral blood, and other organs of the body as the disease progresses, the production of normal hematopoietic stem cells is prevented. If AML is not treated within one year of diagnosis, it progresses quickly, leading to excessive bleeding, life-threatening infections, infiltration of lymph nodes with leukemic cells, or metastases to distant organs (Lowenberg et al., 1999; Shipley & Butera, 2009; Short et al., 2018).

AML patients are treated with intensive chemotherapy, targeted therapy, or bone marrow transplantation, which have overall improved survival outcomes, especially in younger patients (Kantarjian et al., 2021). However, despite the advances in the understanding of AML pathophysiology, mortality rates remain high. For example, in 2020, there were an estimated 60,530 new leukemia cases and 23,100 people died of this disease in the United States (https://www.cancer.gov/newsevents/press-releases/2020 (accessed on 7 June

2021)). Long-term survival rates also appear to be low, and most older patients with AML die from disease relapse (Farag et al., 2005; Klepin & Balducci, 2009; Pulte et al., 2009). The poor outcomes reflect two major challenges in AML management. First, most cases arise without any detectable early symptoms and patients usually present acute complications of bone marrow failure (Abelson et al., 2018). Second, complete hematologic remission (CR), which is a prerequisite for AML cure, cannot be readily achieved, since a significant percentage of AML patients eventually experience relapse after a seemingly successful treatment. This is because, although the therapeutic interventions can eliminate the majority of malignant cells, a few tumor cells inevitably evade therapy and remain persistently in the patient for years after the initial diagnosis (Páez et al., 2012). This condition is termed and is a major cause of AML relapse (Benton & Ravandi, 2017). These obstacles in AML treatment highlight the urgent need for identification of markers for accurate prognosis and estimation of CR.

The rapid progress in '-omics' and HTP technologies has promoted the identification of novel genetic abnormalities and holds promise for accelerating the systematic characterization of AML cancer genomes. Gene expression profiling of AML patients has catalyzed the discovery of novel leukemia subgroups and of prognostic signatures (Döhner et al., 2010). Moreover, despite the poor outcomes of AML, there are a few fortunate cases of patients who manage to survive over exceptionally extended periods, perhaps reflecting a remarkable resistance to MRD and disease progression. A novel rationale proposes that patients with exceptionally prolonged survival rates present distinct molecular profiles (Reifenberger et al., 2014a; Shen et al., 2021a) that could be exploited to unveil favorable prognostic factors that differentiate them from patients with a more typical course of disease (Saner et al., 2019a). Understanding the biological determinants of survival in these extreme outliers may provide a route to improving responses in more typical patients, particularly if these studies identify new biomarkers to guide drug selection or novel pathways that are targetable. Seeking determinants of prolonged survival is particularly important in cancers with generally poor outcomes (Saner et al., 2019a), such as AML. Considering these trends, in the present study, we applied integrative bioinformatics approaches

to process, analyze, and interpret publicly available AML-relevant transcriptome data towards the identification of biomarkers for favorable prognosis of AML.

3. MATERIALS AND METHODOLOGY

3.1. Type of the Study

This study was conducted based on in silico approaches and publicly available data.

3.2. Time and Place of the Study

This study was implemented in Izmir Biomedicine and Genome Institute from March 2018 to June 2022.

3.3. Materials of the Study

- All potential research articles which are related to the associations between HOTAIR expression and cancer survival were extracted from the biomedical literature database PubMed/MEDLINE (Fiorini et al., 2017a).
- RNA sequencing datasets for cancers and their corresponding healthy tissue were retrieved from TCGA (Zhang & Wang, 2015) (https://portal.gdc.cancer.gov/), and NCBI's GEO (Edgar et al., 2002b) (https://www.ncbi.nlm.nih.gov/gds/).
- Reference sequences and gene annotation files were collected from ENSEMBL (Cunningham et al., 2022) (http://www.ensembl.org/) and GENCODE (Frankish et al., 2021) (https://www.gencodegenes.org/).
- The protein-protein associations were collected from STRING (Search Tool for the Retrieval of Interacting Genes) (Szklarczyk et al., 2019a) (https://string-db.org/).

3.4. Variables of the Study

- The low versus high HOTAIR expression in different types of cancers.
- 11 replicates of the pediatric B-precursor acute lymphoblastic leukemia (ALL) cell line.
- 3 replicates of the Dukes' type C colorectal adenocarcinoma (COAD) cell line.
- 2 replicates of the metastatic mammary adenocarcinoma (MDA-MB-231) cell line.
- 2 replicates of the breast adenocarcinoma (MCF-7) cell line.
- 3 replicates of the luminal breast cancer (ZR-751) cell line.

- 76 tissue samples of Acute Myeloid Leukemia (TCGA-LAML).

3.5. Programming Languages and Tools for Data Processing

- Engauge Digitizer (Tierney et al., 2007).
- STATA statistical software version 13.0 (StataCorp. 2013. Stata Statistical Software: Release 13. College Station, TX: StataCorp LP) (conduct of statistical analyses and meta-analyses).
- Linux environment and Bash scripting (data collection and processing).
- SRA Tool Kit (next-generation sequencing data extraction tool) (Alnasir & Shanahan, 2015).
- FastQC (Quality control of sequencing data tool) (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/).
- HISAT2 (sequencing read mapping tool) (Kim et al., 2015).
- SAM Tools (Sequence Alignment/Map format) (Li et al., 2009).
- String Tie (transcript quantification tool) (Pertea et al., 2015).
- WebGestalt (WEB-based GEne SeT AnaLysis Toolkit) (Zhang et al., 2005).
- R programming environment (Gene expression profiling and visualization) (R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.Rproject.org/.)
- 'TCGAbiolinks' R Bioconductor package (Colaprico et al., 2016).
- 'EdgeR' R Bioconductor package (Robinson et al., 2010).
- 'clusterProfiler' R Bioconductor package (gene set enrichment analysis) (Yu et al., 2012).
- GSEA-P 2.0 software (gene set enrichment analysis) (Subramanian et al., 2005).
- STRING (Protein-Protein Interactions) (Szklarczyk et al., 2019b).
- GEPIA2 (Gene Expression Profiling Interactive Analysis) (Tang et al., 2017).

3.6. Data Analysis

3.6.1. Meta-analysis of the prognostic significance of HOTAIR in cancer

3.6.1.1. Search Strategy and Study Eligibility Criteria

This systematic review and meta-analysis were conducted by following strictly the PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) guidelines (Liberati et al., 2009).

The bibliographic database PubMed/MEDLINE (Fiorini et al., 2017b) was manually searched for published scientific studies on the associations between HOTAIR expression and prognosis in different types of cancers by using combinations of the relevant keywords: ("HOTAIR" OR "HOX transcript antisense RNA" or "HOXC cluster antisense RNA 4" or "HOXC-AS4" OR "HOXC11-AS1") and ("cancer" or "carcinoma" or "tumor" or "neoplasm" or "malignancy") and ("prognosis" or "survival" or "outcome" or "mortality" or "death"). The studies had to fulfill the following inclusion criteria so as to be considered eligible: (1) studies of human clinical trials, (2) studies including more than 30 patients in total, (3) the correlation between HOTAIR expression and cancer patients' survival was estimated, (4) availability of HR and 95% confidence interval (CI) or survival curves or sufficient data to calculate HR and 95% CI, (5) quantitative measurement (e.g., qPCR) of HOTAIR expression in cancers was included, and (6) studies published in English. Accordingly, the studies were excluded on the basis of the following exclusion criteria: (1) laboratory studies on animal models or cell lines; (2) reviews, meta-analyses, editorials, case reports, commentaries, unpublished data; (3) lack of sufficient data to estimate HR and 95% CI; and (4) samples other than tissue (e.g., blood, serum).

3.6.1.2. Study Selection, Data Extraction, and Quality

Relevant data were extracted from the included studies and recorded into an ad hoc Excel worksheet. In the case that the HR was not reported in the corresponding article, the data were extracted from the graphical survival plots (i.e., Kaplan-Meier curves) by using the Engauge Digitizer v.10.11 software.

3.6.1.3. Statistical Analyses

All statistical analyses were performed with the STATA statistical software version 13.0 (Stata Corporation, College Station, TX, USA) and Microsoft Excel. The heterogeneity among the included studies was estimated by Higgins I-squared (I^2) statistic as follows: $I^2 < 25\%$; no heterogeneity; $25\% < I^2 < 50\%$: low heterogeneity; $50\% < I^2 < 75\%$: moderate heterogeneity; $I^2 > 75\%$ high heterogeneity (Higgins et al., 2003; Higgins & Thompson, 2002). In the case of statistically significant heterogeneity ($I^2 > 50\%$ and $Ph < 0.05$), a random-effect model was applied, otherwise a fixed-effect model (DerSimonian & Laird, 1986, 2015) was used. Sensitivity analysis was performed by consecutive omission of individual studies to verify the consistency of outcomes. Potential publication bias was detected by Begg's funnel plot (Begg & Mazumdar, 1994) and Egger's test (Egger et al., 1997); a p-value less than 0.05 was indicative of statistically significant publication bias.

3.6.2. Identification of Molecular Determinants of Cancer Cell Resistance to Ionizing Radiation

3.6.2.1. Data acquisition

The public repository NCBI GEO DataSets was searched extensively for gene expression datasets using relevant keywords: ("radiation therapy" or "radiotherapy") and ("cancer" or "tumor") and ("resistance" or "tolerability") and ("sensitivity" or "responsive") and ("human" or "homo sapiens"). A total of three eligible datasets were selected:

The GEO series GSE97543 (Emons et al., 2017) (**Table 1**) includes global gene expression by microarray of both wild-type and radioresistant Dukes' type C colorectal adenocarcinoma (COAD) cell lines that were either non-irradiated or irradiated repeatedly with 2 Gray (Gy) of X-rays in order to acquire a radioresistant phenotype. The Agilent-026652 whole human genome microarray 4x44K, GPL13497 platform was used.

In GSE13280 (Marston et al., 2009a) (**Table 1**), genome-wide gene expression by microarray was performed of cell lines derived from pediatric B-precursor acute lymphoblastic leukemia (ALL) after 8 h in vitro exposure to 5 Gy IR. This dataset contains cell lines both resistant and responsive to radiation. The development of resistance and responsiveness to IR was assessed by measuring apoptosis in cells. The Affymetrix human genome U133A array, GPL96 platform was employed.

In GSE120798 (Gray et al., 2019a) (**Table 1**), three novel radioresistant breast cancer cell lines were established by exposing the corresponding parental cell lines MDA-MB-231 (metastatic mammary adenocarcinoma), MCF-7 (breast adenocarcinoma), and ZR-751 (luminal breast cancer) to increasing doses of X-rays for 2 and 8 h. Genome-wide gene expression analysis of both the parental and radioresistant cell lines (i.e., MDA/MDAR, MCF/MCFR, ZR/ZRR) was performed using high throughput sequencing. The NextSeq 550 (Homo sapiens), GPL21697 platform was used.

Table 1. Samples from each transcriptomic dataset analyzed in this study. The blue is represented Radiosensitive (RS), while red is Radioresistant (RR).

GSE13280		GSE120798		GSE97543	
GSM335244	ALL_responsive_1+IR	GSM3415941	MCF7_2h_2g_S13	GSM2571862	SW1463_wild type_4Gy_rep 1
GSM335245	ALL_responsive_2+IR	GSM3415943	MCF7_8h_2g_S15	GSM2571863	SW1463_wild type_4Gy_rep 2
GSM335246	ALL_responsive_3+IR	GSM3415946	MCF7_RR_2h_2g_S18	GSM2571864	SW1463_wild type_4Gy_rep 3
GSM335247	ALL_responsive_4+IR	GSM3415948	MCF7_RR_8h_2g_S20	GSM2571868	SW1463_radio resistant_4Gy_rep 1
GSM335248	ALL_responsive_5+IR	GSM3415952	ZR_8h_2g_S44	GSM2571869	SW1463_radio resistant_4Gy_rep 2
GSM335249	ALL_responsive_6+IR	GSM3415955	ZR_RR_2h_2g_S47	GSM2571870	SW1463_radio resistant_4Gy_rep 3
GSM335250	ALL_responsive_7+IR	GSM3415957	ZR_RR_8h_2g_S49		
GSM335251	ALL_responsive_8+IR	GSM3415960	MDA_2h_2g_S23		
GSM335252	ALL_responsive_9+IR	GSM3415962	MDA_8h_2g_S25		
GSM335253	ALL_responsive_10+IR	GSM3415965	MDA_RR_2h_2g_S28		
GSM335254	ALL_responsive_11+IR	GSM3415967	MDA_RR_8h_2g_S30		
GSM335266	ALL_ resistant_1+IR				

GSM335267	ALL_ resistant_2+IR
GSM335268	ALL_ resistant_3+IR
GSM335269	ALL_ resistant_4+IR
GSM335270	ALL_ resistant_5+IR
GSM335271	ALL_ resistant_6+IR
GSM335272	ALL_ resistant_7+IR
GSM335273	ALL_ resistant_8+IR
GSM335274	ALL_ resistant_9+IR
GSM335275	ALL_ resistant_10+IR
GSM335276	ALL_ resistant_11+IR

3.6.2.2. Microarray-Based Transcriptomic Data Analysis

For each microarray study, the gene expression data that represent the gene expression summary for every probe and every sample were recorded. In microarrays, many probes can map to the same Gene Symbol for various reasons, and, conversely, a probe may also map to more than one Gene Symbol if the probe sequence is not specific enough. A simple approach would be to use only the probes with one-to-one mapping for further analysis; however, this approach results in the loss of important information. To conduct an analysis based on genes and not probes, the probe identifiers were firstly converted into gene identifiers, according to Ramasamy et al. (Ramasamy et al., 2008) guidelines. To this end, GPL files that contained information about the gene symbols that correspond to probe IDs were used in order to resolve the "many-to-many" relationships between probes and genes by averaging the expression profiles for genes with more than one probe (Ramasamy et al., 2008). We identified the Gene Symbols with the usage of the HUGO Gene Nomenclature Committee (Braschi et al., 2019) and the National Center for Biotechnology Information (NCBI) GENE (Sayers et al., 2021).

The two-sample t-test was employed to identify genes differentially expressed between the case (RR) and control (RS) groups. However, a disadvantage of the t-test

in the analysis of microarray data is that if most of the experiments in a given study contain a relatively small number of samples per group, the assumption of normality is untenable. To resolve this, the statistical method t-test with bootstrapping was used (An Introduction to the Bootstrap - 1st Edition - Bradley Efron - R.J., n.d.) Bootstrapping provides an ideal method to generate accurate estimates of the standard errors when no formula for the sampling distribution is available or when available formulas make inappropriate assumptions (e.g., small sample size, non-normal distribution). In this study, bootstrap analysis was conducted with 1,000 replicates, a relatively high number, in order to generate accurate estimates of the standard errors.

A typical microarray experiment measures the expression of several thousand genes simultaneously across different conditions. When investigating for potential DEGs between two conditions, each gene is treated independently, and the t-test is performed on each gene separately. The incidence of false positives (i.e., genes falsely declared as DEGs) is proportional to the number of tests performed and the critical significance level (p-value cut-off). In order to account for multiple comparisons, a correction method proposed by Benjamini and Hochberg (Benjamini & Hochberg, 1995) which controls False Discovery Rate (FDR) was applied. FDR-controlling procedures have greater power (i.e., they can discover more statistically significant differences), at the expense of increased Type I error rate. Genes with adjusted pvalue (or q-value) less or equal to 0.05 were considered as statistically significant in this study. For all statistical analyses, the Stata 13 statistical software package (StataCorp: Stata Statistical Software: Release 14 - Google Akademik, n.d.) was used.

For the creation of heatmaps from microarray data, the average linkage clustering with Euclidean distance clustering method implemented in Heatmapper (http://heatmapper.ca/) was utilized.

3.6.2.3. RNA-Seq-Based Transcriptomic Data Analysis

For the RNA-Seq based transcriptome analysis, the following pipeline was utilized. The FASTQ files were extracted from the respective Sequence Read Archive (SRA) files containing raw RNA-Seq reads by using the SRA Tool Kit v.2.9.0 with the "fastq-dump

–gzip –skip-technical –readids –dumpbase –clip –split-3" command. The raw RNA-Seq reads in FASTQ files were aligned to the human reference genome GRCh38 (Ensembl version 97) by employing the splice junction aligner HISAT2 v.2.1.0 with "hisat2 -p dta -x {input.index} -U {input.fq} -S {out.sam}" parameters. The generated SAM file was converted to the respective binary BAM file by using SAM Tools v.1.9.0 with "samtools sort -@ 10 -o {output.bam} {input.sam}" commands. String Tie v.1.3.5 was utilized with "stringtie -e -B -p -G {input.gtf} -A {output.tab} -o {output.gtf} -l {input.label}{input.bam}" parameters for the measurement of gene expression levels. The reconstructed transcripts and transcript abundances were reported in the output GTF file. In order to detect differentially expressed genes between RR and RS samples, we utilized the EdgeR package v3.28.0 of the R statistical computation environment v.3.6.1. We firstly applied the trimmed mean of M-values (TMM) normalization (Robinson & Oshlack, 2010) implemented in EdgeR to the count data and we employed generalized linear models with "cell lines," "resistance," and "time" as factors. Then, estimating dispersion was computed with the estimateDisp function, and differential expression analysis between the two RNA-Seq groups (RR and RS) was performed using the glmFit and glmLRT functions of the EdgeR package v3.28.0 of the R statistical computation environment v.3.6.1. In order to detect statistically significant differentially expressed genes, the threshold for the absolute $\log_2$-fold change was set at two ($|\log_2 FC \geq 2|$), and for the FDR (Benjamini & Hochberg, 1995)-corrected p-value at 0.05. All statistical calculations for the RNASeq data were performed by using the R software environment. The pheatmap package of R (https://CRAN.R-project.org/package=pheatmap) was utilized to generate a heatmap from RNA-Seq data.

3.6.2.4. Pathway Enrichment Analysis

To assign biological role(s) to the genes under study that are associated with biological pathways, gene set enrichment analysis (GSEA), or functional enrichment analysis, was conducted. GSEA is a method to identify biological processes or pathways that are over-represented in a large set of genes. To this end, WebGestalt (WEB-based GEne SeT AnaLysis Toolkit) was employed to identify statistically significant overrepresented WikiPathways (Kutmon et al., 2016) cancer terms in the sets of genes;

the threshold for the FDR-corrected p-value was set at 10^{-3}, and hypergeometric distribution analysis was used.

3.6.2.5. Functional Interactions Networks

The associations among the molecules under study were investigated and visualized with the usage of STRING (Search Tool for Retrieval of Interacting Genes/Proteins) version 11.0, a database of either known or predicted, direct or indirect, gene/protein associations derived from diverse resources. The highest confidence interaction score (≥0.9) was chosen to display the associations amongst genes/proteins in the generated network.

3.6.2.6. Survival Analysis

The prognostic potential of the 36 radiogenes was investigated in three types of cancers, namely, breast invasive carcinoma (BRCA), colon adenocarcinoma (COAD), and acute myeloid leukemia (LAML), a type of haematologic cancer like ALL, through the web-based tool GEPIA2, based on data acquired from TCGA. The cancer patient cohort was divided into the high-risk and low-risk categories; the cut-offs for low and high gene expression level patient cohorts were set at 50%.

3.6.2.7. Tissue-Wise Differential Gene Expression Analysis

GEPIA2, which contains gene expression data from cancer and corresponding normal tissues from TCGA and GTEx respectively, was used to investigate the differential expression patterns of the radiogenes under study in BRCA, COAD, and LAML cancer-normal tissue; $|\log FC| \geq 2$ and FDR-corrected p-value ≤ 0.05.

3.6.3. Identification of DEGs Correlated with Long-Term Survival in AML patients

Patient data from the TCGA LAML cohort were retrieved via the 'TCGAbiolinks' package (data.type: Gene Expression Quantification, workflow.type: HTSeq-Counts). Only the genes that were expressed in at least half of the patients were retained and the outliers that were >6 standard deviations (SD) away from the median were

removed. We defined the groups of long-term survivor (LTS) patients (2 × SD = 1567 days, 3 × SD = 2168 days, and 5-year survival = 1825 days) and calculated the respective differentially expressed genes (DEGs) in the population that died within 365 days using edgeR, with log2FC cutoff > 1. Afterwards, we performed gene set enrichment analysis (GSEA) on the DEGs for each comparison, as well as for the common 131 genes that were deregulated in all three analyses. For GSEA, the GSEAP 2.0 software was used. Enriched Gene Ontology (GO) terms were plotted against the negative log10 of their individual FDR value (<0.05).

4. RESULTS

4.1. HOTAIR as an Unfavorable Prognostic Biomarker for Cancer

4.1.1. Study Selection and Characteristics of Eligible Studies

A total of 264 relevant published scientific studies were retrieved from the biomedical literature (up to 31 December 2018). According to the inclusion and exclusion criteria, 53 studies were ultimately included in this meta-analysis, as shown in **Figure 5**. The main characteristics of the included studies are summarized in **Table 2**, where the following information was recorded: first author's surname; year of publication; country of origin; type of cancer; follow-up period (in months); total number of patients; detection assay for HOTAIR expression; hazard ratio (HR) and the corresponding 95% CI for overall survival (OS), recurrence-free survival (RFS), disease-free survival (DFS), progression-free survival (PFS), metastasis-free survival (MFS); survival data extraction method; and specimen type. Collectively, 4873 patients from 55 cohorts between 2010 and 2018 were included. The included studies reported a follow-up period ranging from 36 to 276 months. The level of HOTAIR expression was measured with quantitative reverse transcription polymerase chain reaction (qRTPCR) in all of the included studies, except one where HOTAIR expression was estimated by microarrays (**Table 2**).

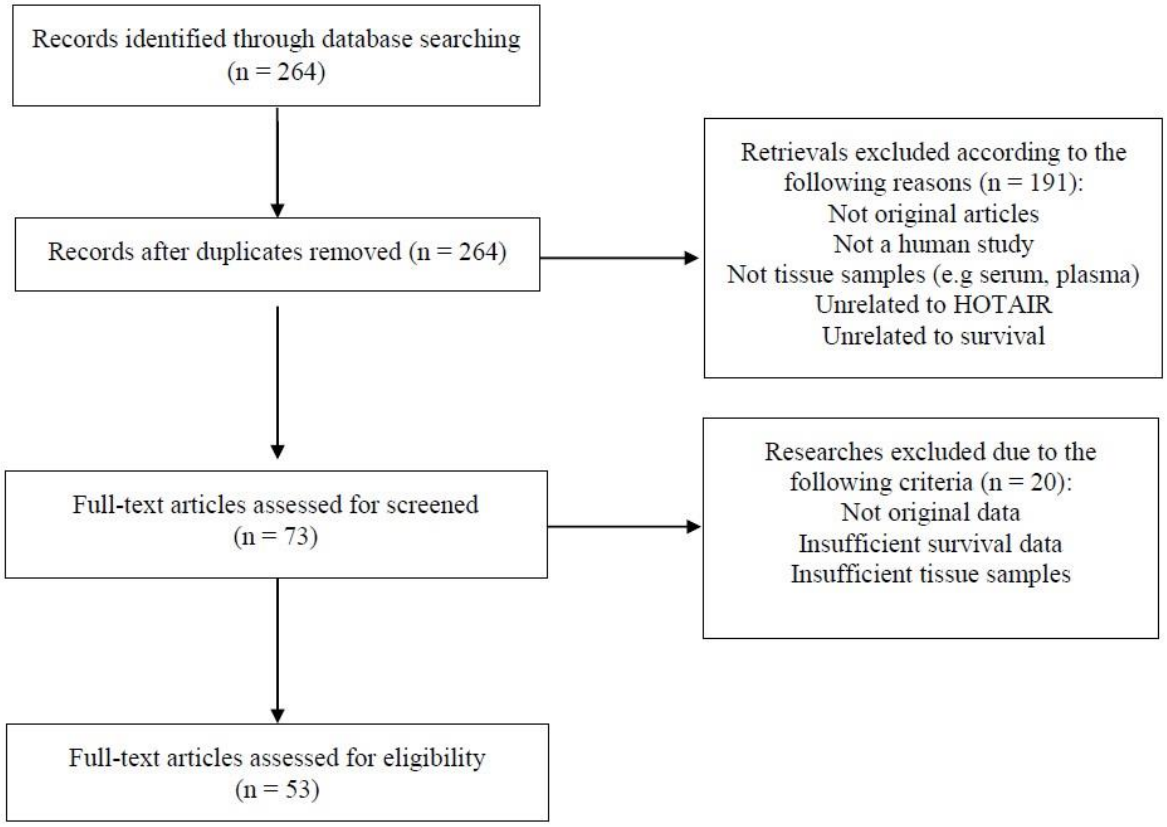


Figure 5. Flow chart of the process for study selection.

Table 2. Main characteristics of the studies included in the meta-analysis.

Author, Year	Country	Cancer	Max. Follow-Up (Months)	Sample	Case Number			OS		DFS/RFS		MFS/PFS		Assay method	Data extraction method
					High expression	Low expression	Total	HR (95% CI)	*p* value	HR (95%CI)	*p* value	HR (95% CI)	*p* value		
Gupta, 2010 [1]	USA	Breast Cancer	240	Tissue	44	88	132	2.76(1.45-3.3)	0.036	NM	NM	3.53 (2.78-4 .89)	0.017	qRT-PCR	K-M
Geng, 2011 [2]	China	HCC	36	Tissue	NM	NM	50	NM	NM	2.24(1.49-3.36)	0,049	NM	NM	qRT-PCR	K-M
Kogo, 2011 [3]	Japan	CRC	60	Tissue	20	80	100	5.62(1.52-9.57)	0.008	NM	NM	NM	NM	qRT-PCR	reported
Yang, 2011 [4]	China	HCC	45	Tissue	32	28	60	NM	NM	3,56(1.67-7.63)	0.001	NM	NM	qRT-PCR	reported
Lu, 2012 [5]	Italy	Breast Cancer	108	Tissue	NM	NM	336	0.43(0.21-0.89)	0.022	0.47(0.26-0.87)	0.016	NM	NM	qRT-PCR	reported
Niinuma, 2012 [6]	Japan	GIST	200	Tissue	11	28	39	3.8(0.7-21.2)	0.123	NM	NM	NM	NM	qRT-PCR	reported
Chen, 2013 [7]	China	ESCC	60	Tissue	27	51	78	2.40(1.35-4.28)	0.003	NM	NM	2.34 (1.22-4.48)	0,01	qRT-PCR	reported
Endo, 2013 [8]	Japan	IGC	68	Tissue	23	13	36	0.63(0.34-1.86)	0.137	NM	NM	NM	NM	qRT-PCR	K-M
Endo, 2013 [8]	Japan	DGC	60	Tissue	20	12	32	3.08(1.77-5.35)	0.0	NM	NM	NM	NM	qRT-PCR	K-M
Ge, 2013 [9]	China	ESCC	100	Tissue	90	47	137	3.16(1.53-6.52)	0.002	NM	NM	4.47(1.99-10.06)	0,001	qRT-PCR	reported
Ishibashi, 2013 [10]	Japan	HCC	36	Tissue	13	51	64	2.84(1.91-4.58)	0.041	NM	NM	NM	NM	qRT-PCR	K-M
Li, 2013 [11]	China	LSCC	60	Tissue	33	39	72	2.86(1.15 -7.07)	0.023	NM	NM	NM	NM	qRT-PCR	reported
Li, 2013 [12]	China	ESCC	60	Tissue	30	70	100	1.91(1.06-3.99)	0.033	NM	NM	NM	NM	qRT-PCR	reported

Liu, 2013 [13]	China	NSCLC	60	Tissue	21	21	42	2.043(0.91-4.58)	0.048	NM	NM	NM	NM	qRT-PCR	K-M
Lv, 2013 [14]	China	ESCC	70	Tissue	49	44	93	1.67(1.02-2.79)	0.049	NM	NM	NM	NM	qRT-PCR	K-M
Nakagawa, 2013 [15]	Japan	NSCLC	50	Tissue	17	60	77	NM	NM	1.81 (1.09-3.74)	0,047	NM	NM	qRT-PCR	K-M
Nie, 2013 [16]	China	NPC	82	Tissue	91	69	160	1.9 (1.13-3.19)	0.012	1.41 (0.95-2.09)	0.47	1.92(1.11-3.31)	0,018	qRT-PCR	K-M
Sorensen, 2013 [17]	Denmark	Breast Cancer	276	Tissue	79	85	164	NM	NM	NM	NM	1.75(1.13-2.71)	0.012	Microarray	reported
Xu, 2013 [18]	China	Gastric cancer	75	Tissue	56	27	83	0.47(0.22-0.99)	0.04	NM	NM	NM	NM	qRT-PCR	reported
He, 2014 [19]	China	EC	48	Tissue	62	83	145	3.04(2.13-4.58)	0.026	NM	NM	NM	NM	qRT-PCR	K-M
Huang, 2014 [20]	China	Cervical cancer	55	Tissue	109	109	218	2.86(1.26-6.49)	0.012	NM	NM	NM	NM	qRT-PCR	reported
Lee, 2014 [21]	Republic of Korea	Gastric cancer	48	Tissue	28	20	48	NM	NM	2.21(0.53-9.16)	0.141	NM	NM	qRT-PCR	reported
Liu, 2014 [22]	China	Gastric cancer	48	Tissue	39	39	78	2.7(1.36-4.34)	0.023	NM	NM	NM	NM	qRT-PCR	K-M
Okugawa, 2014 [23]	Japan	Gastric cancer	60	Tissue	77	73	150	1.77(1.06-2.95)	0.028	NM	NM	NM	NM	qRT-PCR	reported
Qiu, 2014 [24]	China	EOC	79	Tissue	32	32	64	1.87(1.04-5.31)	0.041	2.54 (1.18-5.45)	0.034	NM	NM	qRT-PCR	reported
Svoboda, 2014 [25]	Czech Republic	Colorectal cancer	54	Tissue	36	37	73	4.46(1.02-19.79)	0.048	NM	NM	NM	NM	qRT-PCR	reported
Wu, 2014 [26]	China	Colon Cancer	72	Tissue	40	80	120	3.92(1.23-12.50)	0.021	NM	NM	3.88(1.37-10.98)	0.011	qRT-PCR	K-M
Yan, 2014 [27]	China	Bladder Cancer	60	Tissue	90	20	110	4.71(2.89-8.71)	0,00	NM	NM	NM	NM	qRT-PCR	reported
Heubach, 2015 [28]	Germany	UHC	200	Tissue	27	81	108	2.20(1.23-3.93)	0.008	NM	NM	NM	NM	qRT-PCR	reported
Kim, 2015 [29]	Republic of Korea	Cervical cancer	60	Tissue	89	22	111	NM	NM	5.28(1.01-27.74)	0,049	NM	NM	qRT-PCR	reported

Liu, 2015 [30]	China	Gastric cancer	40	Tissue	24	37	61	NM	NM	2.6(1.74-3.89)	0.0	NM	NM	qRT-PCR	K-M
Ma, 2015 [31]	China	Gastric cancer	60	Tissue	18	53	71	2.10(1.10-4.03)	0.022	NM	NM	NM	NM	qRT-PCR	reported
Martinez-Fernandez, 2015 [32]	Spain	NMIBC	38	Tissue	17	16	33	NM	NM	NM	NM	1.86(0.58-5.96)	0.296	qRT-PCR	K-M
Martinez-Fernandez, 2015 [32]	Spain	NMIBC	38	Tissue	30	33	63	NM	NM	3.78(2.40-5.96	0.000	NM	NM	qRT-PCR	K-M
Qiu, 2015 [33]	China	SOC	96	Tissue	34	34	64	1.90(1.01-3.56)	0.046	NM	NM	NM	NM	qRT-PCR	reported
Wu, 2015 [34]	China	OSCC	60	Tissue	25	25	50	1.91(1.33-2.74)	0.000	NM	NM	NM	NM	qRT-PCR	K-M

Wu, 2015 [35]	China	AML	40	Tissue	52	33	85	3.37(0.99-8.31)	0.008	4.68(2.81-7.79)	0.000	NM	NM	qRT-PCR	reported
Wu, 2015 [36]	China	OSCC	96	Tissue	38	38	76	1.18(0.68-2.84)	0,03	1.11 (0.78-2.54)	0,044	NM	NM	qRT-PCR	reported
Xing, 2015 [37]	China	AML	36	Tissue	68	68	136	2.03(1.16-3.55)	0.007	0.61(0.37-1.00)	0.034	NM	NM	qRT-PCR	reported
Zhang, 2015 [38]	China	Gastric cancer	45	Tissue	35	15	50	1 .87(1.46-2.1)	0.028	NM	NM	NM	NM	qRT-PCR	K-M
Zhao, 2015 [39]	China	Gastric cancer	65	Tissue	84	84	168	1.47(1.04-2.06)	0.027	NM	NM	NM	NM	qRT-PCR	reported
Luczak, 2016 [40]	Poland	EC	96	Tissue	56	100	156	1.44(0.81-3.19)	0,03	NM	NM	NM	NM	qRT-PCR	K-M
Luo, 2016 [41]	China	Colon cancer	70	Tissue	NM	NM	80	1.99(1.4-2.8)	0.0	NM	NM	NM	NM	qRT-PCR	K-M
Sun, 2016 [42]	China	Cervical cancer	50	Tissue	49	10	59	1.31(0.79-2.26)	0,02	NM	NM	NM	NM	qRT-PCR	K-M
Yan, 2016 [43]	China	DLBCL	120	Tissue	25	25	50	3.131.22-8.04)	0.018	NM	NM	NM	NM	qRT-PCR	reported
Zhang, 2016 [44]	China	Acute leukemia	40	Tissue	19	77	96	2.41(1.25-4.62)	0.005	NM	NM	NM	NM	qRT-PCR	K-M
Chen, 2017 [45]	China	Gastric cancer	62	Tissue	33	32	65	1.99(1.06-3.77)	0.033	NM	NM	NM	NM	qRT-PCR	reported
Hu, 2017 [46]	China	RCC	50	Tissue	32	11	43	0.72(0.20-2.55)	0.62	NM	NM	NM	NM	qRT-PCR	K-M
Katayama, 2017 [47]	Japan	RCC	100	Tissue	21	43	64	1.82(1.06-3.88)	0,02	NM	NM	NM	NM	qRT-PCR	K-M
Luan, 2017 [48]	China	MM	60	Tissue	30	30	60	1.36(0.79-2.83)	0,01	NM	NM	NM	NM	qRT-PCR	K-M
Xu, 2017 [49]	China	*EC	36	Tissue	20	20	40	2.69(1.14-6.33)	0.032	NM	NM	NM	NM	qRT-PCR	K-M
Zhang, 2017 [50]	China	Thyroid cancer	60	Tissue	NM	NM	35	2.21(1.38-3.54)	0.001	NM	NM	NM	NM	qRT-PCR	reported
Dong, 2018 [51]	China	Gastric cancer	60	Tissue	22	10	32	2.26(0.74-6.89)	0.158	NM	NM	NM	NM	qRT-PCR	K-M
Huang, 2018 [52]	China	Colorectal cancer	110	Tissue	26	26	52	2.56(0.91-7.35)	<0.01	NM	NM	NM	NM	qRT-PCR	reported
Xiao, 2018 [53]	China	Colorectal cancer	60	Tissue	52	52	104	1.45(0.87-2.43)	0.041	NM	NM	NM	NM	qRT-PCR	K-M

Abbreviations: OS, overall survival; RFS, recurrence-free survival; DFS, disease-free survival; MFS, metastasis-free survival; PFS, progression-free survival; HR, hazard ratio; CI, confidence interval; qRT-PCR, quantitative reverse transcription polymerase chain reaction; NM: not mentioned; K-M, Kaplan-Meier plot; AML, acute myeloid leukemia; CRC, colorectal cancer; DGC, diffuse gastric cancer; DLBCL, diffuse large B cell lymphoma; ESCC, esophageal squamous cell carcinoma; EC, endometrial carcinoma; EOC, epithelial ovarian cancer; * EC, esophageal cancer; GIST, gastrointestinal stromal tumors; HCC, hepatocellular carcinoma; IGC, intestinal gastric cancer; LSCC, laryngeal squamous cell carcinoma; MM, malignant melanoma; NSCLC, non-small cell lung cancer; NPC, nasopharyngeal carcinoma; NMIBC, non-muscle-invasive bladder cancer; OSCC, oral squamous cell carcinoma; RCC, renal cell carcinoma; SOC, serous ovarian cancer; and UHC, urothelial carcinoma.

4.1.2. Association between High HOTAIR Expression and Overall Survival in Diverse Cancers

A total of 45 studies were included for overall survival (OS). We found a statistically significant relationship between elevated HOTAIR expression and poor OS (random-effects model: pooled HR = 2.00; 95% CI: 1.77–2.27; p < 0.001), with marginally moderate heterogeneity (I2 = 50.2%; Ph < 0.001) (**Figure 6a**). Subgroup analyses were performed based on the type of cancers, ethnic group, and data extraction method (**Figure 7**). When the studies were classified based on major cancer types (according to NCBI's medical subject headings (MeSH) (https://www.ncbi.nlm.nih.gov/mesh/)), a significant association was found between HOTAIR overexpression and poorer OS in solid cancers, such as gastrointestinal cancers (fixed-effects model: pooled HR = 1.96; 95% CI: 1.65–2.35; p < 0.001), liver cancers (fixed-effects model: pooled HR = 2.84; 95% CI: 1.83–4.40; p < 0.001), head and neck cancers (fixed-effects model: pooled HR = 1.93; 95% CI: 1.53–2.43; p < 0.001), and urogenital cancers (random-effects model: pooled HR = 2.11; 95% CI: 1.58–2.84; p < 0.001), as well as liquid cancers, including leukemia (fixed-effects model: pooled HR = 2.32; 95% CI: 1.56–3.44; p < 0.001) and lymphoma (fixed-effects model: pooled HR = 3.13; 95% CI: 1.22–8.04; p < 0.001). Of note, the heterogeneity was reduced significantly in the individual cancer types (**Figure 7a**). In the subgroup analysis based on ethnicity, a statistically significant worse OS was observed for Asians (fixed-effects model: pooled HR = 2.04; 95% CI: 1.81–2.31; p < 0.001). Regarding the Caucasian subgroup, despite the relatively high HR, the relationship cannot be considered robust because the p-value is slightly higher that the cutoff value (randomeffects model; pooled HR = 1.65; 95% CI: 0.82–3.33; p = 0.077) (**Figure 7b**). In stratified analysis, according to data extraction method, HOTAIR was found to have a significant prognostic value irrespectively of the data source, that is, the HR reported in the articles (random-effects model: pooled HR = 2.05; 95% CI: 1.64–2.57; p < 0.001) or extracted from the survival curves (fixed-effects model: pooled HR = 2.01; 95% CI: 1.75–2.30; p < 0.001) (**Figure 7c**).

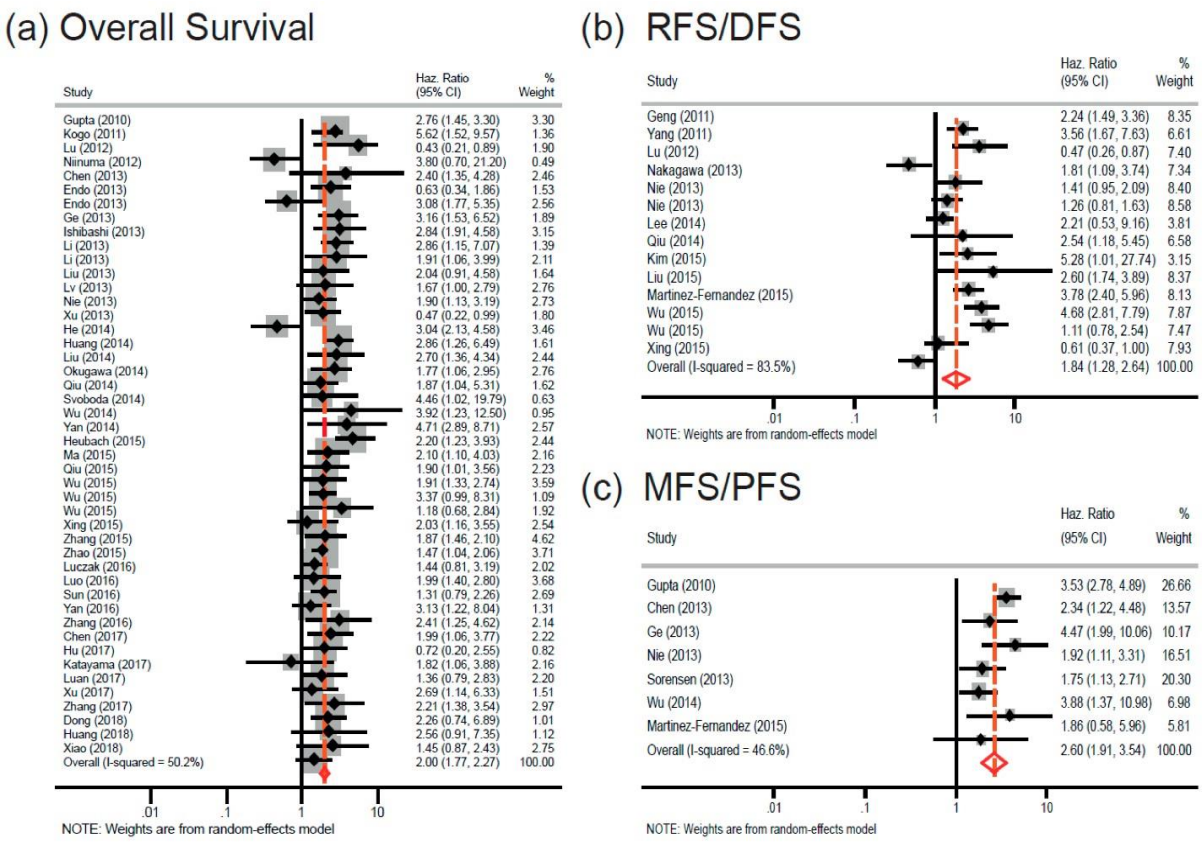


Figure 6. Forest plots of combined analyses on the association of survival with HOTAIR expression. (a) Forest plot of OS analysis, (b) forest plot of RFS/DFS analysis, and (c) forest plot of MFS/PFS analysis. Abbreviations: HR, Hazard ratio; OS, overall survival; RFS, recurrence-free survival; DFS, disease-free survival; MFS, metastasisfree survival; and PFS, progression-free survival.

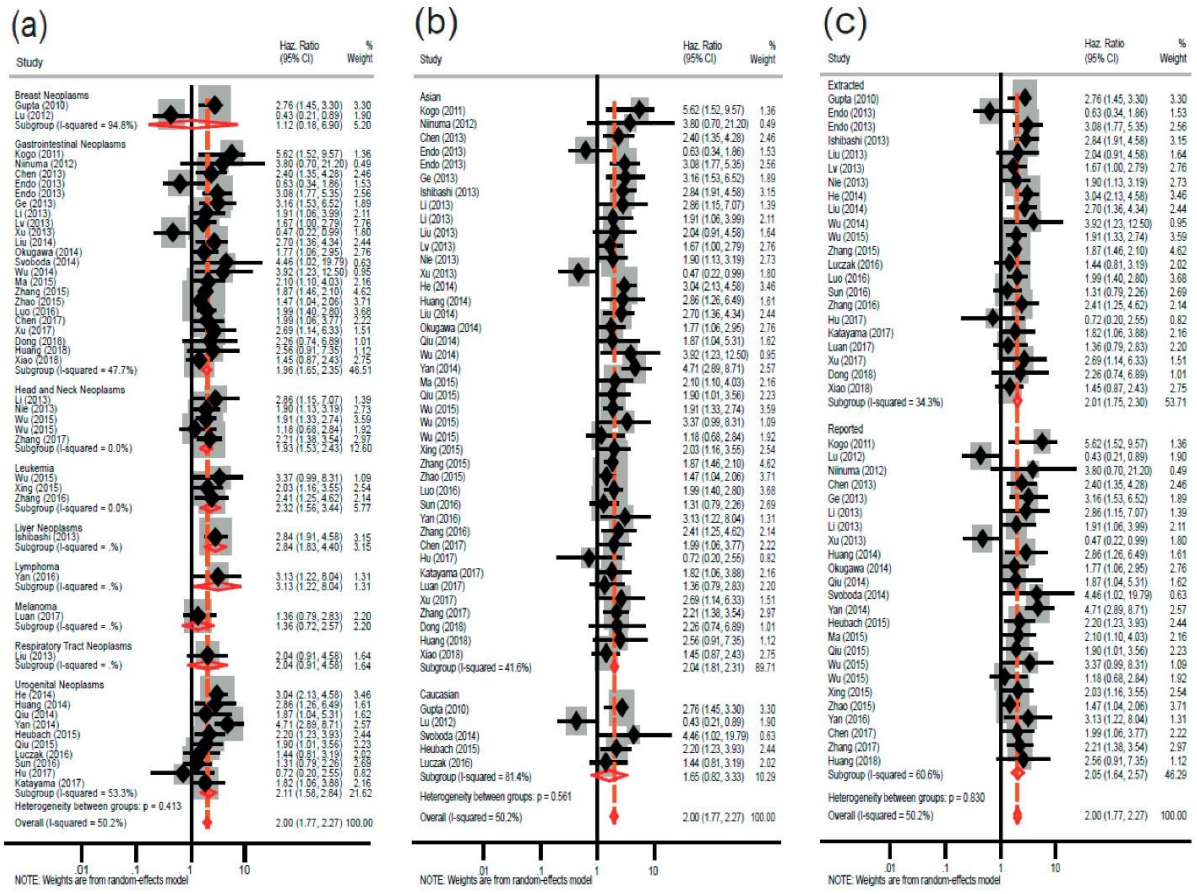


Figure 7. Forest plots of combined analyses for overall survival (OS) associated with HOTAIR expression in different groups. (a) Forest plot for different types of cancers, (b) Forest plot for different ethnic groups, and (c) Forest plot for different data extraction methods.

4.1.3. HOTAIR Overexpression is Associated with Cancer Recurrence and Progression

To investigate the relationship between HOTAIR expression and cancer recurrence or relapse, the recurrence-free survival (RFS) and disease-free survival (DFS) studies were combined; collectively accounting for 14 studies. Increased HOTAIR expression was found to be strongly related to cancer recurrence (pooled HR = 1.84; 95% CI = 1.28–2.64; p = 0.001). A random-effects model was applied because of the high heterogeneity (I2 = 83.5%; Ph < 0.001) across studies (**Figure 6b**).

Furthermore, there are seven studies for combined metastasis-free survival (MFS) and progression-free survival (PFS). Of importance, high HOTAIR expression was predicted to be associated significantly with worse MFS/PFS (pooled HR = 2.60;

95% CI: 1.91–3.54; p < 0.001). A fixed-effects model was used because of the relatively low heterogeneity (I2 = 46.6%; Ph = 0.081) (**Figure 6c**).

4.1.4. Publication Bias

Publication bias was detected by Begg's funnel plot and Egger's test. There was no obvious asymmetry in Begg's funnel plots of OS, RFS/DFS, and MFS/PFS (**Figure 8**). Additionally, the p-values of Egger's tests were all greater than 0.05, indicating no potential publication bias (OS: p = 0.73; RFS/DFS: p = 0.70; MFS/PFS: p = 0.64).

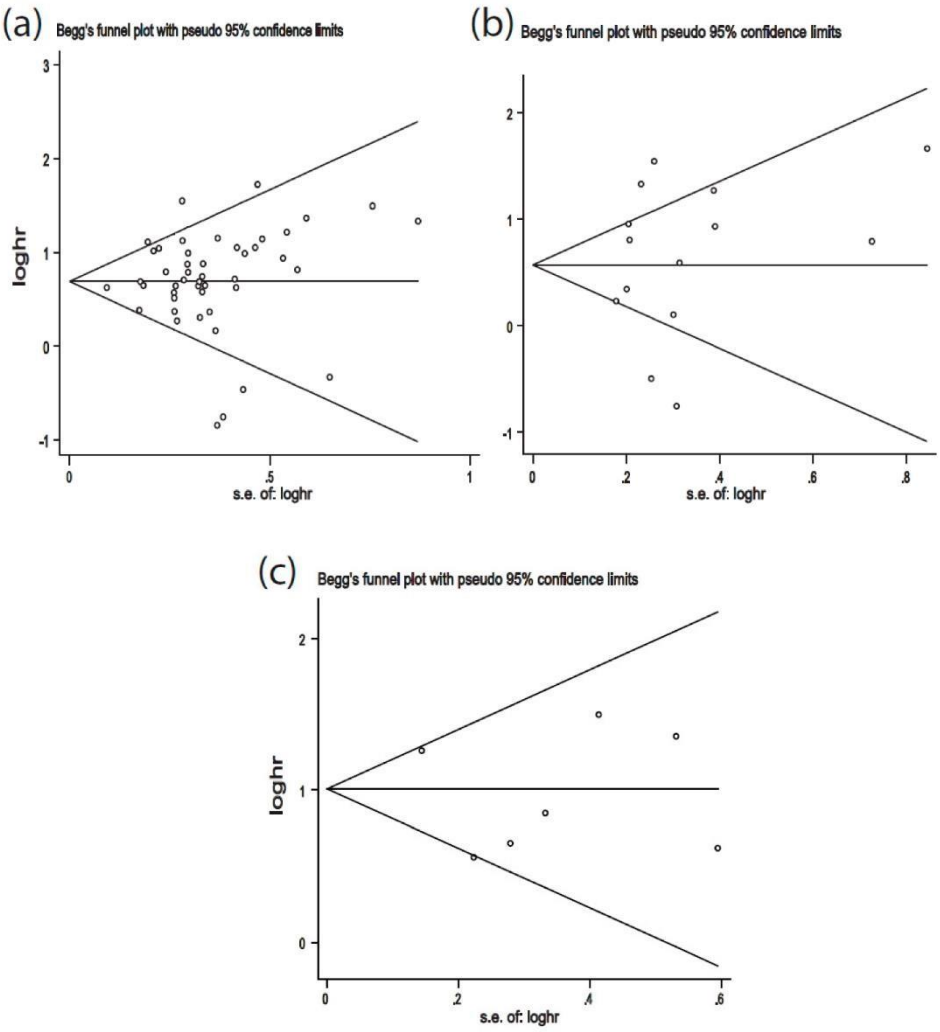


Figure 8. Begg's funnel plots of publication bias. (a) Begg's funnel plot of publication bias for OS; (b) Begg's funnel plot of publication bias for RFS/DFS; (c) Begg's funnel plot of publication bias for MFS/PFS. Each circle represents a separate study.

4.1.5. Sensitivity Analysis

Sensitivity analyses did not indicate alterations in the results due to the inclusion of any individual study (**Figure 9**), that is, no single study affected the pooled HR or 95% CI.

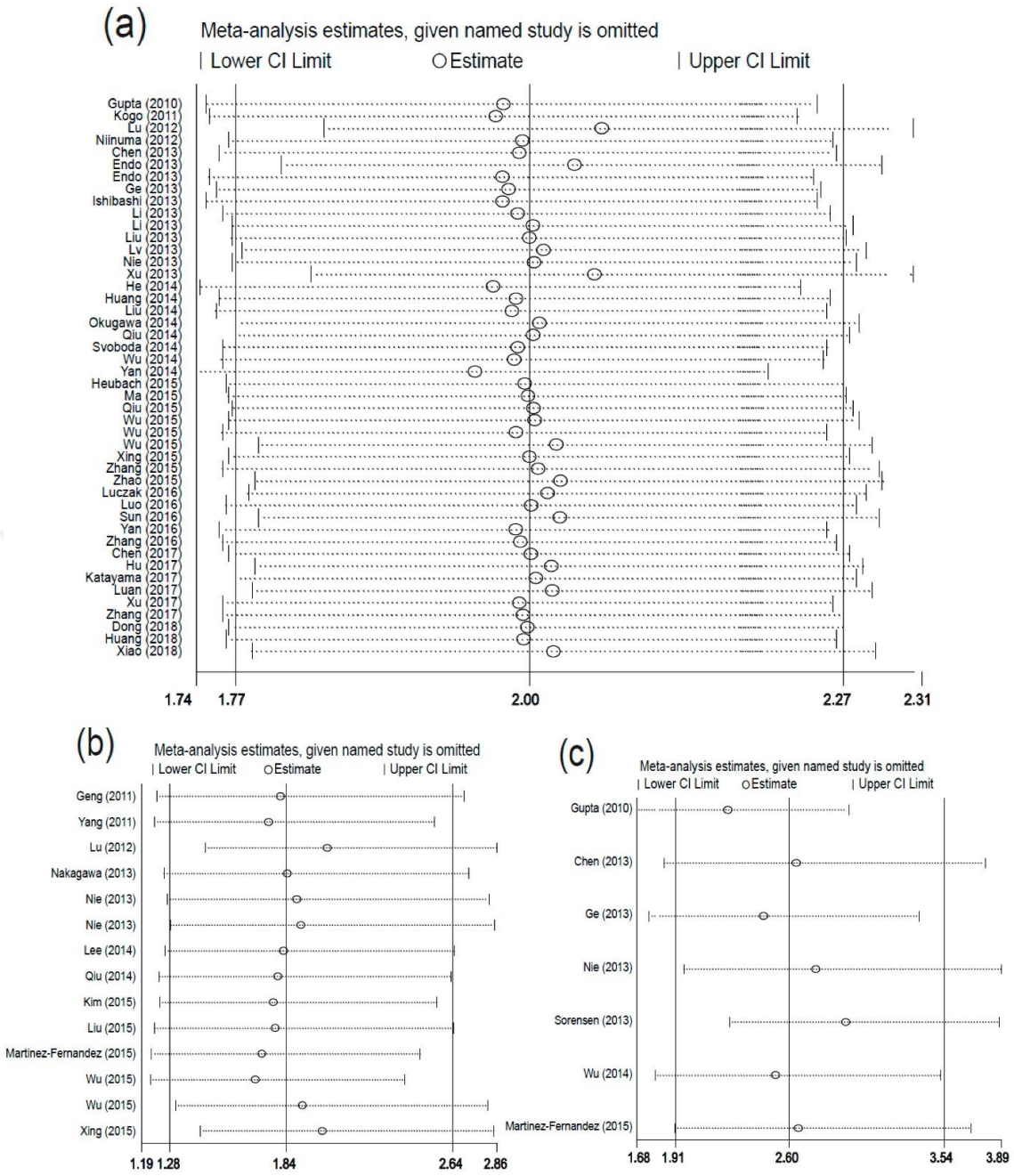


Figure 9. Sensitivity analysis of each eligible study. (a) OS individual studies, (b) RFS/DFS individual studies and (c) MFS/PFS individual studies.

4.2. Identification of Molecular Determinants of Cancer Cell Resistance to Ionizing Radiation

The overall procedure for data collection and analysis followed in the present study is described illustratively in **Figure 10**.

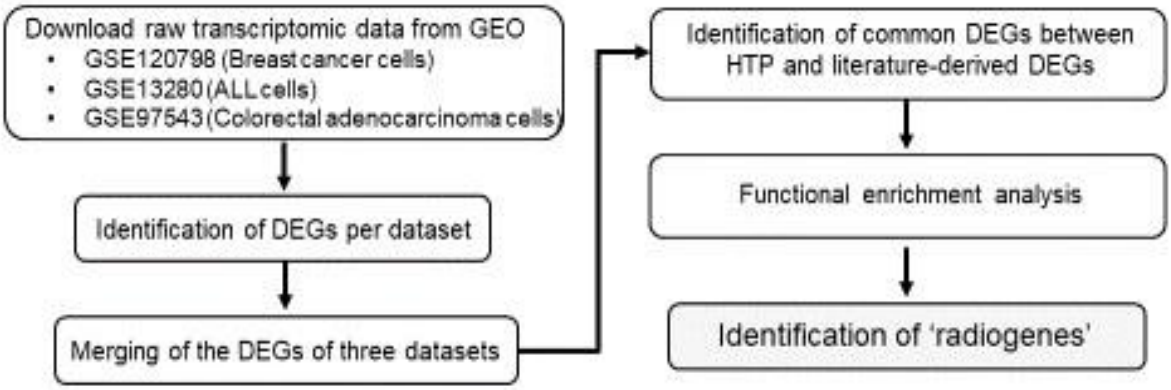


Figure 10. Flowchart diagram of the overall methodology employed in this study. Three eligible transcriptome datasets were retrieved from NCBI's GEO. Genes differentially expressed (DEGs) between the radioresistant and radiosensitive cancer cells were identified per dataset. The DEGs of the three analyzed datasets were integrated and compared with an earlier own study on cancer cell radioresistancerelated genes/proteins derived from literature research. Functional enrichment analysis of the common genes was performed to obtain the so-called "radiogenes".

4.2.1. Identification of Differential Expression Patterns in Radioresistance versus Radiosensitive Cancer Cells

The two techniques for transcriptome profiling, RNA-Seq and microarray, have large inherent differences. RNA-Seq is considered "superior" since it allows the detection of low abundance transcripts and novel transcript isoforms (Marioni et al., 2008). For this reason, we applied different statistical methods for RNA-Seq (D. Li, 2019) and microarray (Kontou et al., 2018) data processing and analysis.

The number of differentially expressed genes (DEGs) found between radioresistant and radiosensitive cancer cells (**Figure 11**) for each dataset is 6372 (GSE120798), 782 (GSE13280), and 541 (GSE97543), respectively. The pathways over-represented in the DEGs of the three datasets are related primarily to DDR/R and cell survival (**Table 3**).

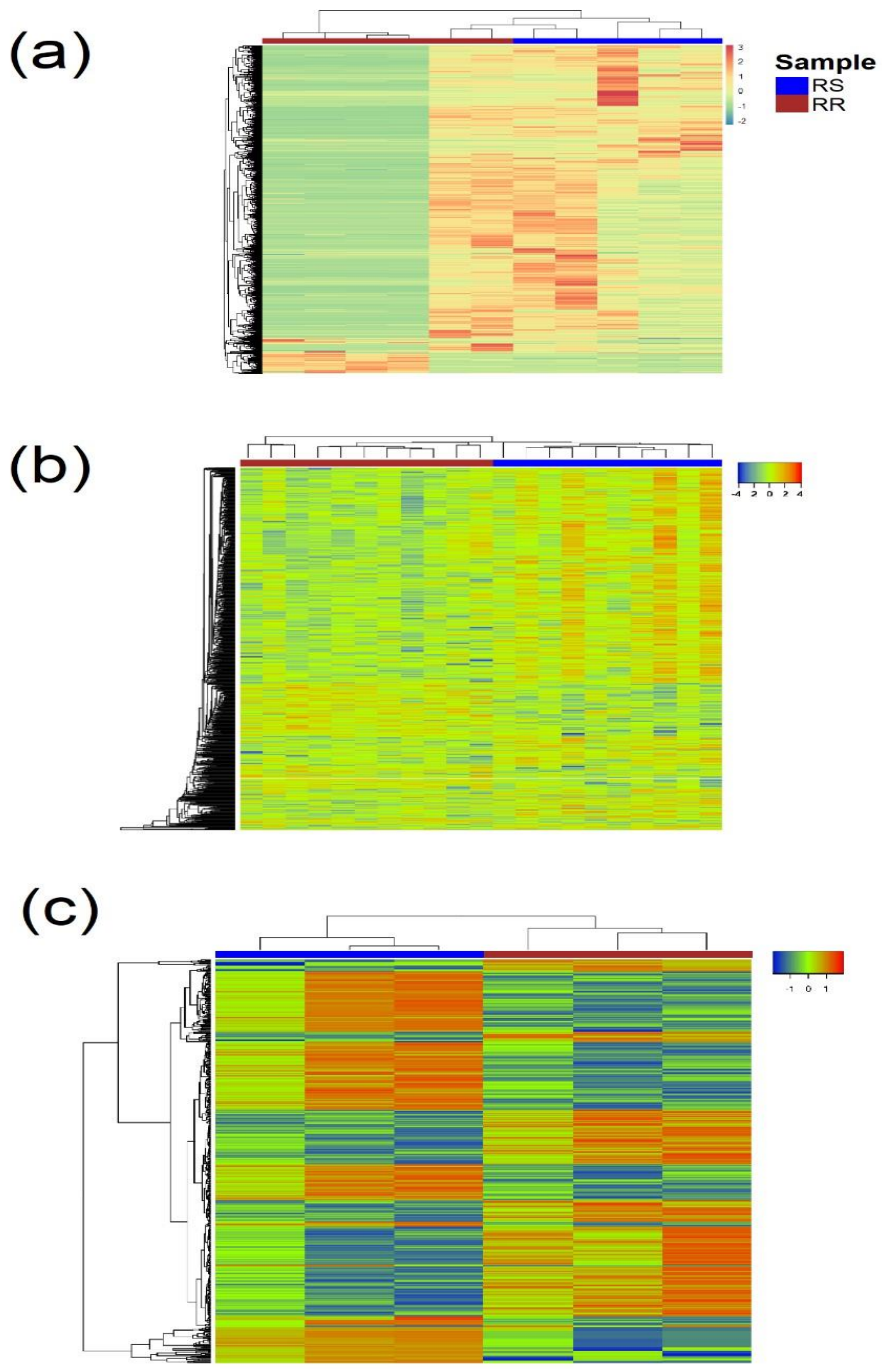


Figure 11. Heatmap representing color-coded expression levels of DEGs for the datasets (a) GSE120798 (breast cancer), (b) GSE13280 (ALL), (c) GSE97543 (COAD); columns correspond to samples and rows correspond to genes. RR, Radioresistant; RS, Radiosensitive.

Table 3. Distribution of the over-represented cancer-related WikiPathways in the DEGs of each transcriptome dataset. The enriched pathways are indicated by gray**.**

Cancer-relevant Pathways	**GSE120798**	**GSE97543**	**GSE13280**

ATM Signaling Network in Development and Disease			
ATM Signaling Pathway			
Cell Cycle			
Chromosomal and microsatellite instability in colorectal cancer			
DNA Damage Response			
DNA IR-damage and cellular response via ATR			
DNA IR-Double Strand Breaks (DSBs) and cellular response via ATM			
G1 to S cell cycle control			
Hereditary leiomyomatosis and renal cell carcinoma pathway			
Metabolic reprogramming in colon cancer			
Mismatch repair			
Notch Signaling Pathway			
PDGF Pathway			
PI3K-AKT-mTOR signaling pathway and therapeutic opportunities			
Retinoblastoma (RB) in Cancer			
Target Of Rapamycin (TOR) Signaling			
TGF-beta Signaling Pathway			

Differential gene expression analysis was performed for the three parental breast cancer cell lines in GSE120798 against their radioresistant derivatives, to identify those genes that are significantly dysregulated in response to radiation stress. A total

of 11 samples were compared; one biological replicate for each condition (**Table 1**). The number of detected DEGs of GSE120798 is remarkably higher as compared to the ones derived from microarray gene expression data. This discrepancy is likely due to the ability of RNA-Seq to detect and quantify, even rare, transcripts without a priori knowledge of a given gene (Metzker, 2010). Accordingly, the number of enriched pathways is also greater in this dataset (**Table 3**). Three DDR/R pathways were found to be enriched in the DEGs, including the generic "DNA Damage Response" pathway. Two pathways are specifically implicated in IR-induced DNA damage response, namely "DNA IR-Double Strand Breaks (DSBs) and cellular response via ATM," and "DNA IRdamage and cellular response via ATR." Ataxia telangiectasia mutated (ATM) and ataxia telangiectasia and Rad3-related (ATR) proteins are evolutionarily conserved proteins that have a critical regulatory role in DDR and maintenance of genome integrity (Awasthi et al., 2015; Maréchal & Zou, 2013a). Furthermore, the "PI3K-AKTmTOR signaling pathway and therapeutic opportunities" was shown to be overrepresented in the set of DEGs; increased activity of PI3K in the radioresistant cells of this transcriptome dataset was also observed by Gray et al. (Gray et al., 2019b). Phosphatidylinositol-3-kinase (PI3K)/AKT/mammalian target of rapamycin (PI3K/AKT/mTOR) signaling is critical to many aspects of tumor cell growth and survival (Porta et al., 2014a) and therefore could be likely involved in the survival of irradiated cancer cells.

Differential expression analysis was also carried out of the irradiated COAD cells in the GSE97543 dataset. A total of six samples were compared; three biological replicates for the wild-type (radiosensitive) cell lines and three for the radioresistant cells (**Table 1**). The notch signaling pathway is significantly over-represented (**Table 3**) in the DEGs of this dataset. Notch signaling is suggested to confer a selective survival advantage on tumors (Capaccione & Pine, 2013a). Hence, the Notch network could be implicated in the resistance and survival of the COAD cells to irradiation. Regarding the over-represented DDR/R pathways, one pathway is related to processing IR-induced DNA lesions through ATR signaling and the other to mismatch repair (MMR), which is responsible for detecting and repairing mismatched nucleotides (Larrea et al., 2010; Lyer et al., 2006). MMR reaction is initiated by binding of the MSH2 (MutS homolog 2)/MSH6 heterodimer to the mismatched DNA; both MSH2 and

MSH6 were found differentially expressed in the radioresistant COAD cells. The MutS homologs MSH2 and MSH6 form a heterodimer that binds to short insertion/deletion DNA mispairs (Edelbrock et al., 2013; Habraken et al., 1996). The MMR proteins also function in signaling DNA damage (Duckett et al., 1996; Modrich, 1997). Earlier studies have shown that MSH2 is also involved in the processing of the biologically significant clustered DNA damages, as well as the execution of apoptosis induced by IR (Holt et al., 2009).

In GSE13280, the gene expression profiles of radioresistant and radiosensitive ALL cell lines (eleven replicates for each condition) were compared (**Table 1**) for the identification of DEGs. The enriched transforming growth factor beta (TGFB) signaling pathway (**Table 3**), like the Notch network (Capaccione & Pine, 2013b), is implicated in several aspects of cancer initiation, promotion, and progression (Syed, 2016). Hence, the Notch- and TGFB-mediated signaling pathways might render ALL cells less vulnerable to IR-induced apoptosis by exerting their cellular pro-survival effect.

4.2.2. Identification of Cancer Cell Radioresistance Related Genes

The procedure we followed to identify an optimal number of cancer cell radioresistance–related genes is illustrated in **Figure 12**. A list of 175 biomolecules was proposed in a previous study by Pavlopoulou and colleagues (Pavlopoulou et al., 2017b) to be implicated in tumor cell radioresistance. These genes/proteins were collected manually through a comprehensive and thorough literature search. The DEGs identified in each of the three transcriptome datasets were merged; a list of 7,185 genes was compiled. Those genes were compared to the literature-derived molecules, and 88 genes were found in common (**Figure 12**). In order to identify an optimal number of genes implicated in radioresistance, we performed functional enrichment analysis of these genes. The pathways over-represented in the 88 genes (**Figure 13**) are related to DDR/R, similar to those detected in the DEGs of the individual datasets (**Table 3**), as well as to apoptosis. Signaling pathways mediated by the cardinal tumor suppressor TP53 are also related to radioresistance (**Figure 13**). Collectively, 36 genes were found to be implicated in cancer-associated biological pathways listed in **Table 4**, 26 of those up-regulated and 10 down-regulated in RR cancer cells. The products of the 36 genes also form a highly connected network (**Figure 14**), with

highconfidence interactions, suggesting that these proteins associate, either functionally or physically, to confer cellular resistance to IR. Therefore, we propose 36 interconnected pivotal genes, henceforth referred to as "radiogenes," which participate in radioresistance-relevant pathways and mechanisms.

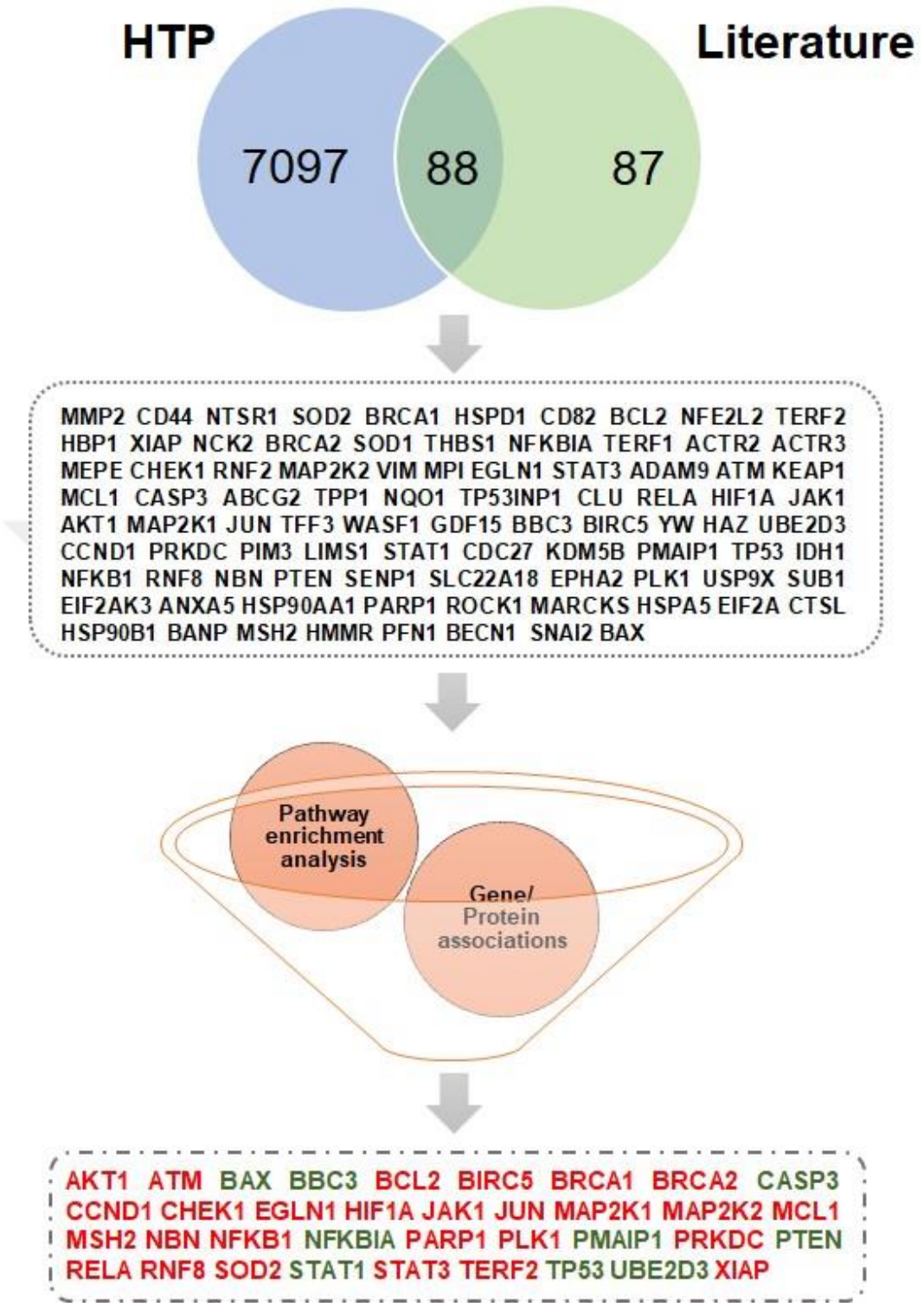


Figure 12. Schematic workflow outlining the selection of radiogenes. Venn diagram comparing the high-throughput (HTP) and literature-derived cancer radioresistanceassociated genes. Functional enrichment and interaction network analysis of the common 88 genes resulted in 36 genes. Red, up-regulated genes; Green, downregulated genes.

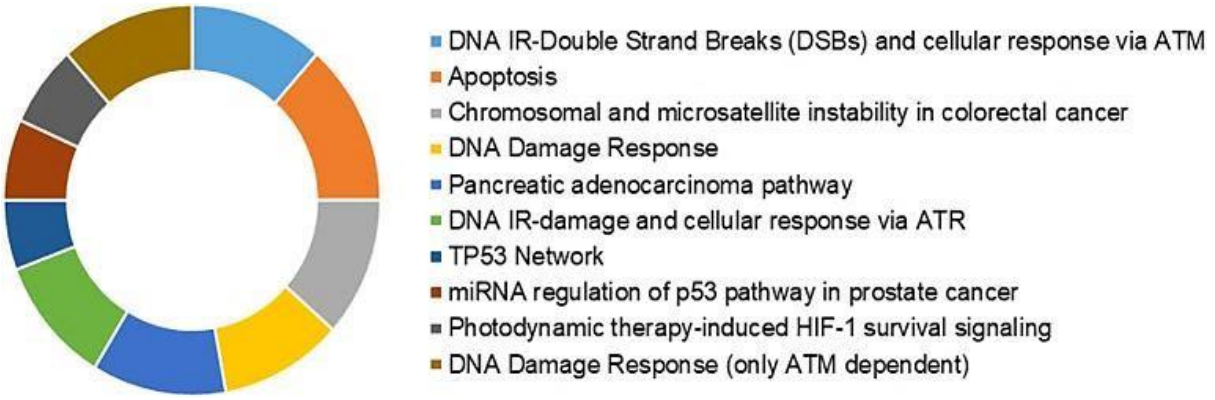


Figure 13. Donut chart depicting the over-represented cancer-relevant pathways across the 88 common genes.

Table 4. Gene symbol and expression status of the 36 differentially expressed radiogenes (radioresistant vs. radiosensitive cancer cells).

Gene	Expression status	Cell line
AKT1	Up	BC
ATM	Up	BC
BAX	Down	BC; ALL
BBC3	Down	BC
BCL2	Up	BC
BIRC5	Up	BC
BRCA1	Up	BC
BRCA2	Up	BC
CASP3	Down	BC
CCND1	Up	BC
CHEK1	UP	BC
EGLN1	Up	BC
HIF1A	Up	BC; ALL
JAK1	Down	BC
JUN	Up	BC
MAP2K1	UP	BC
MAP2K2	Up	BC
MCL1	Up	BC
MSH2	Up	BC; COAD

NBN	Up	BC
NFKB1	Up	BC
NFKBIA	Down	BC
PARP1	Up	BC
PLK1	Up	BC
PMAIP1	Down	BC
PRKDC	Up	BC
PTEN	Down	BC
RELA	Up	BC
RNF8	Up	BC
SOD2	Up	BC; ALL
STAT1	Down	BC
STAT3	Up	BC
TERF2	Up	BC
TP53	Down	BC
UBE2D3	Down	ALL
XIAP	Up	BC

ALL: acute lymphoblastic leukemia; BC: breast cancer; COAD: colorectal adenocarcinoma

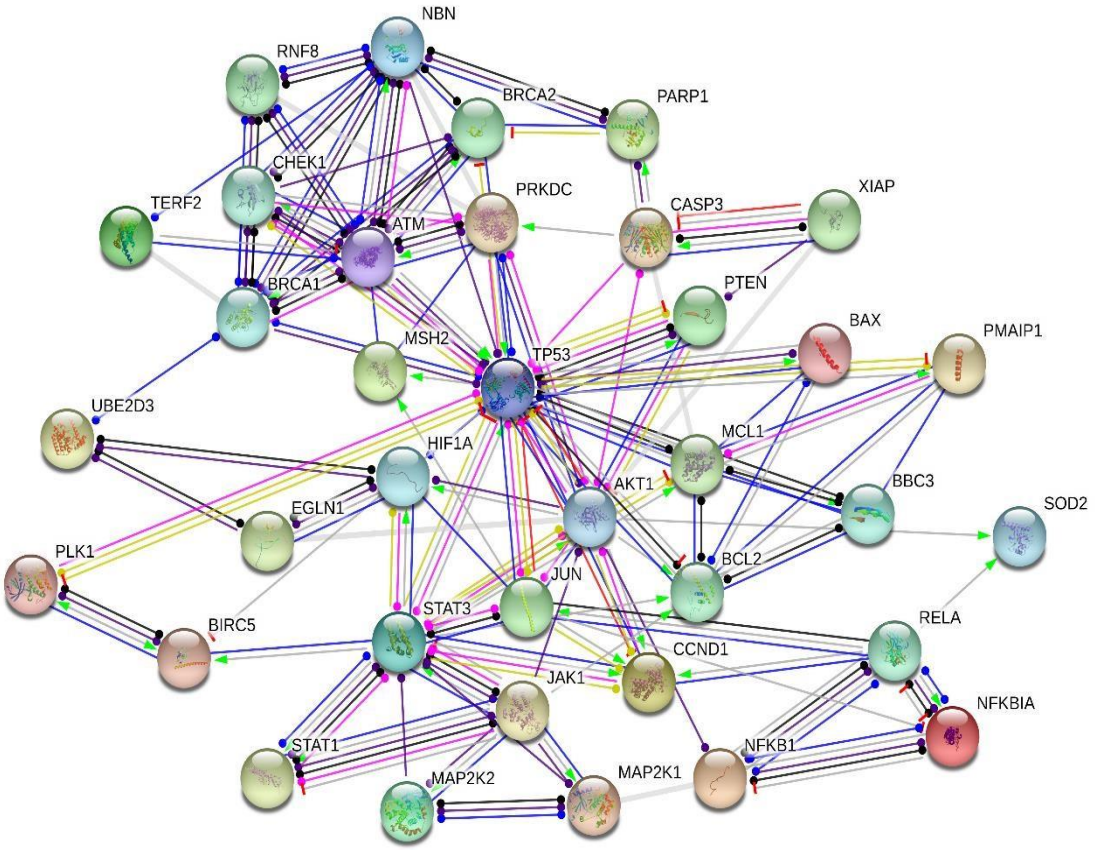


Figure 14. STRING interaction network of the products of the 36 radiogenes. The nodes represent proteins and the edges indicate different modes of interactions with a confidence score ≥ 0.9.

4.2.3. Differentially Expressed Radiogenes in Cancer vs. Normal Tissue

Collectively, 19 radiogenes, found to be differentially expressed in the three radioresistant vs. the radiosensitive cancer cell lines (**Table 4**), are consistently deregulated in their corresponding cancer-matched normal tissues with the same direction (i.e., up- or down-regulated) (**Figure 15**). This finding could be utilized in personalized tumor treatment for the selective eradication of cancer cells by applying radiotherapy without harming the adjacent healthy tissue at the same time.

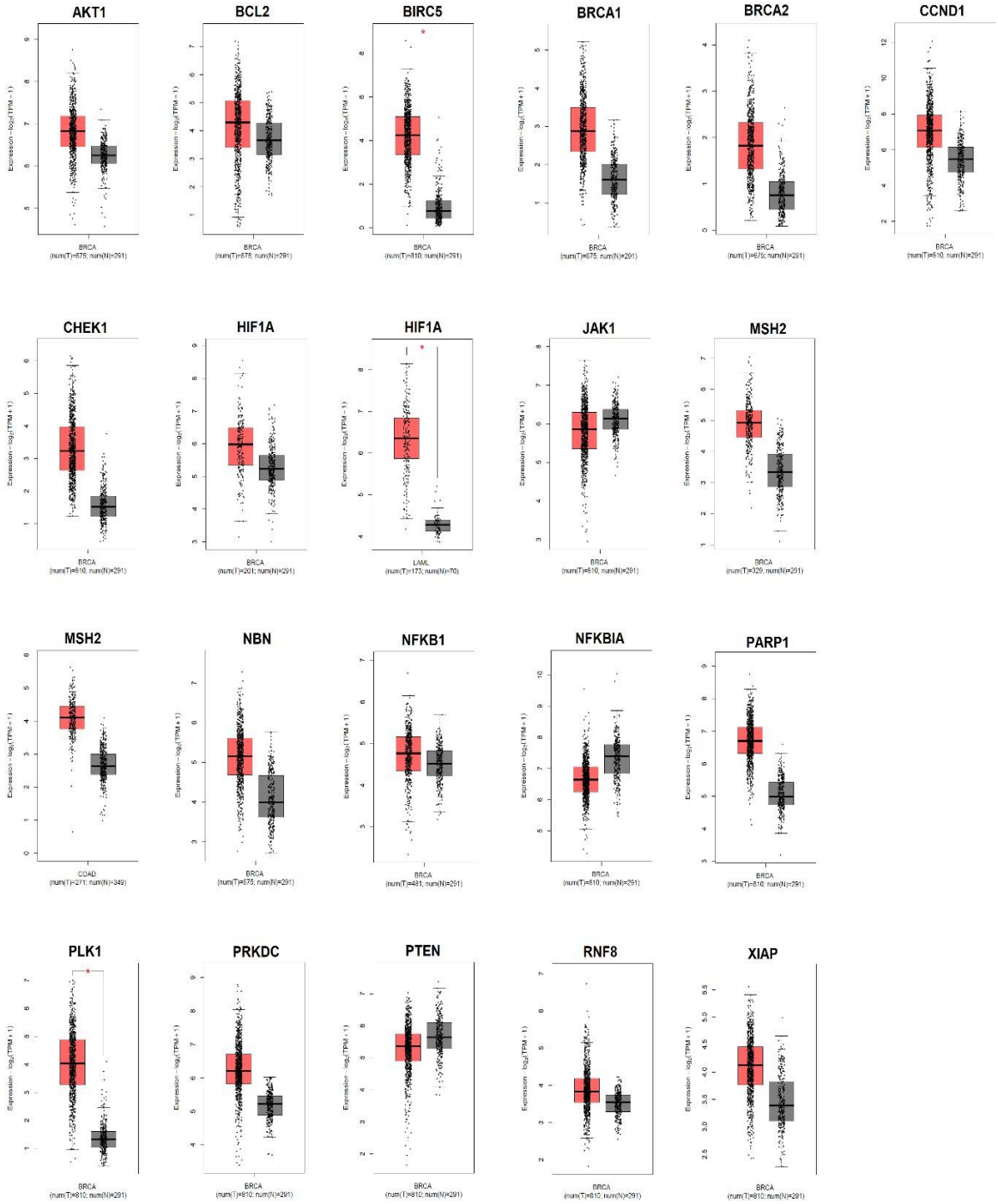


Figure 15. Differential expression patterns of radiogenes in the corresponding cancer tissue; BRCA, breast invasive carcinoma; COAD, colon adenocarcinoma; LAML, acute myeloid leukemia. The red box represents cancer tissue samples and the gray box indicates normal samples.

4.2.4. Radiogenes are Potential Cancer Prognostic Markers

The differential expression of the radiogenes can also predict the clinical outcome of cancer patients. In particular, a statistically significant association was found between CHEK1, MAP2K1, and PLK1 overexpression and worse overall survival in breast invasive carcinoma patients, indicated by pooled hazard ratio (HR) values higher than 1 and p-values lower than 0.05. Conversely, a significant relationship was observed between high expression of NFKBIA, which is otherwise underexpressed in radioresistant breast cancer cells, and favorable prognosis, indicated by an HR value less than 1 (**Figure 16**).

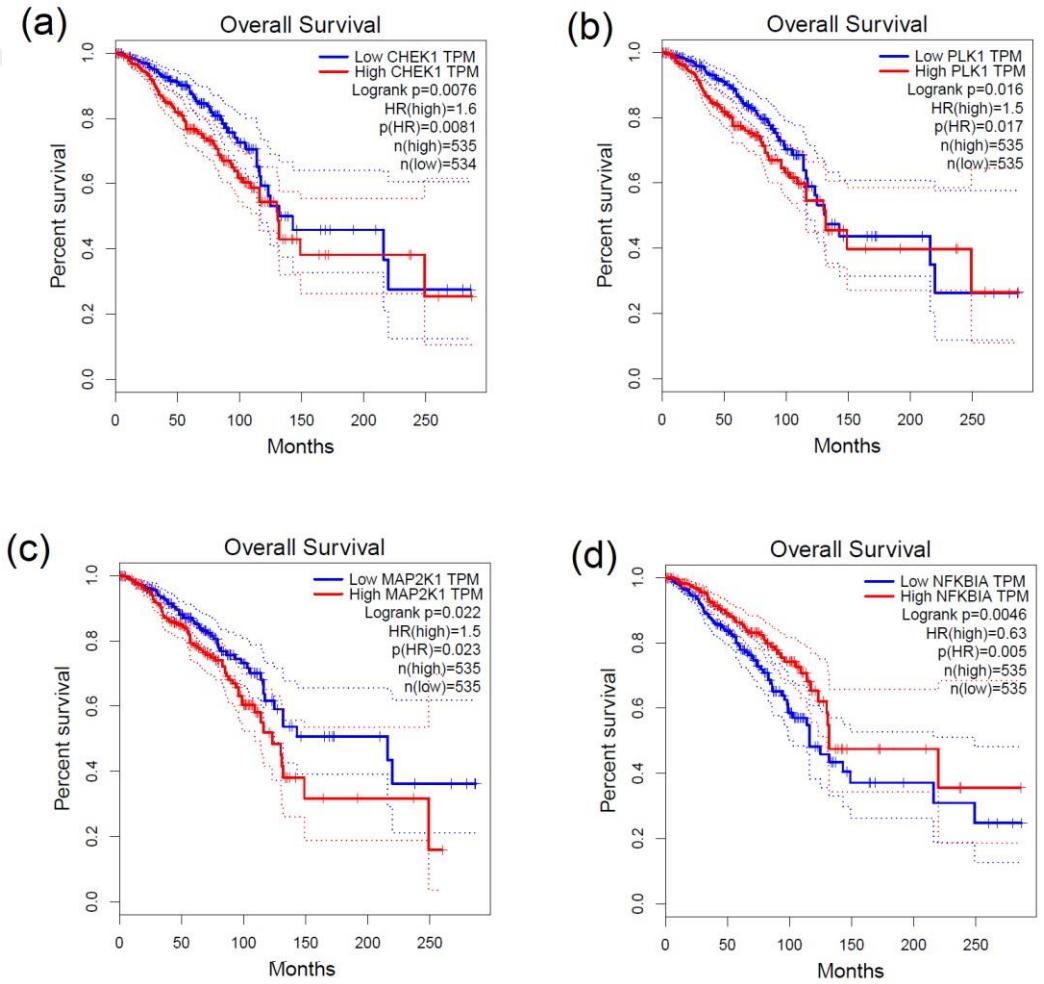


Figure 16. Kaplan-Meier curves depicting the prognostic value of radiogenes for overall survival in invasive breast carcinoma (a) CHEK1, (b) PLK1, (c) MAP2K1 and (d) NFKBIA. The HR "HR(high)" and the corresponding p-values "p(HR)" are shown. The 95% confidence intervals (CI) are denoted by dotted lines. The number of high-risk and low-risk patients are indicated by "n(high)" and "n(low)," respectively.

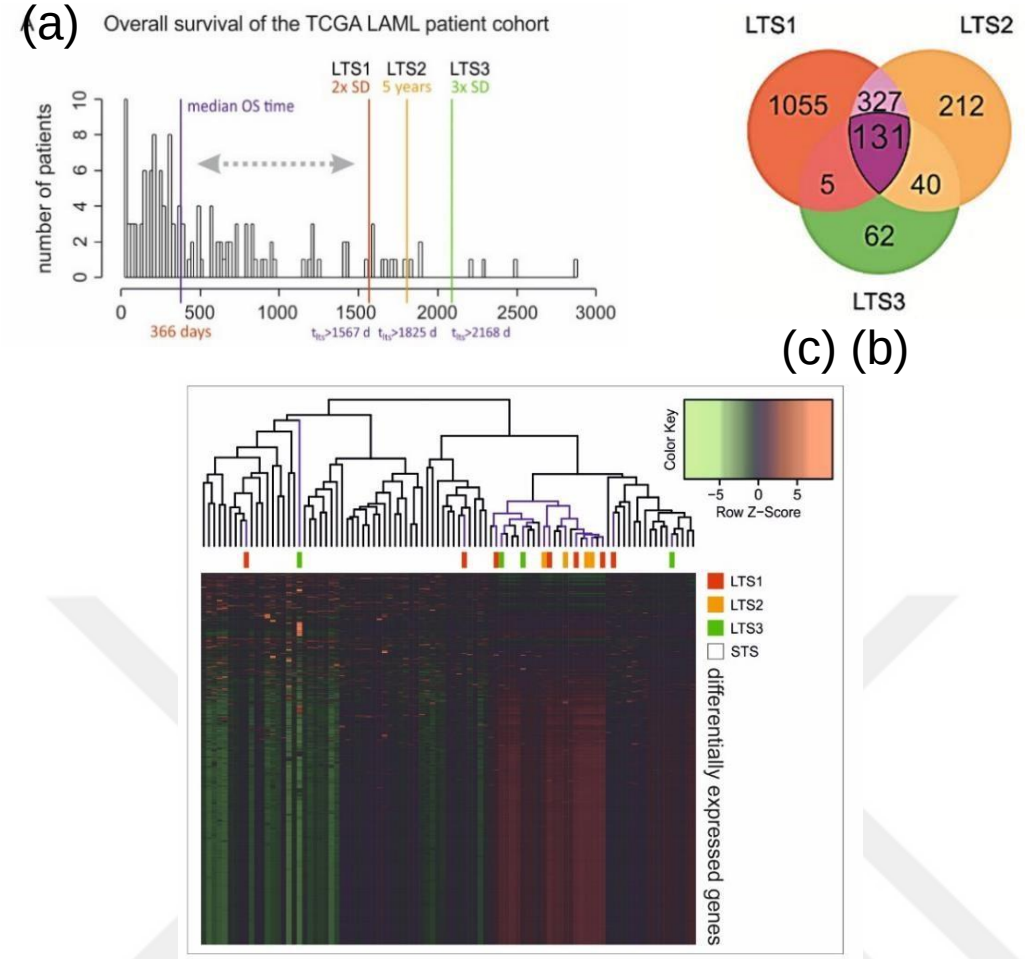
4.3. Identification of DEGs Correlated with Long-Term Survival in AML patients

4.3.1. Inferring Favorable Prognostic Markers from the Transcriptomic Profiles of Long-Term Survivors of AML

Even after a seemingly successful treatment, a significant percentage of AML patients is at risk of disease recurrence (Lauber et al., 2020), as a consequence of minimal residual disease. It has been estimated that of those patients who will achieve initial remission, approximately 50–70% of them will relapse within three years. Nevertheless, a lower percentage of adult AML patients survive to five years (Cheng et al., 2014) while some infrequent cases with survival over ten years have also been reported (Vasu et al., 2018). In general, across several cancer types, there are rare cases of patients who manage to remain progression-free for atypically long durations of time, while in some rare cases patients never relapse (Bloomfield et al., 1997; Saner et al., 2019b). Intriguingly, tumors of these unusual patients present distinct genetic and gene expression profiles as compared to the patients with a more typical course of disease (Reifenberger et al., 2014b; Shen et al., 2021b). Patients with exceptionally prolonged survival rates comprise a rare, but distinct, clinical subpopulation which holds the potential to reveal signatures of favorable prognostic factors that could be informative for the effective management of the general population of cancer patients (Saner et al., 2019b). According to these notions, the patients with unusually prolonged survival (hereafter termed long-term survivors, LTSs) in our study cohort might plausibly better cope with MRD and relapse and, therefore, their distinct transcriptomes, as compared to the general population, can be used to provide prognostic signatures associated with resistance to MRD.

To this end, in the cohort of AML patients of TCGA, we distributed the patients according to their OS and estimated the median OS which would reflect the typical course for the disease for the specific cohort. The median OS for the whole AML population was 366 days (n = 50 patients), while the standard deviation value was estimated as 600.5 days. Then, we defined populations of patients with exceptionally favorable prognosis within the study cohort in a time-dependent manner as the patients who survive i) longer than the median OS + 2 SD (>1567 days, n = 15

patients), or ii) longer than the median OS + 3 SD after initial diagnosis (2168 days, n = 4 patients). Given that in clinical practice the patients who are alive 5+ years past diagnosis are considered LTSs (Arndt et al., 2017), we additionally included the patients who survive past this clinical milestone (1825 days, n = 7 patients) in our analysis (**Figure 17a**). We sought to identify the distinct tumor gene expression profiles across these LTSs versus those of the patients who are deceased by 366 days (**Figure 17b**). Our analysis revealed common deregulated genes between all three comparisons (**Figure 17c**). Subsequent gene set enrichment analysis (GSEA) showed that biological functions such as locomotion, cell migration, and cell junctions are enriched in the DEGs in each LTS subgroup. This is expected, since these traits are linked with a propensity of cancer cells to metastasize. Surprisingly, among the top significantly enriched processes, we also observed deregulations of genes that are involved in developmental processes, in particular neurogenesis, central nervous system development, neuron differentiation, and pattern specification (**Figure 17d**). We further identified a group of 131 DEGs that remain consistently deregulated in all three groups and, hence, could be important for maintaining survival over extended periods of time.


(a)
Overall survival of the TCGA LAML patient cohort
LTS1
2x SD
LTS2
5 years
LTS3
3x SD
median OS time
number of patients
0
2
4
6
8
10
0
500
1000
1500
2000
2500
3000
366 days
t_{80}>1567 d
t_{80}>1825 d
t_{80}>2168 d
LTS1
LTS2
1055
327
212
131
5
40
62
LTS3
(c)
(b)
Color Key
−5
0
5
Row Z−Score
LTS1
LTS2
LTS3
STS
differentially expressed genes

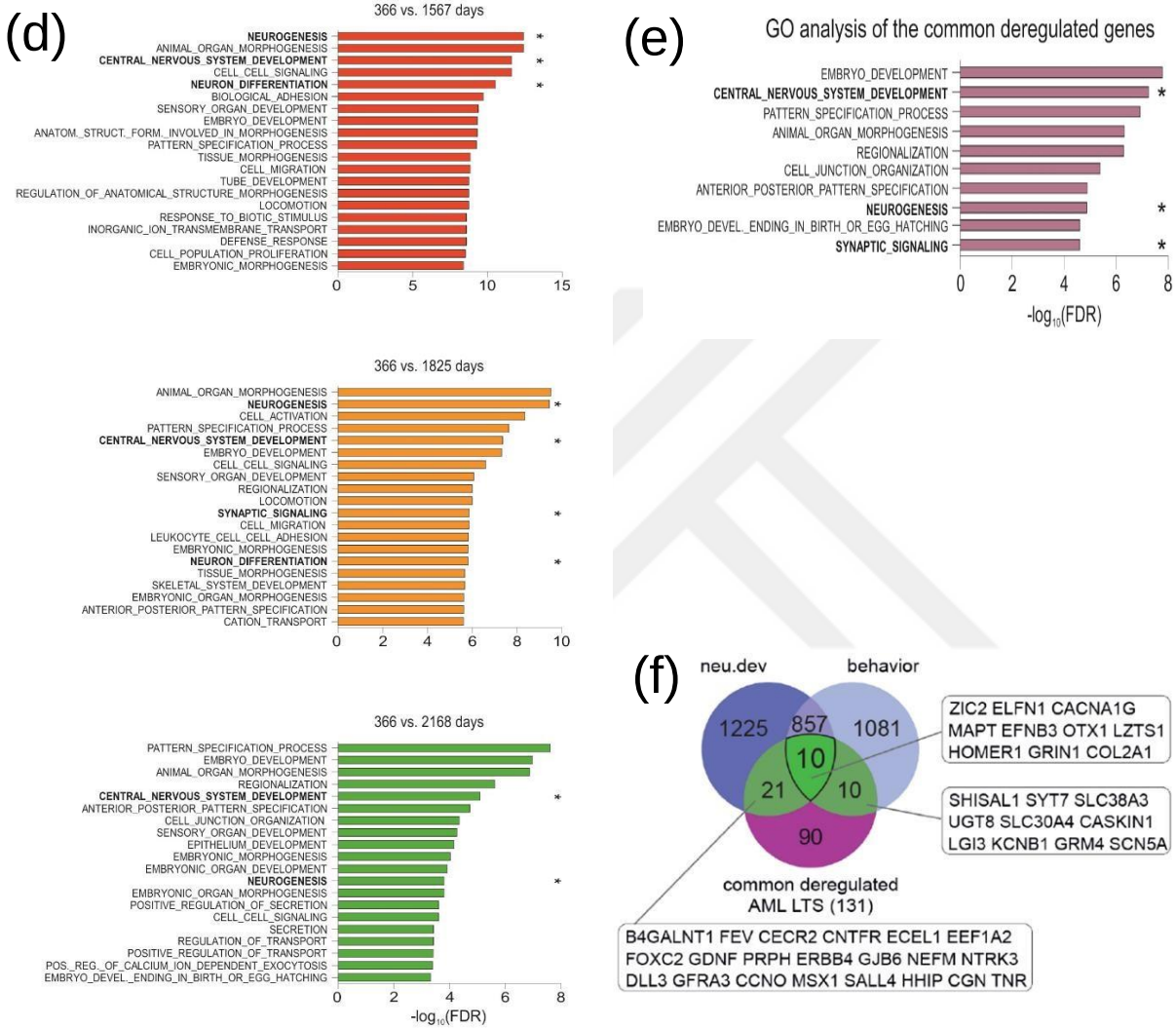


Figure 17. Identification of favorable prognostic factors based on transcriptomic profiles of LTSs. (a) Histogram of the overall survival (OS) of the TCGA LAML cohort with median (366 days, n = 50) and cutoffs for the three long-term survivor (LTS) groups: 2 × SD = 1567 days (n = 15), 5-year survival = 1825 days (n = 7), and 3 × SD = 2168 days (n = 4). (b) Heatmaps of the DEGs in the three LTS subpopulations versus the STSs. (c) Venn diagram of the deregulated genes of all three analyses showing large overlaps between analyses and 131 genes in common. (d) GSEA of GO biological functions associated with the DEGs in each LTS subgroup versus the patients with OS < 366 days. (e) Top 10 enriched GO biological functions for the 131 genes

that are commonly deregulated across all LTS subgroups, ranked by order of significance (−log10 of FDR value). (f) Overlapping of the 131 genes with previously identified nervous system-related genes showed that 31.3% of these genes are indispensable for neuronal development, neurological function, or both (gene abbreviations are indicated in the respective boxes). Nervous system-related pathways are indicated by an asterisk (*).

Strikingly, GSEA showed that the commonly deregulated genes are strongly involved in early and late stages of embryonic development, and particularly in neuronal development and neurological function. In detail, the top 10 most significantly enriched biological processes were embryo development, central nervous system (CNS) development, pattern specification process, animal organ morphogenesis, regionalization, cell junction organization, neurogenesis, anterior– posterior pattern specification, synaptic signaling, and embryo development ending in birth or egg hatching (**Figure 17e**). Among the transcripts found upregulated in LTSs and involved in many of these processes were glial cell-derived neurotrophic factor (GDNF), a neurotrophic molecule that promotes neuronal differentiation and dopaminergic development (Roussa & Krieglstein, 2004), as well as several products of homeobox genes that are master regulators of development, morphogenesis, and cell differentiation, for example, distal-less homeobox 1 (DLX1), msh homeobox 1 (MSX1), and orthodenticle homeobox 1 (OTX1) (Holland et al., 2007). We further considered the cytogenetic background of the patients of the LTS and STS subgroups in regard to the reciprocal translocation t(15;17), which is associated with decreased incidence of relapse and improved survival upon all-trans retinoic acid (ATRA)containing treatments. In particular, six patients across the LTS subgroups versus six patients in the STS subgroup bear a t(15;17) translocation, leading to fusion of the promyelocytic leukemia gene (PML) to retinoic acid receptor alfa (RARA), which is specifically sensitive to ATRA and arsenical derivatives (Adès et al., 2010). To eliminate survival bias in the LTSs caused by this cytogenetic feature, we excluded these patients and performed the comparisons between the transcriptomes of the STSs and LTSs. Strikingly, we found that the DEGs in LTSs remained significantly associated with CNS development, neuronal development, neuronal differentiation, and neurogenesis.

In view of these findings, we sought to further explore the potential link between long-term survival and the nervous system. We have recently performed a

comprehensive, phenotype-driven identification of nervous system-related genes. These include all protein-coding genes that are essential for normal nervous system development (neu.dev) and neurological function (behavior), whereby their knockout leads to malformations of the neural system components or behavioral disorders, respectively. Using system-based approaches, we had found 1246 genes that are specific for neuronal development (neu.dev) and 1086 for neurological function (behavior), while 863 genes are involved in both processes (Logotheti, Marquardt, Richter, Hain, et al., 2020). Overlapping of the 'neu.dev' and 'behavior' lists of genes with the 131 commonly deregulated genes in LTSs showed that, overall, 41 (31.3%) of them are nervous system-related genes, whereby 21 are indispensable for neuronal development, 10 for neurological function, and 10 for both processes (**Figure 17F**). Overall, our analyses indicate a strong association of the nervous system with the exceptionally prolonged survival and patient outcomes in AML.

5. DISCUSSION

5.1. Meta Analysis of the Prognostic Value of HOTAIR in Cancer

HOTAIR exhibits pro-oncogenic activity since it has been shown to be overexpressed in numerous cancers and be implicated in several hallmarks of cancer, such as cellular proliferation, inhibition of apoptosis, genomic instability, angiogenesis, invasion, and metastasis (Hajjari & Salavaty, 2015b; Tang & Hann, 2018b).

In the current study, an updated, comprehensive meta-analysis on the prognostic value of HOTAIR in various human cancers was presented. By applying stringent inclusion and exclusion criteria, we included 53 eligible studies, a relatively large number necessary for a meta-analysis to be considered robust. Previous meta-analyses on the association of HOTAIR with clinical outcome have included a rather limited number of studies with inconclusive and inconsistent findings (Deng et al., 2014b; Zhang et al., 2014b). Other related studies have focused on certain types of cancers, such as head and neck squamous cell carcinoma (Troiano, Alberto, et al., 2017) or digestive system cancers (Abdeahad et al., 2019; Ma et al., 2015; Zhang, Wang, et al., 2018).

In the present study, we showed that there is a statistically significant relationship between elevated HOTAIR expression and poor OS. In the subgroup analysis, based on cancer type, HOTAIR was shown to be a significant predictor for worse prognosis for a variety of cancers, including solid cancers, such as urological cancers, head and neck neoplasms, cancers of the digestive system, and several female cancers (e.g., cervical, ovarian, and endometrial cancers), as well as the blood cancers, lymphoma and leukemia. Moreover, we complemented the findings from meta-analysis and further strengthened our hypotheses with survival information from other types of cancers, for which there were not any available eligible studies, retrieved from TCGA. It was found that there is, also, a strong relationship between HOTAIR overexpression and poor OS in neoplasms of the adrenal cortex, mesothelial neoplasms, and neuroepithelial tumors.

Taken together, the above findings lead to the suggestion that similar HOTAIRmediated pathways might be implicated both in solid and liquid cancers (Bhan

& Mandal, 2015b). In particular, in several solid tumors, HOTAIR has been shown to exert its oncogenic and metastatic potential by mediating a repressive chromatin structure through the recruitment of histone-modifying or chromatin-remodeling complexes, such as PRC2 (Gupta et al., 2010c; Kogo et al., 2011; Wu et al., 2015b). For example, HOTAIR can promote pancreatic cancer cell proliferation by suppressing the expression of miR-663b via remodeling the chromatin structure within the miR663b promoter (Cai et al., 2016). In a recent study, HOTAIR was also found to recruit PRC2 to catalyze H3K27 trimethylation to transcriptionally repress E-cadherin and promote EMT in gastric cancer (Song et al., 2019). Similarly, high expression levels of HOTAIR and PRC2 proteins (H3K27 methylase EZH2, SUZ12, and EED) were found to be positively correlated with lymphomagenesis (Oh et al., 2016). In addition, HOTAIR, through miRNA sponging, contributes to carcinogenesis both in blood (Xing et al., 2015) and solid tumors (Jiang et al., 2018; Z. Zhang et al., 2016). However, there is a rather limited number of studies available on major cancers, such as breast neoplasms and respiratory tract cancers. Thus, more clinical trials on these cancers would enable us to better assess the relationship between HOTAIR expression and cancer patients' survival.

To investigate any possible effect of the genetic background and environment on the overall HRs, analyses were conducted based on the ethnic background of the participants. HOTAIR was found to be a powerful negative prediction biomarker for Asians. In the case of Caucasians, there was a link between HOTAIR overexpression and poor OS, albeit with moderate statistical significance; this is probably due to the relatively low number of available studies on patients of Caucasian origin. There were not, also, any available studies for other major ethnic groups, such as Africans or Indians, which would have further allowed us to estimate the influence of the genetic make-up on the association between HOTAIR and clinical outcome. The overall effect was similar in the stratified analysis according to data source, that is, the estimated HR reported in the articles or extrapolated from survival curves.

Therefore, high HOTAIR expression can predict an unfavorable clinical outcome in different types of cancers and possibly ethnic groups using different extraction

methods. Notably, elevated expression of HOTAIR and prognosis in cancer patients is not particularly affected either by cancer type or even the patients' genetic background.

HOTAIR was found to be a poor predictor for both cancer recurrence and progression. The similar outcomes suggest that there are similar HOTAIR-dependent mechanisms underlying these two phenomena. In particular, HOTAIR was shown to mediate recurrence and progression in bladder cancer via the histone methyltransferase EZH2 (Martínez-Fernández et al., 2015). Similarly, enhanced HOTAIR expression was found to be associated both with progression and tumor recurrence in hepatocellular carcinoma by regulating the Wnt/β-catenin signal transduction pathway (Gao et al., 2016).

HOTAIR has been demonstrated to promote tumor cell invasion and metastasis by modulating epithelial-to-mesenchymal transition (EMT) (Lee et al., 2014; Wu et al., 2015c; Zheng et al., 2017a). Enhanced HOTAIR expression has also been shown to promote metastasis and invasion through different mechanisms including genomewide re-targeting of PRC2 and subsequent epigenetic silencing of multiple antimetastatic genes (Gupta et al., 2010d), inhibition of the expression of the metastasis suppressor gene E-cadherin by recruiting the histone methyltransferase of PRC2, EZH2 (Wu et al., 2015d; Zheng et al., 2017b), targeting of Notch/Wnt signaling pathwayassociated genes (M. Lee et al., 2016), and upregulating chondroitin sulfotransferase CHST15 (Liu et al., 2019), etc. HOTAIR also promotes invasion and migration by acting as a 'miRNA sponge', through targeting the corresponding miRNAs in the miR-1/CCND2 (Di et al., 2017), and miR-148a/SNAIL2 (Xu & Zhang, 2017).

Heterogeneity was observed within the forest plots of OS and RFS/DFS, suggesting that HRs vary across studies. For this reason, the random-effects model was applied, where the overall HR was estimated based on the weighted average of the HRs of the individual studies. Given that the overall effect for OS and RFS/DFS was not affected by any single study, according to sensitivity analyses, we could suggest that, despite heterogeneity, the pooled HR can be considered quite reliable and representative.

Moreover, potential publication bias was not detected in the meta-analysis, probably due to the sufficient representation of eligible studies in this meta-analysis.

5.2. Identification of Molecular Determinants of Cancer Cell Resistance to Ionizing Radiation

Cancer cells confer resistance to irradiation through diverse mechanisms including enhanced DNA damage repair capacity, activation of cell survival signaling pathways, inhibition of apoptotic pathways, and induced autophagy.

DDR/R is a complex entangled process consisting of recognition (or detection, or sensing), signaling, and repair of DNA damage (Rouse & Jackson, 2002). Ionizing radiation usually generates a variety of DNA lesions, including abasic (apurinic/apyrimidinic) sites, oxidized bases, crosslinks between thymine and cytosine bases, DNA single-strand breaks (SSBs), and DNA double-strand breaks (DSBs) (Sutherland et al., 2000). In the case of SSBs, only one strand of the double stranded DNA helix is severed. SSBs are recognized and processed by base excision repair (BER) and nucleotide excision repair (NER) mechanisms (Caldecott, 2008). DSB is the most detrimental type of DNA lesions since both strands of the double helix are broken. DSBs are repaired mainly through homologous recombination (HR) if cells are present in the S/G2 cell-cycle phase or, the less accurate, non-homologous end joining (NHEJ) (Budman & Chu, 2005). Those different types of DNA lesions can be formed either separately or in close vicinity (a few nm), resulting in clustered DNA lesions or locally multiply damaged sites (MDS). Clustered DNA damage, the hallmark of IR, is considered the most severe type of genomic damage because of its complexity. This complex DNA damage includes both DSBs and non-DSBs, usually referred to as oxidatively-clustered DNA lesions (OCDLs), occurring within a DNA region of 15–20 bp. The corresponding DNA damage repair mechanisms are recruited by the cell in response to clustered damaged sites. However, harnessing the corresponding DNA repair machinery to the clustered damaged sites is quite challenging since the presence of a lesion in one strand can delay significantly the simultaneous processing of another lesion on the complementary strand. Furthermore, OCDLs can be rapidly converted into de novo DSBs by a DNA glycosylase during the repair. As a result, unrepaired MDS can lead to increasing levels of genotoxic damage, triggering also systemic responses

(**Figure 18**) (Nikitaki et al., 2016). Clustered DNA lesions, if not properly processed, could contribute to increased genomic instability in the form of chromosomal aberrations (e.g., deletions and inversions) and microsatellite instability, leading eventually to carcinogenesis. Therefore, the induction of clustered DNA damage increases the cytotoxic effect of radiation, especially in highly proliferating cancer cells. Radioresistant cancer cells though can counteract this effect through their ability to respond to and repair complex/clustered DNA damage more efficiently without avoiding necessarily increased genomic instability, as compared to their radiosensitive counterparts (Bukowska & Karwowski, 2018; Georgakilas et al., 2013; Hada & Georgakilas, 2008; Mavragani et al., 2017). The radiogenes identified in the present study are implicated in diverse mechanisms underlying the acquisition of radioresistance in cancer cells.

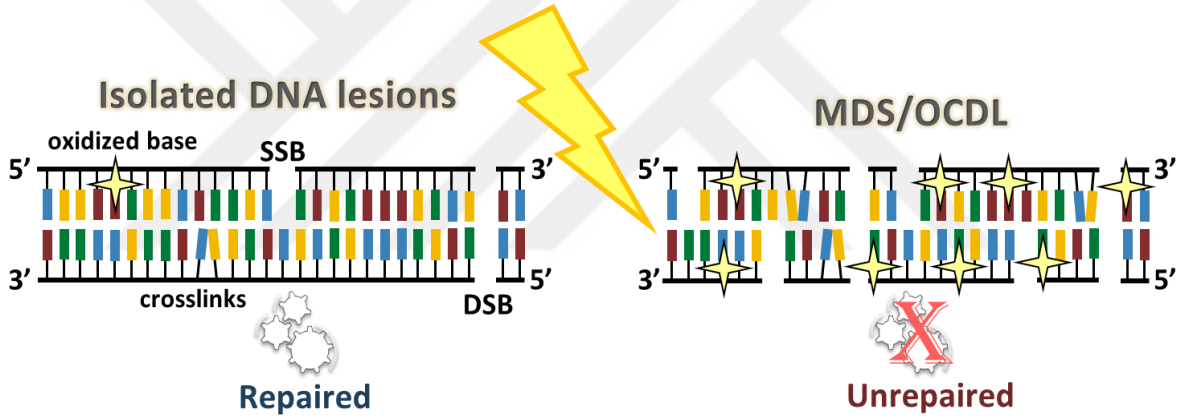


Figure 18. Radiation-induced DNA damage and its repair mechanisms.

Among the 36 "radiogenes", ATM, BRCA1/2, CHEK1, CCND1, MSH2, NBN, PARP1, PLK1, PRKDC, and RNF8, which are consistently up-regulated across the radioresistant cancer cell lines (**Table 4**) and the corresponding cancer tissue (**Figure 15**), are crucially implicated in different stages of DDR/R. In particular, ATM plays a protagonistic role in the initial stage of DDR/R, that is, DNA damage detection and stress-response signaling (Iliakis et al., 2003; Maréchal & Zou, 2013b). ATM signaling is activated by a wide variety of DNA lesions, as well as DNA replication stress (Escamilla et al., 1988; Maréchal & Zou, 2013c). Mutations in ATM result in the genetic disorder ataxia-telangiectasia (AT), which is characterized by the high sensitivity of AT

patients to IR and cancer predisposition (Gatti et al., 1988). Checkpoint kinase 1 (CHEK1), which acts downstream of ATM, is a core regulator of cell cycle checkpoint signaling in DNA damage response (Flaggs et al., 1997; Patil et al., 2013). Cyclin D1 (CCND1) was demonstrated to induce post-DNA damage cell cycle arrest and apoptosis in different types of cancers (Cai et al., 2012; Smith et al., 2016). PRKDC is a serine/threonine DNA-activated protein kinase involved in DSB recognition and DNA damage repair through NHEJ (Soubeyrand et al., 2003). Notably, ATM and PRKDC were found to affect greatly cancer cell response to IR through genome-wide genetic screening in a recent study by (Francica et al., 2020). Moreover, BRCA1 and BRCA2 are largely involved in cell cycle control and maintenance of genomic stability in response to DNA damage (Deng, 2006; Gudmundsdottir & Ashworth, 2006). Another radiogene, NBN (nibrin), encodes one of the three components of the MRN complex (MRE11A-RAD50-NBN), which is implicated in the recognition and repair of DSBs (Lamarche et al., 2010). NBN is mutated in patients with Nijmegen breakage syndrome (NBS), and cells from NBS patients are hypersensitive to IR (Taalman et al., 1983). In addition, a functional link between ATM and NBN proteins has been demonstrated by (Zhao et al., 2000). Also, the MMR protein MSH2 (MutS homolog 2) is suggested to contribute to radioresistance via SSB processing (Martin et al., 2010). Moreover, RNF8 (ring finger protein 8) protein catalyzes the mono-ubiquitination of histones H2A and H2B during DNA damage, thereby facilitating DNA damage repair and activation of cell cycle checkpoint (Kolas et al., 2007; Ma et al., 2011); RNF8 is associated with radioresistance in human nasopharyngeal cancer cells (Wang et al., 2015). The protein encoded by the radiogene PLK1 (Polo-like kinase 1) is involved in cell cycle resumption following DNA damage-induced checkpoint arrest (Bruinsma et al., 2017; Hyun et al., 2014). It has been demonstrated that inhibition of PLK1 renders glioblastoma and nonsmall cell lung cancer cells sensitive to IR (Pezuk et al., 2013; van den Bossche et al., 2019). Of particular note, pharmacological inhibitors of the BER enzyme PARP1 [poly(ADP-ribose) polymerase 1], such as niraparib, olaparib, and rucaparib, are widely used for targeted cancer therapy (Patel et al., 2020a; Sulai & Tan, 2018). Importantly, CHEK1 and PLK1 were found to be poor prognostic biomarkers for the survival of breast cancer patients (**Figure 16**), further supporting that enhanced DNA damage repair mechanisms in cancer cells play a catastaltic role in efficient radiotherapy (Patel

et al., 2020b; Pavlopoulou et al., 2017c). Therefore, DDR/R might represent a primary "danger" signal, leading eventually to the activation of downstream signaling cascades and pro-survival mechanisms (Nikitaki et al., 2015).

The apoptotic pathway is also affected by the cellular response to radiationinduced genomic damage in cancers, as we suggest in this study. Radioresistant cancer cells have developed the ability to evade apoptosis prompted by their response to extreme and repair-resistant DNA damage, mainly due to deregulation of key proapoptotic molecules like TP53 (Fridman & Lowe, 2003; Haupt et al., 2003), PTEN (Lu et al., 2016), PMAIP1 (mediator of damage-induced p53-mediated apoptosis) (Oda et al., 2000), BBC3 (TP53-upregulated modulator of apoptosis) (Han et al., 2001; Nakano & Vousden, 2001), BAX (BCL2-associated X protein) (Pawlowski & Kraft, 2000), as well as anti-apoptotic proteins like MCL1 (Fujise et al., 2000; Glaser et al., 2012), BCL2 (Akl et al., 2014), BIRC5 (Chiou et al., 2003), and XIAP (X-linked inhibitor of apoptosis) (Colin et al., 1996; Deveraux & Reed, 1999). In a consistent manner, in this study, the pro-apoptotic genes were shown to be up-regulated, whereas the anti-apoptotic genes are down-regulated in radioresistant cancer cells (**Table 4**). Of those, BCL2, which is down-regulated in radioresistant breast cancer cells and tissue (**Table 4, Figure 15**), can suppress apoptosis by inhibiting the activity of caspases indispensable for apoptosis, such as caspase-3 (CASP3) (Porter & Jänicke, 1999; Swanton et al., 1999). Of note, aberrant activation of the Notch signaling pathway was demonstrated to inhibit TP53-mediated damage response and promote breast carcinogenesis by preventing apoptosis (Stylianou et al., 2006), suggesting a link between pro-survival and apoptotic mechanisms. PTEN was also found to promote apoptosis and cell cycle arrest via PI3K/AKT-dependent and -independent signaling (Weng et al., 2001).

The radiogene NFKB1, which transactivates several pro-inflammatory genes (Liu et al., 2017), was found to be overexpressed in radioresistant breast cancer cells and tissue (**Table 4, Figure 15**). Furthermore, the members of the NFKB1 family, NFKBIA and RELA, can act either as inducers or inhibitors of apoptosis, respectively (Barkett & Gilmore, 1999; Sonenshein, 1997), consistent with their expression status in RR cells (**Table 4**; over- and under-expressed). The well-known NFKB1 inhibitor alpha

(NFKBIA) was shown to be a favorable predictor in breast cancer patients (**Figure 16**). Moreover, NFKB1, together with the inflammatory factors HIF1A (hypoxiainducible factor 1) and STAT3, both of which were found to be up-regulated in radioresistant breast cancer cells (**Table 4**) and HIF1A overexpressed in breast cancer tissue (**Figure 15**), are critically implicated in cancer radioresistance and radiationinduced inflammatory responses (Multhoff & Radons, 2012). On the other hand, STAT1, found down-regulated in the same cell lines (**Table 4**), as opposed to STAT3, elicits pro-apoptotic and anti-proliferative responses and promotes anti-cancer immunity (Avalle et al., 2012).

Suppression of UBE2D3, which is down-regulated in radioresistant ALL cells (**Table 4**), was demonstrated in a study by (Wang et al., 2013) to decrease radiosensitivity of human breast cancer cells by promoting telomere maintenance. In addition, UBE2D3 is negatively correlated with TERF2 (telomeric repeat binding factor 2), the latter of which is primarily involved in telomere maintenance (Kim et al., 2009) and is down-regulated in radioresistance (**Table 4**).

Radiation can also exert its genotoxic and cytotoxic effects through the indirect and systemic induction of severe oxidative stress and the production of ROS in the organism (Kryston et al., 2011). The radiogene HIF1A plays a pivotal cytoprotective role against oxidative stress (Li et al., 2019) by inhibiting autophagy and cell death (Pouysségur et al., 2006). Moreover, the superoxide scavenging enzyme SOD2 (superoxide dismutase 2), found overexpressed both in radioresistant breast cancer and ALL cells (**Table 4**), at normal expression levels provides a cytoprotective effect. Thus, SOD2 can likely exert a protective effect on RR cancer cells by controlling potential ROS-mediated DNA damage via catalyzing the reduction of superoxide into less genotoxic molecules like oxygen (Wang et al., 2018).

Notably, in this study, pro-survival pathways (**Table 3**) like Notch signaling, were found to be implicated both in solid and blood cancers, probably by mediating survival of cancer cells to radiation-induced clustered/complex DNA damage (Capaccione & Pine, 2013c; Marston et al., 2009b). Moreover, the PI3K/AKT/mTOR signaling pathway is suggested to be important in regulating cell survival in response to different types of cellular stress (Hung et al., 2012; Porta et al., 2014b), including genotoxic stress.

Hence, pro-survival pathways could be considered as potential therapeutic targets in cancer (Porta et al., 2014c).

A major limitation of this study, particularly for the two solid cancers datasets, is the lack of patient-derived tumor tissue, as well as cancer stem cells, which constitute a subpopulation in solid tumors that display stem-like characteristics (Pavlopoulou et al., 2016c). Instead, the respective experimental studies relied on the use of commercial cancer cell lines; of particular note, the ALL cells were derived from "real" patients. However, the extent to which the individual cancer cell lines can capture the cellular and genomic complexity of tumors is questioned. Further research is needed to determine whether the results derived from the cancer cell lines investigated in this study or other studies could be extrapolated to their corresponding tissues of origin, like breast tumor tissue and colorectal adenocarcinoma. Nevertheless, it is suggested that large panels of cancer cell lines can faithfully capture the genomic heterogeneity of cancers (Sakellaropoulos et al., 2019; Vougas et al., 2019). Beyond the discussed limitations, the over- or under-expressions of the radiogenes in radioresistant phenotypes have been verified, to a great extent, by independent experimental biochemical studies available in the literature. In addition, the clinical implications of those genes are further supported by the patient survival results (**Figure 16**).

5.3. Identification of Favorable Prognostic Markers from the Transcriptomic Profiles of Long-Term Survivors of AML

The incidence of AML increases with age, and the survival rate is exceptionally poor when diagnosed after age 65. In the present study, we employed integrative bioinformatics approaches on AML-relevant transcriptomic data to identify biomarkers of favorable prognosis.

The asymptomatic nature of AML in combination with MRD and frequent disease relapse hamper AML prognostics, but a few patients manage to remain progressionfree for exceptional durations of time. Despite their sparsity, unusual patients can inform cancer biology (Saner et al., 2019c). For example, families with rare and high penetrant germline mutations (Li–Fraumeni syndrome) have catalyzed understanding of the function of tumor suppressor genes (Malkin et al., 1990), while populations with diminished cancer incidence (subjects with Laron syndrome in Ecuador) provided

insights on cancer metabolism and the insulin-like growth factor (IGF) pathway in tumor progression (Guevara-Aguirre et al., 2021). In an analogous manner, cases with exceptionally prolonged survival may bear distinct molecular characteristics, the elucidation of which could ultimately improve the outcome of individuals with more typical disease trajectories (Saner et al., 2019c). Long-term survivorship has, at least in part, a genetic and immunological basis (Balachandran et al., 2017; Reifenberger et al., 2014c; Shen et al., 2021c). Molecular factors contributing to favorable survival in a patient, who may otherwise have had a more typical course of disease, may include germline genetics, endogenous antitumor immune responses, tumor mutational load, and characteristics of the microenvironment (Saner et al., 2019c). Putting this novel rationale under testing, we defined subgroups of LTSs in our study population and investigated their distinct transcriptomic profiles as compared to patients below the median OS in order to infer favorable prognostic signatures which could be associated with resistance to disease progression. Surprisingly, this approach revealed that malignant cells of LTS AML patients exhibit prominent deregulations of genes involved in neuronal processes, thereby providing hints for associations between exceptionally prolonged survival and neurogenic cell traits.

The effect of the nervous system on tumor progression has recently come into the spotlight as an enigmatic phenomenon with a paramount effect in disease outcome. We recently demonstrated that all cancer types can activate neuronal networks 'off-context' and this ability largely influences the acquisition of aggressive characteristics. This trend seems to have a universal character across several cancers, since nervous system-related gene transcripts are upregulated in invasive stages and predict clinical outcomes in all cancer types, including AML. The prognostic potential of each nervous system-related gene depends on the cancer type (Logotheti, Marquardt, Richter, Sophie Hain, et al., 2020). Depending on the nervous systemrelated genes that are ectopically activated in a cancer cell context, patient survival can be affected either positively or negatively. Furthermore, different cancer types coopt genes essential for the development of distinct anatomical parts of the nervous system. In the case of AML, genes indispensable for the somatic sensory system, the cerebellum fissum, the sympathetic nervous system, and the dentate gyrus granule cell layer are associated with AML patient outcomes (Logotheti, Marquardt, Richter, Sophie Hain, et

al., 2020). Herein, we show that deregulations of nervous systemrelated genes are strongly associated with exceptionally long survival periods in AML. The relevant underlying mechanism awaits to be unveiled. On the one hand, it has been recently described that leukemia stem cells can evade the arachnoid space, migrate, proliferate in the cerebrospinal fluid (CSF), and consequently infiltrate perivascular spaces and brain parenchyma. The CNS is an immune-privileged site that can protect leukemic cells from chemotherapy. Overexpression of the neural cell adhesion molecule NCAM/CD56 in the surface of leukemic stem cells increases their ability to infiltrate the CNS (Deak et al., 2021). In this regard, deterministic deregulation of selected neuronal signaling pathways may hypothetically enable leukemic cells to 'mingle' within the neuronal tissue microenvironment and rewire with the electrically active neurons (Deak et al., 2021). On the other hand, Hanoun et al. (2014) demonstrated that adrenergic signals regulate AML and that when leukemic cell injections in mice are followed by sympathetic denervation, the course of disease is accelerated. They further provided compelling evidence that sympathetic neuropathy represents a mechanism by which AML co-opts the microenvironment to its own advantage in order to deplete niche cells that maintain healthy hematopoietic stem cells and expand leukemia-supportive mesenchymal progenitors (Hanoun et al., 2014). The interactions between leukemic and neuronal cells in the context of AML recurrence, MDR, and/or metastasis remain a terra incognita. Future research can unveil how 'hijacking' of networks of the nervous system might offer several selective advantages in leukemic cells, especially those that persist in the circulation after initial AML management, and represent the 'seeds' of MRD. Conversely, it would be interesting to investigate whether activation of other neuronal signaling pathways can be protective against MRD in AML patients. Deciphering the neuronal signaling pathways that are beneficial for leukemic cell survival over those with a detrimental effect is anticipated to shed more light on AML pathophysiology and might have important translational implications in disease prevention and management. Such studies could open up novel avenues for therapeutic interventions, for example, by repurposing already approved neurological drugs to target the cancer–neuronal crosstalk and prevent progression to aggressive stages (Dolgin, 2020).

Last, but not least, the profiles of the rare LTS patients who either never regress or manage to remain relapse-free for extended periods of time might be particularly meaningful in the context of AML prognosis. Their alleged resistance to disease progression or relapse could be an indicator of better abilities to cope both with MRD, which may be associated not only with epidemiological, environmental, or lifestyle factors, but also with their inherent molecular traits (Saner et al., 2019c). Hints supporting the influence of the hereditable genetic characteristics on the resistance to recurrence arise from studies demonstrating that ethnicities with inherently different genetic backgrounds, such as Caucasians and African Americans, exhibit differences in the AML prognosis, whereby Caucasian males have better CR and overall survival than African American males (Sekeres et al., 2004). It would be interesting to investigate whether patients with unusually long survival periods share favorable intrinsic characteristics that overall predispose them to better detect and self-eliminate residual cancer cells during or after treatment. Checking for additional associations between the inherent characteristics of the LTSs and several clinical traits that are important for treatment personalization, such as cytogenetic characteristics and therapy regimens, could also be highly informative for the management of AML. This endeavor would require appropriate study designs that are specifically focused on LTS patient subgroups, where patients surviving over 5, 10, or more years will be oversampled and compared to those with more conventional disease trajectories (Saner et al., 2019c). Unveiling the cancer genomes and intrinsic genetic traits of this novel clinical population as opposed to relapsing cancer patients or even the general non-cancer population could reveal predisposing factors for resistance to recurrence. Future clinical trials could aim at the characterization of the molecular factors that are associated with LTSs. Such clinical study endpoints would be invaluable for the prognosis and management of AML and could enable not only the identification of patients with unfavorable prognosis who will need special treatments, but also the design of appropriate targeted strategies that prolong survival.

6. CONCLUSIONS AND FUTURE ASPECTS

In the investigation of Meta analysis of the prognostic value of HOTAIR in cancer, we have performed a meta-analysis towards investigating the prognostic potential of the prominent lncRNA HOTAIR in cancer. On the basis of our findings, HOTAIR represents a potential powerful predictor of prognosis of overall survival, cancer recurrence, progression, and metastasis in multiple and diverse types of cancers. Therefore, HOTAIR could be applied in the clinical setting as a universal biomarker for monitoring cancer patient survival.

In the study of the identification of molecular determinants of cancer cell resistance to ionizing radiation, we employed an integrative bioinformatics approach to analyze transcriptomic data regarding the molecular determinants of cancer cell radioresistance. On the basis of our findings, both solid and hematologic cancer cells likely depend on similar mechanisms to confer resistance to IR (i.e., DDR/R and cell survival). Moreover, we identified 36 functionally associated radiogenes that participate in radioresistance-associated pathways. Most of those radiogenes were also shown to be differentially regulated in the corresponding cancer tissues. Moreover, several of the radiogenes were found to have potential prognostic value for the clinical outcome of cancer patients. However, the availability of clinically derived cancer tissues would provide a more reliable source for conducting research on the response of cancer patients to radiation. The overall data presented herein can be particularly useful for clinicians in selecting suitable targets (e.g., DDR/R inhibitors) for appropriate combination therapy using IR. In conclusion, we suggest that this bioinformatics premise can be harnessed as a first step in the rational design of in vivo experimental studies or in personalized medicine for optimizing tumor response and cancer cell susceptibility to therapeutic ionizing radiation and reduction of the total effective radiation dose administered to the patient.

In the study regarding the identification of favorable prognostic markers in AML, we applied in silico workflows on publicly available AML-relevant transcriptomic data. Taking advantage of the fact that transcriptomic data of AML cases with unusually extended survival periods already exist in well-established databases, we tested a novel notion, which proposes that rare patients at the extremes can be informative for

prognosis and management of the general cancer population. Comparisons of the transcriptomes of these exceptional cases with those of more typical disease trajectories in our study cohort revealed an association of deregulations of nervous system-related genes with favorable AML prognosis. Future experimental and clinical studies can provide further insights on the underlying molecular mechanisms, inform clinical decisions, and catalyze the translation of these findings to therapeutic solutions.

7. REFERENCES

Abdeahad, H., Avan, A., Pashirzad, M., Khazaei, M., Soleimanpour, S., Ferns, G. A., Fiuji, H., Ryzhikov, M., Bahrami, A., & Hassanian, S. M. (2019). The prognostic potential of long noncoding RNA HOTAIR expression in human digestive system carcinomas: A meta-analysis. Journal of Cellular Physiology, 234(7), 10926–10933. https://doi.org/10.1002/jcp.27918

Abelson, S., Collord, G., Ng, S. W. K., Weissbrod, O., Mendelson Cohen, N., Niemeyer, E., Barda, N., Zuzarte, P. C., Heisler, L., Sundaravadanam, Y., Luben, R., Hayat, S., Wang, T. T., Zhao, Z., Cirlan, I., Pugh, T. J., Soave, D., Ng, K., Latimer, C., … Shlush, L. I. (2018). Prediction of acute myeloid leukaemia risk in healthy individuals. Nature, 559(7714), 400–404. https://doi.org/10.1038/S41586-0180317-6

Adès, L., Guerci, A., Raffoux, E., Sanz, M., Chevallier, P., Lapusan, S., Recher, C., Thomas, X., Rayon, C., Castaigne, S., Tournilhac, O., de Botton, S., Ifrah, N., Cahn, J. Y., Solary, E., Gardin, C., Fegeux, N., Bordessoule, D., Ferrant, A., … Fenaux, P. (2010). Very long-term outcome of acute promyelocytic leukemia after treatment with all-trans retinoic acid and chemotherapy: The European APL Group experience. Blood, 115(9), 1690–1696. https://doi.org/10.1182/blood-2009-07233387

Akl, H., Vervloessem, T., Kiviluoto, S., Bittremieux, M., Parys, J. B., de Smedt, H., & Bultynck, G. (2014). A dual role for the anti-apoptotic Bcl-2 protein in cancer: Mitochondria versus endoplasmic reticulum. Biochimica et Biophysica Acta - Molecular Cell Research, 1843(10), 2240–2252. https://doi.org/10.1016/j.bbamcr.2014.04.017

Alnasir, J., & Shanahan, H. P. (2015). Investigation into the annotation of protocol sequencing steps in the sequence read archive. GigaScience, 4(1). https://doi.org/10.1186/S13742-015-0064-7

An Introduction to the Bootstrap - 1st Edition - Bradley Efron - R.J. (n.d.). Retrieved June 24, 2022, from https://www.routledge.com/An-Introduction-to-the-Bootstrap/Efron-Tibshirani/p/book/9780412042317

Arndt, V., Koch-Gallenkamp, L., Jansen, L., Bertram, H., Eberle, A., Holleczek, B., Schmid-Höpfner, S., Waldmann, A., Zeissig, S. R., & Brenner, H. (2017). Quality of life in long-term and very long-term cancer survivors versus population controls in Germany. Acta Oncologica (Stockholm, Sweden), 56(2), 190–197. https://doi.org/10.1080/0284186X.2016.1266089

Avalle, L., Pensa, S., Regis, G., Novelli, F., & Poli, V. (2012). STAT1 and STAT3 in tumorigenesis: A matter of balance. JAK-STAT, 1(2), 65–72. https://doi.org/10.4161/jkst.20045

Awasthi, P., Foiani, M., & Kumar, A. (2015). ATM and ATR signaling at a glance. Journal of Cell Science, 128(23), 4255–4262. https://doi.org/10.1242/jcs.169730

Balachandran, V. P., Łuksza, M., Zhao, J. N., Makarov, V., Moral, J. A., Remark, R., Herbst, B., Askan, G., Bhanot, U., Senbabaoglu, Y., Wells, D. K., Cary, C. I. O., Grbovic-Huezo, O., Attiyeh, M., Medina, B., Zhang, J., Loo, J., Saglimbeni, J., AbuAkeel, M., … Bassi, C. (2017). Identification of unique neoantigen qualities in long-

term survivors of pancreatic cancer. Nature, 551(7681), S12–S16. https://doi.org/10.1038/nature24462

Barkett, M., & Gilmore, T. D. (1999). Control of apoptosis by Rel/NF-κB transcription factors. Oncogene, 18(49), 6910–6924. https://doi.org/10.1038/sj.onc.1203238

Barrett, T., Wilhite, S. E., Ledoux, P., Evangelista, C., Kim, I. F., Tomashevsky, M., Marshall, K. A., Phillippy, K. H., Sherman, P. M., Holko, M., Yefanov, A., Lee, H., Zhang, N., Robertson, C. L., Serova, N., Davis, S., & Soboleva, A. (2013). NCBI GEO: Archive for functional genomics data sets - Update. Nucleic Acids Research, 41(D1). https://doi.org/10.1093/nar/gks1193

Baskar, R., Lee, K. A., Yeo, R., & Yeoh, K. W. (2012a). Cancer and radiation therapy: current advances and future directions. International Journal of Medical Sciences,

9(3), 193–199. https://doi.org/10.7150/IJMS.3635

Baskar, R., Lee, K. A., Yeo, R., & Yeoh, K. W. (2012b). Cancer and radiation therapy: current advances and future directions. International Journal of Medical Sciences,

9(3), 193–199. https://doi.org/10.7150/IJMS.3635

Begg, A. C., Stewart, F. A., & Vens, C. (2011a). Strategies to improve radiotherapy with targeted drugs. Nature Reviews. Cancer, 11(4), 239–253. https://doi.org/10.1038/NRC3007

Begg, A. C., Stewart, F. A., & Vens, C. (2011b). Strategies to improve radiotherapy with targeted drugs. Nature Reviews Cancer, 11(4), 239–253. https://doi.org/10.1038/nrc3007

Begg, C. B., & Mazumdar, M. (1994). Operating Characteristics of a Rank Correlation Test for Publication Bias. Biometrics, 50(4), 1088. https://doi.org/10.2307/2533446

Beheshti, A., McDonald, J. T., Miller, J., Grabham, P., & Costes, S. v. (2019). Genelab database analyses suggest long-term impact of space radiation on the cardiovascular system by the activation of FYN through reactive oxygen species. International Journal of Molecular Sciences, 20(3). https://doi.org/10.3390/ijms20030661

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society: Series B (Methodological), 57(1), 289–300. https://doi.org/10.1111/J.25176161.1995.TB02031.X

Benton, C. B., & Ravandi, F. (2017). A mind map for managing minimal residual disease in acute myeloid leukemia. Clinical Advances in Hematology & Oncology : H&O, 15(11), 859–867. https://pubmed.ncbi.nlm.nih.gov/29200419/

Bhan, A., & Mandal, S. S. (2015a). LncRNA HOTAIR: A master regulator of chromatin dynamics and cancer. Biochimica et Biophysica Acta - Reviews on Cancer, 1856(1), 151–164. https://doi.org/10.1016/j.bbcan.2015.07.001

Bhan, A., & Mandal, S. S. (2015b). LncRNA HOTAIR: A master regulator of chromatin dynamics and cancer. Biochimica et Biophysica Acta, 1856(1), 151–164. https://doi.org/10.1016/J.BBCAN.2015.07.001

Bloomfield, C. D., Shuma, C., Regal, L., Philip, P. P., Hossfeld, D. K., Hagemeijer, A. M., Garson, O. M., Peterson, B. A., Sakurai, M., Alimena, G., Berger, R., Rowley, J. D., Ruutu, T., Mitelman, F., Dewald, G. W., & Swansbury, J. (1997). Long-term

survival of patients with acute myeloid leukemia: A third follow-up of Fourth International Workshop on Chromosomes in Leukemia. Cancer, 80(11 SUPPL.), 2191–2198. https://doi.org/10.1002/(sici)1097-0142(19971201)80:11+<2191::aid-cncr5>3.0.co;2-l

Braschi, B., Denny, P., Gray, K., Jones, T., Seal, R., Tweedie, S., Yates, B., & Bruford, E. (2019). Genenames.org: The HGNC and VGNC resources in 2019. Nucleic Acids Research, 47(D1), D786–D792. https://doi.org/10.1093/nar/gky930

Brennan, D. J., O'Connor, D. P., Rexhepaj, E., Ponten, F., & Gallagher, W. M. (2010a). Antibody-based proteomics: fast-tracking molecular diagnostics in oncology. Nature Reviews. Cancer, 10(9), 605–617. https://doi.org/10.1038/NRC2902

Brennan, D. J., O'Connor, D. P., Rexhepaj, E., Ponten, F., & Gallagher, W. M. (2010b). Antibody-based proteomics: fast-tracking molecular diagnostics in oncology. Nature Reviews. Cancer, 10(9), 605–617. https://doi.org/10.1038/NRC2902

Bruinsma, W., Aprelia, M., Garciá-Santisteban, I., Kool, J., Xu, Y. J., & Medema, R. H. (2017). Inhibition of Polo-like kinase 1 during the DNA damage response is mediated through loss of Aurora A recruitment by Bora. Oncogene, 36(13), 1840–1848. https://doi.org/10.1038/onc.2016.347

Buckley, A. M., Lynam-Lennon, N., O'Neill, H., & O'Sullivan, J. (2020). Targeting hallmarks of cancer to enhance radiosensitivity in gastrointestinal cancers. Nature Reviews. Gastroenterology & Hepatology, 17(5), 298–313. https://doi.org/10.1038/S41575-019-0247-2

Budman, J., & Chu, G. (2005). Processing of DNA for nonhomologous end-joining by cell-free extract. EMBO Journal, 24(4), 849–860. https://doi.org/10.1038/sj.emboj.7600563

Bukowska, B., & Karwowski, B. T. (2018). The Clustered DNA Lesions – Types, Pathways of Repair and Relevance to Human Health. Current Medicinal Chemistry, 25(23), 2722–2735. https://doi.org/10.2174/0929867325666180226110502

Cai, B., Song, X. Q., Cai, J. P., & Zhang, S. (2014). HOTAIR: a cancer-related long noncoding RNA. Neoplasma, 61(4), 379–391. https://doi.org/10.4149/NEO_2014_075

Cai, C. K., Zhao, G. Y., Tian, L. Y., Liu, L., Yan, K., Ma, Y. L., Ji, Z. W., Li, X. X., Han,

K., Gao, J., Qiu, X. C., Fan, Q. Y., Yang, T. T., & Ma, B. A. (2012). miR-15a and miR-16-1 downregulate CCND1 and induce apoptosis and cell cycle arrest in osteosarcoma. Oncology Reports, 28(5), 1764–1770. https://doi.org/10.3892/or.2012.1995

Cai, H., An, Y., Chen, X., Sun, D., Chen, T., Peng, Y., Zhu, F., Jiang, Y., & He, X. (2016). Epigenetic inhibition of miR-663b by long non-coding RNA HOTAIR promotes pancreatic cancer cell proliferation via upregulation of insulin-like growth factor 2. Oncotarget, 7(52), 86857–86870. https://doi.org/10.18632/oncotarget.13490

Caldecott, K. W. (2008). Single-strand break repair and genetic disease. Nature Reviews Genetics, 9(8), 619–631. https://doi.org/10.1038/nrg2380

Capaccione, K. M., & Pine, S. R. (2013a). The Notch signaling pathway as a mediator of tumor survival. Carcinogenesis, 34(7), 1420–1430. https://doi.org/10.1093/CARCIN/BGT127

Capaccione, K. M., & Pine, S. R. (2013b). The Notch signaling pathway as a mediator of tumor survival. Carcinogenesis, 34(7), 1420–1430. https://doi.org/10.1093/carcin/bgt127

Capaccione, K. M., & Pine, S. R. (2013c). The Notch signaling pathway as a mediator of tumor survival. Carcinogenesis, 34(7), 1420–1430. https://doi.org/10.1093/carcin/bgt127

Chen, F. J., Sun, M., Li, S. Q., Wu, Q. Q., Ji, L., Liu, Z. L., Zhou, G. Z., Cao, G., Jin, L., Xie, H. W., Wang, C. M., Lv, J., De, W., Wu, M., & Cao, X. F. (2013). Upregulation of the long non-coding rna hotair promotes esophageal squamous cell carcinoma metastasis and poor prognosis. Molecular Carcinogenesis, 52(11), 908–915. https://doi.org/10.1002/mc.21944

Cheng, M. J., Hourigan, C. S., & Smith, T. J. (2014). Adult Acute Myeloid Leukemia Long-term Survivors. Journal of Leukemia (Los Angeles, Calif.), 2(2). https://doi.org/10.4172/2329-6917.1000135

Chiou, S. K., Jones, M. K., & Tarnawski, A. S. (2003). Survivin - An anti-apoptosis protein: Its biological roles and implications for cancer and beyond. Medical Science Monitor, 9(4).

Clough, E., & Barrett, T. (2016). The Gene Expression Omnibus database. Methods in Molecular Biology, 1418, 93–110. https://doi.org/10.1007/978-1-4939-3578-9_5

Colaprico, A., Silva, T. C., Olsen, C., Garofano, L., Cava, C., Garolini, D., Sabedot, T. S., Malta, T. M., Pagnotta, S. M., Castiglioni, I., Ceccarelli, M., Bontempi, G., & Noushmehr, H. (2016). TCGAbiolinks: An R/Bioconductor package for integrative analysis of TCGA data. Nucleic Acids Research, 44(8), e71. https://doi.org/10.1093/nar/gkv1507

Colin, S. D., Nava, V. E., Gedrich, R. W., Clem, R. J., van Dongen, J. L., Gilfillan, M. C., Shiels, H., Hardwick, J. M., & Thompson, C. B. (1996). A conserved family of cellular genes related to the baculovirus iap gene and encoding apoptosis inhibitors. EMBO Journal, 15(11), 2685–2694. https://doi.org/10.1002/j.14602075.1996.tb00629.x

Cunningham, F., Achuthan, P., Akanni, W., Allen, J., Amode, M. R., Armean, I. M., Bennett, R., Bhai, J., Billis, K., Boddu, S., Cummins, C., Davidson, C., Dodiya, K. J., Gall, A., Girón, C. G., Gil, L., Grego, T., Haggerty, L., Haskell, E., … Flicek, P. (2019). Ensembl 2019. Nucleic Acids Research, 47(D1), D745–D751. https://doi.org/10.1093/NAR/GKY1113

Cunningham, F., Allen, J. E., Allen, J., Alvarez-Jarreta, J., Amode, M. R., Armean, I. M., Austine-Orimoloye, O., Azov, A. G., Barnes, I., Bennett, R., Berry, A., Bhai, J., Bignell, A., Billis, K., Boddu, S., Brooks, L., Charkhchi, M., Cummins, C., da Rin Fioretto, L., … Flicek, P. (2022). Ensembl 2022. Nucleic Acids Research, 50(D1), D988–D995. https://doi.org/10.1093/NAR/GKAB1049

Deak, D., Gorcea-Andronic, N., Sas, V., Teodorescu, P., Constantinescu, C., Iluta, S., Pasca, S., Hotea, I., Turcas, C., Moisoiu, V., Zimta, A.-A., Galdean, S., Steinheber, J., Rus, I., Rauch, S., Richlitzki, C., Munteanu, R., Jurj, A., Petrushev, B., … Tomuleasa, C. (2021). A narrative review of central nervous system involvement in acute leukemias. Annals of Translational Medicine, 9(1), 68. https://doi.org/10.21037/atm-20-3140

Delaney, G., Jacob, S., Featherstone, C., & Barton, M. (2005a). The role of radiotherapy in cancer treatment: estimating optimal utilization from a review of evidence-based clinical guidelines. Cancer, 104(6), 1129–1137. https://doi.org/10.1002/CNCR.21324

Delaney, G., Jacob, S., Featherstone, C., & Barton, M. (2005b). The role of radiotherapy in cancer treatment: Estimating optimal utilization from a review of evidence-based clinical guidelines. Cancer, 104(6), 1129–1137. https://doi.org/10.1002/cncr.21324

Deng, C. X. (2006). BRCA1: Cell cycle checkpoint, genetic instability, DNA damage response and cancer evolution. Nucleic Acids Research, 34(5), 1416–1426. https://doi.org/10.1093/nar/gkl010

Deng, Q., Sun, H., He, B., Pan, Y., Gao, T., Chen, J., Ying, H., Liu, X., Wang, F., Xu, Y., & Wang, S. (2014a). Prognostic value of long non-coding RNA HOTAIR in various cancers. PLoS ONE, 9(10). https://doi.org/10.1371/journal.pone.0110059 Deng, Q., Sun, H., He, B., Pan, Y., Gao, T., Chen, J., Ying, H., Liu, X., Wang, F., Xu, Y., & Wang, S. (2014b). Prognostic value of long non-coding RNA HOTAIR in various cancers. PLoS ONE, 9(10). https://doi.org/10.1371/journal.pone.0110059

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. Controlled Clinical Trials, 7(3), 177–188. https://doi.org/10.1016/0197-2456(86)90046-2

DerSimonian, R., & Laird, N. (2015). Meta-analysis in clinical trials revisited. Contemporary Clinical Trials, 45(Pt A), 139–145. https://doi.org/10.1016/J.CCT.2015.09.002

Deveraux, Q. L., & Reed, J. C. (1999). IAP family proteins - Suppressors of apoptosis. Genes and Development, 13(3), 239–252. https://doi.org/10.1101/gad.13.3.239

Di, W., Li, Q., Shen, W., Guo, H., & Zhao, S. (2017). The long non-coding RNA HOTAIR promotes thyroid cancer cell growth, invasion and migration through the miR-1CCND2 axis. American Journal of Cancer Research, 7(6), 1298–1309.

Döhner, H., Estey, E. H., Amadori, S., Appelbaum, F. R., Büchner, T., Burnett, A. K., Dombret, H., Fenaux, P., Grimwade, D., Larson, R. A., Lo-Coco, F., Naoe, T., Niederwieser, D., Ossenkoppele, G. J., Sanz, M. A., Sierra, J., Tallman, M. S., Löwenberg, B., & Bloomfield, C. D. (2010). Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet. Blood, 115(3), 453–474. https://doi.org/10.1182/BLOOD-2009-07-235358

Dolgin, E. (2020). Cancer-neuronal crosstalk and the startups working to silence it.

Nature Biotechnology, 38(2), 115–117. https://doi.org/10.1038/s41587-0200411-9

Duckett, D. R., Drummond, J. T., Murchie, A. I. H., Reardon, J. T., Sancar, A., Lilley, D. M. J., & Modrich, P. (1996). Humnan MutSα recognizes damaged DNA base pairs containing O6-methylguanine, O4-methylthymine, or the cisplatin-d(GpG) adduct. Proceedings of the National Academy of Sciences of the United States of America, 93(13), 6443–6447. https://doi.org/10.1073/pnas.93.13.6443

Edelbrock, M. A., Kaliyaperumal, S., & Williams, K. J. (2013). Structural, molecular and cellular functions of MSH2 and MSH6 during DNA mismatch repair, damage signaling and other noncanonical activities. Mutation Research - Fundamental and

Molecular Mechanisms of Mutagenesis, 743–744, 53–66. https://doi.org/10.1016/j.mrfmmm.2012.12.008

Edgar, R., Domrachev, M., & Lash, A. E. (2002a). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Research, 30(1), 207–210. https://doi.org/10.1093/NAR/30.1.207

Edgar, R., Domrachev, M., & Lash, A. E. (2002b). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Research, 30(1), 207–210. https://doi.org/10.1093/NAR/30.1.207

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. British Medical Journal, 315(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629

Emons, G., Spitzner, M., Reineke, S., Möller, J., Auslander, N., Kramer, F., Hu, Y., Beissbarth, T., Wolff, H. A., Rave-Fränk, M., Heßmann, E., Gaedcke, J., Ghadimi, B. M., Johnsen, S. A., Ried, T., & Grade, M. (2017). Chemoradiotherapy resistance in colorectal cancer cells is mediated by Wnt/β-catenin signaling. Molecular Cancer Research, 15(11), 1481–1490. https://doi.org/10.1158/1541-7786.MCR-17-0205

Endo, H., Shiroki, T., Nakagawa, T., Yokoyama, M., Tamai, K., Yamanami, H., Fujiya, T., Sato, I., Yamaguchi, K., Tanaka, N., Iijima, K., Shimosegawa, T., Sugamura, K., & Satoh, K. (2013). Enhanced Expression of Long Non-Coding RNA HOTAIR Is Associated with the Development of Gastric Cancer. PLoS ONE, 8(10). https://doi.org/10.1371/journal.pone.0077070

Escamilla, J. E., Barquera, B., Ramírez, R., García-Horsman, A., & del Arenal, P. (1988). Role of menaquinone in inactivation and activation of the Bacillus cereus forespore respiratory system. Journal of Bacteriology, 170(12), 5908–5912. https://doi.org/10.1128/jb.170.12.5908-5912.1988

Farag, S. S., Ruppert, A. S., Mrózek, K., Mayer, R. J., Stone, R. M., Carroll, A. J., Powell, B. L., Moore, J. O., Pettenati, M. J., Koduru, P. R. K., Stamberg, J., Baer, M. R., Block, A. W., Vardiman, J. W., Kolitz, J. E., Schiffer, C. A., Larson, R. A., & Bloomfield, C. D. (2005). Outcome of induction and postremission therapy in younger adults with acute myeloid leukemia with normal karyotype: A cancer and leukemia group B study. Journal of Clinical Oncology, 23(3), 482–493. https://doi.org/10.1200/JCO.2005.06.090

Fiorini, N., Lipman, D. J., & Lu, Z. (2017a). Towards PubMed 2.0. ELife, 6. https://doi.org/10.7554/ELIFE.28801

Fiorini, N., Lipman, D. J., & Lu, Z. (2017b). Towards PubMed 2.0. ELife, 6. https://doi.org/10.7554/eLife.28801

Flaggs, G., Plug, A. W., Dunks, K. M., Mundt, K. E., Ford, J. C., Quiggle, M. R. E., Taylor, E. M., Westphal, C. H., Ashley, T., Hoekstra, M. F., & Carr, A. M. (1997). Atm-dependent interactions of a mammalian Chk1 homolog with meiotic chromosomes. Current Biology, 7(12), 977–986. https://doi.org/10.1016/S09609822(06)00417-9

Francica, P., Mutlu, M., Blomen, V. A., Oliveira, C., Nowicka, Z., Trenner, A., Gerhards, N. M., Bouwman, P., Stickel, E., Hekkelman, M. L., Lingg, L., Klebic, I., van de Ven, M., de Korte-Grimmerink, R., Howald, D., Jonkers, J., Sartori, A. A., Fendler, W., Chapman, J. R., … Rottenberg, S. (2020). Functional Radiogenetic Profiling Implicates ERCC6L2 in Non-homologous End Joining. Cell Reports, 32(8). https://doi.org/10.1016/j.celrep.2020.108068

Frankish, A., Diekhans, M., Jungreis, I., Lagarde, J., Loveland, J. E., Mudge, J. M., Sisu, C., Wright, J. C., Armstrong, J., Barnes, I., Berry, A., Bignell, A., Boix, C., Sala, S. C., Cunningham, F., Domenico, T. di, Donaldson, S., Fiddes, I. T., Girón, C. G., … Flicek, P. (2021). GENCODE 2021. Nucleic Acids Research, 49(D1), D916–D923. https://doi.org/10.1093/NAR/GKAA1087

Fridman, J. S., & Lowe, S. W. (2003). Control of apoptosis by p53. Oncogene, 22(56 REV. ISS. 8), 9030–9040. https://doi.org/10.1038/sj.onc.1207116

Fujise, K., Zhang, D., Liu, J. L., & Yeh, E. T. H. (2000). Regulation of apoptosis and cell cycle progression by MCL1; Differential role of proliferating cell nuclear antigen. Journal of Biological Chemistry, 275(50), 39458–39465. https://doi.org/10.1074/jbc.M006626200

Gao, J.-Z., Li, J., DU, J.-L., & Li, X.-L. (2016). Long non-coding RNA HOTAIR is a marker for hepatocellular carcinoma progression and tumor recurrence. Oncology Letters, 11(3), 1791–1798. https://doi.org/10.3892/ol.2016.4130

Gatti, R. A., Berkel, I., Boder, E., Braedt, G., Charmley, P., Concannon, P., Ersoy, F., Foroud, T., Jaspers, N. G. J., Lange, K., Lathrop, G. M., Leppert, M., Nakamura, Y., O'Connell, P., Paterson, M., Salser, W., Sanal, O., Silver, J., Sparkes, R. S., … Yoder, F. (1988). Localization of an ataxia-telangiectasia gene to chromosome 11q22-23. Nature, 336(6199), 577–580. https://doi.org/10.1038/336577a0

Geisler, S., & Coller, J. (2013). RNA in unexpected places: long non-coding RNA functions in diverse cellular contexts. Nature Reviews. Molecular Cell Biology, 14(11), 699–712. https://doi.org/10.1038/NRM3679

Georgakilas, A. G., O'Neill, P., & Stewart, R. D. (2013). Induction and repair of clustered DNA lesions: What do we know so far? Radiation Research, 180(1), 100– 109. https://doi.org/10.1667/RR3041.1

Glaser, S. P., Lee, E. F., Trounson, E., Bouillet, P., Wei, A., Fairlie, W. D., Izon, D. J., Zuber, J., Rappaport, A. R., Herold, M. J., Alexander, W. S., Lowe, S. W., Robb, L., & Strasser, A. (2012). Anti-apoptotic mcl-1 is essential for the development and sustained growth of acute myeloid leukemia. Genes and Development, 26(2), 120–125. https://doi.org/10.1101/gad.182980.111

Gong, L., Zhang, Y., Liu, C., Zhang, M., & Han, S. (2021). Application of radiosensitizers in cancer radiotherapy. International Journal of Nanomedicine, 16, 1083–1102. https://doi.org/10.2147/IJN.S290438

Gray, M., Turnbull, A. K., Ward, C., Meehan, J., Martínez-Pérez, C., Bonello, M., Pang, L. Y., Langdon, S. P., Kunkler, I. H., Murray, A., & Argyle, D. (2019a). Development and characterisation of acquired radioresistant breast cancer cell lines. Radiation Oncology, 14(1). https://doi.org/10.1186/s13014-019-1268-2

Gray, M., Turnbull, A. K., Ward, C., Meehan, J., Martínez-Pérez, C., Bonello, M., Pang, L. Y., Langdon, S. P., Kunkler, I. H., Murray, A., & Argyle, D. (2019b). Development and characterisation of acquired radioresistant breast cancer cell lines. Radiation Oncology, 14(1). https://doi.org/10.1186/s13014-019-1268-2

Gudmundsdottir, K., & Ashworth, A. (2006). The roles of BRCA1 and BRCA2 and associated proteins in the maintenance of genomic stability. Oncogene, 25(43), 5864–5874. https://doi.org/10.1038/sj.onc.1209874

Guevara-Aguirre, J., Bautista, C., Torres, C., Peña, G., Guevara, C., Palacios, C., Guevara, A., & Gavilanes, A. W. D. (2021). Insights from the clinical phenotype of subjects with Laron syndrome in Ecuador. Reviews in Endocrine and Metabolic Disorders, 22(1), 59–70. https://doi.org/10.1007/s11154-020-09602-4

Gupta, R. A., Shah, N., Wang, K. C., Kim, J., Horlings, H. M., Wong, D. J., Tsai, M. C., Hung, T., Argani, P., Rinn, J. L., Wang, Y., Brzoska, P., Kong, B., Li, R., West, R. B., van de Vijver, M. J., Sukumar, S., & Chang, H. Y. (2010a). Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature, 464(7291), 1071–1076. https://doi.org/10.1038/nature08975

Gupta, R. A., Shah, N., Wang, K. C., Kim, J., Horlings, H. M., Wong, D. J., Tsai, M. C., Hung, T., Argani, P., Rinn, J. L., Wang, Y., Brzoska, P., Kong, B., Li, R., West, R. B., van de Vijver, M. J., Sukumar, S., & Chang, H. Y. (2010b). Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature, 464(7291), 1071–1076. https://doi.org/10.1038/nature08975

Gupta, R. A., Shah, N., Wang, K. C., Kim, J., Horlings, H. M., Wong, D. J., Tsai, M. C., Hung, T., Argani, P., Rinn, J. L., Wang, Y., Brzoska, P., Kong, B., Li, R., West, R. B., van de Vijver, M. J., Sukumar, S., & Chang, H. Y. (2010c). Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature, 464(7291), 1071–1076. https://doi.org/10.1038/nature08975

Gupta, R. A., Shah, N., Wang, K. C., Kim, J., Horlings, H. M., Wong, D. J., Tsai, M. C., Hung, T., Argani, P., Rinn, J. L., Wang, Y., Brzoska, P., Kong, B., Li, R., West, R. B., van de Vijver, M. J., Sukumar, S., & Chang, H. Y. (2010d). Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature, 464(7291), 1071–1076. https://doi.org/10.1038/NATURE08975

Habraken, Y., Sung, P., Prakash, L., & Prakash, S. (1996). Binding of insertion/deletion DNA mismatches by the heterodimer of yeast mismatch repair proteins MSH2 and MSH3. Current Biology, 6(9), 1185–1187. https://doi.org/10.1016/S09609822(02)70686-6

Hada, M., & Georgakilas, A. G. (2008). Formation of clustered DNA damage after highLET irradiation: A review. Journal of Radiation Research, 49(3), 203–210. https://doi.org/10.1269/jrr.07123

Hajjari, M., & Salavaty, A. (2015a). HOTAIR: an oncogenic long non-coding RNA in different cancers. Cancer Biology and Medicine, 12(1), 1–9. https://doi.org/10.7497/j.issn.2095-3941.2015.0006

Hajjari, M., & Salavaty, A. (2015b). HOTAIR: an oncogenic long non-coding RNA in different cancers. Cancer Biology and Medicine, 12(1), 1–9. https://doi.org/10.7497/j.issn.2095-3941.2015.0006

Han, J. W., Flemington, C., Houghton, A. B., Gu, Z., Zambetti, G. P., Lutz, R. J., Zhu, L., & Chittenden, T. (2001). Expression of bbc3, a pro-apoptotic BH3-only gene, is regulated by diverse cell death and survival signals. Proceedings of the National Academy of Sciences of the United States of America, 98(20), 11318–11323. https://doi.org/10.1073/pnas.201208798

Hanoun, M., Zhang, D., Mizoguchi, T., Pinho, S., Pierce, H., Kunisaki, Y., Lacombe, J., Armstrong, S. A., Dührsen, U., & Frenette, P. S. (2014). Acute myelogenous leukemia-induced sympathetic neuropathy promotes malignancy in an altered hematopoietic stem cell Niche. Cell Stem Cell, 15(3), 365–375. https://doi.org/10.1016/j.stem.2014.06.020

Hartwell, L., Mankoff, D., Paulovich, A., Ramsey, S., & Swisher, E. (2006a). Cancer biomarkers: a systems approach. Nature Biotechnology, 24(8), 905–908. https://doi.org/10.1038/NBT0806-905

Hartwell, L., Mankoff, D., Paulovich, A., Ramsey, S., & Swisher, E. (2006b). Cancer biomarkers: a systems approach. Nature Biotechnology, 24(8), 905–908. https://doi.org/10.1038/NBT0806-905

Haupt, S., Berger, M., Goldberg, Z., & Haupt, Y. (2003). Apoptosis - The p53 network. Journal of Cell Science, 116(20), 4077–4085. https://doi.org/10.1242/jcs.00739

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-

analysis. Statistics in Medicine, 21(11), 1539–1558. https://doi.org/10.1002/sim.1186

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. British Medical Journal, 327(7414), 557–560. https://doi.org/10.1136/bmj.327.7414.557

Holland, P. W. H., Booth, H. A. F., & Bruford, E. A. (2007). Classification and nomenclature of all human homeobox genes. BMC Biology, 5. https://doi.org/10.1186/1741-7007-5-47

Holt, S. M., Scemama, J. L., Panayiotidis, M. I., & Georgakilas, A. G. (2009). Compromised repair of clustered DNA damage in the human acute lymphoblastic leukemia MSH2-deficient NALM-6 cells. Mutation Research, 674(1–2), 123–130. https://doi.org/10.1016/J.MRGENTOX.2008.09.014

HOTAIR is upregulated in acute myeloid leukemia and that indicates a poor prognosis - PubMed. (n.d.). Retrieved June 24, 2022, from https://pubmed.ncbi.nlm.nih.gov/26261618/

Hung, C. M., Garcia-Haro, L., Sparks, C. A., & Guertin, D. A. (2012). mTOR-dependent cell survival mechanisms. Cold Spring Harbor Perspectives in Biology, 4(12). https://doi.org/10.1101/cshperspect.a008771

Hyun, S. Y., Hwan, H. I., & Jang, Y. J. (2014). Polo-like kinase-1 in DNA damage response. BMB Reports, 47(5), 249–255. https://doi.org/10.5483/BMBRep.2014.47.5.061

Iliakis, G., Wang, Y., Guan, J., & Wang, H. (2003). DNA damage checkpoint control in cells exposed to ionizing radiation. Oncogene, 22(37 REV. ISS. 3), 5834–5847. https://doi.org/10.1038/sj.onc.1206682

Jiang, C., Yang, Y., Yang, Y., Guo, L., Huang, J., Liu, X., Wu, C., & Zou, J. (2018). Long noncoding RNA (lncRNA) HOTAIR affects tumorigenesis and metastasis of non-small cell lung cancer by upregulating miR-613. Oncology Research, 26(5), 725–734. https://doi.org/10.3727/096504017X15119467381615

Kanakoglou, D. S., Michalettou, T. D., Vasileiou, C., Gioukakis, E., Maneta, D., Kyriakidis, K. v., Georgakilas, A. G., & Michalopoulos, I. (2020). Effects of highdose

ionizing radiation in human gene expression: A meta-analysis. International Journal of Molecular Sciences, 21(6). https://doi.org/10.3390/ijms21061938

Kantarjian, H., Kadia, T., DiNardo, C., Daver, N., Borthakur, G., Jabbour, E., GarciaManero, G., Konopleva, M., & Ravandi, F. (2021). Acute myeloid leukemia: current progress and future directions. Blood Cancer Journal, 11(2). https://doi.org/10.1038/S41408-021-00425-3

Kazimierczyk, M., Kasprowicz, M. K., Kasprzyk, M. E., & Wrzesinski, J. (2020). Human Long Noncoding RNA Interactome: Detection, Characterization and Function. International Journal of Molecular Sciences, 21(3). https://doi.org/10.3390/IJMS21031027

Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. Nature Methods, 12(4), 357–360. https://doi.org/10.1038/NMETH.3317

Kim, H., Lee, O. H., Xin, H., Chen, L. Y., Qin, J., Chae, H. K., Lin, S. Y., Safari, A., Liu, D., & Songyang, Z. (2009). TRF2 functions as a protein hub and regulates telomere maintenance by recognizing specific peptide motifs. Nature Structural and Molecular Biology, 16(4), 372–379. https://doi.org/10.1038/nsmb.1575

Klepin, H. D., & Balducci, L. (2009). Acute Myelogenous Leukemia in Older Adults. The Oncologist, 14(3), 222–232. https://doi.org/10.1634/theoncologist.2008-0224

Kogo, R., Shimamura, T., Mimori, K., Kawahara, K., Imoto, S., Sudo, T., Tanaka, F., Shibata, K., Suzuki, A., Komune, S., Miyano, S., & Mori, M. (2011). Long noncoding RNA HOTAIR regulates polycomb-dependent chromatin modification and is associated with poor prognosis in colorectal cancers. Cancer Research, 71(20), 6320–6326. https://doi.org/10.1158/0008-5472.CAN-11-1021

Kolas, N. K., Chapman, J. R., Nakada, S., Ylanko, J., Chahwan, R., Sweeney, F. D., Panier, S., Mendez, M., Wildenhain, J., Thomson, T. M., Pelletier, L., Jackson, S. P., & Durocher, D. (2007). Orchestration of the DNA-damage response by the RNF8 ubiquitin ligase. Science, 318(5856), 1637–1640. https://doi.org/10.1126/science.1150034

Kontou, P. I., Pavlopoulou, A., & Bagos, P. G. (2018). Methods of analysis and metaanalysis for identifying differentially expressed genes. Methods in Molecular Biology, 1793, 183–210. https://doi.org/10.1007/978-1-4939-7868-7_12

Kryston, T. B., Georgiev, A. B., Pissis, P., & Georgakilas, A. G. (2011). Role of oxidative stress and DNA damage in human carcinogenesis. Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis, 711(1–2), 193–201. https://doi.org/10.1016/j.mrfmmm.2010.12.016

Kutmon, M., Riutta, A., Nunes, N., Hanspers, K., Willighagen, E. L., Bohler, A., Mélius, J., Waagmeester, A., Sinha, S. R., Miller, R., Coort, S. L., Cirillo, E., Smeets, B., Evelo, C. T., & Pico, A. R. (2016). WikiPathways: capturing the full diversity of pathway knowledge. Nucleic Acids Research, 44(D1), D488–D494. https://doi.org/10.1093/NAR/GKV1024

Lamarche, B. J., Orazio, N. I., & Weitzman, M. D. (2010). The MRN complex in doublestrand break repair and telomere maintenance. FEBS Letters, 584(17), 3682–
3695. https://doi.org/10.1016/J.FEBSLET.2010.07.029

Larrea, A. A., Lujan, S. A., & Kunkel, T. A. (2010). SnapShot: DNA Mismatch Repair. Cell, 141(4). https://doi.org/10.1016/j.cell.2010.05.002

Lauber, C., Correia, N., Trumpp, A., Rieger, M. A., Dolnik, A., Bullinger, L., Roeder, I., & Seifert, M. (2020). Survival differences and associated molecular signatures of DNMT3A-mutant acute myeloid leukemia patients. Scientific Reports, 10(1). https://doi.org/10.1038/s41598-020-69691-8

Lee, M., Kim, H. J., Kim, S. W., Park, S. A., Chun, K. H., Cho, N. H., Song, Y. S., & Kim, Y. T. (2016). The long non-coding RNA HOTAIR increases tumour growth and invasion in cervical cancer by targeting the Notch pathway. Oncotarget, 7(28), 44558–44571. https://doi.org/10.18632/oncotarget.10065

Lee, N. K., Lee, J. H., Park, C. H., Yu, D., Lee, Y. C., Cheong, J. H., Noh, S. H., & Lee, S. K. (2014). Long non-coding RNA HOTAIR promotes carcinogenesis and invasion of gastric adenocarcinoma. Biochemical and Biophysical Research Communications, 451(2), 171–178. https://doi.org/10.1016/j.bbrc.2014.07.067

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. PLoS Medicine, 6(7). https://doi.org/10.1371/journal.pmed.1000100

Li, D. (2019). Statistical Methods for RNA Sequencing Data Analysis. Computational Biology. http://www.ncbi.nlm.nih.gov/pubmed/31815396

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics, 25(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Li, H. S., Zhou, Y. N., Li, L., Li, S. F., Long, D., Chen, X. L., Zhang, J. B., Feng, L., & Li, Y. P. (2019). HIF-1α protects against oxidative stress by directly targeting mitochondria. Redox Biology, 25. https://doi.org/10.1016/j.redox.2019.101109

Li, J., Wen, W., Zhao, S., Wang, J., Chen, J., Wang, Y., & Zhang, Q. (2015). Prognostic role of HOTAIR in four estrogen-dependent malignant tumors: a meta-analysis. OncoTargets and Therapy, 8, 1471–1482. https://doi.org/10.2147/OTT.S84687

Liu, L. C., Wang, Y. L., Lin, P. le, Zhang, X., Cheng, W. C., Liu, S. H., Chen, C. J., Hung, Y., Jan, C. I., Chang, L. C., Qi, X., Hsieh-Wilson, L. C., & Wang, S. C. (2019). Long noncoding RNA HOTAIR promotes invasion of breast cancer cells through chondroitin sulfotransferase CHST15. International Journal of Cancer, 145(9), 2478–2487. https://doi.org/10.1002/ijc.32319

Liu, T., Zhang, L., Joo, D., & Sun, S. C. (2017). NF-κB signaling in inflammation. Signal Transduction and Targeted Therapy, 2. https://doi.org/10.1038/sigtrans.2017.23

Liu, X. H., Sun, M., Nie, F. qi, Ge, Y. bin, Zhang, E. bao, Yin, D. dan, Kong, R., Xia, R., Lu, K. hua, Li, J. hai, De, W., Wang, K. ming, & Wang, Z. xia. (2014). Lnc RNA HOTAIR functions as a competing endogenous RNA to regulate HER2 expression by sponging miR-331-3p in gastric cancer. Molecular Cancer, 13(1). https://doi.org/10.1186/1476-4598-13-92

Logotheti, S., Marquardt, S., Richter, C., Hain, R. S., Murr, N., Takan, I., Pavlopoulou, A., & Pützer, B. M. (2020). Neural Networks Recapitulation by Cancer Cells Promotes Disease Progression: A Novel Role of p73 Isoforms in Cancer-Neuronal Crosstalk. Cancers, 12(12), 1–12. https://doi.org/10.3390/CANCERS12123789

Logotheti, S., Marquardt, S., Richter, C., Sophie Hain, R., Murr, N., Takan, I., Pavlopoulou, A., & Pützer, B. M. (2020). Neural Networks Recapitulation by Cancer Cells Promotes Disease Progression: A Novel Role of p73 Isoforms in Cancer-

Neuronal Crosstalk. Cancers, 12(12), 1–12. https://doi.org/10.3390/cancers12123789

Lowenberg, B., Downing, J. R., & Burnett, A. (1999). Acute Myeloid Leukemia. New England Journal of Medicine, 341(14), 1051–1062. https://doi.org/10.1056/NEJM199909303411407

Lu, X. X., Cao, L. Y., Chen, X., Xiao, J., Zou, Y., & Chen, Q. (2016). PTEN inhibits cell proliferation, promotes cell apoptosis, and induces cell cycle arrest via downregulating the PI3K/AKT/ hTERT Pathway in Lung Adenocarcinoma A549 Cells. BioMed Research International, 2016. https://doi.org/10.1155/2016/2476842

Lyer, R. R., Pluciennik, A., Burdett, V., & Modrich, P. L. (2006). DNA mismatch repair: Functions and mechanisms. Chemical Reviews, 106(2), 302–323. https://doi.org/10.1021/cr0404794

Ma, G., Wang, Q., Lv, C., Qiang, F., Hua, Q., Chu, H., Du, M., Tong, N., Jiang, Y., Wang, M., Zhang, Z., Wang, J., & Gong, W. (2015). The prognostic significance of HOTAIR for predicting clinical outcome in patients with digestive system tumors. Journal of Cancer Research and Clinical Oncology, 141(12), 2139–2145. https://doi.org/10.1007/s00432-015-1980-8

Malkin, D., Li, F. P., Strong, L. C., Fraumeni, J. F., Nelson, C. E., Kim, D. H., Kassel, J., Gryka, M. A., Bischoff, F. Z., Tainsky, M. A., & Friend, S. H. (1990). Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. Science, 250(4985), 1233–1238. https://doi.org/10.1126/science.1978757

Maréchal, A., & Zou, L. (2013a). DNA damage sensing by the ATM and ATR kinases. Cold Spring Harbor Perspectives in Biology, 5(9). https://doi.org/10.1101/cshperspect.a012716

Maréchal, A., & Zou, L. (2013b). DNA damage sensing by the ATM and ATR kinases. Cold Spring Harbor Perspectives in Biology, 5(9). https://doi.org/10.1101/cshperspect.a012716

Maréchal, A., & Zou, L. (2013c). DNA damage sensing by the ATM and ATR kinases. Cold Spring Harbor Perspectives in Biology, 5(9). https://doi.org/10.1101/cshperspect.a012716

Marioni, J. C., Mason, C. E., Mane, S. M., Stephens, M., & Gilad, Y. (2008). RNA-seq: An assessment of technical reproducibility and comparison with gene expression arrays. Genome Research, 18(9), 1509–1517. https://doi.org/10.1101/gr.079558.108

Marston, E., Weston, V., Jesson, J., Maina, E., McConville, C., Agathanggelou, A., Skowronska, A., Mapp, K., Sameith, K., Powell, J. E., Lawson, S., Kearns, P., Falciani, F., Taylor, M., & Stankovic, T. (2009a). Stratification of pediatric ALL by in vitro cellular responses to DNA double-strand breaks provides insight into the molecular mechanisms underlying clinical response. Blood, 113(1), 117–126. https://doi.org/10.1182/blood-2008-03-142950

Marston, E., Weston, V., Jesson, J., Maina, E., McConville, C., Agathanggelou, A., Skowronska, A., Mapp, K., Sameith, K., Powell, J. E., Lawson, S., Kearns, P., Falciani, F., Taylor, M., & Stankovic, T. (2009b). Stratification of pediatric ALL by in vitro cellular responses to DNA double-strand breaks provides insight into the molecular mechanisms underlying clinical response. Blood, 113(1), 117–126. https://doi.org/10.1182/blood-2008-03-142950

Martínez-Fernández, M., Feber, A., Dueñas, M., Segovia, C., Rubio, C., Fernandez, M., Villacampa, F., Duarte, J., López-Calderón, F. F., Gómez-Rodriguez, M. J., Castellano, D., Rodriguez-Peralto, J. L., de La Rosa, F., Beck, S., & Paramio, J. M. (2015). Analysis of the polycomb-related lncRNAs HOTAIR and ANRIL in bladder cancer. Clinical Epigenetics, 7(1). https://doi.org/10.1186/s13148-015-0141-x

Martin, L. M., Marples, B., Coffey, M., Lawler, M., Lynch, T. H., Hollywood, D., & Marignol, L. (2010). DNA mismatch repair and the DNA damage response to ionizing radiation: making sense of apparently conflicting data. Cancer Treatment Reviews, 36(7), 518–527. https://doi.org/10.1016/J.CTRV.2010.03.008

Ma, T., Keller, J. A., & Yu, X. (2011). RNF8-dependent histone ubiquitination during DNA damage response and spermatogenesis. Acta Biochimica et Biophysica Sinica, 43(5), 339–345. https://doi.org/10.1093/abbs/gmr016

Mavragani, I. v, Nikitaki, Z., Souli, M. P., Aziz, A., Nowsheen, S., Aziz, K., Rogakou, E., & Georgakilas, A. G. (2017). Complex DNA Damage: A Route to Radiation-Induced Genomic Instability and Carcinogenesis. Cancers, 9(7). https://doi.org/10.3390/cancers9070091

McFadden, E. J., & Hargrove, A. E. (2016). Biochemical Methods To Investigate lncRNA and the Influence of lncRNA:Protein Complexes on Chromatin. Biochemistry, 55(11), 1615–1630. https://doi.org/10.1021/ACS.BIOCHEM.5B01141

Metzker, M. L. (2010). Sequencing technologies - the next generation. Nature Reviews. Genetics, 11(1), 31–46. https://doi.org/10.1038/NRG2626

Mikkelsen, R. B., & Wardman, P. (2003). Biological chemistry of reactive oxygen and nitrogen and radiation-induced signal transduction mechanisms. Oncogene, 22(37 REV. ISS. 3), 5734–5754. https://doi.org/10.1038/sj.onc.1206663

Modrich, P. (1997). Strand-specific mismatch repair in mammalian cells. Journal of Biological Chemistry, 272(40), 24727–24730. https://doi.org/10.1074/jbc.272.40.24727

Multhoff, G., & Radons, J. (2012). Radiation, inflammation, and immune responses in cancer. Frontiers in Oncology, 2, 58. https://doi.org/10.3389/fonc.2012.00058

Nagano, T., & Fraser, P. (2011). No-nonsense functions for long noncoding RNAs. Cell, 145(2), 178–181. https://doi.org/10.1016/J.CELL.2011.03.014

Nakagawa, T., Endo, H., Yokoyama, M., Abe, J., Tamai, K., Tanaka, N., Sato, I., Takahashi, S., Kondo, T., & Satoh, K. (2013). Large noncoding RNA HOTAIR enhances aggressive biological behavior and is associated with short disease-free survival in human non-small cell lung cancer. Biochemical and Biophysical Research Communications, 436(2), 319–324. https://doi.org/10.1016/j.bbrc.2013.05.101

Nakano, K., & Vousden, K. H. (2001). PUMA, a novel proapoptotic gene, is induced by p53. Molecular Cell, 7(3), 683–694. https://doi.org/10.1016/S10972765(01)00214-3

Nikitaki, Z., Mavragani, I. v., Laskaratou, D. A., Gika, V., Moskvin, V. P., Theofilatos, K., Vougas, K., Stewart, R. D., & Georgakilas, A. G. (2016). Systemic mechanisms and effects of ionizing radiation: A new "old" paradigm of how the bystanders and distant can become the players. Seminars in Cancer Biology, 37–38, 77–95. https://doi.org/10.1016/J.SEMCANCER.2016.02.002

Nikitaki, Z., Michalopoulos, I., & Georgakilas, A. G. (2015). Molecular inhibitors of DNA repair: searching for the ultimate tumor killing weapon. Future Medicinal

Chemistry, 7(12), 1543–1558. https://doi.org/10.4155/FMC.15.95

Oda, E., Ohki, R., Murasawa, H., Nemoto, J., Shibue, T., Yamashita, T., Tokino, T., Taniguchi, T., & Tanaka, N. (2000). Noxa, a BH3-only member of the Bcl-2 family and candidate mediator of p53-induced apoptosis. Science, 288(5468), 1053–1058. https://doi.org/10.1126/science.288.5468.1053

Oh, E. J., Kim, S. H., Yang, W. I., Ko, Y. H., & Yoon, S. O. (2016). Long Non-coding RNA HOTAIR Expression in Diffuse Large B-Cell Lymphoma: In Relation to Polycomb Repressive Complex Pathway Proteins and H3K27 Trimethylation. Journal of Pathology and Translational Medicine, 50(5), 369–376. https://doi.org/10.4132/jptm.2016.06.06

Páez, D., Labonte, M. J., Bohanes, P., Zhang, W., Benhanim, L., Ning, Y., Wakatsuki, T., Loupakis, F., & Lenz, H. J. (2012). Cancer dormancy: a model of early dissemination and late cancer recurrence. Clinical Cancer Research : An Official Journal of the American Association for Cancer Research, 18(3), 645–653. https://doi.org/10.1158/1078-0432.CCR-11-2186

Patel, M., Nowsheen, S., Maraboyina, S., & Xia, F. (2020a). The role of poly(ADPribose) polymerase inhibitors in the treatment of cancer and methods to overcome resistance: a review. Cell & Bioscience, 10(1), 35. https://doi.org/10.1186/s13578-020-00390-7

Patel, M., Nowsheen, S., Maraboyina, S., & Xia, F. (2020b). The role of poly(ADPribose) polymerase inhibitors in the treatment of cancer and methods to overcome resistance: a review. Cell & Bioscience, 10(1), 35. https://doi.org/10.1186/s13578-020-00390-7

Patil, M., Pabla, N., & Dong, Z. (2013). Checkpoint kinase 1 in DNA damage response and cell cycle regulation. Cellular and Molecular Life Sciences, 70(21), 4009–4021. https://doi.org/10.1007/s00018-013-1307-3

Pavlopoulou, A., Bagos, P. G., Koutsandrea, V., & Georgakilas, A. G. (2017a). Molecular determinants of radiosensitivity in normal and tumor tissue: A bioinformatic approach. Cancer Letters, 403, 37–47. https://doi.org/10.1016/j.canlet.2017.05.023

Pavlopoulou, A., Bagos, P. G., Koutsandrea, V., & Georgakilas, A. G. (2017b). Molecular determinants of radiosensitivity in normal and tumor tissue: A bioinformatic

approach. Cancer Letters, 403, 37–47. https://doi.org/10.1016/j.canlet.2017.05.023

Pavlopoulou, A., Bagos, P. G., Koutsandrea, V., & Georgakilas, A. G. (2017c). Molecular determinants of radiosensitivity in normal and tumor tissue: A bioinformatic approach. Cancer Letters, 403, 37–47. https://doi.org/10.1016/j.canlet.2017.05.023

Pavlopoulou, A., Oktay, Y., Vougas, K., Louka, M., Vorgias, C. E., & Georgakilas, A. G. (2016a). Determinants of resistance to chemotherapy and ionizing radiation in breast cancer stem cells. Cancer Letters, 380(2), 485–493. https://doi.org/10.1016/J.CANLET.2016.07.018

Pavlopoulou, A., Oktay, Y., Vougas, K., Louka, M., Vorgias, C. E., & Georgakilas, A. G. (2016b). Determinants of resistance to chemotherapy and ionizing radiation in breast cancer stem cells. Cancer Letters, 380(2), 485–493. https://doi.org/10.1016/j.canlet.2016.07.018

Pavlopoulou, A., Oktay, Y., Vougas, K., Louka, M., Vorgias, C. E., & Georgakilas, A. G. (2016c). Determinants of resistance to chemotherapy and ionizing radiation in breast cancer stem cells. Cancer Letters, 380(2), 485–493. https://doi.org/10.1016/j.canlet.2016.07.018

Pavlopoulou, A., Spandidos, D. A., & Michalopoulos, I. (2015). Human cancer databases (review). Oncology Reports, 33(1), 3–18. https://doi.org/10.3892/OR.2014.3579

Pawlowski, J., & Kraft, A. S. (2000). Bax-induced apoptotic cell death. Proceedings of the National Academy of Sciences of the United States of America, 97(2), 529–531. https://doi.org/10.1073/pnas.97.2.529

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T. C., Mendell, J. T., & Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nature Biotechnology, 33(3), 290–295. https://doi.org/10.1038/nbt.3122

Pezuk, J. A., Brassesco, M. S., Morales, A. G., de Oliveira, J. C., de Oliveira, H. F., Scrideli, C. A., & Tone, L. G. (2013). Inhibition of polo-like kinase 1 induces cell cycle arrest and sensitizes glioblastoma cells to ionizing radiation. Cancer

Biotherapy and Radiopharmaceuticals, 28(7), 516–522. https://doi.org/10.1089/cbr.2012.1415

Porta, C., Paglino, C., & Mosca, A. (2014a). Targeting PI3K/Akt/mTOR signaling in cancer. Frontiers in Oncology, 4 APR. https://doi.org/10.3389/fonc.2014.00064

Porta, C., Paglino, C., & Mosca, A. (2014b). Targeting PI3K/Akt/mTOR signaling in cancer. Frontiers in Oncology, 4 APR. https://doi.org/10.3389/fonc.2014.00064

Porta, C., Paglino, C., & Mosca, A. (2014c). Targeting PI3K/Akt/mTOR signaling in cancer. Frontiers in Oncology, 4 APR. https://doi.org/10.3389/fonc.2014.00064

Porter, A. G., & Jänicke, R. U. (1999). Emerging roles of caspase-3 in apoptosis. Cell Death and Differentiation, 6(2), 99–104. https://doi.org/10.1038/sj.cdd.4400476

Pouysségur, J., Dayan, F., & Mazure, N. M. (2006). Hypoxia signalling in cancer and approaches to enforce tumour regression. Nature, 441(7092), 437–443. https://doi.org/10.1038/nature04871

Prensner, J. R., & Chinnaiyan, A. M. (2011a). The emergence of lncRNAs in cancer biology. Cancer Discovery, 1(5), 391–407. https://doi.org/10.1158/2159-8290.CD-11-0209

Prensner, J. R., & Chinnaiyan, A. M. (2011b). The emergence of lncRNAs in cancer biology. Cancer Discovery, 1(5), 391–407. https://doi.org/10.1158/21598290.CD-11-0209

Pulte, D., Gondos, A., & Brenner, H. (2009). Improvement in survival in younger patients with acute lymphoblastic leukemia from the 1980s to the early 21st century. Blood, 113(7), 1408–1411. https://doi.org/10.1182/blood-2008-06164863

Qu, X., Alsager, S., Zhuo, Y., & Shan, B. (2019). HOX transcript antisense RNA (HOTAIR) in cancer. Cancer Letters, 454, 90–97. https://doi.org/10.1016/j.canlet.2019.04.016

Ramasamy, A., Mondry, A., Holmes, C. C., & Altman, D. G. (2008). Key issues in conducting a meta-analysis of gene expression microarray datasets. PLoS Medicine, 5(9), 1320–1332. https://doi.org/10.1371/journal.pmed.0050184

Reifenberger, G., Weber, R. G., Riehmer, V., Kaulich, K., Willscher, E., Wirth, H., Gietzelt, J., Hentschel, B., Westphal, M., Simon, M., Schackert, G., Schramm, J.,

Matschke, J., Sabel, M. C., Gramatzki, D., Felsberg, J., Hartmann, C., Steinbach, J. P., Schlegel, U., … Loeffler, M. (2014a). Molecular characterization of long-term survivors of glioblastoma using genome- and transcriptome-wide profiling. International Journal of Cancer, 135(8), 1822–1831. https://doi.org/10.1002/ijc.28836

Reifenberger, G., Weber, R. G., Riehmer, V., Kaulich, K., Willscher, E., Wirth, H., Gietzelt, J., Hentschel, B., Westphal, M., Simon, M., Schackert, G., Schramm, J., Matschke, J., Sabel, M. C., Gramatzki, D., Felsberg, J., Hartmann, C., Steinbach, J. P., Schlegel, U., … Loeffler, M. (2014b). Molecular characterization of long-term survivors of glioblastoma using genome- and transcriptome-wide profiling. International Journal of Cancer, 135(8), 1822–1831. https://doi.org/10.1002/ijc.28836

Reifenberger, G., Weber, R. G., Riehmer, V., Kaulich, K., Willscher, E., Wirth, H., Gietzelt, J., Hentschel, B., Westphal, M., Simon, M., Schackert, G., Schramm, J., Matschke, J., Sabel, M. C., Gramatzki, D., Felsberg, J., Hartmann, C., Steinbach, J. P., Schlegel, U., … Loeffler, M. (2014c). Molecular characterization of long-term survivors of glioblastoma using genome- and transcriptome-wide profiling. International Journal of Cancer, 135(8), 1822–1831. https://doi.org/10.1002/ijc.28836

Rinn, J. L., & Chang, H. Y. (2012). Genome regulation by long noncoding RNAs. Annual Review of Biochemistry, 81, 145–166. https://doi.org/10.1146/ANNUREVBIOCHEM-051410-092902

Rinn, J. L., Kertesz, M., Wang, J. K., Squazzo, S. L., Xu, X., Brugmann, S. A., Goodnough, L. H., Helms, J. A., Farnham, P. J., Segal, E., & Chang, H. Y. (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell, 129(7), 1311–1323. https://doi.org/10.1016/J.CELL.2007.05.022

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics (Oxford, England), 26(1), 139–140. https://doi.org/10.1093/BIOINFORMATICS/BTP616

Robinson, M. D., & Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biology, 11(3). https://doi.org/10.1186/GB-2010-11-3-R25

Rouse, J., & Jackson, S. P. (2002). Interfaces between the detection, signaling, and repair of DNA damage. Science, 297(5581), 547–551. https://doi.org/10.1126/science.1074740

Roussa, E., & Krieglstein, K. (2004). GDNF promotes neuronal differentiation and dopaminergic development of mouse mesencephalic neurospheres. Neuroscience Letters, 361(1–3), 52–55. https://doi.org/10.1016/j.neulet.2003.12.106

Sakellaropoulos, T., Vougas, K., Narang, S., Koinis, F., Kotsinas, A., Polyzos, A., Moss, T. J., Piha-Paul, S., Zhou, H., Kardala, E., Damianidou, E., Alexopoulos, L. G., Aifantis, I., Townsend, P. A., Panayiotidis, M. I., Sfikakis, P., Bartek, J., Fitzgerald, R. C., Thanos, D., … Gorgoulis, V. G. (2019). A Deep Learning Framework for Predicting Response to Therapy in Cancer. Cell Reports, 29(11), 3367-3373.e4. https://doi.org/10.1016/j.celrep.2019.11.017

Saner, F. A. M., Herschtal, A., Nelson, B. H., deFazio, A., Goode, E. L., Ramus, S. J., Pandey, A., Beach, J. A., Fereday, S., Berchuck, A., Lheureux, S., Pearce, C. L., Pharoah, P. D., Pike, M. C., Garsed, D. W., & Bowtell, D. D. L. (2019a). Going to extremes: determinants of extraordinary response and survival in patients with cancer. Nature Reviews. Cancer, 19(6), 339–348. https://doi.org/10.1038/S41568-019-0145-5

Saner, F. A. M., Herschtal, A., Nelson, B. H., deFazio, A., Goode, E. L., Ramus, S. J., Pandey, A., Beach, J. A., Fereday, S., Berchuck, A., Lheureux, S., Pearce, C. L., Pharoah, P. D., Pike, M. C., Garsed, D. W., & Bowtell, D. D. L. (2019b). Going to extremes: determinants of extraordinary response and survival in patients with cancer. Nature Reviews Cancer, 19(6), 339–348. https://doi.org/10.1038/s41568019-0145-5

Saner, F. A. M., Herschtal, A., Nelson, B. H., deFazio, A., Goode, E. L., Ramus, S. J., Pandey, A., Beach, J. A., Fereday, S., Berchuck, A., Lheureux, S., Pearce, C. L., Pharoah, P. D., Pike, M. C., Garsed, D. W., & Bowtell, D. D. L. (2019c). Going to extremes: determinants of extraordinary response and survival in patients with

cancer. Nature Reviews Cancer, 19(6), 339–348. https://doi.org/10.1038/s41568019-0145-5

Saville, J. T., McDermott, B. K., Fletcher, J. M., & Fuller, M. (2019). Disease and subtype specific signatures enable precise diagnosis of the mucopolysaccharidoses. Genetics in Medicine : Official Journal of the American College of Medical Genetics, 21(3), 753–757. https://doi.org/10.1038/S41436018-0136-Z

Sayers, E. W., Beck, J., Bolton, E. E., Bourexis, D., Brister, J. R., Canese, K., Comeau, D. C., Funk, K., Kim, S., Klimke, W., Marchler-Bauer, A., Landrum, M., Lathrop, S., Lu, Z., Madden, T. L., O'Leary, N., Phan, L., Rangwala, S. H., Schneider, V. A., … Sherry, S. T. (2021). Database resources of the National Center for Biotechnology Information. Nucleic Acids Research, 49(D1), D10–D17. https://doi.org/10.1093/nar/gkaa892

Schmitt, A. M., & Chang, H. Y. (2016). Long Noncoding RNAs in Cancer Pathways. Cancer Cell, 29(4), 452–463. https://doi.org/10.1016/J.CCELL.2016.03.010

Schoenhals, J. E., Seyedin, S. N., Tang, C., Cortez, M. A., Niknam, S., Tsouko, E., Chang, J. Y., Hahn, S. M., & Welsh, J. W. (2016a). Preclinical Rationale and Clinical Considerations for Radiotherapy Plus Immunotherapy: Going Beyond Local Control. Cancer Journal (Sudbury, Mass.), 22(2), 130–137. https://doi.org/10.1097/PPO.0000000000000181

Schoenhals, J. E., Seyedin, S. N., Tang, C., Cortez, M. A., Niknam, S., Tsouko, E., Chang, J. Y., Hahn, S. M., & Welsh, J. W. (2016b). Preclinical rationale and clinical considerations for radiotherapy plus immunotherapy going beyond local control. Cancer Journal (United States), 22(2), 130–137. https://doi.org/10.1097/PPO.0000000000000181

Sekeres, M. A., Peterson, B., Dodge, R. K., Mayer, R. J., Moore, J. O., Lee, E. J., Kolitz, J., Baer, M. R., Schiffer, C. A., Carroll, A. J., Vardiman, J. W., Davey, F. R., Bloomfield, C. D., Larson, R. A., & Stone, R. M. (2004). Differences in prognostic factors and outcomes in African Americans and whites with acute myeloid leukemia. Blood, 103(11), 4036–4042. https://doi.org/10.1182/BLOOD-2003-093118

Shen, J., Hu, J., Wu, J., Luo, X., Li, Y., & Li, J. (2021a). Molecular characterization of long-term survivors of hepatocellular carcinoma. Aging, 13(5), 7517–7537. https://doi.org/10.18632/aging.202615

Shen, J., Hu, J., Wu, J., Luo, X., Li, Y., & Li, J. (2021b). Molecular characterization of long-term survivors of hepatocellular carcinoma. Aging, 13(5), 7517–7537. https://doi.org/10.18632/aging.202615

Shen, J., Hu, J., Wu, J., Luo, X., Li, Y., & Li, J. (2021c). Molecular characterization of long-term survivors of hepatocellular carcinoma. Aging, 13(5), 7517–7537. https://doi.org/10.18632/aging.202615

Shipley, J. L., & Butera, J. N. (2009). Acute myelogenous leukemia. Experimental Hematology, 37(6), 649–658. https://doi.org/10.1016/j.exphem.2009.04.002

Short, N. J., Rytting, M. E., & Cortes, J. E. (2018). Acute myeloid leukaemia. The Lancet, 392(10147), 593–606. https://doi.org/10.1016/S0140-6736(18)31041-9

Smith, D., Mann, D., & Yong, K. (2016). Cyclin D type does not influence cell cycle response to DNA damage caused by ionizing radiation in multiple myeloma tumours. British Journal of Haematology, 173(5), 693–704. https://doi.org/10.1111/bjh.13982

Sonenshein, G. E. (1997). Rel/NF-κB transcription factors and the control of apoptosis. Seminars in Cancer Biology, 8(2), 113–119. https://doi.org/10.1006/scbi.1997.0062

Song, Y., Wang, R., Li, L. W., Liu, X., Wang, Y. F., Wang, Q. X., & Zhang, Q. (2019). Long non-coding RNA HOTAIR mediates the switching of histone H3 lysine 27 acetylation to methylation to promote epithelial-to-mesenchymal transition in gastric cancer. International Journal of Oncology, 54(1), 77–86. https://doi.org/10.3892/ijo.2018.4625

Soubeyrand, S., Pope, L., Pakuts, B., & Haché, R. J. G. (2003). Threonines 2638/2647 in DNA-PK are essential for cellular resistance to ionizing radiation. Cancer Research, 63(6), 1198–1201.

Spratt, D. E., & Speers, C. (2019). RadiogX: A new preclinical tool to model intrinsic radiosensitivity. Cancer Research, 79(24), 6076–6078. https://doi.org/10.1158/0008-5472.CAN-19-3277

StataCorp: Stata statistical software: release 14 - Google Akademik. (n.d.). Retrieved June 24, 2022, from https://scholar.google.com/scholar_lookup?publication_year=2013&title=Stata+Statistical+Software%3A+Release+13

Stylianou, S., Clarke, R. B., & Brennan, K. (2006). Aberrant activation of Notch signaling in human breast cancer. Cancer Research, 66(3), 1517–1525. https://doi.org/10.1158/0008-5472.CAN-05-3054

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., & Mesirov, J. P. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences of the United States of America, 102(43), 15545–15550. https://doi.org/10.1073/pnas.0506580102

Sulai, N. H., & Tan, A. R. (2018). Development of poly(ADP-ribose) polymerase inhibitors in the treatment of BRCA-mutated breast cancer. Clinical Advances in Hematology and Oncology, 16(7), 491–501.

Sutherland, B. M., Bennett, P. v., Sidorkina, O., & Laval, J. (2000). Clustered DNA damages induced in isolated DNA and in human cells by low doses of ionizing radiation. Proceedings of the National Academy of Sciences of the United States of America, 97(1), 103–108. https://doi.org/10.1073/pnas.97.1.103

Swanton, E., Savory, P., Cosulich, S., Clarke, P., & Woodman, P. (1999). Bcl-2 regulates a caspase-3/caspase-2 apoptotic cascade in cytosolic extracts. Oncogene, 18(10), 1781–1787. https://doi.org/10.1038/sj.onc.1202490

Syed, V. (2016). TGF-β Signaling in Cancer. Journal of Cellular Biochemistry, 117(6), 1279–1287. https://doi.org/10.1002/jcb.25496

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N. T., Morris, J. H., Bork, P., Jensen, L. J., & von Mering, C. (2019a). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Research, 47(D1), D607–D613. https://doi.org/10.1093/NAR/GKY1131

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N. T., Morris, J. H., Bork, P., Jensen, L. J., & von Mering, C. (2019b). STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Research, 47(D1), D607–D613. https://doi.org/10.1093/nar/gky1131

Taalman, R. D. F. M., Jaspers, N. G. J., Scheres, J. M. J. C., de Wit, J., & Hustinx, T. W. J. (1983). Hypersensitivity to ionizing radiation, in vitro, in a new chromosomal breakage disorder, the Nijmegen Breakage Syndrome. Mutation Research, 112(1), 23–32. https://doi.org/10.1016/0167-8817(83)90021-4

Tang, C., Wang, X., Soh, H., Seyedin, S., Cortez, M. A. ngelica, Krishnan, S., Massarelli, E., Hong, D., Naing, A., Diab, A., Gomez, D., Ye, H., Heymach, J., Komaki, R., Allison, J. P., Sharma, P., & Welsh, J. W. (2014a). Combining radiation and immunotherapy: a new systemic therapy for solid tumors? Cancer Immunology Research, 2(9), 831–838. https://doi.org/10.1158/2326-6066.CIR-14-0069

Tang, C., Wang, X., Soh, H., Seyedin, S., Cortez, M. A. ngelica, Krishnan, S., Massarelli, E., Hong, D., Naing, A., Diab, A., Gomez, D., Ye, H., Heymach, J., Komaki, R., Allison, J. P., Sharma, P., & Welsh, J. W. (2014b). Combining radiation and immunotherapy: a new systemic therapy for solid tumors? Cancer Immunology Research, 2(9), 831–838. https://doi.org/10.1158/2326-6066.CIR-14-0069

Tang, Q., & Hann, S. S. (2018a). HOTAIR: An oncogenic long non-coding RNA in human cancer. Cellular Physiology and Biochemistry, 47(3), 893–913. https://doi.org/10.1159/000490131

Tang, Q., & Hann, S. S. (2018b). HOTAIR: An oncogenic long non-coding RNA in human cancer. Cellular Physiology and Biochemistry, 47(3), 893–913. https://doi.org/10.1159/000490131

Tang, Z., Li, C., Kang, B., Gao, G., Li, C., & Zhang, Z. (2017). GEPIA: A web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Research, 45(W1), W98–W102. https://doi.org/10.1093/nar/gkx247

Tan, S. K., Pastori, C., Penas, C., Komotar, R. J., Ivan, M. E., Wahlestedt, C., & Ayad, N. G. (2018). Serum long noncoding RNA HOTAIR as a novel diagnostic and

prognostic biomarker in glioblastoma multiforme. Molecular Cancer, 17(1). https://doi.org/10.1186/s12943-018-0822-0

Tierney, J. F., Stewart, L. A., Ghersi, D., Burdett, S., & Sydes, M. R. (2007). Practical methods for incorporating summary time-to-event data into meta-analysis. Trials, 8. https://doi.org/10.1186/1745-6215-8-16

Troiano, G., Alberto, V. C., Boldrup, L., Gu, X., Muzio, L. lo, Sgaramella, N., Wang, L., & Nylander, K. (2017). Expression of the long non-coding RNA HOTAIR as a prognostic factor in squamous cell carcinoma of the head and neck: a systematic review and meta-analysis. Oncotarget, 8(42), 73029–73036. https://doi.org/10.18632/ONCOTARGET.20373

Troiano, G., Caponio, V. C. A., Boldrup, L., Gu, X., Muzio, L. lo, Sgaramella, N., Wang, L., & Nylander, K. (2017). Expression of the long non-coding RNA HOTAIR as a prognostic factor in squamous cell carcinoma of the head and neck: a systematic review and meta-analysis. Oncotarget, 8(42), 73029–73036. https://doi.org/10.18632/oncotarget.20373

Tsai, M. C., Manor, O., Wan, Y., Mosammaparast, N., Wang, J. K., Lan, F., Shi, Y., Segal, E., & Chang, H. Y. (2010). Long noncoding RNA as modular scaffold of histone modification complexes. Science (New York, N.Y.), 329(5992), 689–693. https://doi.org/10.1126/SCIENCE.1192002

Unger, K. (2014). Integrative radiation systems biology. Radiation Oncology, 9(1). https://doi.org/10.1186/1748-717X-9-21

van den Bossche, J., Domen, A., Peeters, M., Deben, C., de Pauw, I., Jacobs, J., de Bruycker, S., Specenier, P., Pauwels, P., Vermorken, J. B., Lardon, F., & Wouters, A. (2019). Radiosensitization of Non-Small Cell Lung Cancer Cells by the Plk1 Inhibitor Volasertib Is Dependent on the p53 Status. Cancers, 11(12). https://doi.org/10.3390/cancers11121893

van 't Veer, L. J., & Bernards, R. (2008a). Enabling personalized cancer medicine through analysis of gene-expression patterns. Nature, 452(7187), 564–570. https://doi.org/10.1038/NATURE06915

van 't Veer, L. J., & Bernards, R. (2008b). Enabling personalized cancer medicine through analysis of gene-expression patterns. Nature, 452(7187), 564–570. https://doi.org/10.1038/NATURE06915

Vasu, S., Kohlschmidt, J., Mrózek, K., Eisfeld, A. K., Nicolet, D., Sterling, L. J., Becker, H., Metzeler, K. H., Papaioannou, D., Powell, B. L., Kolitz, J. E., Moore, J. O., Baer, M. R., Roboz, G. J., Stone, R. M., Byrd, J. C., Carroll, A. J., & Bloomfield, C. D. (2018). Ten-year outcome of patients with acute myeloid leukemia not treated with allogeneic transplantation in first complete remission. Blood Advances, 2(13), 1645–1650. https://doi.org/10.1182/bloodadvances.2017015222

Vougas, K., Sakellaropoulos, T., Kotsinas, A., Foukas, G. R. P., Ntargaras, A., Koinis, F., Polyzos, A., Myrianthopoulos, V., Zhou, H., Narang, S., Georgoulias, V., Alexopoulos, L., Aifantis, I., Townsend, P. A., Sfikakis, P., Fitzgerald, R., Thanos, D., Bartek, J., Petty, R., … Gorgoulis, V. G. (2019). Machine learning and data mining frameworks for predicting drug response in cancer: An overview and a novel in silico screening process based on association rule mining. Pharmacology and Therapeutics, 203. https://doi.org/10.1016/j.pharmthera.2019.107395

Wang, M., Chen, X., Chen, H., Zhang, X., Li, J., Gong, H., Shiyan, C., & Yang, F. (2015). RNF8 plays an important role in the radioresistance of human nasopharyngeal cancer cells in vitro. Oncology Reports, 34(1), 341–349. https://doi.org/10.3892/OR.2015.3958

Wang, W., Yang, L., Hu, L., Li, F., Ren, L., Yu, H., Liu, Y., Xia, L., Lei, H., Liao, Z., Zhou, F., Xie, C., & Zhou, Y. (2013). Inhibition of UBE2D3 Expression Attenuates Radiosensitivity of MCF-7 Human Breast Cancer Cells by Increasing hTERT Expression and Activity. PLoS ONE, 8(5). https://doi.org/10.1371/journal.pone.0064660

Wang, Y., Branicky, R., Noë, A., & Hekimi, S. (2018). Superoxide dismutases: Dual roles in controlling ROS damage and regulating ROS signaling. Journal of Cell Biology, 217(6), 1915–1928. https://doi.org/10.1083/jcb.201708007

Weinstein, J. N., Collisson, E. A., Mills, G. B., Shaw, K. R. M., Ozenberger, B. A., Ellrott, K., Sander, C., Stuart, J. M., Chang, K., Creighton, C. J., Davis, C., Donehower, L., Drummond, J., Wheeler, D., Ally, A., Balasundaram, M., Birol, I., Butterfield, Y. S. N., Chu, A., … Kling, T. (2013a). The Cancer Genome Atlas Pan-Cancer analysis project. Nature Genetics, 45(10), 1113–1120. https://doi.org/10.1038/NG.2764

Weinstein, J. N., Collisson, E. A., Mills, G. B., Shaw, K. R. M., Ozenberger, B. A., Ellrott,

K., Sander, C., Stuart, J. M., Chang, K., Creighton, C. J., Davis, C., Donehower, L., Drummond, J., Wheeler, D., Ally, A., Balasundaram, M., Birol, I., Butterfield, Y. S. N., Chu, A., … Kling, T. (2013b). The Cancer Genome Atlas Pan-Cancer analysis project. Nature Genetics, 45(10), 1113–1120. https://doi.org/10.1038/NG.2764

Weng, L. P., Brown, J. L., & Eng, C. (2001). PTEN induces apoptosis and cell cycle arrest through phosphoinositol-3-kinase/Akt-dependent and -independent pathways. Human Molecular Genetics, 10(3), 237–242. https://doi.org/10.1093/hmg/10.3.237

Woo, C. J., & Kingston, R. E. (2007). HOTAIR lifts noncoding RNAs to new levels. Cell, 129(7), 1257–1259. https://doi.org/10.1016/J.CELL.2007.06.014

Wu, Y., Zhang, L., Zhang, L., Wang, Y., Li, H., Ren, X., Wei, F., Yu, W. W., Liu, T., Wang, X., Zhou, X., Yu, J., & Hao, X. (2015a). Long non-coding RNA HOTAIR promotes tumor cell invasion and metastasis by recruiting EZH2 and repressing Ecadherin in oral squamous cell carcinoma. International Journal of Oncology, 46(6), 2586–2594. https://doi.org/10.3892/ijo.2015.2976

Wu, Y., Zhang, L., Zhang, L., Wang, Y., Li, H., Ren, X., Wei, F., Yu, W. W., Liu, T., Wang, X., Zhou, X., Yu, J., & Hao, X. (2015b). Long non-coding RNA HOTAIR promotes tumor cell invasion and metastasis by recruiting EZH2 and repressing Ecadherin in oral squamous cell carcinoma. International Journal of Oncology, 46(6), 2586–2594. https://doi.org/10.3892/ijo.2015.2976

Wu, Y., Zhang, L., Zhang, L., Wang, Y., Li, H., Ren, X., Wei, F., Yu, W. W., Liu, T., Wang, X., Zhou, X., Yu, J., & Hao, X. (2015c). Long non-coding RNA HOTAIR promotes tumor cell invasion and metastasis by recruiting EZH2 and repressing Ecadherin in oral squamous cell carcinoma. International Journal of Oncology, 46(6), 2586–2594. https://doi.org/10.3892/ijo.2015.2976

Wu, Y., Zhang, L., Zhang, L., Wang, Y., Li, H., Ren, X., Wei, F., Yu, W. W., Liu, T., Wang, X., Zhou, X., Yu, J., & Hao, X. (2015d). Long non-coding RNA HOTAIR promotes tumor cell invasion and metastasis by recruiting EZH2 and repressing Ecadherin in oral squamous cell carcinoma. International Journal of Oncology, 46(6), 2586–2594. https://doi.org/10.3892/ijo.2015.2976

Xing, C. yun, Hu, X. qu, Xie, F. yan, Yu, Z. jie, Li, H. ying, Bin-Zhou, Wu, J. bo, Tang, L. yuan, & Gao, S. meng. (2015). Long non-coding RNA HOTAIR modulates c-KIT expression through sponging miR-193a in acute myeloid leukemia. FEBS Letters, 589(15), 1981–1987. https://doi.org/10.1016/j.febslet.2015.04.061

Xiong, Y., Wang, T., Wang, M., Zhao, J., Li, X., Zhang, Z., Zhou, Y., Liu, J., Jia, L., & Han, Y. (2018). Long non-coding RNAs function as novel predictors and targets of non-small cell lung cancer: a systematic review and meta-analysis. Oncotarget, 9(13), 11377–11386. https://doi.org/10.18632/oncotarget.23994

Xu, F., & Zhang, J. (2017). Long non-coding RNA HOTAIR functions as miRNA sponge to promote the epithelial to mesenchymal transition in esophageal cancer. Biomedicine and Pharmacotherapy, 90, 888–896. https://doi.org/10.1016/j.biopha.2017.03.103

Yamamori, T., Yasui, H., Yamazumi, M., Wada, Y., Nakamura, Y., Nakamura, H., & Inanami, O. (2012). Ionizing radiation induces mitochondrial reactive oxygen species production accompanied by upregulation of mitochondrial electron transport chain function and mitochondrial content under control of the cell cycle checkpoint. Free Radical Biology and Medicine, 53(2), 260–270. https://doi.org/10.1016/j.freeradbiomed.2012.04.033

Yu, G., Wang, L. G., Han, Y., & He, Q. Y. (2012). ClusterProfiler: An R package for comparing biological themes among gene clusters. OMICS A Journal of Integrative Biology, 16(5), 284–287. https://doi.org/10.1089/omi.2011.0118

Zhang, B., Kirov, S., & Snoddy, J. (2005). WebGestalt: An integrated system for exploring gene sets in various biological contexts. Nucleic Acids Research, 33(SUPPL. 2). https://doi.org/10.1093/nar/gki475

Zhang, K., & Wang, H. (2015). [Cancer Genome Atlas Pan-cancer Analysis Project]. Zhongguo Fei Ai Za Zhi = Chinese Journal of Lung Cancer, 18(4), 219–223. https://doi.org/10.3779/J.ISSN.1009-3419.2015.04.02

Zhang, S., Chen, S., Yang, G., Gu, F., Li, M., Zhong, B., Hu, J., Hoffman, A., & Chen, M. (2014a). Long noncoding RNA HOTAIR as an independent prognostic marker in cancer: A meta-analysis. PLoS ONE, 9(8). https://doi.org/10.1371/journal.pone.0105538

Zhang, S., Chen, S., Yang, G., Gu, F., Li, M., Zhong, B., Hu, J., Hoffman, A., & Chen, M. (2014b). Long noncoding RNA HOTAIR as an independent prognostic marker in cancer: A meta-analysis. PLoS ONE, 9(8). https://doi.org/10.1371/journal.pone.0105538

Zhang, Y., Wang, L. J., Li, W. F., Zhang, X., & Yang, X. J. (2018). The prognostic value of HOTAIR for predicting long-Term prognosis of patients with gastrointestinal cancers. Medicine (United States), 97(26). https://doi.org/10.1097/MD.0000000000011139

Zhang, Y., Zhou, Y., Xu, T., Tian, W., Yang, C., Wang, X., Zhong, S., Ran, Q., Yang, H., & Zhu, S. (2018). Clinical value of long noncoding RNA HOTAIR as a novel biomarker in digestive cancers: A meta-analysis. Technology in Cancer Research and Treatment, 17. https://doi.org/10.1177/1533034618756783

Zhang, Z., Cheng, J., Wu, Y., Qiu, J., Sun, Y., & Tong, X. (2016). LncRNA HOTAIR controls the expression of Rab22a by sponging miR-373 in ovarian cancer. Molecular Medicine Reports, 14(3), 2465–2472. https://doi.org/10.3892/mmr.2016.5572

Zhao, S., Wong, Y. C., Yuan, S. S. F., Lin, Y. T., Hsu, H. C., Lin, S. C. J., Gerbino, E., Song, M. H., Zdzlenlcka, M. Z., Gatti, R. A., Shay, J. W., Ziv, Y., Shilch, Y., & Lee, E. Y. H. P. (2000). Functional link between ataxia-telangiectasia and Nijmegen breakage syndrome gene products. Nature, 405(6785), 473–477. https://doi.org/10.1038/35013083

Zheng, J., Xiao, X., Wu, C., Huang, J., Zhang, Y., Xie, M., Zhang, M., & Zhou, L. (2017a). The role of long non-coding RNA HOTAIR in the progression and development of laryngeal squamous cell carcinoma interacting with EZH2. Acta Oto-Laryngologica, 137(1), 90–98. https://doi.org/10.1080/00016489.2016.1214982

Zheng, J., Xiao, X., Wu, C., Huang, J., Zhang, Y., Xie, M., Zhang, M., & Zhou, L. (2017b). The role of long non-coding RNA HOTAIR in the progression and development of laryngeal squamous cell carcinoma interacting with EZH2. Acta Oto-Laryngologica, 137(1), 90–98. https://doi.org/10.1080/00016489.2016.1214982

Zheng, P., Yin, Z., Wu, Y., Xu, Y., Luo, Y., & Zhang, T. C. (2018). LncRNA HOTAIR promotes cell migration and invasion by regulating MKL1 via inhibition miR206 expression in HeLa cells. Cell Communication and Signaling, 16(1). https://doi.org/10.1186/s12964-018-0216-3

8. ADDENDA

8.1. References (Table 1)

1. Gupta, R.A., Shah, N., Wang, K.C., Kim, J., Horlings, H.M., Wong, D.J., Tsai, M.C., Hung, T., Argani, P., Rinn, J.L., Wang, Y., Brzoska, P., Kong, B., Li, R., West, R.B., van de Vijver, M.J., Sukumar, S., & Chang, H.Y. (2010). Long noncoding RNA HOTAIR reprograms chromatin state to promote cancer metastasis, Nature, 464 (7291), 1071-1076.
2. Geng, Y.J., Xie, S.L., Li, Q., Ma, J., & Wang, G.Y. (2011). Large intervening noncoding RNA HOTAIR is associated with hepatocellular carcinoma progression, J Int Med Res, 39 (6), 2119-2128.
3. Kogo, R., Shimamura, T., Mimori, K., Kawahara, K., Imoto, S., Sudo, T., Tanaka, F., Shibata, K., Suzuki, A., Komune, S., Miyano, S., & Mori, M. (2011). Long noncoding RNA HOTAIR regulates polycomb-dependent chromatin modification and is associated with poor prognosis in colorectal cancers, Cancer Res, 71 (20), 6320-6326.
4. Yang, Z., Zhou, L., Wu, L.M., Lai, M.C., Xie, H.Y., Zhang, F., & Zheng, S.S. (2011). Overexpression of long non-coding RNA HOTAIR predicts tumor recurrence in hepatocellular carcinoma patients following liver transplantation, Ann Surg Oncol, 18 (5), 1243-1250.
5. Lu, L., Zhu, G., Zhang, C., Deng, Q., Katsaros, D., Mayne, S.T., Risch, H.A., Mu, L., Canuto, E.M., Gregori, G., Benedetto, C., & Yu, H. (2012). Association of large noncoding RNA HOTAIR expression and its downstream intergenic CpG island methylation with survival in breast cancer, Breast Cancer Res Treat, 136

(3), 875-883.

6. Niinuma, T., Suzuki, H., Nojima, M., Nosho, K., Yamamoto, H., Takamaru, H., Yamamoto, E., Maruyama, R., Nobuoka, T., Miyazaki, Y., Nishida, T., Bamba, T., Kanda, T., Ajioka, Y., Taguchi, T., Okahara, S., Takahashi, H., Nishida, Y., Hosokawa, M., Hasegawa, T., Tokino, T., Hirata, K., Imai, K., Toyota, M., & Shinomura, Y. (2012). Upregulation of miR-196a and HOTAIR drive malignant character in gastrointestinal stromal tumors, Cancer Res, 72 (5), 1126-1136.
7. Chen, F.J., Sun, M., Li, S.Q., Wu, Q.Q., Ji, L., Liu, Z.L., Zhou, G.Z., Cao, G., Jin, L., Xie, H.W., Wang, C.M., Lv, J., De, W., Wu, M., & Cao, X.F. (2013). Upregulation of the long non-coding RNA HOTAIR promotes esophageal squamous cell carcinoma metastasis and poor prognosis, Mol Carcinog, 52 (11), 908-915.
8. Endo, H., Shiroki, T., Nakagawa, T., Yokoyama, M., Tamai, K., Yamanami, H., Fujiya, T., Sato, I., Yamaguchi, K., Tanaka, N., Iijima, K., Shimosegawa, T., Sugamura, K., & Satoh, K. (2013). Enhanced expression of long non-coding RNA HOTAIR is associated with the development of gastric cancer, PLoS One, 8 (10), e77070.
9. Ge, X.S., Ma, H.J., Zheng, X.H., Ruan, H.L., Liao, X.Y., Xue, W.Q., Chen, Y.B., Zhang, Y., & Jia, W.H. (2013). HOTAIR, a prognostic factor in esophageal squamous cell carcinoma, inhibits WIF-1 expression and activates Wnt pathway, Cancer Sci, 104 (12), 1675-1682.
10. Ishibashi, M., Kogo, R., Shibata, K., Sawada, G., Takahashi, Y., Kurashige, J., Akiyoshi, S., Sasaki, S., Iwaya, T., Sudo, T., Sugimachi, K., Mimori, K., Wakabayashi, G., & Mori, M. (2013). Clinical significance of the expression of long non-coding RNA HOTAIR in primary hepatocellular carcinoma, Oncol Rep, 29 (3), 946-950.
11. Li, D., Feng, J., Wu, T., Wang, Y., Sun, Y., Ren, J., & Liu, M. (2013). Long intergenic noncoding RNA HOTAIR is overexpressed and regulates PTEN methylation in laryngeal squamous cell carcinoma, Am J Pathol, 182 (1), 64-70.
12. Li, X., Wu, Z., Mei, Q., Li, X., Guo, M., Fu, X., & Han, W. (2013). Long noncoding RNA HOTAIR, a driver of malignancy, predicts negative prognosis and exhibits

oncogenic activity in oesophageal squamous cell carcinoma, Br J Cancer, 109 (8), 2266-2278.

13. Liu, X.H., Liu, Z.L., Sun, M., Liu, J., Wang, Z.X., & De, W. (2013). The long noncoding RNA HOTAIR indicates a poor prognosis and promotes metastasis in non-small cell lung cancer, BMC Cancer, 13, 464.
14. Lv, X.B., Lian, G.Y., Wang, H.R., Song, E., Yao, H., & Wang, M.H. (2013). Long noncoding RNA HOTAIR is a prognostic marker for esophageal squamous cell carcinoma progression and survival, PLoS One, 8 (5), e63516.
15. Nakagawa, T., Endo, H., Yokoyama, M., Abe, J., Tamai, K., Tanaka, N., Sato, I., Takahashi, S., Kondo, T., & Satoh, K. (2013). Large noncoding RNA HOTAIR enhances aggressive biological behavior and is associated with short diseasefree survival in human non-small cell lung cancer, Biochem Biophys Res Commun, 436 (2), 319-324.
16. Nie, Y., Liu, X., Qu, S., Song, E., Zou, H., & Gong, C. (2013). Long non-coding RNA HOTAIR is an independent prognostic marker for nasopharyngeal carcinoma progression and survival, Cancer Sci, 104 (4), 458-464.
17. Sorensen, K.P., Thomassen, M., Tan, Q., Bak, M., Cold, S., Burton, M., Larsen, M.J., & Kruse, T.A. (2013). Long non-coding RNA HOTAIR is an independent prognostic marker of metastasis in estrogen receptor-positive primary breast cancer, Breast Cancer Res Treat, 142 (3), 529-536.
18. Xu, Z.Y., Yu, Q.M., Du, Y.A., Yang, L.T., Dong, R.Z., Huang, L., Yu, P.F., & Cheng, X.D. (2013). Knockdown of long non-coding RNA HOTAIR suppresses tumor invasion and reverses epithelial-mesenchymal transition in gastric cancer, Int J Biol Sci, 9 (6), 587-597.
19. He, X., Bao, W., Li, X., Chen, Z., Che, Q., Wang, H., & Wan, X.P. (2014). The long non-coding RNA HOTAIR is upregulated in endometrial carcinoma and correlates with poor prognosis, Int J Mol Med, 33 (2), 325-332.
20. Huang, L., Liao, L.M., Liu, A.W., Wu, J.B., Cheng, X.L., Lin, J.X., & Zheng, M. (2014). Overexpression of long noncoding RNA HOTAIR predicts a poor prognosis in patients with cervical cancer, Arch Gynecol Obstet, 290 (4), 717723.

21. Lee, N.K., Lee, J.H., Park, C.H., Yu, D., Lee, Y.C., Cheong, J.H., Noh, S.H., & Lee, S.K. (2014). Long non-coding RNA HOTAIR promotes carcinogenesis and invasion of gastric adenocarcinoma, Biochem Biophys Res Commun, 451 (2), 171-178.
22. Liu, X.H., Sun, M., Nie, F.Q., Ge, Y.B., Zhang, E.B., Yin, D.D., Kong, R., Xia, R., Lu, K.H., Li, J.H., De, W., Wang, K.M., & Wang, Z.X. (2014). Lnc RNA HOTAIR functions as a competing endogenous RNA to regulate HER2 expression by sponging miR-331-3p in gastric cancer, Mol Cancer, 13, 92.
23. Okugawa, Y., Toiyama, Y., Hur, K., Toden, S., Saigusa, S., Tanaka, K., Inoue, Y., Mohri, Y., Kusunoki, M., Boland, C.R., & Goel, A. (2014). Metastasisassociated long non-coding RNA drives gastric cancer development and promotes peritoneal metastasis, Carcinogenesis, 35 (12), 2731-2739.
24. Qiu, J.J., Lin, Y.Y., Ye, L.C., Ding, J.X., Feng, W.W., Jin, H.Y., Zhang, Y., Li, Q., & Hua, K.Q. (2014). Overexpression of long non-coding RNA HOTAIR predicts poor patient prognosis and promotes tumor metastasis in epithelial ovarian cancer, Gynecol Oncol, 134 (1), 121-128.
25. Svoboda, M., Slyskova, J., Schneiderova, M., Makovicky, P., Bielik, L., Levy, M., Lipska, L., Hemmelova, B., Kala, Z., Protivankova, M., Vycital, O., Liska, V., Schwarzova, L., Vodickova, L., & Vodicka, P. (2014). HOTAIR long non-coding RNA is a negative prognostic factor not only in primary tumors, but also in the blood of colorectal cancer patients, Carcinogenesis, 35 (7), 1510-1515.
26. Wu, Z.H., Wang, X.L., Tang, H.M., Jiang, T., Chen, J., Lu, S., Qiu, G.Q., Peng, Z.H., & Yan, D.W. (2014). Long non-coding RNA HOTAIR is a powerful predictor of metastasis and poor prognosis and is associated with epithelial-mesenchymal transition in colon cancer, Oncol Rep, 32 (1), 395-402.
27. Yan, T.H., Lu, S.W., Huang, Y.Q., Que, G.B., Chen, J.H., Chen, Y.P., Zhang, H.B., Liang, X.L., & Jiang, J.H. (2014). Upregulation of the long noncoding RNA HOTAIR predicts recurrence in stage Ta/T1 bladder cancer, Tumour Biol, 35 (10), 10249-10257.
28. Heubach, J., Monsior, J., Deenen, R., Niegisch, G., Szarvas, T., Niedworok, C., Schulz, W.A., & Hoffmann, M.J. (2015). The long noncoding RNA HOTAIR has

tissue and cell type-dependent effects on HOX gene expression and phenotype of urothelial cancer cells, Mol Cancer, 14, 108.

29. Kim, H.J., Lee, D.W., Yim, G.W., Nam, E.J., Kim, S., Kim, S.W., & Kim, Y.T. (2015). Long non-coding RNA HOTAIR is associated with human cervical cancer progression, Int J Oncol, 46 (2), 521-530.
30. Liu, Y.W., Sun, M., Xia, R., Zhang, E.B., Liu, X.H., Zhang, Z.H., Xu, T.P., De, W., Liu, B.R., & Wang, Z.X. (2015). LincHOTAIR epigenetically silences miR34a by binding to PRC2 to promote the epithelial-to-mesenchymal transition in human gastric cancer, Cell Death Dis, 6, e1802.
31. Ma, G., Wang, Q., Lv, C., Qiang, F., Hua, Q., Chu, H., Du, M., Tong, N., Jiang, Y., Wang, M., Zhang, Z., Wang, J., & Gong, W. (2015). The prognostic significance of HOTAIR for predicting clinical outcome in patients with digestive system tumors, J Cancer Res Clin Oncol, 141 (12), 2139-2145.
32. Martinez-Fernandez, M., Feber, A., Duenas, M., Segovia, C., Rubio, C., Fernandez, M., Villacampa, F., Duarte, J., Lopez-Calderon, F.F., GomezRodriguez, M.J., Castellano, D., Rodriguez-Peralto, J.L., de la Rosa, F., Beck, S., & Paramio, J.M. (2015). Analysis of the Polycomb-related lncRNAs HOTAIR and ANRIL in bladder cancer, Clin Epigenetics, 7, 109.
33. Qiu, J.J., Wang, Y., Ding, J.X., Jin, H.Y., Yang, G., & Hua, K.Q. (2015). The long non-coding RNA HOTAIR promotes the proliferation of serous ovarian cancer cells through the regulation of cell cycle arrest and apoptosis, Exp Cell Res, 333 (2), 238-248.
34. Wu, J.Xie, H. (2015). Expression of long noncoding RNA-HOX transcript antisense intergenic RNA in oral squamous cell carcinoma and effect on cell growth, Tumour Biol, 36 (11), 8573-8578.
35. Wu, S., Zheng, C., Chen, S., Cai, X., Shi, Y., Lin, B., & Chen, Y. (2015). Overexpression of long non-coding RNA HOTAIR predicts a poor prognosis in patients with acute myeloid leukemia, Oncol Lett, 10 (4), 2410-2414.
36. Wu, Y., Zhang, L., Zhang, L., Wang, Y., Li, H., Ren, X., Wei, F., Yu, W., Liu, T., Wang, X., Zhou, X., Yu, J., & Hao, X. (2015). Long non-coding RNA HOTAIR promotes tumor cell invasion and metastasis by recruiting EZH2 and repressing E-cadherin in oral squamous cell carcinoma, Int J Oncol, 46 (6), 2586-2594.

37. Xing, C.Y., Hu, X.Q., Xie, F.Y., Yu, Z.J., Li, H.Y., Bin, Z., Wu, J.B., Tang, L.Y., & Gao, S.M. (2015). Long non-coding RNA HOTAIR modulates c-KIT expression through sponging miR-193a in acute myeloid leukemia, FEBS Lett, 589 (15), 1981-1987.
38. Zhang, Z.Z., Shen, Z.Y., Shen, Y.Y., Zhao, E.H., Wang, M., Wang, C.J., Cao, H., & Xu, J. (2015). HOTAIR Long Noncoding RNA Promotes Gastric Cancer Metastasis through Suppression of Poly r(C)-Binding Protein (PCBP) 1, Mol Cancer Ther, 14 (5), 1162-1170.
39. Zhao, W., Dong, S., Duan, B., Chen, P., Shi, L., Gao, H., & Qi, H. (2015). HOTAIR is a predictive and prognostic biomarker for patients with advanced gastric adenocarcinoma receiving fluorouracil and platinum combination chemotherapy, Am J Transl Res, 7 (7), 1295-1302.
40. Luczak, A., Supernat, A., Lapinska-Szumczyk, S., Jachimowicz, D., Majewska, H., Gulczynski, J., & Zaczek, A.J. (2016). HOTAIR in relation to epithelialmesenchymal transition and cancer stem cells in molecular subtypes of endometrial cancer, Int J Biol Markers, 31 (3), e245-251.
41. Luo, Z.F., Zhao, D., Li, X.Q., Cui, Y.X., Ma, N., Lu, C.X., Liu, M.Y., & Zhou, Y. (2016). Clinical significance of HOTAIR expression in colon cancer, World J Gastroenterol, 22 (22), 5254-5259.
42. Sun, J., Chu, H., Ji, J., Huo, G., Song, Q., & Zhang, X. (2016). Long non-coding RNA HOTAIR modulates HLA-G expression by absorbing miR-148a in human cervical cancer, Int J Oncol, 49 (3), 943-952.
43. Yan, Y., Han, J., Li, Z., Yang, H., Sui, Y., & Wang, M. (2016). Elevated RNA expression of long noncoding HOTAIR promotes cell proliferation and predicts a poor prognosis in patients with diffuse large B cell lymphoma, Mol Med Rep, 13 (6), 5125-5131.
44. Zhang, Y.Y., Huang, S.H., Zhou, H.R., Chen, C.J., Tian, L.H., & Shen, J.Z. (2016). Role of HOTAIR in the diagnosis and prognosis of acute leukemia, Oncol Rep, 36 (6), 3113-3122.
45. Chen, W.M., Chen, W.D., Jiang, X.M., Jia, X.F., Wang, H.M., Zhang, Q.J., Shu,

Y.Q., & Zhao, H.B. (2017). HOX transcript antisense intergenic RNA represses E-cadherin expression by binding to EZH2 in gastric cancer, World J Gastroenterol, 23 (33), 6100-6110.

46. Hu, G., Dong, B., Zhang, J., Zhai, W., Xie, T., Huang, B., Huang, C., Yao, X., Zheng, J., Che, J., & Xu, Y.F. (2017). The long noncoding RNA HOTAIR activates the Hippo pathway by directly binding to SAV1 in renal cell carcinoma, Oncotarget, 8 (35), 58654-58667.
47. Katayama, H., Tamai, K., Shibuya, R., Nakamura, M., Mochizuki, M., Yamaguchi, K., Kawamura, S., Tochigi, T., Sato, I., Okanishi, T., Sakurai, K., Fujibuchi, W., Arai, Y., & Satoh, K. (2017). Long non-coding RNA HOTAIR promotes cell migration by upregulating insulin growth factor-binding protein 2 in renal cell carcinoma, Sci Rep, 7 (1), 12016.
48. Luan, W., Li, R., Liu, L., Ni, X., Shi, Y., Xia, Y., Wang, J., Lu, F., & Xu, B. (2017). Long non-coding RNA HOTAIR acts as a competing endogenous RNA to promote malignant melanoma progression by sponging miR-152-3p, Oncotarget, 8 (49), 85401-85414.
49. Xu, F.Zhang, J. (2017). Long non-coding RNA HOTAIR functions as miRNA sponge to promote the epithelial to mesenchymal transition in esophageal cancer, Biomed Pharmacother, 90, 888-896.
50. Zhang, Y., Yu, S., Jiang, L., Wang, X., & Song, X. (2017). HOTAIR is a promising novel biomarker in patients with thyroid cancer, Exp Ther Med, 13 (5), 22742278.
51. Dong, X., He, X., Guan, A., Huang, W., Jia, H., Huang, Y., Chen, S., Zhang, Z., Gao, J., & Wang, H. (2018). Long non-coding RNA Hotair promotes gastric cancer progression via miR-217-GPC5 axis, Life Sci.
52. Huang, K.B., Zhang, S.P., Zhu, Y.J., Guo, C.H., Yang, M., Liu, J., Xia, L.G., & Zhang, J.F. (2018). Hotair mediates tumorigenesis through recruiting EZH2 in colorectal cancer, J Cell Biochem.
53. Xiao, Z., Qu, Z., Chen, Z., Fang, Z., Zhou, K., Huang, Z., Guo, X., & Zhang, Y. (2018). LncRNA HOTAIR is a Prognostic Biomarker for the Proliferation and Chemoresistance of Colorectal Cancer via MiR-203a-3p-Mediated Wnt/ssCatenin Signaling Pathway, Cell Physiol Biochem, 46 (3), 1275-1285.

8.2. Ethics Committee Report



İZMİR BİYOTIP VE GENOM MERKEZİ
GİRİŞİMSEL OLMAYAN ARAŞTIRMALAR ETİK KURULU (İBG-GOEK)
KARARI

Toplantı Tarihi	: 02/06/2022	**Toplantı Günü**	: Perşembe
Toplantı Sayısı	: 8	**Toplantı Saati**	: 10:30

Sayın None Athanasia PAVLOPOULOU,

2020-036 Protokol No'lu; sorumlusu olduğunuz "Sistem biyolojisi yaklaşımlarıyla kanser'in tanısı, prognozu ve radyoterapiye yanıt için gen ekspresyon imzalarının tanımlanması" başlıklı araştırmanın uygulanmasında etik açıdan sakınca olmadığına oy birliği ile karar verilmiştir.

Bilgilerinizi ve gereğini rica ederiz.

Prof. Dr. H. Alper BAĞRIYANIK **Başkan**	**Dr. Serap ERKEK ÖZHAN** **Başkan Yardımcısı**
Prof. Dr. Sedef AKGÜNGÖR **Üye**	**Prof. Dr. İnci ALACIOĞLU** **Üye**
Prof. Dr. Gülgün OKTAY **Üye**	**Prof. Dr. Hilmi ORHAN** **Üye**
Prof. Dr. Şermin GENÇ **Üye**	**Doç. Dr. Canan Aslı YILDIRIM** **Üye** (Katılmadı)
Doç. Dr. Can KÜÇÜK **Üye** (Katılmadı)	**Dr. Öğr. Üyesi Yavuz OKTAY** **Üye** (Katılmadı)

8.3. Curriculum Vitae

HALİL İBRAHİM TOY

Personal Information

Contact Information

Contact Address	[redacted]
Telephone	[redacted]
Email	[redacted]
Web Page	

Education Information

07 September 2017 - Now (4 Year 10 Month)
Doktora, Doktora DOKUZ EYLUL UNIVERSITESI, TURKIYE
İZMİR ULUSLARARASI BİYOTIP VE GENOM ENSTİTÜSÜ, MOLEKÜLER BİYOLOJİ VE GENETİK (DR) (İNGİLİZCE)
Diploma Number: -
Cumulative Grade Point Average: 3.1 / 4.0

20 September 2013 - 26 October 2016 (3 Year 2 Month)
Yüksek Lisans, Tezli Program, GAZİANTEP ÜNİVERSİTESİ, TÜRKİYE
SAĞLIK BİLİMLERİ ENSTİTÜSÜ, TEMEL SOLUNUM BİYOLOJİSİ (YL)
Diploma Number: 55530
Cumulative Grade Point Average: 3.36 / 4.0

12 September 2005 - 01 February 2013 (7 Year 5 Month)
Lisans, İkinci Öğretim, DİCLE ÜNİVERSİTESİ, TÜRKİYE
FEN FAKÜLTESİ, BİYOLOJİ PR. (İÖ)
Diploma Number: 752
Cumulative Grade Point Average: 65.44 / 100.0

Foreign Language Information

İNGİLİZCE (Reading: İyi, Writing: İyi, Speaking: İyi)

Scientific Technological Activity Fields

Scientific Technological Activity Field Information

Temel Bilimler -- Yaşam Bilimleri -- Biyoinformatik -- Biyolojik Veritabanları

Temel Bilimler -- Yaşam Bilimleri -- Moleküler Biyoloji ve Genetik -- Kanser Moleküler Biyolojisi

Temel Bilimler -- Yaşam Bilimleri -- Biyoinformatik -- Biyoenformasyon

R&D Competency

Articles	H. YİLMAZ, H. I. TOY, S. MARQUARDT, G. KARAKULAH, C. KUCUK, P. KONTOU, S. LOGOTHETİ & A. PAVLOPOULOU, In Silico Methods for the Identification of Diagnostic and Favorable Prognostic Markers in Acute Myeloid Leukemia, INTERNATIONAL JOURNAL OF MOLECULAR SCIENCES, 2021, 1422-0067, 22, 17.
	H. I. TOY, G. KARAKULAH, P. I. KONTOU, H. ALOTAİBİ, A. G. GEORGAKİLAS & A. PAVLOPOULOU, Investigating Molecular Determinants of Cancer Cell Resistance to Ionizing Radiation Through an Integrative Bioinformatics Approach, FRONTIERS IN CELL AND DEVELOPMENTAL BIOLOGY, 2021, 2296-634X, 9.
	H. I. TOY, D. OKMEN, P. KONTOU, A. G. GEORGAKİLAS & A. PAVLOPOULOU, HOTAIR as a Prognostic Predictor for Diverse Human Cancers: A Meta- and Bioinformatics Analysis, CANCERS, 2019, 2072-6694, 11, 6.
Conference Papers	H. İ. TOY, D. ÖKMEN & A. PAVLOPOULOU, HOTAIR as a prognostic predictor for multiple types of human cancers: a meta-analysis, Sözlü Sunum, 22. Türk Toraks Derneği Kongresi, 10 April 2019, 14 April 2019, 633 - 638.
	H. İ. TOY, D. TAŞDEMİR, S. ILHAN, Ö. DİKENSOY & H. BAYRAM, Göz Yaşartıcı Gazların KOAH'da Primer Bronş Epitel Hücre Toksisitesi ve İnflamatuar Parametrelere Etkisi, Sözlü Sunum, 20. Türk Toraks Derneği Kongresi, 05 April 2017, 09 April 2017, 98 - 98.
	H. İ. TOY, D. TAŞDEMİR, S. İLHAN, Ö. DİKENSOY & H. BAYRAM, BİBER GAZININ KOAH' TA PRİMER BRONŞ EPİTEL HÜCRE KÜLTÜR PERMABİLİTESİNE VE SİLYA TİTREŞİM FREKANSINA ETKİSİ, Sözlü Sunum, SANKO Üniversity Innovation in Medicine Summit-2, 05 May 2016, 07 May 2016.
	D. TAŞDEMİR, H. İ. TOY, M. ULAŞLI, S. KIRKBEŞ, E. TEMİZ ADIGÜZEL, S. DEMİRAL, S. ILHAN, S. ÖZTUZCU & H. BAYRAM, PRİMER BRONŞ EPİTEL HÜCRELERİ ÜZERİNE ETKİ EDEN DİZEL EKZOZ PARTİKÜLERİNİN MİKRORNA İFADESİ VE HAVAYOLU ANTİ-İNFLAMASYON YOLAKLARIYLA İLİŞKİLENDİRİLMESİ, Sözlü Sunum, SANKO Üniversity Innovation in Medicine Summit-2, 05 May 2016, 07 May 2016.
Thesis Supervised	H. İ. TOY (Thesis Author) , A. PAVLOPOULOU (Thesis Advisor) , G. KARAKÜLAH (Thesis Co-Advisor) , A systems biology approach for the identification of gene expression signature for cancer diagnosis, prognosis, and response to radiotherapy., DOKUZ EYLÜL ÜNİVERSİTESİ, TURKEY, 2020.

TÜBİTAK Scholarships and Fundings

BIDEB Supports	HALİL İBRAHİM TOY, Etkinlik Destekleri ve Eğitim Bursları Müdürlüğü, 2211-C Öncelikli Alanlara Yönelik Yurt İçi Doktora Burs Programı, Destekleniyor, 2020 - 2, 10/1/20 - 9/30/23.

Number of Panelist/Observer/Reporter Jobs

Refereeing/Panelist/Outer Consultant Count	ARDEB/BİDEB 0	TEYDEB 0	Total 0

8.4. Publications from the Thesis

1. Toy HI, Okmen D, Kontou PI, Georgakilas AG, Pavlopoulou A (2019) "HOTAIR as a Prognostic Predictor for Diverse Human Cancers: A Meta- and Bioinformatics Analysis" Cancers 11:778

2. Toy HI, Karakülah G, Kontou PI, Alotaibi H, Georgakilas AG, Pavlopoulou A (2021) "Investigating Molecular Determinants of Cancer Cell Resistance to Ionizing Radiation Through an Integrative Bioinformatics Approach" Frontiers in Cell and Developmental Biology 9: 770
3. Yılmaz H, Toy HI, Marquardt S, Karakülah G, Küçük C, Kontou PI, Logotheti S, Pavlopoulou A (2021) "In Silico Methods for the Identification of Diagnostic and Favorable Prognostic Markers in Acute Myeloid Leukemia." International Journal of Molecular Science 22:9601

cancers

MDPI

Article

HOTAIR as a Prognostic Predictor for Diverse Human Cancers: A Meta- and Bioinformatics Analysis

Halil Ibrahim Toy [1], Didem Okmen [1], Panagiota I. Kontou [2], Alexandros G. Georgakilas [3] and Athanasia Pavlopoulou [1,*]

1 Izmir International Biomedicine and Genome Institute, Dokuz Eylül University, Balcova 35340, Turkey; ibrahim.toy@msfr.ibg.edu.tr (H.I.T.); didem.okmen@msfr.ibg.edu.tr (D.O.)
2 Department of Computer Science and Biomedical Informatics, University of Thessaly, Lamia 35131, Greece; pankontou@gmail.com
3 DNA Damage Laboratory, Department of Physics, School of Applied Mathematical and Physical Sciences, Zografou Campus, National Technical University of Athens (NTUA), 15780 Athens, Greece; alexg@mail.ntua.gr
* Correspondence: athanasia.pavlopoulou@deu.edu.tr; Tel.: +90-232-412-6549

Received: 26 April 2019; Accepted: 1 June 2019; Published: 5 June 2019



Abstract: Several studies suggest that upregulated expression of the long non-coding RNA *HOX transcript antisense RNA* (*HOTAIR*) is a negative predictive biomarker for numerous cancers. Herein, we performed a meta-analysis to further investigate the prognostic value of *HOTAIR* expression in diverse human cancers. To this end, a systematic literature review was conducted in order to select scientific studies relevant to the association between *HOTAIR* expression and clinical outcomes, including overall survival (OS), recurrence-free survival (RFS)/disease-free survival (DFS), and progression-free survival (PFS)/metastasis-free survival (MFS) of cancer patients. Collectively, 53 eligible studies including a total of 4873 patients were enrolled in the current meta-analysis. Pooled hazard ratios (HRs) with their corresponding 95% confidence intervals (CIs) were calculated to assess the relationship between *HOTAIR* and cancer patients' survival. Elevated *HOTAIR* expression was found to be significantly associated with OS, RFS/DFS and PFS/MFS in diverse types of cancers. These findings were also corroborated by the results of bioinformatics analysis on overall survival. Therefore, based on our findings, *HOTAIR* could serve as a potential biomarker for the prediction of cancer patient survival in many different types of human cancers.

Keywords: *HOTAIR*; prognostic biomarker; survival; meta-analysis; cancer

1. Introduction

The long non-coding RNAs (lncRNAs) are non-protein-coding RNAs ≥ 200 bp in length, transcribed by RNA polymerase II. LncRNAs can be capped, polyadenylated and spliced, but they lack a functional open reading frame. It is estimated that approximately 27% (i.e., up to 60,000) of the annotated genes in the human genome encode lncRNAs, while the number of protein-coding genes ranges from 20,000 to 25,000 [1,2]. They are largely involved in a myriad of cellular functions, regulating gene expression at the transcriptional, post-transcriptional, and epigenetic level [1,3]. LncRNAs have emerged as critical components of cancer pathophysiology, being involved in one or more hallmarks of cancer, such as proliferation and metastasis [4,5]. They can act either as oncogenes or tumor suppressors, or indirectly through interaction with oncogenes and tumor suppressors, such as MYC proto-oncogene (MYC) and tumor protein p53 (TP53), respectively [4,5].

One of the most well-studied lncRNAs is *HOX transcript antisense RNA* (*HOTAIR*) which is located within the *HOMEOBOX C* (*HOXC*) gene cluster on chromosome 12q13.13 [6]. *HOTAIR* is 2158 bp

long and consists of six exons. *HOTAIR* orthologs are restricted to eutherian mammals [7]. *HOTAIR* is known to bind to the Polycomb Repressive Complex 2 (PRC2) and the histone H3K4 demethylase LSD1, and serves as a scaffold to assemble these regulators at the *HOXD* gene cluster, where it establishes a transcriptionally repressive chromatin structure, thereby resulting in epigenetic repression of the *HOXD* gene locus [8]. *HOTAIR* has been shown to function as an oncogene since its expression is dysregulated in multiple types of cancers, including breast, lung, liver, renal, hepatocellular, gastric, nasopharyngeal, cervical, colorectal, bladder, pancreatic cancer, as well as melanoma, leukemia, etc. [9–13]. Furthermore, *HOTAIR* is suggested to promote cancer progression and contribute largely to cancer cell invasion and metastasis [14–17]. The multifunctional *HOTAIR* is implicated in the different aspects of cancer pathophysiology by regulating gene expression at the transcriptional, post-transcriptional, and epigenetic level [14,18–20]. Of note, several studies suggest that *HOTAIR* expression is highly predictive of cancer patient survival rates in diverse cancer types [21–29].

Herein, we conducted a comprehensive and updated meta-analysis to further investigate the prognostic value of *HOTAIR* expression for cancer patients. The potential clinical applications of our findings are also discussed towards the prognostic application of *HOTAIR* to multiple and different types of cancers.

2. Results

2.1. Study Selection and Charasteristics of Eligible Studies

A total of 264 relevant published scientific studies were retrieved from the biomedical literature (up to 31 December 2018). According to the inclusion and exclusion criteria, 53 studies were ultimately included in this meta-analysis, as shown in Figure 1. The main characteristics of the included studies are summarized in Table 1, where the following information was recorded: first author's surname; year of publication; country of origin; type of cancer; follow-up period (in months); total number of patients; detection assay for *HOTAIR* expression; HR and the corresponding 95% CI for overall survival (OS), recurrence-free survival (RFS), disease-free survival (DFS), progression-free survival (PFS), metastasis-free survival (MFS); survival data extraction method; and specimen type. Collectively, 4873 patients from 55 cohorts between 2010 and 2018 were included. The included studies reported a follow-up period ranging from 36 to 276 months. The level of *HOTAIR* expression was measured with quantitative reverse transcription polymerase chain reaction (qRT-PCR) in all of the included studies, except one where *HOTAIR* expression was estimated by microarrays (Table 1).

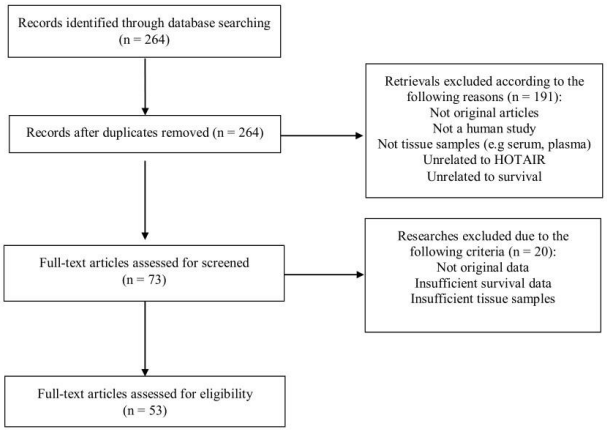


Figure 1. Flow chart of the process for study selection.

Table 1. Main characteristics of the studies included in the meta-analysis.

Author, Year	Country	Cancer	Max. Follow-Up (Months)	Sample	Case Number			OS		DFS/RFS		MFS/PFS		Assay Method	Data Extraction Method
					High Expression	Low Expression	Total	HR (95% CI)	*p*-Value	HR (95%CI)	*p*-Value	HR (95% CI)	*p*-Value		
Gupta, 2010 [14]	USA	Breast Cancer	240	Tissue	44	88	132	2.76 (1.45–3.3)	0.036	NM	NM	3.53 (2.78–4.89)	0.017	qRT-PCR	K-M
Geng, 2011 [30]	China	HCC	36	Tissue	NM	NM	50	NM	NM	2.24 (1.49–3.36)	0,049	NM	NM	qRT-PCR	K-M
Kogo, 2011 [31]	Japan	CRC	60	Tissue	20	80	100	5.62 (1.52–9.57)	0.008	NM	NM	NM	NM	qRT-PCR	reported
Yang, 2011 [32]	China	HCC	45	Tissue	32	28	60	NM	NM	3.56 (1.67–7.63)	0.001	NM	NM	qRT-PCR	reported
Lu, 2012 [33]	Italy	Breast Cancer	108	Tissue	NM	NM	336	0.43 (0.21–0.89)	0.022	0.47 (0.26–0.87)	0.016	NM	NM	qRT-PCR	reported
Niinuma, 2012 [34]	Japan	GIST	200	Tissue	11	28	39	3.8 (0.7–21.2)	0.123	NM	NM	NM	NM	qRT-PCR	reported
Chen, 2013 [24]	China	ESCC	60	Tissue	27	51	78	2.40 (1.35–4.28)	0.003	NM	NM	2.34 (1.22–4.48)	0.01	qRT-PCR	reported
Endo, 2013 [17]	Japan	IGC	68	Tissue	23	13	36	0.63 (0.34–1.86)	0.137	NM	NM	NM	NM	qRT-PCR	K-M
Endo, 2013 [17]	Japan	DGC	60	Tissue	20	12	32	3.08 (1.77–5.35)	<0.01	NM	NM	NM	NM	qRT-PCR	K-M
Ge, 2013 [35]	China	ESCC	100	Tissue	90	47	137	3.16 (1.53–6.52)	0.002	NM	NM	4.47 (1.99–10.06)	0.001	qRT-PCR	reported
Ishibashi, 2013 [36]	Japan	HCC	36	Tissue	13	51	64	2.84 (1.91–4.58)	0.041	NM	NM	NM	NM	qRT-PCR	K-M
Li, 2013 [37]	China	LSCC	60	Tissue	33	39	72	2.86 (1.15–7.07)	0.023	NM	NM	NM	NM	qRT-PCR	reported
Li, 2013 [38]	China	ESCC	60	Tissue	30	70	100	1.91 (1.06–3.99)	0.033	NM	NM	NM	NM	qRT-PCR	reported
Liu, 2013 [39]	China	NSCLC	60	Tissue	21	21	42	2.043 (0.91–4.58)	0.048	NM	NM	NM	NM	qRT-PCR	K-M
Lv, 2013 [40]	China	ESCC	70	Tissue	49	44	93	1.67 (1.02–2.79)	0.049	NM	NM	NM	NM	qRT-PCR	K-M
Nakagawa, 2013 [21]	Japan	NSCLC	50	Tissue	17	60	77	NM	NM	1.81 (1.09–3.74)	0,047	NM	NM	qRT-PCR	K-M
Nie, 2013 [41]	China	NPC	82	Tissue	91	69	160	1.9 (1.13–3.19)	0.012	1.41 (0.95–2.09)	0.47	1.92 (1.11–3.31)	0.018	qRT-PCR	K-M
Sorensen, 2013 [42]	Denmark	Breast Cancer	276	Tissue	79	85	164	NM	NM	NM	NM	1.75 (1.13–2.71)	0.012	Microarray	reported
Xu, 2013 [43]	China	Gastric cancer	75	Tissue	56	27	83	0.47 (0.22–0.99)	0.04	NM	NM	NM	NM	qRT-PCR	reported
He, 2014 [44]	China	EC	48	Tissue	62	83	145	3.04 (2.13–4.58)	0.026	NM	NM	NM	NM	qRT-PCR	K-M
Huang, 2014 [45]	China	Cervical cancer	55	Tissue	109	109	218	2.86 (1.26–6.49)	0.012	NM	NM	NM	NM	qRT-PCR	reported
Lee, 2014 [46]	Korea	Gastric cancer	48	Tissue	28	20	48	NM	NM	2.21 (0.53–9.16)	0.141	NM	NM	qRT-PCR	reported
Liu, 2014 [18]	China	Gastric cancer	48	Tissue	39	39	78	2.7 (1.36–4.34)	0.023	NM	NM	NM	NM	qRT-PCR	K-M
Okugawa, 2014 [47]	Japan	Gastric cancer	60	Tissue	77	73	150	1.77 (1.06–2.95)	0.028	NM	NM	NM	NM	qRT-PCR	reported
Qiu, 2014 [48]	China	EOC	79	Tissue	32	32	64	1.87 (1.04–5.31)	0.041	2.54 (1.18–5.45)	0.034	NM	NM	qRT-PCR	reported
Svoboda, 2014 [49]	Czech Republic	Colorectal cancer	54	Tissue	36	37	73	4.46 (1.02–19.79)	0.048	NM	NM	NM	NM	qRT-PCR	reported
Wu, 2014 [50]	China	Colon Cancer	72	Tissue	40	80	120	3.92 (1.23–12.50)	0.021	NM	NM	3.88 (1.37–10.98)	0.011	qRT-PCR	K-M
Yan, 2014 [51]	China	Bladder Cancer	60	Tissue	90	20	110	4.71 (2.89–8.71)	<0.001	NM	NM	NM	NM	qRT-PCR	reported
Heubach, 2015 [52]	Germany	UHC	200	Tissue	27	81	108	2.20 (1.23–3.93)	0.008	NM	NM	NM	NM	qRT-PCR	reported
Kim, 2015 [53]	Korea	Cervical cancer	60	Tissue	89	22	111	NM	NM	5.28 (1.01–27.74)	0,049	NM	NM	qRT-PCR	reported
Liu, 2015 [54]	China	Gastric cancer	40	Tissue	24	37	61	NM	NM	2.6 (1.74–3.89)	<0.001	NM	NM	qRT-PCR	K-M
Ma, 2015 [55]	China	Gastric cancer	60	Tissue	18	53	71	2.10 (1.10–4.03)	0.022	NM	NM	NM	NM	qRT-PCR	reported
Martinez-Fernandez, 2015 [56]	Spain	NMIBC	38	Tissue	17	16	33	NM	NM	NM	NM	1.86 (0.58–5.96)	0.296	qRT-PCR	K-M
Martinez-Fernandez, 2015 [56]	Spain	NMIBC	38	Tissue	30	33	63	NM	NM	3.78 (2.40–5.96)	<0.001	NM	NM	qRT-PCR	K-M
Qiu, 2015 [57]	China	SOC	96	Tissue	34	34	64	1.90 (1.01–3.56)	0.046	NM	NM	NM	NM	qRT-PCR	reported
Wu, 2015 [58]	China	OSCC	60	Tissue	25	25	50	1.91 (1.33–2.74)	<0.001	NM	NM	NM	NM	qRT-PCR	K-M
Wu, 2015 [59]	China	AML	40	Tissue	52	33	85	3.37 (0.99–8.31)	0.008	4.68 (2.81–7.79)	<0.001	NM	NM	qRT-PCR	reported
Wu, 2015 [16]	China	OSCC	96	Tissue	38	38	76	1.18 (0.68–2.84)	0.03	1.11 (0.78–2.54)	0.044	NM	NM	qRT-PCR	reported
Xing, 2015 [60]	China	AML	36	Tissue	68	68	136	2.03 (1.16–3.55)	0.007	0.61 (0.37–1.00)	0.034	NM	NM	qRT-PCR	reported
Zhang, 2015 [61]	China	Gastric cancer	45	Tissue	35	15	50	1.87 (1.46–2.1)	0.028	NM	NM	NM	NM	qRT-PCR	K-M
Zhao, 2015 [62]	China	Gastric cancer	65	Tissue	84	84	168	1.47 (1.04–2.06)	0.027	NM	NM	NM	NM	qRT-PCR	reported
Luczak, 2016 [63]	Poland	EC	96	Tissue	56	100	156	1.44 (0.81–3.19)	0.03	NM	NM	NM	NM	qRT-PCR	K-M
Luo, 2016 [64]	China	Colon cancer	70	Tissue	NM	NM	80	1.99 (1.4–2.8)	<0.001	NM	NM	NM	NM	qRT-PCR	K-M
Sun, 2016 [65]	China	Cervical cancer	50	Tissue	49	10	59	1.31 (0.79–2.26)	0.02	NM	NM	NM	NM	qRT-PCR	K-M

Table 1. *Cont.*

Author, Year	Country	Cancer	Max. Follow-Up (Months)	Sample	Case Number			OS		DFS/RFS		MFS/PFS		Assay Method	Data Extraction Method
					High Expression	Low Expression	Total	HR (95% CI)	*p*-Value	HR (95%CI)	*p*-Value	HR (95% CI)	*p*-Value		
Yan, 2016 [66]	China	DLBCL	120	Tissue	25	25	50	3.13 (1.22–8.04)	0.018	NM	NM	NM	NM	qRT-PCR	reported
Zhang, 2016 [67]	China	Acute leukemia	40	Tissue	19	77	96	2.41 (1.25–4.62)	0.005	NM	NM	NM	NM	qRT-PCR	K-M
Chen, 2017 [68]	China	Gastric cancer	62	Tissue	33	32	65	1.99 (1.06–3.77)	0.033	NM	NM	NM	NM	qRT-PCR	reported
Hu, 2017 [69]	China	RCC	50	Tissue	32	11	43	0.72 (0.20–2.55)	0.62	NM	NM	NM	NM	qRT-PCR	K-M
Katayama, 2017 [70]	Japan	RCC	100	Tissue	21	43	64	1.82 (1.06–3.88)	0.02	NM	NM	NM	NM	qRT-PCR	K-M
Luan, 2017 [71]	China	MM	60	Tissue	30	30	60	1.36 (0.79–2.83)	0.01	NM	NM	NM	NM	qRT-PCR	K-M
Xu, 2017 [72]	China	* EC	36	Tissue	20	20	40	2.69 (1.14–6.33)	0.032	NM	NM	NM	NM	qRT-PCR	K-M
Zhang, 2017 [73]	China	Thyroid cancer	60	Tissue	NM	NM	35	2.21 (1.38–3.54)	0.001	NM	NM	NM	NM	qRT-PCR	reported
Dong, 2018 [74]	China	Gastric cancer	60	Tissue	22	10	32	2.26 (0.74–6.89)	0.158	NM	NM	NM	NM	qRT-PCR	K-M
Huang, 2018 [75]	China	Colorectal cancer	110	Tissue	26	26	52	2.56 (0.91–7.35)	<0.01	NM	NM	NM	NM	qRT-PCR	reported
Xiao, 2018 [76]	China	Colorectal cancer	60	Tissue	52	52	104	1.45 (0.87–2.43)	0.041	NM	NM	NM	NM	qRT-PCR	K-M

Abbreviations: OS, overall survival; RFS, recurrence-free survival; DFS, disease-free survival; MFS, metastasis-free survival; PFS, progression-free survival; HR, hazard ratio; CI, confidence interval; qRT-PCR, quantitative reverse transcription polymerase chain reaction; NM: not mentioned; K-M, Kaplan-Meier plot; AML, acute myeloid leukemia; CRC, colorectal cancer; DGC, diffuse gastric cancer; DLBCL, diffuse large B cell lymphoma; ESCC, esophageal squamous cell carcinoma; EC, endometrial carcinoma; EOC, epithelial ovarian cancer; * EC, esophageal cancer; GIST, gastrointestinal stromal tumors; HCC, hepatocellular carcinoma; IGC, intestinal gastric cancer; LSCC, laryngeal squamous cell carcinoma; MM, malignant melanoma; NSCLC, non-small cell lung cancer; NPC, nasopharyngeal carcinoma; NMIBC, non-muscle-invasive bladder cancer; OSCC, oral squamous cell carcinoma; RCC, renal cell carcinoma; SOC, serous ovarian cancer; and UHC, urothelial carcinoma.

2.2. Association between High HOTAIR Expression and Overall Survival in Diverse Cancers

A total of 45 studies were included for overall survival (OS). We found a statistically significant relationship between elevated *HOTAIR* expression and poor OS (random-effects model: pooled HR = 2.00; 95% CI: 1.77–2.27; $p < 0.001$), with marginally moderate heterogeneity ($I^2 = 50.2\%$; $P_h < 0.001$) (Figure 2a). Subgroup analyses were performed based on the type of cancers, ethnic group, and data extraction method (Figure 3). When the studies were classified based on major cancer types (according to NCBI's medical subject headings (MeSH) [77]), a significant association was found between *HOTAIR* overexpression and poorer OS in solid cancers, such as gastrointestinal cancers (fixed-effects model: pooled HR = 1.96; 95% CI: 1.65–2.35; $p < 0.001$), liver cancers (fixed-effects model: pooled HR = 2.84; 95% CI: 1.83–4.40; $p < 0.001$), head and neck cancers (fixed-effects model: pooled HR = 1.93; 95% CI: 1.53–2.43; $p < 0.001$), and urogenital cancers (random-effects model: pooled HR = 2.11; 95% CI: 1.58–2.84; $p < 0.001$), as well as liquid cancers, including leukemia (fixed-effects model: pooled HR = 2.32; 95% CI: 1.56–3.44; $p < 0.001$) and lymphoma (fixed-effects model: pooled HR = 3.13; 95% CI: 1.22–8.04; $p < 0.001$). Of note, the heterogeneity was reduced significantly in the individual cancer types (Figure 3a). In the subgroup analysis based on ethnicity, a statistically significant worse OS was observed for Asians (fixed-effects model: pooled HR = 2.04; 95% CI: 1.81–2.31; $p < 0.001$). Regarding the Caucasian subgroup, despite the relatively high HR, the relationship cannot be considered robust because the *p*-value is slightly higher that the cutoff value (random-effects model; pooled HR = 1.65; 95% CI: 0.82–3.33; $p = 0.077$) (Figure 3b). In stratified analysis, according to data extraction method, *HOTAIR* was found to have a significant prognostic value irrespectively of the data source. that is, the HR reported in the articles (random-effects model: pooled HR = 2.05; 95% CI: 1.64–2.57; $p < 0.001$) or extracted from the survival curves (fixed-effects model: pooled HR = 2.01; 95% CI: 1.75–2.30; $p < 0.001$) (Figure 3c).

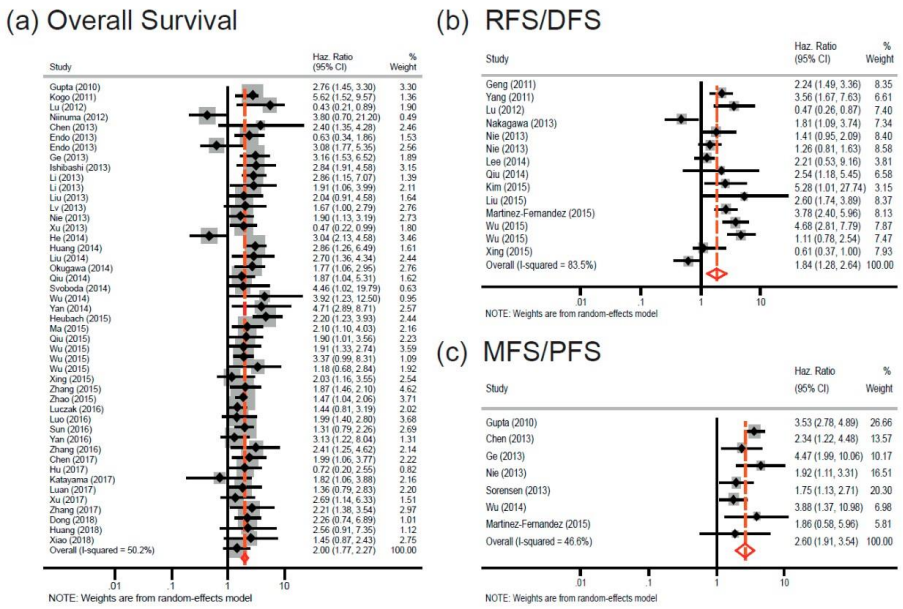


Figure 2. Forest plots of combined analyses on the association of survival with *HOTAIR* expression. (**a**) Forest plot of OS analysis, (**b**) forest plot of RFS/DFS analysis, and (**c**) forest plot of MFS/PFS analysis. Abbreviations: HR, Hazard ratio; OS, overall survival; RFS, recurrence-free survival; DFS, disease-free survival; MFS, metastasis-free survival; and PFS, progression-free survival.

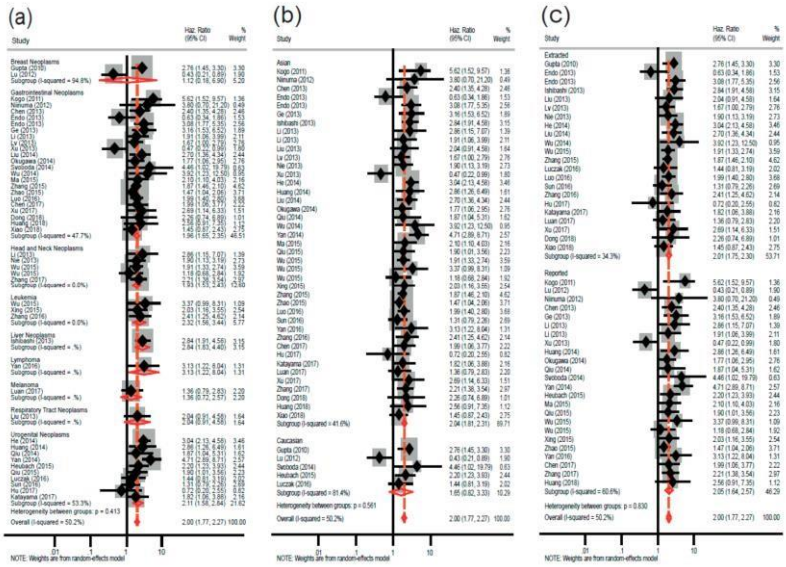


Figure 3. Forest plots of combined analyses for overall survival (OS) associated with *HOTAIR* expression in different groups. (**a**) Forest plot for different types of cancers, (**b**) forest plot for different ethnic groups, and (**c**) forest plot for different data extraction methods.

2.3. HOTAIR Overexpression Is Associated with Cancer Recurrence and Progression

To investigate the relationship between *HOTAIR* expression and cancer recurrence or relapse, the recurrence-free survival (RFS) and disease-free survival (DFS) studies were combined; collectively accounting for 14 studies. Increased *HOTAIR* expression was found to be strongly related to cancer recurrence (pooled HR = 1.84; 95% CI = 1.28–2.64; $p = 0.001$). A random-effects model was applied because of the high heterogeneity ($I^2 = 83.5\%$; $P_h < 0.001$) across studies (Figure 2b).

Furthermore, there are seven studies for combined metastasis-free survival (MFS) and progression-free survival (PFS). Of importance, high *HOTAIR* expression was predicted to be associated significantly with worse MFS/PFS (pooled HR = 2.60; 95% CI: 1.91–3.54; $p < 0.001$). A fixed-effects model was used because of the relatively low heterogeneity ($I^2 = 46.6\%$; $P_h = 0.081$) (Figure 2c).

2.4. Publication Bias

Publication bias was detected by Begg's funnel plot and Egger's test. There was no obvious asymmetry in Begg's funnel plots of OS, RFS/DFS, and MFS/PFS (Figure 4). Additionally, the *p*-values of Egger's tests were all greater than 0.05, indicating no potential publication bias (OS: $p = 0.73$; RFS/DFS: $p = 0.70$; MFS/PFS: $p = 0.64$).

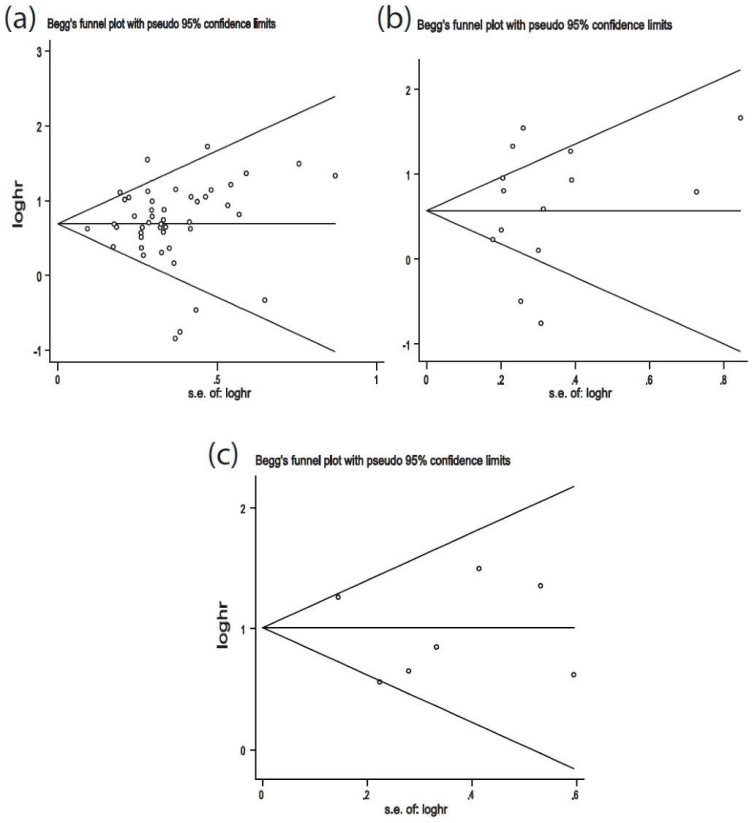


Figure 4. Begg's funnel plots of publication bias. (**a**) Begg's funnel plot of publication bias for OS; (**b**) Begg's funnel plot of publication bias for RFS/DFS; (**c**) Begg's funnel plot of publication bias for MFS/PFS. Each circle represents a separate study.

2.5. Sensitivity Analysis

Sensitivity analyses did not indicate alterations in the results due to the inclusion of any individual study (Figure 5), that is, no single study affected the pooled HR or 95% CI.

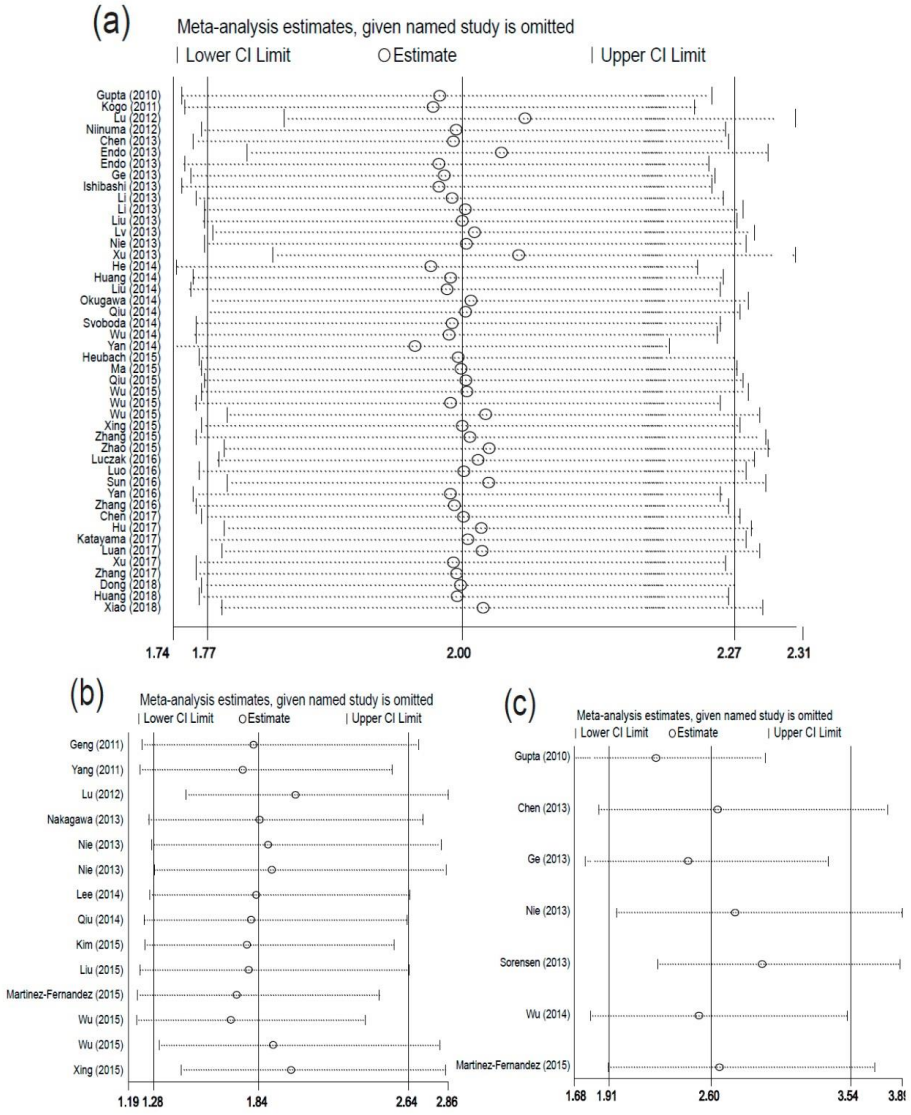


Figure 5. Sensitivity analysis of each eligible study. (**a**) OS individual studies, (**b**) RFS/DFS individual studies and (**c**) MFS/PFS individual studies.

2.6. TCGA-Derived Survival Curves

To further the clinical relevance of our work and *HOTAIR* importance, we explored the possibility for any association of the *HOTAIR* expression to overall cancer survival. It was found that *HOTAIR* overexpression was significantly associated with worse OS in adrenocortical carcinoma (ACC), mesothelioma (MESO), and glioblastoma multiforme (GBM) (Figure S1).

3. Discussion

HOTAIR exhibits pro-oncogenic activity since it has been shown to be overexpressed in numerous cancers and be implicated in several hallmarks of cancer, such as cellular proliferation, inhibition of apoptosis, genomic instability, angiogenesis, invasion, and metastasis [19,20].

In the current study, an updated, comprehensive meta-analysis on the prognostic value of *HOTAIR* in various human cancers was presented. By applying stringent inclusion and exclusion criteria, we included 53 eligible studies, a relatively large number necessary for a meta-analysis to be considered robust. Previous meta-analyses on the association of *HOTAIR* with clinical outcome have included a rather limited number of studies with inconclusive and inconsistent findings [28,29]. Other related studies have focused on certain types of cancers, such as head and neck squamous cell carcinoma [22] or digestive system cancers [55,78,79].

In the present study, we showed that there is a statistically significant relationship between elevated *HOTAIR* expression and poor OS. In the subgroup analysis, based on cancer type, *HOTAIR* was shown to be a significant predictor for worse prognosis for a variety of cancers, including solid cancers, such as urological cancers, head and neck neoplasms, cancers of the digestive system, and several female cancers (e.g., cervical, ovarian, and endometrial cancers), as well as the blood cancers, lymphoma and leukemia. Moreover, we complemented the findings from meta-analysis and further strengthened our hypotheses with survival information from other types of cancers, for which there were not any available eligible studies, retrieved from TCGA. It was found that there is, also, a strong relationship between *HOTAIR* overexpression and poor OS in neoplasms of the adrenal cortex, mesothelial neoplasms, and neuroepithelial tumors.

Taken together, the above findings lead to the suggestion that similar *HOTAIR*-mediated pathways might be implicated both in solid and liquid cancers [13]. In particular, in several solid tumors, *HOTAIR* has been shown to exert its oncogenic and metastatic potential by mediating a repressive chromatin structure through the recruitment of histone-modifying or chromatin-remodeling complexes, such as PRC2 [14,16,31]. For example, *HOTAIR* can promote pancreatic cancer cell proliferation by suppressing the expression of miR-663b via remodeling the chromatin structure within the miR-663b promoter [80]. In a recent study, *HOTAIR* was also found to recruit PRC2 to catalyze H3K27 trimethylation to transcriptionally repress *E-cadherin* and promote EMT in gastric cancer [81]. Similarly, high expression levels of *HOTAIR* and PRC2 proteins (H3K27 methylase EZH2, SUZ12, and EED) were found to be positively correlated with lymphomagenesis [82]. In addition, *HOTAIR*, through miRNA sponging, contributes to carcinogenesis both in blood [60] and solid tumors [83,84]. However, there is a rather limited number of studies available on major cancers, such as breast neoplasms and respiratory tract cancers. Thus, more clinical trials on these cancers would enable us to better assess the relationship between *HOTAIR* expression and cancer patients' survival.

A positive correlation between *HOTAIR* and *CDKN1A* (*p21*) expression levels was also found (Figure S2), suggesting a possible functional and/or physical association between *HOTAIR* and *CDKN1A* (*p21*) in cancer pathophysiology. From a clinical perspective, there is an emerging role of *CDKNIA* (*p21*), especially in cases where p53 is mutated like in many different solid tumors. The role of *p21* has been extensively viewed as an indicator of wildtype p53 activity [85]. However, recent evidence suggests that upregulated *p21* can also act as an oncogenic factor in a p53-deficient environment, thereby driving a subset of atypical cancerous cells to more chemoresistant and aggressive phenotypes [86]. Therefore, we cannot exclude a possible mechanistic association between *HOTAIR* and *p21* towards the negative regulation of target genes and a potential role in OS. Interestingly, recent studies have

shown that *HOTAIR* expression was significantly higher in non-small-cell lung cancer (NSCLC) tissues compared to the adjacent normal tissues, and *HOTAIR* was negatively associated with p53 functionality rather than *p53* expression [87]. In addition, *HOTAIR*, *p21*, and *p53* mRNA expression in doxorubicin- or γ rays-treated oral squamous cell carcinoma (OSCC) cells was up-regulated, indicating that the DNA damage response includes *HOTAIR* upregulation and may be closely connected to *p53* and *p21* expression and/or functionality [88].

To investigate any possible effect of the genetic background and environment on the overall HRs, analyses were conducted based on the ethnic background of the participants. *HOTAIR* was found to be a powerful negative prediction biomarker for Asians. In the case of Caucasians, there was a link between *HOTAIR* overexpression and poor OS, albeit with moderate statistical significance; this is probably due to the relatively low number of available studies on patients of Caucasian origin. There were not, also, any available studies for other major ethnic groups, such as Africans or Indians, which would have further allowed us to estimate the influence of the genetic make-up on the association between *HOTAIR* and clinical outcome. The overall effect was similar in the stratified analysis according to data source, that is, the estimated HR reported in the articles or extrapolated from survival curves.

Therefore, high *HOTAIR* expression can predict an unfavorable clinical outcome in different types of cancers and possibly ethnic groups using different extraction methods. Notably, elevated expression of *HOTAIR* and prognosis in cancer patients is not particularly affected either by cancer type or even the patients' genetic background.

HOTAIR was found to be a poor predictor for both cancer recurrence and progression. The similar outcomes suggest that there are similar *HOTAIR*-dependent mechanisms underlying these two phenomena. In particular, *HOTAIR* was shown to mediate recurrence and progression in bladder cancer via the histone methyltransferase EZH2 [56]. Similarly, enhanced *HOTAIR* expression was found to be associated both with progression and tumor recurrence in hepatocellular carcinoma by regulating the Wnt/β-catenin signal transduction pathway [89].

HOTAIR has been demonstrated to promote tumor cell invasion and metastasis by modulating epithelial-to-mesenchymal transition (EMT) [16,46,90]. Enhanced *HOTAIR* expression has also been shown to promote metastasis and invasion through different mechanisms including genome-wide re-targeting of PRC2 and subsequent epigenetic silencing of multiple anti-metastatic genes [14], inhibition of the expression of the metastasis suppressor gene *E-cadherin* by recruiting the histone methyltransferase of PRC2, EZH2 [16,90], targeting of Notch/Wnt signaling pathway-associated genes [91], and upregulating chondroitin sulfotransferase CHST15 [92], etc. *HOTAIR* also promotes invasion and migration by acting as a 'miRNA sponge', through targeting the corresponding miRNAs in the miR-1/CCND2 [93], miR-148a/SNAIL2 [72], and miR-23b/MAPK1 [94] axes.

Heterogeneity was observed within the forest plots of OS and RFS/DFS, suggesting that HRs vary across studies. For this reason, the random-effects model was applied, where the overall HR was estimated based on the weighted average of the HRs of the individual studies. Given that the overall effect for OS and RFS/DFS was not affected by any single study, according to sensitivity analyses, we could suggest that, despite heterogeneity, the pooled HR can be considered quite reliable and representative.

Moreover, potential publication bias was not detected in the present meta-analysis, probably due to the sufficient representation of eligible studies in this meta-analysis.

4. Materials and Methods

4.1. Search Strategy and Study Eligibility Criteria

This systematic review and meta-analysis was conducted by following strictly the PRISMA (preferred reporting items for systematic reviews and meta-analyses) guidelines [95].

The bibliographic database PubMed/MEDLINE [96] was manually searched for published scientific studies on the associations between *HOTAIR* expression and prognosis in different types of cancers by

using combinations of the relevant keywords: ("HOTAIR" OR "HOX transcript antisense RNA" or "HOXC cluster antisense RNA 4" or "HOXC-AS4" OR "HOXC11-AS1") and ("cancer" or "carcinoma" or "tumor" or "neoplasm" or "malignancy") and ("prognosis" or "survival" or "outcome" or "mortality" or "death"). The studies had to fulfill the following inclusion criteria so as to be considered eligible: (1) studies of human clinical trials, (2) studies including more than 30 patients in total, (3) the correlation between *HOTAIR* expression and cancer patients' survival was estimated, (4) availability of HR and 95% confidence interval (CI) or survival curves or sufficient data to calculate HR and 95% CI, (5) quantitative measurement (e.g., qPCR) of *HOTAIR* expression in cancers was included, and (6) studies published in English. Accordingly, the studies were excluded on the basis of the following exclusion criteria: (1) laboratory studies on animal models or cell lines; (2) reviews, meta-analyses, editorials, case reports, commentaries, unpublished data; (3) lack of sufficient data to estimate HR and 95% CI; and (4) samples other than tissue (e.g., blood, serum).

4.2. Study Selection, Data Extraction, and Quality Assessment

All potential studies were independently retrieved from the literature by two of the authors (H.I.T. and D.O.). Quality assessment of the studies was performed by H.I.T. and D.O. independently. Any disagreement was resolved by a third investigator (A.P.). Relevant data were extracted from the included studies and recorded into an ad hoc Excel worksheet. In the case that the HR was not reported in the corresponding article, the data were extracted from the graphical survival plots (i.e., Kaplan-Meier curves) by using the Engauge Digitizer v10.11 software, as previously described [97].

4.3. Statistical Analyses

All statistical analyses were performed with STATA statistical software version 13.0 (Stata Corporation, College Station, TX, USA) and Microsoft Excel. The heterogeneity among the included studies was estimated by Higgins I-squared (I^2) statistic as follows: $I^2 < 25\%$; no heterogeneity; $25\% < I^2 < 50\%$: low heterogeneity; $50\% < I^2 < 75\%$: moderate heterogeneity; $I^2 > 75\%$ high heterogeneity [98,99]. In the case of statistically significant heterogeneity ($I^2 > 50\%$ and $P_h < 0.05$), a random-effect model was applied, otherwise a fixed-effect model [100,101] was used. Sensitivity analysis was performed by consecutive omission of individual studies to verify the consistency of outcomes. Potential publication bias was detected by Begg's funnel plot [102] and Egger's test [103]; a *p*-value less than 0.05 was indicative of statistically significant publication bias.

4.4. Bioinformatics Analysis

4.4.1. Survival Analysis

Overall survival curves for different types of cancers were retrieved through the online tool GEPIA (Gene Expression Profiling Interactive Analysis) [104], which provides survival analysis based on datasets obtained from The Cancer Genome Atlas (TCGA) (https://tcga-data.nci.nih.gov).

4.4.2. Correlation Analysis

Correlation analysis between gene expression levels was performed through the web-based tool GEPIA [104] which analyzes gene expression based on RNA sequencing (RNA-Seq) data from TCGA.

5. Conclusions

In this study, we have performed a meta-analysis complemented with bioinformatics analyses towards investigating the prognostic potential of the prominent lncRNA *HOTAIR* in cancer. On the basis of our findings, *HOTAIR* represents a potential powerful predictor of prognosis of overall survival, cancer recurrence, progression, and metastasis in multiple and diverse types of cancers. Therefore, *HOTAIR* could be applied in the clinical setting as a universal biomarker for monitoring cancer patient survival.

Supplementary Materials: The following are available online at http://www.mdpi.com/2072-6694/11/6/778/s1, Figure S1: Kaplan-Meier plots depicting the prognostic potential of *HOTAIR* for OS in various types of cancers. (A) ACC; (B) MESO and (C) GBM. The corresponding HRs and *p*-values are indicated. The CIs are denoted by dashed lines, Figure S2: Correlation between *HOTAIR* and *CDKN1A* expression.

Author Contributions: Conceptualization, A.P.; methodology, H.I.T., D.O and A.P.; software, H.I.T, D.O. and P.I.K.; validation, H.I.T, D.O. and A.P.; formal analysis, H.I.T, D.O. and A.P.; investigation, H.I.T, D.O. and A.P.; data curation, H.I.T, D.O., P.I.K. and A.P.; writing—original draft preparation, H.I.T, D.O., P.I.K., A.G.G., and A.P.; writing—review and editing, H.I.T, D.O., P.I.K., A.G.G. and A.P.; supervision, A.P.; project administration, A.P.

Funding: P.I.K. acknowledges support of this work by the project "ELIXIR-GR: The Greek Research Infrastructure for Data Management and Analysis in Life Sciences", Grant Number (MIS) 5002780.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Nagano, T.; Fraser, P. No-nonsense functions for long noncoding RNAs. *Cell* **2011**, *145*, 178–181. [CrossRef]
2. Rinn, J.L.; Chang, H.Y. Genome regulation by long noncoding RNAs. *Annu. Rev. Biochem.* **2012**, *81*, 145–166. [CrossRef] [PubMed]
3. Geisler, S.; Coller, J. RNA in unexpected places: Long non-coding RNA functions in diverse cellular contexts. *Nat. Rev. Mol. Cell Biol.* **2013**, *14*, 699–712. [CrossRef] [PubMed]
4. Prensner, J.R.; Chinnaiyan, A.M. The emergence of lncRNAs in cancer biology. *Cancer Discov.* **2011**, *1*, 391–407. [CrossRef] [PubMed]
5. Schmitt, A.M.; Chang, H.Y. Long noncoding RNAs in cancer pathways. *Cancer Cell* **2016**, *29*, 452–463. [CrossRef] [PubMed]
6. Rinn, J.L.; Kertesz, M.; Wang, J.K.; Squazzo, S.L.; Xu, X.; Brugmann, S.A.; Goodnough, L.H.; Helms, J.A.; Farnham, P.J.; Segal, E.; et al. Functional demarcation of active and silent chromatin domains in human hox loci by noncoding RNAs. *Cell* **2007**, *129*, 1311–1323. [CrossRef] [PubMed]
7. Cunningham, F.; Achuthan, P.; Akanni, W.; Allen, J.; Amode, M.R.; Armean, I.M.; Bennett, R.; Bhai, J.; Billis, K.; Boddu, S.; et al. Ensembl 2019. *Nucleic Acids Res.* **2019**, *47*, D745–D751. [CrossRef]
8. Tsai, M.C.; Manor, O.; Wan, Y.; Mosammaparast, N.; Wang, J.K.; Lan, F.; Shi, Y.; Segal, E.; Chang, H.Y. Long noncoding RNA as modular scaffold of histone modification complexes. *Science* **2010**, *329*, 689–693. [CrossRef]
9. Cai, B.; Song, X.Q.; Cai, J.P.; Zhang, S. Hotair: A cancer-related long non-coding RNA. *Neoplasma* **2014**, *61*, 379–391. [CrossRef]
10. Woo, C.J.; Kingston, R.E. Hotair lifts noncoding RNAs to new levels. *Cell* **2007**, *129*, 1257–1259. [CrossRef]
11. Hao, S.; Shao, Z. Hotair is upregulated in acute myeloid leukemia and that indicates a poor prognosis. *Int. J. Clin. Exp. Pathol.* **2015**, *8*, 7223–7228. [PubMed]
12. Qu, X.; Alsager, S.; Zhuo, Y.; Shan, B. Hox transcript antisense RNA (hotair) in cancer. *Cancer Lett.* **2019**, *454*, 90–97. [CrossRef] [PubMed]
13. Bhan, A.; Mandal, S.S. LncRNA hotair: A master regulator of chromatin dynamics and cancer. *Biochim. Biophys. Acta* **2015**, *1856*, 151–164. [CrossRef] [PubMed]
14. Gupta, R.A.; Shah, N.; Wang, K.C.; Kim, J.; Horlings, H.M.; Wong, D.J.; Tsai, M.C.; Hung, T.; Argani, P.; Rinn, J.L.; et al. Long non-coding RNA hotair reprograms chromatin state to promote cancer metastasis. *Nature* **2010**, *464*, 1071–1076. [CrossRef] [PubMed]
15. Zheng, P.; Yin, Z.; Wu, Y.; Xu, Y.; Luo, Y.; Zhang, T.C. LncRNA hotair promotes cell migration and invasion by regulating mkl1 via inhibition mir206 expression in hela cells. *Cell Commun. Signal.* **2018**, *16*, 5. [CrossRef] [PubMed]
16. Wu, Y.; Zhang, L.; Zhang, L.; Wang, Y.; Li, H.; Ren, X.; Wei, F.; Yu, W.; Liu, T.; Wang, X.; et al. Long non-coding RNA hotair promotes tumor cell invasion and metastasis by recruiting ezh2 and repressing e-cadherin in oral squamous cell carcinoma. *Int. J. Oncol.* **2015**, *46*, 2586–2594. [CrossRef] [PubMed]
17. Endo, H.; Shiroki, T.; Nakagawa, T.; Yokoyama, M.; Tamai, K.; Yamanami, H.; Fujiya, T.; Sato, I.; Yamaguchi, K.; Tanaka, N.; et al. Enhanced expression of long non-coding RNA hotair is associated with the development of gastric cancer. *PLoS ONE* **2013**, *8*, e77070. [CrossRef] [PubMed]

18. Liu, X.H.; Sun, M.; Nie, F.Q.; Ge, Y.B.; Zhang, E.B.; Yin, D.D.; Kong, R.; Xia, R.; Lu, K.H.; Li, J.H.; et al. Lnc RNA hotair functions as a competing endogenous RNA to regulate her2 expression by sponging mir-331-3p in gastric cancer. *Mol. Cancer* **2014**, *13*, 92. [CrossRef]
19. Hajjari, M.; Salavaty, A. HOTAIR: An oncogenic long non-coding RNA in different cancers. *Cancer Biol. Med.* **2015**, *12*, 1–9.
20. Tang, Q.; Hann, S.S. Hotair: An oncogenic long non-coding RNA in human cancer. *Cell. Physiol. Biochem.* **2018**, *47*, 893–913. [CrossRef]
21. Nakagawa, T.; Endo, H.; Yokoyama, M.; Abe, J.; Tamai, K.; Tanaka, N.; Sato, I.; Takahashi, S.; Kondo, T.; Satoh, K. Large noncoding RNA hotair enhances aggressive biological behavior and is associated with short disease-free survival in human non-small cell lung cancer. *Biochem. Biophys. Res. Commun.* **2013**, *436*, 319–324. [CrossRef] [PubMed]
22. Troiano, G.; Caponio, V.C.A.; Boldrup, L.; Gu, X.; Muzio, L.L.; Sgaramella, N.; Wang, L.; Nylander, K. Expression of the long non-coding RNA hotair as a prognostic factor in squamous cell carcinoma of the head and neck: A systematic review and meta-analysis. *Oncotarget* **2017**, *8*, 73029–73036. [CrossRef] [PubMed]
23. Zhang, Y.; Zhou, Y.; Xu, T.; Tian, W.; Yang, C.; Wang, X.; Zhong, S.; Ran, Q.; Yang, H.; Zhu, S. Clinical value of long noncoding RNA hotair as a novel biomarker in digestive cancers: A meta-analysis. *Technol. Cancer Res. Treat.* **2018**, *17*. [CrossRef] [PubMed]
24. Chen, F.J.; Sun, M.; Li, S.Q.; Wu, Q.Q.; Ji, L.; Liu, Z.L.; Zhou, G.Z.; Cao, G.; Jin, L.; Xie, H.W.; et al. Upregulation of the long non-coding RNA hotair promotes esophageal squamous cell carcinoma metastasis and poor prognosis. *Mol. Carcinog.* **2013**, *52*, 908–915. [CrossRef] [PubMed]
25. Li, J.; Wen, W.; Zhao, S.; Wang, J.; Chen, J.; Wang, Y.; Zhang, Q. Prognostic role of hotair in four estrogen-dependent malignant tumors: A meta-analysis. *Oncotargets Ther.* **2015**, *8*, 1471–1482. [CrossRef] [PubMed]
26. Tan, S.K.; Pastori, C.; Penas, C.; Komotar, R.J.; Ivan, M.E.; Wahlestedt, C.; Ayad, N.G. Serum long noncoding RNA hotair as a novel diagnostic and prognostic biomarker in glioblastoma multiforme. *Mol. Cancer* **2018**, *17*, 74. [CrossRef] [PubMed]
27. Xiong, Y.; Wang, T.; Wang, M.; Zhao, J.; Li, X.; Zhang, Z.; Zhou, Y.; Liu, J.; Jia, L.; Han, Y. Long non-coding RNAs function as novel predictors and targets of non-small cell lung cancer: A systematic review and meta-analysis. *Oncotarget* **2018**, *9*, 11377–11386. [CrossRef] [PubMed]
28. Deng, Q.; Sun, H.; He, B.; Pan, Y.; Gao, T.; Chen, J.; Ying, H.; Liu, X.; Wang, F.; Xu, Y.; et al. Prognostic value of long non-coding RNA hotair in various cancers. *PLoS ONE* **2014**, *9*, e110059. [CrossRef] [PubMed]
29. Zhang, S.; Chen, S.; Yang, G.; Gu, F.; Li, M.; Zhong, B.; Hu, J.; Hoffman, A.; Chen, M. Long noncoding RNA hotair as an independent prognostic marker in cancer: A meta-analysis. *PLoS ONE* **2014**, *9*, e105538. [CrossRef]
30. Geng, Y.J.; Xie, S.L.; Li, Q.; Ma, J.; Wang, G.Y. Large intervening non-coding RNA hotair is associated with hepatocellular carcinoma progression. *J. Int. Med. Res.* **2011**, *39*, 2119–2128. [CrossRef]
31. Kogo, R.; Shimamura, T.; Mimori, K.; Kawahara, K.; Imoto, S.; Sudo, T.; Tanaka, F.; Shibata, K.; Suzuki, A.; Komune, S.; et al. Long noncoding RNA hotair regulates polycomb-dependent chromatin modification and is associated with poor prognosis in colorectal cancers. *Cancer Res.* **2011**, *71*, 6320–6326. [CrossRef] [PubMed]
32. Yang, Z.; Zhou, L.; Wu, L.M.; Lai, M.C.; Xie, H.Y.; Zhang, F.; Zheng, S.S. Overexpression of long non-coding RNA hotair predicts tumor recurrence in hepatocellular carcinoma patients following liver transplantation. *Ann. Surg. Oncol.* **2011**, *18*, 1243–1250. [CrossRef] [PubMed]
33. Lu, L.; Zhu, G.; Zhang, C.; Deng, Q.; Katsaros, D.; Mayne, S.T.; Risch, H.A.; Mu, L.; Canuto, E.M.; Gregori, G.; et al. Association of large noncoding RNA hotair expression and its downstream intergenic cpg island methylation with survival in breast cancer. *Breast Cancer Res. Treat.* **2012**, *136*, 875–883. [CrossRef] [PubMed]
34. Niinuma, T.; Suzuki, H.; Nojima, M.; Nosho, K.; Yamamoto, H.; Takamaru, H.; Yamamoto, E.; Maruyama, R.; Nobuoka, T.; Miyazaki, Y.; et al. Upregulation of mir-196a and hotair drive malignant character in gastrointestinal stromal tumors. *Cancer Res.* **2012**, *72*, 1126–1136. [CrossRef] [PubMed]
35. Ge, X.S.; Ma, H.J.; Zheng, X.H.; Ruan, H.L.; Liao, X.Y.; Xue, W.Q.; Chen, Y.B.; Zhang, Y.; Jia, W.H. Hotair, a prognostic factor in esophageal squamous cell carcinoma, inhibits wif-1 expression and activates wnt pathway. *Cancer Sci.* **2013**, *104*, 1675–1682. [CrossRef] [PubMed]

36. Ishibashi, M.; Kogo, R.; Shibata, K.; Sawada, G.; Takahashi, Y.; Kurashige, J.; Akiyoshi, S.; Sasaki, S.; Iwaya, T.; Sudo, T.; et al. Clinical significance of the expression of long non-coding RNA hotair in primary hepatocellular carcinoma. *Oncol. Rep.* **2013**, *29*, 946–950. [CrossRef]
37. Li, D.; Feng, J.; Wu, T.; Wang, Y.; Sun, Y.; Ren, J.; Liu, M. Long intergenic noncoding RNA hotair is overexpressed and regulates pten methylation in laryngeal squamous cell carcinoma. *Am. J. Pathol.* **2013**, *182*, 64–70. [CrossRef]
38. Li, X.; Wu, Z.; Mei, Q.; Li, X.; Guo, M.; Fu, X.; Han, W. Long non-coding RNA hotair, a driver of malignancy, predicts negative prognosis and exhibits oncogenic activity in oesophageal squamous cell carcinoma. *Br. J. Cancer* **2013**, *109*, 2266–2278. [CrossRef]
39. Liu, X.H.; Liu, Z.L.; Sun, M.; Liu, J.; Wang, Z.X.; De, W. The long non-coding RNA hotair indicates a poor prognosis and promotes metastasis in non-small cell lung cancer. *BMC Cancer* **2013**, *13*, 464. [CrossRef]
40. Lv, X.B.; Lian, G.Y.; Wang, H.R.; Song, E.; Yao, H.; Wang, M.H. Long noncoding RNA hotair is a prognostic marker for esophageal squamous cell carcinoma progression and survival. *PLoS ONE* **2013**, *8*, e63516. [CrossRef]
41. Nie, Y.; Liu, X.; Qu, S.; Song, E.; Zou, H.; Gong, C. Long non-coding RNA hotair is an independent prognostic marker for nasopharyngeal carcinoma progression and survival. *Cancer Sci.* **2013**, *104*, 458–464. [CrossRef]
42. Sorensen, K.P.; Thomassen, M.; Tan, Q.; Bak, M.; Cold, S.; Burton, M.; Larsen, M.J.; Kruse, T.A. Long non-coding RNA hotair is an independent prognostic marker of metastasis in estrogen receptor-positive primary breast cancer. *Breast Cancer Res. Treat.* **2013**, *142*, 529–536. [CrossRef] [PubMed]
43. Xu, Z.Y.; Yu, Q.M.; Du, Y.A.; Yang, L.T.; Dong, R.Z.; Huang, L.; Yu, P.F.; Cheng, X.D. Knockdown of long non-coding RNA hotair suppresses tumor invasion and reverses epithelial-mesenchymal transition in gastric cancer. *Int. J. Biol. Sci.* **2013**, *9*, 587–597. [CrossRef] [PubMed]
44. He, X.; Bao, W.; Li, X.; Chen, Z.; Che, Q.; Wang, H.; Wan, X.P. The long non-coding RNA hotair is upregulated in endometrial carcinoma and correlates with poor prognosis. *Int. J. Mol. Med.* **2014**, *33*, 325–332. [CrossRef] [PubMed]
45. Huang, L.; Liao, L.M.; Liu, A.W.; Wu, J.B.; Cheng, X.L.; Lin, J.X.; Zheng, M. Overexpression of long noncoding RNA hotair predicts a poor prognosis in patients with cervical cancer. *Arch. Gynecol. Obstet.* **2014**, *290*, 717–723. [CrossRef] [PubMed]
46. Lee, N.K.; Lee, J.H.; Park, C.H.; Yu, D.; Lee, Y.C.; Cheong, J.H.; Noh, S.H.; Lee, S.K. Long non-coding RNA hotair promotes carcinogenesis and invasion of gastric adenocarcinoma. *Biochem. Biophys. Res. Commun.* **2014**, *451*, 171–178. [CrossRef] [PubMed]
47. Okugawa, Y.; Toiyama, Y.; Hur, K.; Toden, S.; Saigusa, S.; Tanaka, K.; Inoue, Y.; Mohri, Y.; Kusunoki, M.; Boland, C.R.; et al. Metastasis-associated long non-coding RNA drives gastric cancer development and promotes peritoneal metastasis. *Carcinogenesis* **2014**, *35*, 2731–2739. [CrossRef]
48. Qiu, J.J.; Lin, Y.Y.; Ye, L.C.; Ding, J.X.; Feng, W.W.; Jin, H.Y.; Zhang, Y.; Li, Q.; Hua, K.Q. Overexpression of long non-coding RNA hotair predicts poor patient prognosis and promotes tumor metastasis in epithelial ovarian cancer. *Gynecol. Oncol.* **2014**, *134*, 121–128. [CrossRef]
49. Svoboda, M.; Slyskova, J.; Schneiderova, M.; Makovicky, P.; Bielik, L.; Levy, M.; Lipska, L.; Hemmelova, B.; Kala, Z.; Protivankova, M.; et al. Hotair long non-coding RNA is a negative prognostic factor not only in primary tumors, but also in the blood of colorectal cancer patients. *Carcinogenesis* **2014**, *35*, 1510–1515. [CrossRef]
50. Wu, Z.H.; Wang, X.L.; Tang, H.M.; Jiang, T.; Chen, J.; Lu, S.; Qiu, G.Q.; Peng, Z.H.; Yan, D.W. Long non-coding RNA hotair is a powerful predictor of metastasis and poor prognosis and is associated with epithelial-mesenchymal transition in colon cancer. *Oncol. Rep.* **2014**, *32*, 395–402. [CrossRef]
51. Yan, T.H.; Lu, S.W.; Huang, Y.Q.; Que, G.B.; Chen, J.H.; Chen, Y.P.; Zhang, H.B.; Liang, X.L.; Jiang, J.H. Upregulation of the long noncoding RNA hotair predicts recurrence in stage ta/t1 bladder cancer. *Tumour Biol.* **2014**, *35*, 10249–10257. [CrossRef] [PubMed]
52. Heubach, J.; Monsior, J.; Deenen, R.; Niegisch, G.; Szarvas, T.; Niedworok, C.; Schulz, W.A.; Hoffmann, M.J. The long noncoding RNA hotair has tissue and cell type-dependent effects on hox gene expression and phenotype of urothelial cancer cells. *Mol. Cancer* **2015**, *14*, 108. [CrossRef] [PubMed]
53. Kim, H.J.; Lee, D.W.; Yim, G.W.; Nam, E.J.; Kim, S.; Kim, S.W.; Kim, Y.T. Long non-coding RNA hotair is associated with human cervical cancer progression. *Int. J. Oncol.* **2015**, *46*, 521–530. [CrossRef] [PubMed]

54. Liu, Y.W.; Sun, M.; Xia, R.; Zhang, E.B.; Liu, X.H.; Zhang, Z.H.; Xu, T.P.; De, W.; Liu, B.R.; Wang, Z.X. Linchotair epigenetically silences mir34a by binding to prc2 to promote the epithelial-to-mesenchymal transition in human gastric cancer. *Cell Death Dis.* **2015**, *6*, e1802. [CrossRef] [PubMed]
55. Ma, G.; Wang, Q.; Lv, C.; Qiang, F.; Hua, Q.; Chu, H.; Du, M.; Tong, N.; Jiang, Y.; Wang, M.; et al. The prognostic significance of hotair for predicting clinical outcome in patients with digestive system tumors. *J. Cancer Res. Clin. Oncol.* **2015**, *141*, 2139–2145. [CrossRef] [PubMed]
56. Martinez-FeRNAndez, M.; Feber, A.; Duenas, M.; Segovia, C.; Rubio, C.; FeRNAndez, M.; Villacampa, F.; Duarte, J.; Lopez-Calderon, F.F.; Gomez-Rodriguez, M.J.; et al. Analysis of the polycomb-related lncRNAs hotair and anril in bladder cancer. *Clin. Epigenet.* **2015**, *7*, 109. [CrossRef]
57. Qiu, J.J.; Wang, Y.; Ding, J.X.; Jin, H.Y.; Yang, G.; Hua, K.Q. The long non-coding RNA hotair promotes the proliferation of serous ovarian cancer cells through the regulation of cell cycle arrest and apoptosis. *Exp. Cell Res.* **2015**, *333*, 238–248. [CrossRef]
58. Wu, J.; Xie, H. Expression of long noncoding RNA-hox transcript antisense intergenic RNA in oral squamous cell carcinoma and effect on cell growth. *Tumour Biol.* **2015**, *36*, 8573–8578. [CrossRef]
59. Wu, S.; Zheng, C.; Chen, S.; Cai, X.; Shi, Y.; Lin, B.; Chen, Y. Overexpression of long non-coding RNA hotair predicts a poor prognosis in patients with acute myeloid leukemia. *Oncol. Lett.* **2015**, *10*, 2410–2414. [CrossRef]
60. Xing, C.Y.; Hu, X.Q.; Xie, F.Y.; Yu, Z.J.; Li, H.Y.; Bin, Z.; Wu, J.B.; Tang, L.Y.; Gao, S.M. Long non-coding RNA hotair modulates c-kit expression through sponging mir-193a in acute myeloid leukemia. *FEBS Lett.* **2015**, *589*, 1981–1987. [CrossRef]
61. Zhang, Z.Z.; Shen, Z.Y.; Shen, Y.Y.; Zhao, E.H.; Wang, M.; Wang, C.J.; Cao, H.; Xu, J. Hotair long noncoding RNA promotes gastric cancer metastasis through suppression of poly r(c)-binding protein (pcbp) 1. *Mol. Cancer Ther.* **2015**, *14*, 1162–1170. [CrossRef] [PubMed]
62. Zhao, W.; Dong, S.; Duan, B.; Chen, P.; Shi, L.; Gao, H.; Qi, H. Hotair is a predictive and prognostic biomarker for patients with advanced gastric adenocarcinoma receiving fluorouracil and platinum combination chemotherapy. *Am. J. Transl. Res.* **2015**, *7*, 1295–1302. [PubMed]
63. Luczak, A.; SupeRNAt, A.; Lapinska-Szumczyk, S.; Jachimowicz, D.; Majewska, H.; Gulczynski, J.; Zaczek, A.J. Hotair in relation to epithelial-mesenchymal transition and cancer stem cells in molecular subtypes of endometrial cancer. *Int. J. Biol. Markers* **2016**, *31*, e245–e251. [CrossRef] [PubMed]
64. Luo, Z.F.; Zhao, D.; Li, X.Q.; Cui, Y.X.; Ma, N.; Lu, C.X.; Liu, M.Y.; Zhou, Y. Clinical significance of hotair expression in colon cancer. *World J. Gastroenterol.* **2016**, *22*, 5254–5259. [CrossRef] [PubMed]
65. Sun, J.; Chu, H.; Ji, J.; Huo, G.; Song, Q.; Zhang, X. Long non-coding RNA hotair modulates hla-g expression by absorbing mir-148a in human cervical cancer. *Int. J. Oncol.* **2016**, *49*, 943–952. [CrossRef] [PubMed]
66. Yan, Y.; Han, J.; Li, Z.; Yang, H.; Sui, Y.; Wang, M. Elevated RNA expression of long noncoding hotair promotes cell proliferation and predicts a poor prognosis in patients with diffuse large b cell lymphoma. *Mol. Med. Rep.* **2016**, *13*, 5125–5131. [CrossRef] [PubMed]
67. Zhang, Y.Y.; Huang, S.H.; Zhou, H.R.; Chen, C.J.; Tian, L.H.; Shen, J.Z. Role of hotair in the diagnosis and prognosis of acute leukemia. *Oncol. Rep.* **2016**, *36*, 3113–3122. [CrossRef]
68. Chen, W.M.; Chen, W.D.; Jiang, X.M.; Jia, X.F.; Wang, H.M.; Zhang, Q.J.; Shu, Y.Q.; Zhao, H.B. Hox transcript antisense intergenic RNA represses e-cadherin expression by binding to ezh2 in gastric cancer. *World J. Gastroenterol.* **2017**, *23*, 6100–6110. [CrossRef]
69. Hu, G.; Dong, B.; Zhang, J.; Zhai, W.; Xie, T.; Huang, B.; Huang, C.; Yao, X.; Zheng, J.; Che, J.; et al. The long noncoding RNA hotair activates the hippo pathway by directly binding to sav1 in renal cell carcinoma. *Oncotarget* **2017**, *8*, 58654–58667. [CrossRef]
70. Katayama, H.; Tamai, K.; Shibuya, R.; Nakamura, M.; Mochizuki, M.; Yamaguchi, K.; Kawamura, S.; Tochigi, T.; Sato, I.; Okanishi, T.; et al. Long non-coding RNA hotair promotes cell migration by upregulating insulin growth factor-binding protein 2 in renal cell carcinoma. *Sci. Rep.* **2017**, *7*, 12016. [CrossRef]
71. Luan, W.; Li, R.; Liu, L.; Ni, X.; Shi, Y.; Xia, Y.; Wang, J.; Lu, F.; Xu, B. Long non-coding RNA hotair acts as a competing endogenous RNA to promote malignant melanoma progression by sponging mir-152-3p. *Oncotarget* **2017**, *8*, 85401–85414. [CrossRef] [PubMed]
72. Xu, F.; Zhang, J. Long non-coding RNA hotair functions as miRNA sponge to promote the epithelial to mesenchymal transition in esophageal cancer. *Biomed. Pharmacother.* **2017**, *90*, 888–896. [CrossRef] [PubMed]

73. Zhang, Y.; Yu, S.; Jiang, L.; Wang, X.; Song, X. Hotair is a promising novel biomarker in patients with thyroid cancer. *Exp. Ther. Med.* **2017**, *13*, 2274–2278. [CrossRef] [PubMed]
74. Dong, X.; He, X.; Guan, A.; Huang, W.; Jia, H.; Huang, Y.; Chen, S.; Zhang, Z.; Gao, J.; Wang, H. Long non-coding RNA hotair promotes gastric cancer progression via mir-217-gpc5 axis. *Life Sci.* **2019**, *217*, 271–282. [CrossRef] [PubMed]
75. Huang, K.B.; Zhang, S.P.; Zhu, Y.J.; Guo, C.H.; Yang, M.; Liu, J.; Xia, L.G.; Zhang, J.F. Hotair mediates tumorigenesis through recruiting ezh2 in colorectal cancer. *J. Cell. Biochem.* **2019**, *120*, 6071–6077. [CrossRef] [PubMed]
76. Xiao, Z.; Qu, Z.; Chen, Z.; Fang, Z.; Zhou, K.; Huang, Z.; Guo, X.; Zhang, Y. LncRNA hotair is a prognostic biomarker for the proliferation and chemoresistance of colorectal cancer via mir-203a-3p-mediated wnt/ss-catenin signaling pathway. *Cell. Physiol. Biochem.* **2018**, *46*, 1275–1285. [CrossRef] [PubMed]
77. Coordinators, N.R. Database resources of the national center for biotechnology information. *Nucleic Acids Res.* **2018**, *46*, D8–D13. [CrossRef]
78. Zhang, Y.; Wang, L.J.; Li, W.F.; Zhang, X.; Yang, X.J. The prognostic value of hotair for predicting long-term prognosis of patients with gastrointestinal cancers. *Medicine* **2018**, *97*, e11139. [CrossRef] [PubMed]
79. Abdeahad, H.; Avan, A.; Pashirzad, M.; Khazaei, M.; Soleimanpour, S.; Ferns, G.A.; Fiuji, H.; Ryzhikov, M.; Bahrami, A.; Hassanian, S.M. The prognostic potential of long noncoding RNA hotair expression in human digestive system carcinomas: A meta-analysis. *J. Cell. Physiol.* **2019**, *234*, 10926–10933. [CrossRef] [PubMed]
80. Cai, H.; An, Y.; Chen, X.; Sun, D.; Chen, T.; Peng, Y.; Zhu, F.; Jiang, Y.; He, X. Epigenetic inhibition of mir-663b by long non-coding RNA hotair promotes pancreatic cancer cell proliferation via up-regulation of insulin-like growth factor 2. *Oncotarget* **2016**, *7*, 86857–86870. [CrossRef]
81. Song, Y.; Wang, R.; Li, L.W.; Liu, X.; Wang, Y.F.; Wang, Q.X.; Zhang, Q. Long non-coding RNA hotair mediates the switching of histone h3 lysine 27 acetylation to methylation to promote epithelial-to-mesenchymal transition in gastric cancer. *Int. J. Oncol.* **2019**, *54*, 77–86. [CrossRef] [PubMed]
82. Oh, E.J.; Kim, S.H.; Yang, W.I.; Ko, Y.H.; Yoon, S.O. Long non-coding RNA hotair expression in diffuse large b-cell lymphoma: In relation to polycomb repressive complex pathway proteins and h3k27 trimethylation. *J. Pathol. Transl. Med.* **2016**, *50*, 369–376. [CrossRef] [PubMed]
83. Jiang, C.; Yang, Y.; Yang, Y.; Guo, L.; Huang, J.; Liu, X.; Wu, C.; Zou, J. Long noncoding RNA (lncRNA) hotair affects tumorigenesis and metastasis of non-small cell lung cancer by upregulating mir-613. *Oncol. Res.* **2018**, *26*, 725–734. [CrossRef] [PubMed]
84. Zhang, Z.; Cheng, J.; Wu, Y.; Qiu, J.; Sun, Y.; Tong, X. LncRNA hotair controls the expression of rab22a by sponging mir-373 in ovarian cancer. *Mol. Med. Rep.* **2016**, *14*, 2465–2472. [CrossRef] [PubMed]
85. Georgakilas, A.G.; Martin, O.A.; Bonner, W.M. P21: A two-faced genome guardian. *Trends Mol. Med.* **2017**, *23*, 310–319. [CrossRef] [PubMed]
86. Galanos, P.; Vougas, K.; Walter, D.; Polyzos, A.; Maya-Mendoza, A.; Haagensen, E.J.; Kokkalis, A.; Roumelioti, F.M.; Gagos, S.; Tzetis, M.; et al. Chronic p53-independent p21 expression causes genomic instability by deregulating replication licensing. *Nat. Cell Biol.* **2016**, *18*, 777–789. [CrossRef] [PubMed]
87. Zhai, N.; Xia, Y.; Yin, R.; Liu, J.; Gao, F. A negative regulation loop of long noncoding RNA hotair and p53 in non-small-cell lung cancer. *Oncotargets Ther.* **2016**, *9*, 5713–5720.
88. Liu, H.; Li, Z.; Wang, C.; Feng, L.; Huang, H.; Liu, C.; Li, F. Expression of long non-coding RNA-hotair in oral squamous cell carcinoma tca8113 cells and its associated biological behavior. *Am. J. Transl. Res.* **2016**, *8*, 4726–4734.
89. Gao, J.Z.; Li, J.; Du, J.L.; Li, X.L. Long non-coding RNA hotair is a marker for hepatocellular carcinoma progression and tumor recurrence. *Oncol. Lett.* **2016**, *11*, 1791–1798. [CrossRef]
90. Zheng, J.; Xiao, X.; Wu, C.; Huang, J.; Zhang, Y.; Xie, M.; Zhang, M.; Zhou, L. The role of long non-coding RNA hotair in the progression and development of laryngeal squamous cell carcinoma interacting with ezh2. *Acta Oto-Laryngol.* **2017**, *137*, 90–98. [CrossRef]
91. Lee, M.; Kim, H.J.; Kim, S.W.; Park, S.A.; Chun, K.H.; Cho, N.H.; Song, Y.S.; Kim, Y.T. The long non-coding RNA hotair increases tumour growth and invasion in cervical cancer by targeting the notch pathway. *Oncotarget* **2016**, *7*, 44558–44571. [CrossRef] [PubMed]
92. Liu, L.C.; Wang, Y.L.; Lin, P.L.; Zhang, X.; Cheng, W.C.; Liu, S.H.; Chen, C.J.; Hung, Y.; Jan, C.I.; Chang, L.C.; et al. Long noncoding RNA hotair promotes invasion of breast cancer cells through chondroitin sulfotransferase chst15. *Int. J. Cancer* **2019**. [CrossRef] [PubMed]

93. Di, W.; Li, Q.; Shen, W.; Guo, H.; Zhao, S. The long non-coding RNA hotair promotes thyroid cancer cell growth, invasion and migration through the mir-1-ccnd2 axis. *Am. J. Cancer Res.* **2017**, *7*, 1298–1309. [PubMed]
94. Li, Q.; Feng, Y.; Chao, X.; Shi, S.; Liang, M.; Qiao, Y.; Wang, B.; Wang, P.; Zhu, Z. Hotair contributes to cell proliferation and metastasis of cervical cancer via targetting mir-23b/mapk1 axis. *Biosci. Rep.* **2018**, *38*. [CrossRef] [PubMed]
95. Liberati, A.; Altman, D.G.; Tetzlaff, J.; Mulrow, C.; Gotzsche, P.C.; Ioannidis, J.P.; Clarke, M.; Devereaux, P.J.; Kleijnen, J.; Moher, D. The prisma statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *J. Clin. Epidemiol.* **2009**, *62*, e1–e34. [CrossRef]
96. Fiorini, N.; Lipman, D.J.; Lu, Z. Towards pubmed 2.0. *eLife* **2017**, *6*, e28801. [CrossRef]
97. Tierney, J.F.; Stewart, L.A.; Ghersi, D.; Burdett, S.; Sydes, M.R. Practical methods for incorporating summary time-to-event data into meta-analysis. *Trials* **2007**, *8*, 16. [CrossRef]
98. Higgins, J.P.; Thompson, S.G.; Deeks, J.J.; Altman, D.G. Measuring inconsistency in meta-analyses. *BMJ* **2003**, *327*, 557–560. [CrossRef]
99. Higgins, J.P.; Thompson, S.G. Quantifying heterogeneity in a meta-analysis. *Stat. Med.* **2002**, *21*, 1539–1558. [CrossRef]
100. DerSimonian, R.; Laird, N. Meta-analysis in clinical trials. *Control. Clin. Trials* **1986**, *7*, 177–188. [CrossRef]
101. DerSimonian, R.; Laird, N. Meta-analysis in clinical trials revisited. *Contemp. Clin. Trials* **2015**, *45*, 139–145. [CrossRef] [PubMed]
102. Begg, C.B.; Mazumdar, M. Operating characteristics of a rank correlation test for publication bias. *Biometrics* **1994**, *50*, 1088–1101. [CrossRef] [PubMed]
103. Egger, M.; Davey Smith, G.; Schneider, M.; Minder, C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* **1997**, *315*, 629–634. [CrossRef] [PubMed]
104. Tang, Z.; Li, C.; Kang, B.; Gao, G.; Li, C.; Zhang, Z. Gepia: A web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res.* **2017**, *45*, W98–W102. [CrossRef] [PubMed]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

HOTAIR as a Prognostic Predictor for Diverse Human Cancers: A Meta- and Bioinformatics Analysis

Halil Ibrahim Toy, Didem Okmen, Panagiota I. Kontou, Alexandros G. Georgakilas and Athanasia Pavlopoulou

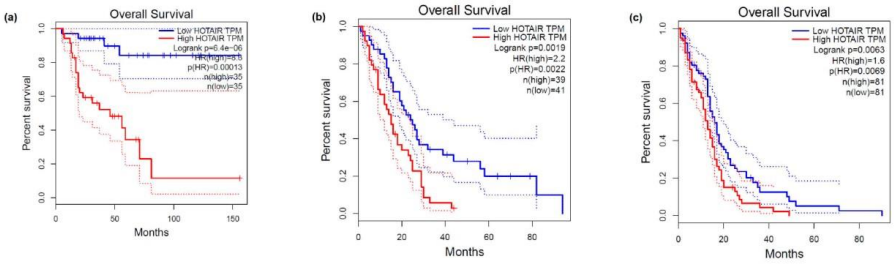


Figure S1. Kaplan–Meier plots depicting the prognostic potential of *HOTAIR* for OS in various types of cancers. (**a**) ACC; (**b**) MESO and (**c**) GBM. The corresponding HRs and *p*-values are indicated. The CIs are denoted by dashed lines.

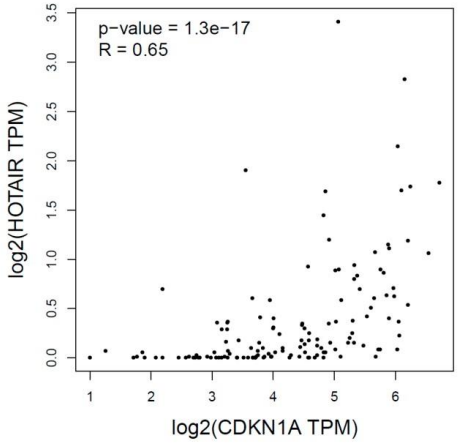


Figure S2. Correlation between *HOTAIR* and *CDKN1A* expression.

© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

frontiers in Cell and Developmental Biology

ORIGINAL RESEARCH
published: 07 April 2021
doi: 10.3389/fcell.2021.620248

Investigating Molecular Determinants of Cancer Cell Resistance to Ionizing Radiation Through an Integrative Bioinformatics Approach

Halil Ibrahim Toy[1,2], *Gökhan Karakülah*[1,2], *Panagiota I. Kontou*[3], *Hani Alotaibi*[1,2], *Alexandros G. Georgakilas*[4] *and Athanasia Pavlopoulou*[1,2*]

[1] Izmir Biomedicine and Genome Center, Izmir, Turkey, [2] Izmir International Biomedicine and Genome Institute, Dokuz Eylül University, Izmir, Turkey, [3] Department of Computer Science and Biomedical Informatics, University of Thessaly, Lamia, Greece, [4] DNA Damage Laboratory, Department of Physics, School of Applied Mathematical and Physical Sciences, Zografou, National Technical University of Athens, Athens, Greece

Eradication of cancer cells through exposure to high doses of ionizing radiation (IR) is a widely used therapeutic strategy in the clinical setting. However, in many cases, cancer cells can develop remarkable resistance to radiation. Radioresistance represents a prominent obstacle in the effective treatment of cancer. Therefore, elucidation of the molecular mechanisms and pathways related to radioresistance in cancer cells is of paramount importance. In the present study, an integrative bioinformatics approach was applied to three publicly available RNA sequencing and microarray transcriptome datasets of human cancer cells of different tissue origins treated with ionizing radiation. These data were investigated in order to identify genes with a significantly altered expression between radioresistant and corresponding radiosensitive cancer cells. Through rigorous statistical and biological analyses, 36 genes were identified as potential biomarkers of radioresistance. These genes, which are primarily implicated in DNA damage repair, oxidative stress, cell pro-survival, and apoptotic pathways, could serve as potential diagnostic/prognostic markers cancer cell resistance to radiation treatment, as well as for therapy outcome and cancer patient survival. In addition, our findings could be potentially utilized in the laboratory and clinical setting for enhancing cancer cell susceptibility to radiation therapy protocols.

Keywords: ionizing radiation, DNA damage repair, cancer cell radioresistance, bioinformatics, gene expression profiles, biomarkers

OPEN ACCESS

Edited by:
Brigitte M. Pützer,
University Hospital Rostock, Germany

Reviewed by:
Andrew Kellett,
Dublin City University, Ireland
Anna Dubrovska,
Technische Universität Dresden, Germany

***Correspondence:**
Athanasia Pavlopoulou
athanasia.pavlopoulou@ibg.edu.tr

Specialty section:
This article was submitted to Molecular Medicine, a section of the journal Frontiers in Cell and Developmental Biology

Received: *22 October 2020*
Accepted: *15 March 2021*
Published: *07 April 2021*

Citation:
Toy HI, Karakülah G, Kontou PI, Alotaibi H, Georgakilas AG and Pavlopoulou A (2021) Investigating Molecular Determinants of Cancer Cell Resistance to Ionizing Radiation Through an Integrative Bioinformatics Approach. Front. Cell Dev. Biol. 9:620248. doi: 10.3389/fcell.2021.620248

INTRODUCTION

Radiation therapy or radiotherapy (RT) represents one of the optimal, most widely used modalities in the treatment of multiple cancers, either alone or combined with other curative anti-cancer modalities like chemotherapy (Delaney et al., 2005; Begg et al., 2011) or immunotherapy (Tang et al., 2014; Schoenhals et al., 2016). It is estimated that approximately 50% of all cancer patients worldwide undergo radiotherapy throughout their illness trajectory (Baskar et al., 2012).

Advances in radiotherapy contribute greatly to cancer patients' improvement of overall survival and quality of life (Baskar et al., 2012). The aim of radiotherapeutic regimens is to specifically and efficiently sensitize cancer cells to IR in order to eliminate them and

prevent cancer recurrence and relapse, minimizing at the same time the adverse effects of radiation on healthy tissue. RT affects cancer cells either directly, by inducing genomic (DNA) lesions, or indirectly, through the generation of DNA damaging intermediates through the interaction with water, like reactive oxygen/nitrogen species (ROS/RNS) and free radicals (e.g., hydrogen ion, hydroxide, etc.) (Mikkelsen and Wardman, 2003; Yamamori et al., 2012).

However, cancer cells have the capacity to develop incredible tolerability and resistance to RT, thereby evading death. Radioresistance represents a major limiting factor in the effective treatment of different types of cancers. The response of tumor cells to radiation depends both on the resistance mechanisms of the cells and also on the accelerated repopulation of the tumor bulk by cells that have developed further radioresistance (Pavlopoulou et al., 2016, 2017). As noted in previous studies, the genes that are differentially expressed (either up- or down-regulated) between radioresistant (RR) and radiosensitive (RR) cancer cells are generally implicated in DNA damage response and repair (DDR/R) pathways, apoptosis, hypoxia, or response to oxidative stress, etc. (Pavlopoulou et al., 2016, 2017). The complexity of radiation resistance mechanisms suggests the involvement of different and diverse biological mechanisms.

During the last decade, the advances in high-throughput (HTP) "omics" technologies (e.g., RNA-Seq and microarrays) enabled the generation of an enormous amount of gene expression data. Data produced with HTP technologies are stored in international public repositories such as NCBI's GEO (Gene Expression Omnibus) (https://www.ncbi.nlm.nih.gov/gds/) (Barrett et al., 2013; Clough and Barrett, 2016). GEO DataSets contains both original records and curated datasets.

Accumulated knowledge over years of research on the biological effects of radiation points toward the development of holistic approaches to "big data" analysis by employing systems biology methodologies (Unger, 2014; Beheshti et al., 2019; Spratt and Speers, 2019; Kanakoglou et al., 2020). Herein, we employed a rigorous systems biology approach to unravel the molecular determinants of resistance of cancer cells to IR, based solely on HTP data. To this end, publicly available transcriptome datasets relevant to cancer cell response to radiation were retrieved from GEO, and specifically, cancer cell lines that displayed enhanced resistance to radiation. Statistical analyses were carried out to identify the differentially expressed genes (DEGs) between radioresistant and radiosensitive tumor cells. Furthermore, functional annotation of genes allowed us to identify specific biological pathways implicated in cancer cell resistance to radiation. Our findings could be applied in the laboratory and clinical setting as biomarkers for the design of targeted and personalized radiotherapy regimens in order to effectively sensitize cancer cells to radiation, enhance tumor control and thereby minimize tumor recurrence and metastasis.

METHODS

Data Retrieval

The public repository NCBI GEO DataSets was searched extensively for gene expression datasets using relevant keywords: ("radiation therapy" or "radiotherapy") and ("cancer" or "tumor") and ("resistance" or "tolerability") and ("sensitivity" or "responsive") and ("human" or "homo sapiens"). A total of three eligible datasets were selected:

The GEO series GSE97543 (Emons et al., 2017) (**Supplementary Table 1**) includes global gene expression by microarray of both wild-type and radioresistant Dukes' type C colorectal adenocarcinoma (COAD) cell lines that were either non-irradiated or irradiated repeatedly with 2 Gray (Gy) of X-rays in order to acquire a radioresistant phenotype. The Agilent-026652 whole human genome microarray 4x44K, GPL13497 platform was used.

In GSE13280 (Marston et al., 2009) (**Supplementary Table 1**), genome-wide gene expression by microarray was performed of cell lines derived from pediatric B-precursor acute lymphoblastic leukemia (ALL) after 8 h *in vitro* exposure to 5 Gy IR. This dataset contains cell lines both resistant and responsive to radiation. The development of resistance and responsiveness to IR was assessed by measuring apoptosis in cells. The Affymetrix human genome U133A array, GPL96 platform was employed.

In GSE120798 (Gray et al., 2019) (**Supplementary Table 1**), three novel radioresistant breast cancer cell lines were established by exposing the corresponding parental cell lines MDA-MB-231 (metastatic mammary adenocarcinoma), MCF-7 (breast adenocarcinoma), and ZR-751 (luminal breast cancer) to increasing doses of X-rays for 2 and 8 h. Genome-wide gene expression analysis of both the parental and radioresistant cell lines (i.e., MDA/MDAR, MCF/MCFR, ZR/ZRR) was performed using high throughput sequencing. The NextSeq 550 (Homo sapiens), GPL21697 platform was used.

Microarray-Based Transcriptomic Data Analysis

For each microarray study, the gene expression data that represent the gene expression summary for every probe and every sample were recorded. In microarrays, many probes can map to the same Gene Symbol for various reasons, and, conversely, a probe may also map to more than one Gene Symbol if the probe sequence is not specific enough. A simple approach would be to use only the probes with one-to-one mapping for further analysis; however, this approach results in the loss of important information. To conduct an analysis based on genes and not probes, the probe identifiers were firstly converted into gene identifiers, according to Ramasamy et al. (2008) guidelines. To this end, GPL files that contained information about the gene symbols that correspond to probe IDs were used in order to resolve the "many-to-many" relationships between probes and genes by averaging the expression profiles for genes with more than one probe (Ramasamy et al., 2008). We identified the Gene Symbols with the usage of the HUGO Gene Nomenclature Committee (Braschi et al., 2019) and the National Center for Biotechnology Information (NCBI) GENE (Sayers et al., 2020).

The two-sample *t*-test was employed to identify genes differentially expressed between the case (RR) and control (RS) groups. However, a disadvantage of the *t*-test in the analysis of microarray data is that if most of the experiments in a given

study contain a relatively small number of samples per group, the assumption of normality is untenable. To resolve this, the statistical method *t*-test with bootstrap was used (Efron and Tibshirani, 1993). Bootstrap provides an ideal method to generate accurate estimates of the standard errors when no formula for the sampling distribution is available or when available formulas make inappropriate assumptions (e.g., small sample size, non-normal distribution). In this study, bootstrap analysis was conducted with 1,000 replicates, a relatively high number, in order to generate accurate estimates of the standard errors.

A typical microarray experiment measures the expression of several thousand genes simultaneously across different conditions. When investigating for potential DEGs between two conditions, each gene is treated independently, and the *t*-test is performed on each gene separately. The incidence of false positives (i.e., genes falsely declared as DEGs) is proportional to the number of tests performed and the critical significance level (*p*-value cut-off). In order to account for multiple comparisons, a correction method proposed by Benjamini and Hochberg (1995) which controls False Discovery Rate (FDR) was applied. FDR-controlling procedures have greater power (i.e., they can discover more statistically significant differences), at the expense of increased Type I error rate. Genes with adjusted *p*-value (or *q*-value) less or equal to 0.05 were considered as statistically significant in this study. For all statistical analyses, the Stata 13 statistical software package (StataCorp, 2013) was used. For the creation of heatmaps from microarray data, the average linkage clustering with Euclidean distance clustering method implemented in Heatmapper (http://heatmapper.ca/) was utilized.

RNA-Seq-Based Transcriptomic Data Analysis

For the RNA-Seq based transcriptome analysis, the following pipeline was utilized. The FASTQ files were extracted from the respective Sequence Read Archive (SRA) files containing raw RNA-Seq reads by using the SRA Tool Kit v.2.9.0 (Alnasir and Shanahan, 2015) with the *"fastq-dump –gzip –skip-technical –readids –dumpbase –clip –split-3"* command. The raw RNA-Seq reads in FASTQ files were aligned to the human reference genome GRCh38 (Ensembl version 97) by employing the splice junction aligner HISAT2 v.2.1.0 (Kim et al., 2015) with *"hisat2 -p -dta -x {input.index} -U {input.fq} -S {out.sam}"* parameters. The generated SAM file was converted to the respective binary BAM file by using SAM Tools v.1.9.0 (Li et al., 2009) with *"samtools sort -@ 10 -o {output.bam} {input.sam}"* commands. String Tie v.1.3.5 (Pertea et al., 2015) was utilized with *"stringtie -e -B -p -G {input.gtf} -A {output.tab} -o {output.gtf} -l {input.label}{input.bam}"* parameters for the measurement of gene expression levels. The reconstructed transcripts and transcript abundances were reported in the output GTF file. In order to detect differentially expressed genes between RR and RS samples, we utilized the EdgeR package v3.28.0 (Robinson et al., 2010) of the R statistical computation environment v.3.6.1 (https://www.r-project.org). We firstly applied the trimmed mean of M-values (TMM) normalization (Robinson and Oshlack, 2010) implemented in EdgeR to the count data and we employed generalized linear models with "cell lines," "resistance," and "time" as factors. Then, estimating dispersion was computed with the *estimateDisp* function, and differential expression analysis between the two RNA-Seq groups (RR and RS) was performed using the *glmFit* and *glmLRT* functions of the EdgeR package v3.28.0 (Robinson et al., 2010) of the R statistical computation environment v.3.6.1 (https://www.r-project.org). In order to detect statistically significant differentially expressed genes, the threshold for the absolute $\log_2$-fold change was set at two ($|\log_2 FC \geq 2|$), and for the FDR (Benjamini and Hochberg, 1995)-corrected *p*-value at 0.05. All statistical calculations for the RNASeq data were performed by using the R software environment. The *pheatmap* package of R (https://CRAN.R-project.org/package=pheatmap) was utilized to generate a heatmap from RNA-Seq data.

Pathway Enrichment Analysis

To assign biological role(s) to the genes under study that are associated with biological pathways, gene set enrichment analysis (GSEA), or functional enrichment analysis, was conducted. GSEA is a method to identify biological processes or pathways that are over-represented in a large set of genes. To this end, WebGestalt (WEB-based GEne SeT AnaLysis Toolkit) (Zhang et al., 2005; Liao et al., 2019) was employed to identify statistically significant over-represented WikiPathways (Kutmon et al., 2016) cancer terms in the sets of genes; the threshold for the FDR-corrected *p*-value was set at 10^{-3}, and hypergeometric distribution analysis was used.

Functional Interactions Networks

The associations among the molecules under study were investigated and visualized with the usage of STRING (Search Tool for Retrieval of Interacting Genes/Proteins) v11.0 (Szklarczyk et al., 2019), a database of either known or predicted, direct or indirect, gene/protein associations derived from diverse resources. The highest confidence interaction score ($\geq$0.9) was chosen to display the associations amongst genes/proteins in the generated network.

Survival Analysis

The prognostic potential of the 36 radiogenes was investigated in three types of cancers, namely, breast invasive carcinoma (BRCA), colon adenocarcinoma (COAD), and acute myeloid leukemia (LAML), a type of haematologic cancer like ALL, through the web-based tool GEPIA (Gene Expression Profiling Interactive Analysis) (Tang et al., 2017) version 2 (http://gepia2.cancer-pku.cn/#index), based on data acquired from The Cancer Genome Atlas (TCGA) (https://tcga-data.nci.nih.gov). The cancer patient cohort is divided into the high-risk and low-risk categories; the cut-offs for low and high gene expression level patient cohorts were set at 50%.

Tissue-Wise Differential Gene Expression Analysis

GEPIA2 (Tang et al., 2017), which contains gene expression data from cancer and corresponding normal tissues from

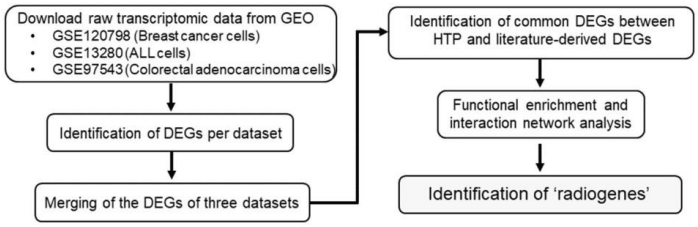


FIGURE 1 | Flowchart diagram of the overall methodology employed in this study. Three eligible transcriptome datasets were retrieved from NCBI's GEO. Genes differentially expressed (DEGs) between the radioresistant and radiosensitive cancer cells were identified per dataset. The DEGs of the three analyzed datasets were integrated and compared with an earlier own study on cancer cell radioresistance-related genes/proteins derived from literature research. Functional enrichment analysis of the common genes was performed to obtain the so-called "radiogenes".

TCGA (https://tcga-data.nci.nih.gov) and the Genotype-Tissue Expression (GTEx) (https://gtexportal.org/home/), respectively, was used to investigate the differential expression patterns of the radiogenes under study in BRCA, COAD, and LAML cancer-normal tissue; $|\log FC| \geq 2$ and FDR-corrected p-value ≤ 0.05.

RESULTS

The overall procedure for data collection and analysis followed in the present study is described illustratively in **Figure 1**.

Identification of Differential Expression Patterns in Radioresistance vs. Radiosensitive Cancer Cells

The two techniques for transcriptome profiling, RNA-Seq and microarray, have large inherent differences. RNA-Seq is considered "superior" since it allows the detection of low abundance transcripts and novel transcript isoforms (Marioni et al., 2008). For this reason, we applied different statistical methods for RNA-Seq (Li, 2019) and microarray (Kontou et al., 2018) data processing and analysis.

The number of differentially expressed genes (DEGs) found between radioresistant and radiosensitive cancer cells (**Figure 2**) for each dataset is 6372 (GSE120798), 782 (GSE13280), and 541 (GSE97543), respectively (**Supplementary Table 2**). The pathways over-represented in the DEGs of the three datasets are related primarily to DDR/R and cell survival (**Figure 3**).

Differential gene expression analysis was performed for the three parental breast cancer cell lines in GSE120798 (Gray et al., 2019) against their radioresistant derivatives, to identify those genes that are significantly dysregulated in response to radiation stress. A total of 11 samples were compared; one biological replicate for each condition (**Supplementary Table 1**). The number of detected DEGs (Gray et al., 2019) is remarkably higher as compared to the ones derived from microarray gene expression data (**Supplementary Table 2**). This discrepancy is likely due to the ability of RNA-Seq to detect and quantify, even rare, transcripts without *a priori* knowledge of a given gene (Metzker, 2010). Accordingly, the number of enriched pathways is also greater in this dataset (**Figure 3**). Three DDR/R pathways were found to be enriched in the DEGs, including the generic "DNA Damage Response" pathway. Two pathways are specifically implicated in IR-induced DNA damage response, namely "DNA IR-Double Strand Breaks (DSBs) and cellular response via ATM," and "DNA IR-damage and cellular response via ATR." Ataxia telangiectasia mutated (ATM) and ataxia telangiectasia and Rad3-related (ATR) proteins are evolutionarily conserved proteins that have a critical regulatory role in DDR and maintenance of genome integrity (Marechal and Zou, 2013; Awasthi et al., 2015). Furthermore, the "PI3K-AKT-mTOR signaling pathway and therapeutic opportunities" was shown to be over-represented in the set of DEGs; increased activity of PI3K in the radioresistant cells of this transcriptome dataset was also observed by Gray et al. (2019). Phosphatidylinositol-3-kinase (PI3K)/AKT/mammalian target of rapamycin (PI3K/AKT/mTOR) signaling is critical to many aspects of tumor cell growth and survival (Porta et al., 2014) and therefore could be likely involved in the survival of irradiated cancer cells.

Differential expression analysis was also carried out of the irradiated COAD cells in the GSE97543 dataset (Emons et al., 2017). A total of six samples were compared; three biological replicates for the wild-type (radiosensitive) cell lines and three for the radioresistant cells (**Supplementary Table 1**). The notch signaling pathway is significantly over-represented (**Figure 3**) in the DEGs of this dataset (**Supplementary Table 2**). Notch signaling is suggested to confer a selective survival advantage on tumors (Capaccione and Pine, 2013). Hence, the Notch network could be implicated in the resistance and survival of the COAD cells to irradiation. Regarding the over-represented DDR/R pathways, one pathway is related to processing IR-induced DNA lesions through ATR signaling and the other to mismatch repair (MMR), which is responsible for detecting and repairing mismatched nucleotides (Iyer et al., 2006; Larrea et al., 2010). MMR reaction is initiated by binding of the MSH2 (MutS homolog 2)/MSH6 heterodimer to the mismatched DNA; both *MSH2* and *MSH6* were found differentially expressed in the radioresistant COAD cells. The MutS homologs MSH2 and MSH6 form a heterodimer that binds to short insertion/deletion

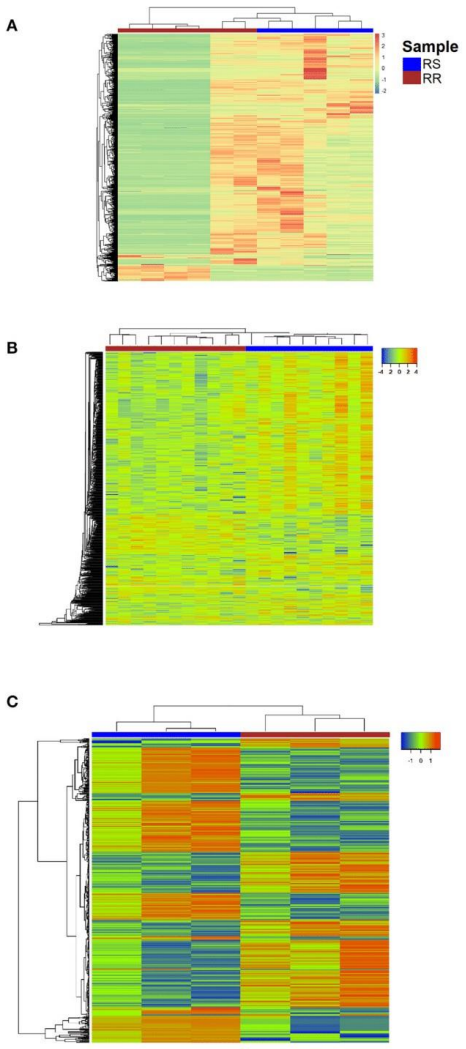


FIGURE 2 | Heatmap representing color-coded expression levels of DEGs for the datasets **(A)** GSE120798 (breast cancer), **(B)** GSE13280 (ALL), **(C)** GSE97543 (COAD); columns correspond to samples and rows correspond to genes. RR, Radioresistant; RS, Radiosensitive.

DNA mispairs (Habraken et al., 1996; Edelbrock et al., 2013). The MMR proteins also function in signaling DNA damage (Duckett et al., 1996b; Modrich, 1997). Earlier studies have shown that MSH2 is also involved in the processing of the biologically significant clustered DNA damages, as well as the execution of apoptosis induced by IR (Holt et al., 2009).

In GSE13280 (Marston et al., 2009), the gene expression profiles of radioresistant and radiosensitive ALL cell lines (eleven replicates for each condition) were compared (**Supplementary Table 1**) for the identification of DEGs (**Supplementary Table 2**). The enriched transforming growth factor beta (TGFB) signaling pathway (**Figure 3**), like the Notch network (Capaccione and Pine, 2013), is implicated in several aspects of cancer initiation, promotion, and progression (Syed, 2016). Hence, the Notch- and TGFB-mediated signaling pathways might render ALL cells less vulnerable to IR-induced apoptosis by exerting their cellular pro-survival effect.

Identification of Cancer Cell Radioresistance–Related Genes

The procedure we followed to identify an optimal number of cancer cell radioresistance–related genes is illustrated in **Supplementary Figure 1**. A list of 175 bio molecules (**Supplementary Table 3**) was proposed in a previous study by Pavlopoulou et al. (2017) to be implicated in tumor cell radioresistance. These genes/proteins were manually collected through a comprehensive and thorough literature search. The DEGs identified in each of the three transcriptome datasets were merged; a list of 7,185 genes was compiled (**Supplementary Table 3**). Those genes were compared to the literature-derived molecules, and 88 genes were found in common (**Supplementary Table 3**). In order to identify an optimal number of genes implicated in radioresistance, we performed functional enrichment analysis of these genes. The pathways over-represented in the 88 genes (**Figure 4**) are related to DDR/R, similar to those detected in the DEGs of the individual datasets (**Figure 3**), as well as to apoptosis. Signaling pathways mediated by the cardinal tumor suppressor TP53 are also related to radioresistance (**Figure 4**). Collectively, 36 genes were found to be implicated in cancer-associated biological pathways listed in **Table 1**, 26 of those up-regulated and 10 down-regulated in RR cancer cells. The products of the 36 genes also form a highly connected network (**Figure 5**), with high-confidence interactions, suggesting that these proteins associate, either functionally or physically, to confer cellular resistance to IR. Therefore, we propose 36 interconnected pivotal genes, henceforth referred to as "radiogenes," which participate in radioresistance-relevant pathways and mechanisms.

Differentially Expressed Radiogenes in Cancer vs. Normal Tissue

Collectively, 19 radiogenes, found to be differentially expressed in the three radioresistant vs. the radiosensitive cancer cell lines (**Table 1**), are consistently deregulated in their corresponding cancer-matched normal tissues with the same direction (i.e., up- or down-regulated) (**Supplementary Figure 2**). This finding could be utilized in personalized tumor treatment for the selective eradication of cancer cells by applying radiotherapy without harming the adjacent healthy tissue at the same time.

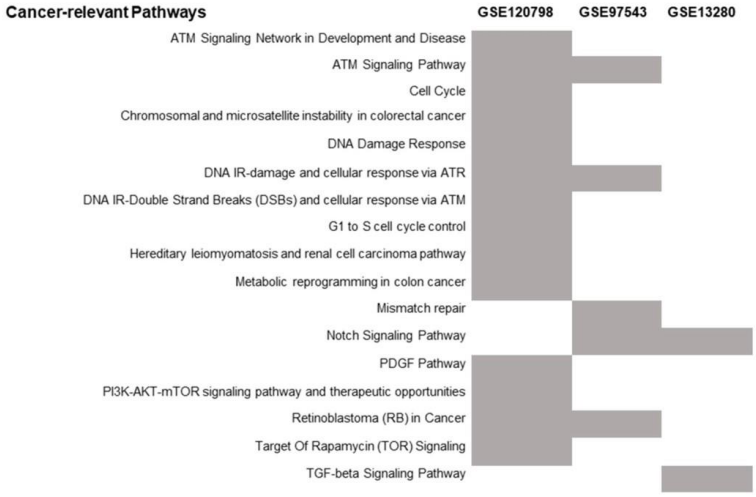


FIGURE 3 | Distribution of the over-represented cancer-related WikiPathways in the DEGs of each transcriptome dataset. The enriched pathways are indicated by gray.

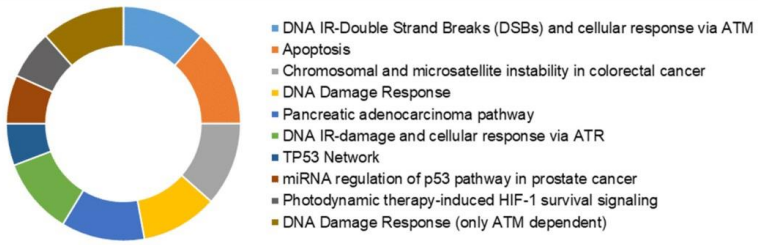


FIGURE 4 | Donut chart depicting the over-represented cancer-relevant pathways across the 88 common genes.

Radiogenes Are Potential Cancer Prognostic Markers

The differential expression of the radiogenes can also predict the clinical outcome of cancer patients. In particular, a statistically significant association was found between *CHEK1*, *MAP2K1*, and *PLK1* overexpression and worse overall survival in breast invasive carcinoma patients, indicated by pooled hazard ratio (HR) values higher than 1 and *p*-values lower than 0.05. Conversely, a significant relationship was observed between high expression of *NFKBIA*, which is otherwise underexpressed in radioresistant breast cancer cells, and favorable prognosis, indicated by an HR value < 1 (**Supplementary Figure 3**).

DISCUSSION

Cancer cells confer resistance to irradiation through diverse mechanisms including enhanced DNA damage repair capacity, activation of cell survival signaling pathways, inhibition of apoptotic pathways, and induced autophagy.

DDR/R is a complex entangled process consisting of recognition (or detection, or sensing), signaling, and repair of DNA damage (Rouse and Jackson, 2002). Ionizing radiation usually generates a variety of DNA lesions, including abasic (apurinic/apyrimidinic) sites, oxidized bases, crosslinks between thymine and cytosine bases, DNA single-strand breaks (SSBs), and DNA double-strand breaks (DSBs) (Sutherland et al., 2000). In the case of SSBs, only one strand of the double stranded DNA helix is severed. SSBs are recognized and processed by base excision repair (BER) and nucleotide excision repair (NER) mechanisms (Caldecott, 2008). DSB is the most detrimental type of DNA lesions since both strands of the double helix are broken. DSBs are repaired mainly through homologous recombination (HR) if cells are present in the S/G2 cell-cycle phase or, the less accurate, non-homologous end joining (NHEJ) (Budman and Chu, 2005). Those different types of DNA lesions can be formed

TABLE 1 | Gene symbol and expression status of the 36 differentially expressed radiogenes (radioresistant vs. radiosensitive cancer cells).

Gene	Expression status	Cell line
AKT1	Up	BC
ATM	Up	BC
BAX	Down	BC; ALL
BBC3	Down	BC
BCL2	Up	BC
BIRC5	Up	BC
BRCA1	Up	BC
BRCA2	Up	BC
CASP3	Down	BC
CCND1	Up	BC
CHEK1	UP	BC
EGLN1	Up	BC
HIF1A	Up	BC; ALL
JAK1	Down	BC
JUN	Up	BC
MAP2K1	UP	BC
MAP2K2	Up	BC
MCL1	Up	BC
MSH2	Up	BC; COAD
NBN	Up	BC
NFKB1	Up	BC
NFKBIA	Down	BC
PARP1	Up	BC
PLK1	Up	BC
PMAIP1	Down	BC
PRKDC	Up	BC
PTEN	Down	BC
RELA	Up	BC
RNF8	Up	BC
SOD2	Up	BC; ALL
STAT1	Down	BC
STAT3	Up	BC
TERF2	Up	BC
TP53	Down	BC
UBE2D3	Down	ALL
XIAP	Up	BC

ALL, acute lymphoblastic leukemia; BC, breast cancer; COAD, colorectal adenocarcinoma.
Those radiogenes found to be differentially expressed between TCGA-derived cancer tissue and corresponding GTEx matched normal tissue are italicized.

either separately or in close vicinity (a few nm), resulting in clustered DNA lesions or locally multiply damaged sites (MDS). Clustered DNA damage, the hallmark of IR, is considered the most severe type of genomic damage because of its complexity. This complex DNA damage includes both DSBs and non-DSBs, usually referred to as oxidatively-clustered DNA lesions (OCDLs), occurring within a DNA region of 15–20 bp. The corresponding DNA damage repair mechanisms are recruited by the cell in response to clustered damaged sites. However, harnessing the corresponding DNA repair machinery to the clustered damaged sites is quite challenging since the presence of a lesion in one strand can delay significantly the simultaneous processing of another lesion on the complementary strand. Furthermore, OCDLs can be rapidly converted into *de novo* DSBs by a DNA glycosylase during the repair. As a result, unrepaired MDS can lead to increasing levels of genotoxic damage, triggering also systemic responses (Nikitaki et al., 2016). Clustered DNA lesions, if not properly processed, could contribute to increased genomic instability in the form of chromosomal aberrations (e.g., deletions and inversions) and microsatellite instability, leading eventually to carcinogenesis. Therefore, the induction of clustered DNA damage increases the cytotoxic effect of radiation, especially in highly proliferating cancer cells. Radioresistant cancer cells though can counteract this effect through their ability to respond to and repair complex/clustered DNA damage more efficiently without avoiding necessarily increased genomic instability, as compared to their radiosensitive counterparts (Hada and Georgakilas, 2008; Georgakilas et al., 2013; Mavragani et al., 2017, 2019; Bukowska and Karwowski, 2018). The radiogenes identified in the present study are implicated in diverse mechanisms underlying the acquisition of radioresistance in cancer cells.

Among the 36 "radiogenes", *ATM*, *BRCA1/2*, *CHEK1*, *CCND1*, *MSH2*, *NBN*, *PARP1*, *PLK1*, *PRKDC,* and *RNF8*, which are consistently up-regulated across the radioresistant cancer cell lines (**Table 1**) and the corresponding cancer tissue (**Supplementary Figure 2**), are crucially implicated in different stages of DDR/R. In particular, ATM plays a protagonistic role in the initial stage of DDR/R, that is, DNA damage detection and stress-response signaling (Iliakis et al., 2003; Yang et al., 2004). ATM signaling is activated by a wide variety of DNA lesions, as well as DNA replication stress (Marechal and Zou, 2013; Burger et al., 2019). Mutations in ATM result in the genetic disorder ataxia-telangiectasia (AT), which is characterized by the high sensitivity of AT patients to IR and cancer predisposition (Gatti et al., 1988). Checkpoint kinase 1 (CHEK1), which acts downstream of ATM, is a core regulator of cell cycle checkpoint signaling in DNA damage response (Flaggs et al., 1997; Patil et al., 2013). Cyclin D1 (CCND1) was demonstrated to induce post-DNA damage cell cycle arrest and apoptosis in different types of cancers (Cai et al., 2012; Smith et al., 2016). PRKDC is a serine/threonine DNA-activated protein kinase involved in DSB recognition and DNA damage repair through NHEJ (Soubeyrand et al., 2003). Notably, ATM and PRKDC were found to affect greatly cancer cell response to IR through genome-wide genetic screening in a recent study by Francica et al. (2020). Moreover, BRCA1 and BRCA2 are largely involved in cell cycle control and maintenance of genomic stability in response to DNA damage (Deng, 2006; Gudmundsdottir and Ashworth, 2006). Another radiogene, *NBN (nibrin)*, encodes one of the three components of the MRN complex (MRE11A-RAD50-NBN), which is implicated in the recognition and repair of DSBs (Lamarche et al., 2010). NBN is mutated in patients with Nijmegen breakage syndrome (NBS), and cells from NBS patients are hypersensitive to IR (Taalman et al., 1983). In addition, a functional link between ATM and NBN proteins has been demonstrated by Zhao et al. (2000). Also, the MMR protein MSH2 (MutS homolog 2) is

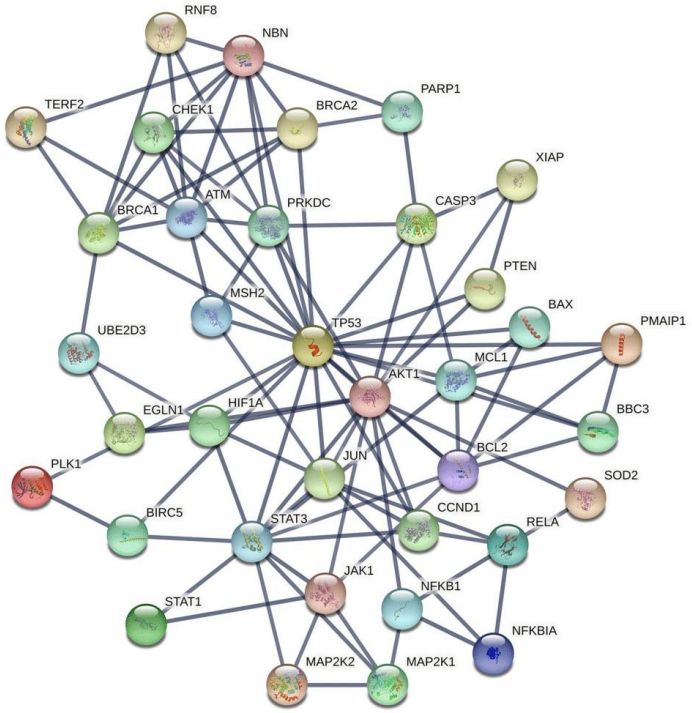


FIGURE 5 | STRING interaction network of the products of the 36 radiogenes. The nodes represent proteins and the edges indicate different modes of interactions with a confidence score ≥ 0.9.

suggested to contribute to radioresistance via SSB processing (Li et al., 2016). Moreover, RNF8 (ring finger protein 8) protein catalyzes the mono-ubiquitination of histones H2A and H2B during DNA damage, thereby facilitating DNA damage repair and activation of cell cycle checkpoint (Kolas et al., 2007; Ma et al., 2011); RNF8 is associated with radioresistance in human nasopharyngeal cancer cells (Wang et al., 2015). The protein encoded by the radiogene *PLK1 (Polo-like kinase 1)* is involved in cell cycle resumption following DNA damage-induced checkpoint arrest (Hyun et al., 2014; Bruinsma et al., 2017). It has been demonstrated that inhibition of *PLK1* renders glioblastoma and non-small cell lung cancer cells sensitive to IR (Pezuk et al., 2013; Van den Bossche et al., 2019). Of particular note, pharmacological inhibitors of the BER enzyme PARP1 [poly(ADP-ribose) polymerase 1], such as niraparib, olaparib, and rucaparib, are widely used for targeted cancer therapy (Sulai and Tan, 2018; Patel et al., 2020). Importantly, *CHEK1* and *PLK1* were found to be poor prognostic biomarkers for the survival of breast cancer patients (**Supplementary Figure 3**), further supporting that enhanced DNA damage repair mechanisms in cancer cells play a catastaltic role in efficient radiotherapy (Pavlopoulou et al., 2016, 2017). Therefore, DDR/R might represent a primary "danger" signal, leading eventually to the activation of downstream signaling cascades and pro-survival mechanisms (Nikitaki et al., 2015).

The apoptotic pathway is also affected by the cellular response to radiation-induced genomic damage in cancers, as we suggest in this study. Radioresistant cancer cells have developed the ability to evade apoptosis prompted by their response to extreme and repair-resistant DNA damage, mainly due to deregulation of key pro-apoptotic molecules like TP53 (Fridman and Lowe, 2003; Haupt et al., 2003), PTEN (Lu et al., 2016), PMAIP1 (mediator of damage-induced p53-mediated apoptosis) (Oda et al., 2000), BBC3 (TP53-upregulated modulator of apoptosis) (Han et al., 2001; Nakano and Vousden, 2001), BAX (BCL2-associated X protein) (Pawlowski and Kraft, 2000), as well as anti-apoptotic proteins like MCL1 (Fujise et al., 2000; Glaser et al., 2012), BCL2 (Akl et al., 2014), BIRC5 (Chiou et al., 2003), and XIAP (X-linked inhibitor of apoptosis) (Duckett et al., 1996a; Deveraux and Reed, 1999). In a consistent manner, in this study, the pro-apoptotic genes were shown to be up-regulated, whereas the anti-apoptotic genes are down-regulated

in radioresistant cancer cells (**Table 1**). Of those, BCL2, which is down-regulated in radioresistant breast cancer cells and tissue (**Table 1**, **Supplementary Figure 2**), can suppress apoptosis by inhibiting the activity of caspases indispensable for apoptosis, such as caspase-3 (CASP3) (Porter and Janicke, 1999; Swanton et al., 1999). Of note, aberrant activation of the Notch signaling pathway was demonstrated to inhibit TP53-mediated damage response and promote breast carcinogenesis by preventing apoptosis (Stylianou et al., 2006), suggesting a link between pro-survival and apoptotic mechanisms. PTEN was also found to promote apoptosis and cell cycle arrest via PI3K/AKT-dependent and -independent signaling (Weng et al., 2001).

The radiogene *NFKB1*, which transactivates several pro-inflammatory genes (Liu et al., 2017), was found to be overexpressed in radioresistant breast cancer cells and tissue (**Table 1**, **Supplementary Figure 2**). Furthermore, the members of the NFKB1 family, NFKBIA and RELA, can act either as inducers or inhibitors of apoptosis, respectively (Sonenshein, 1997; Barkett and Gilmore, 1999), consistent with their expression status in RR cells (**Table 1**; over- and under-expressed). The well-known NFKB1 inhibitor alpha (NFKBIA) was shown to be a favorable predictor in breast cancer patients (**Supplementary Figure 3**). Moreover, NFKB1, together with the inflammatory factors HIF1A (hypoxia-inducible factor 1) and STAT3, both of which were found to be up-regulated in radioresistant breast cancer cells (**Table 1**) and HIF1A overexpressed in breast cancer tissue (**Supplementary Figure 2**), are critically implicated in cancer radioresistance and radiation-induced inflammatory responses (Multhoff and Radons, 2012). On the other hand, STAT1, found down-regulated in the same cell lines (**Table 1**), as opposed to STAT3, elicits pro-apoptotic and anti-proliferative responses and promotes anti-cancer immunity (Avalle et al., 2012).

Suppression of *UBE2D3*, which is down-regulated in radioresistant ALL cells (**Table 1**), was demonstrated in a study by Wang et al. (2013) to decrease radiosensitivity of human breast cancer cells by promoting telomere maintenance. In addition, UBE2D3 is negatively correlated with TERF2 (telomeric repeat binding factor 2), the latter of which is primarily involved in telomere maintenance (Kim et al., 2009) and is down-regulated in radioresistance (**Table 1**).

Radiation can also exert its genotoxic and cytotoxic effects through the indirect and systemic induction of severe oxidative stress and the production of ROS in the organism (Kryston et al., 2011). The radiogene *HIF1A* plays a pivotal cytoprotective role against oxidative stress (Li et al., 2019) by inhibiting autophagy and cell death (Pouyssegur et al., 2006). Moreover, the superoxide scavenging enzyme SOD2 (superoxide dismutase 2), found overexpressed both in radioresistant breast cancer and ALL cells (**Table 1**), at normal expression levels provides a cytoprotective effect. Thus, SOD2 can likely exert a protective effect on RR cancer cells by controlling potential ROS-mediated DNA damage via catalyzing the reduction of superoxide into less genotoxic molecules like oxygen (Wang et al., 2018).

Notably, in this study, pro-survival pathways (**Figure 3**) like Notch signaling, were found to be implicated both in solid and blood cancers, probably by mediating survival of cancer cells to radiation-induced clustered/complex DNA damage (Marston et al., 2009; Capaccione and Pine, 2013). Moreover, the PI3K/AKT/mTOR signaling pathway is suggested to be important in regulating cell survival in response to different types of cellular stress (Hung et al., 2012; Porta et al., 2014), including genotoxic stress. Hence, pro-survival pathways could be considered as potential therapeutic targets in cancer (Porta et al., 2014).

A major limitation of this study, particularly for the two solid cancers datasets, is the lack of patient-derived tumor tissue, as well as cancer stem cells, which constitute a subpopulation in solid tumors that display stem-like characteristics (Pavlopoulou et al., 2016). Instead, the respective experimental studies relied on the use of commercial cancer cell lines; of particular note, the ALL cells were derived from "real" patients. However, the extent to which the individual cancer cell lines can capture the cellular and genomic complexity of tumors is questioned. Further research is needed to determine whether the results derived from the cancer cell lines investigated in this study or other studies could be extrapolated to their corresponding tissues of origin, like breast tumor tissue and colorectal adenocarcinoma. Nevertheless, it is suggested that large panels of cancer cell lines can faithfully capture the genomic heterogeneity of cancers (Sakellaropoulos et al., 2019; Vougas et al., 2019). Beyond the discussed limitations, the over- or under-expressions of the radiogenes in radioresistant phenotypes have been verified, to a great extent, by independent experimental biochemical studies available in the literature. In addition, the clinical implications of those genes are further supported by the patient survival results (**Supplementary Figure 3**).

CONCLUSION

In the present study, we employed an integrative bioinformatics approach to analyze transcriptomic data regarding the molecular determinants of cancer cell radioresistance. On the basis of our findings, both solid and hematologic cancer cells likely depend on similar mechanisms to confer resistance to IR (i.e., DDR/R and cell survival). Moreover, we identified 36 functionally associated radiogenes that participate in radioresistance-associated pathways. Most of those radiogenes were also shown to be differentially regulated in the corresponding cancer tissues. Moreover, several of the radiogenes were found to have potential prognostic value for the clinical outcome of cancer patients. However, the availability of clinically derived cancer tissues would provide a more reliable source for conducting research on the response of cancer patients to radiation. The overall data presented herein can be particularly useful for clinicians in selecting suitable targets (e.g., DDR/R inhibitors) for appropriate combination therapy using IR. In conclusion, we suggest that this bioinformatics premise can be harnessed as a first step in the rational design of *in vivo* experimental studies or in personalized medicine for optimizing tumor response and cancer cell susceptibility to therapeutic ionizing radiation and

reduction of the total effective radiation dose administered to the patient.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**; further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

AP conceived the study. AP and AGG designed and supervised the study. HIT, GK, PK, and AP performed the experiments and analyzed the data. HIT, GK, PK, HA, AGG, and AP wrote the manuscript. All authors reviewed and approved of the final manuscript.

FUNDING

HIT acknowledges YÖK (Yükseköğretim Kurulu) and TÜBİTAK 2211-C scholarship. PK was co-financed by Greece and the European Union (European Social Fund- ESF) through the Operational Programme "Human Resources Development, Education and Lifelong Learning" in the context of the project "Reinforcement of Postdoctoral Researchers - 2 nd Cycle" (MIS-5033021), implemented by the State Scholarships Foundation (IKY).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2021.620248/full#supplementary-material

Supplementary Figure 1 | Schematic workflow outlining the selection of radiogenes. Venn diagram comparing the high-throughput (HTP) and literature-derived cancer radioresistance-associated genes. Functional enrichment and interaction network analysis of the common 88 genes resulted in 36 genes. Red, up-regulated genes; Green, down-regulated genes.

Supplementary Figure 2 | Differential expression patterns of radiogenes in the corresponding cancer tissue; BRCA, breast invasive carcinoma; COAD, colon adenocarcinoma; LAML, acute myeloid leukemia. The red box represents cancer tissue samples and the gray box indicates normal samples.

Supplementary Figure 3 | Kaplan-Meier curves depicting the prognostic value of radiogenes for overall survival in invasive breast carcinoma **(A)** *CHEK1*, **(B)** *PLK1*, **(C)** *MAP2K1* and **(D)** *NFKBIA*. The HR "HR(high)" and the corresponding p-values "p(HR)" are shown. The 95% confidence intervals (CI) are denoted by dotted lines. The number of high-risk and low-risk patients are indicated by "n(high)" and "n(low)," respectively.

Supplementary Table 1 | Samples from each transcriptomic dataset analyzed in this study.

Supplementary Table 2 | List of the differentially expressed genes (DEGs) of each of the three transcriptome datasets and the DEGs common among datasets. Log_2FC of radioresistant vs. radiosensitive differentially expressed genes in GSE120798 (breast cancer) cell lines; cut-off $|log_2FC \geq 2|$.

Supplementary Table 3 | List of the merged DEGs from each dataset, literature-derived genes/proteins and common molecules.

REFERENCES

Akl, H., Vervloessem, T., Kiviluoto, S., Bittremieux, M., Parys, J. B., De Smedt, H., et al. (2014). A dual role for the anti-apoptotic Bcl-2 protein in cancer: mitochondria versus endoplasmic reticulum. *Biochim. Biophys. Acta* 1843, 2240–2252. doi: 10.1016/j.bbamcr.2014.04.017

Alnasir, J., and Shanahan, H. P. (2015). Investigation into the annotation of protocol sequencing steps in the sequence read archive. *Gigascience* 4:23. doi: 10.1186/s13742-015-0064-7

Avalle, L., Pensa, S., Regis, G., Novelli, F., and Poli, V. (2012). STAT1 and STAT3 in tumorigenesis: a matter of balance. *JAKSTAT* 1, 65–72. doi: 10.4161/jkst.20045

Awasthi, P., Foiani, M., and Kumar, A. (2015). ATM and ATR signaling at a glance. *J. Cell Sci.* 128, 4255–4262. doi: 10.1242/jcs.169730

Barkett, M., and Gilmore, T. D. (1999). Control of apoptosis by Rel/NF-kappaB transcription factors. *Oncogene* 18, 6910–6924. doi: 10.1038/sj.onc.1203238

Barrett, T., Wilhite, S. E., Ledoux, P., Evangelista, C., Kim, I. F., Tomashevsky, M., et al. (2013). NCBI GEO: archive for functional genomics data sets–update. *Nucleic Acids Res.* 41, D991–D995. doi: 10.1093/nar/gks1193

Baskar, R., Lee, K. A., Yeo, R., and Yeoh, K. W. (2012). Cancer and radiation therapy: current advances and future directions. *Int. J. Med. Sci.* 9, 193–199. doi: 10.7150/ijms.3635

Begg, A. C., Stewart, F. A., and Vens, C. (2011). Strategies to improve radiotherapy with targeted drugs. *Nat. Rev. Cancer* 11, 239–253. doi: 10.1038/nrc3007

Beheshti, A., McDonald, J. T., Miller, J., Grabham, P., and Costes, S. V. (2019). GeneLab database analyses suggest long-term impact of space radiation on the cardiovascular system by the activation of FYN through reactive oxygen species. *Int. J. Mol. Sci.* 20:661. doi: 10.3390/ijms20030661

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* 57, 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x

Braschi, B., Denny, P., Gray, K., Jones, T., Seal, R., Tweedie, S., et al. (2019). Genenames.org: the HGNC and VGNC resources in 2019. *Nucleic Acids Res.* 47, D786–D792. doi: 10.1093/nar/gky930

Bruinsma, W., Aprelia, M., Garcia-Santisteban, I., Kool, J., Xu, Y. J., and Medema, R. H. (2017). Inhibition of Polo-like kinase 1 during the DNA damage response is mediated through loss of Aurora A recruitment by Bora. *Oncogene* 36, 1840–1848. doi: 10.1038/onc.2016.347

Budman, J., and Chu, G. (2005). Processing of DNA for nonhomologous end-joining by cell-free extract. *EMBO J.* 24, 849–860. doi: 10.1038/sj.emboj.7600563

Bukowska, B., and Karwowski, B. T. (2018). The clustered DNA lesions - types, pathways of repair and relevance to human health. *Curr. Med. Chem.* 25, 2722–2735. doi: 10.2174/0929867325666180226110502

Burger, K., Ketley, R. F., and Gullerova, M. (2019). Beyond the trinity of ATM, ATR, and DNA-PK: multiple kinases shape the DNA damage response in concert with RNA metabolism. *Front. Mol. Biosci.* 6:61. doi: 10.3389/fmolb.2019.00061

Cai, C. K., Zhao, G. Y., Tian, L. Y., Liu, L., Yan, K., Ma, Y. L., et al. (2012). miR-15a and miR-16-1 downregulate CCND1 and induce apoptosis and cell cycle arrest in osteosarcoma. *Oncol. Rep.* 28, 1764–1770. doi: 10.3892/or.2012.1995

Caldecott, K. W. (2008). Single-strand break repair and genetic disease. *Nat. Rev. Genet.* 9, 619–631. doi: 10.1038/nrg2380

Capaccione, K. M., and Pine, S. R. (2013). The Notch signaling pathway as a mediator of tumor survival. *Carcinogenesis* 34, 1420–1430. doi: 10.1093/carcin/bgt127

Chiou, S. K., Jones, M. K., and Tarnawski, A. S. (2003). Survivin - an anti-apoptosis protein: its biological roles and implications for cancer and beyond. *Med. Sci. Monit.* 9, PI25–PI29.

Clough, E., and Barrett, T. (2016). The gene expression omnibus database. *Methods Mol. Biol.* 1418, 93–110. doi: 10.1007/978-1-4939-3578-9_5

Delaney, G., Jacob, S., Featherstone, C., and Barton, M. (2005). The role of radiotherapy in cancer treatment: estimating optimal utilization from a review of evidence-based clinical guidelines. *Cancer* 104, 1129–1137. doi: 10.1002/cncr.21324

Deng, C. X. (2006). BRCA1: cell cycle checkpoint, genetic instability, DNA damage response and cancer evolution. *Nucleic Acids Res.* 34, 1416–1426. doi: 10.1093/nar/gkl010

Deveraux, Q. L., and Reed, J. C. (1999). IAP family proteins–suppressors of apoptosis. *Genes Dev.* 13, 239–252. doi: 10.1101/gad.13.3.239

Duckett, C. S., Nava, V. E., Gedrich, R. W., Clem, R. J., Van Dongen, J. L., Gilfillan, M. C., et al. (1996a). A conserved family of cellular genes related to the baculovirus iap gene and encoding apoptosis inhibitors. *EMBO J.* 15, 2685–2694 doi: 10.1002/j.1460-2075.1996.tb00629.x

Duckett, D. R., Drummond, J. T., Murchie, A. I., Reardon, J. T., Sancar, A., Lilley, D. M., et al. (1996b). Human MutSalpha recognizes damaged DNA base pairs containing O6-methylguanine, O4-methylthymine, or the cisplatin-d(GpG) adduct. *Proc. Natl. Acad. Sci. U. S. A.* 93, 6443–6447. doi: 10.1073/pnas.93.13.6443

Edelbrock, M. A., Kaliyaperumal, S., and Williams, K. J. (2013). Structural, molecular and cellular functions of MSH2 and MSH6 during DNA mismatch repair, damage signaling and other noncanonical activities. *Mutat. Res.* 743–744, 53–66. doi: 10.1016/j.mrfmmm.2012.12.008

Efron, B., and Tibshirani, R. (1993). *An Introduction to the Bootstrap*. Boca Raton, FL: Chapman & Hall/CRC.

Emons, G., Spitzner, M., Reineke, S., Moller, J., Auslander, N., Kramer, F., et al. (2017). Chemoradiotherapy resistance in colorectal cancer cells is mediated by Wnt/beta-catenin signaling. *Mol. Cancer Res.* 15, 1481–1490. doi: 10.1158/1541-7786.MCR-17-0205

Flaggs, G., Plug, A. W., Dunks, K. M., Mundt, K. E., Ford, J. C., Quiggle, M. R., et al. (1997). Atm-dependent interactions of a mammalian chk1 homolog with meiotic chromosomes. *Curr. Biol.* 7, 977–986. doi: 10.1016/S0960-9822(06)00417-9

Francica, P., Mutlu, M., Blomen, V. A., Oliveira, C., Nowicka, Z., Trenner, A., et al. (2020). Functional radiogenetic profiling implicates ERCC6L2 in non-homologous end joining. *Cell Rep.* 32:108068. doi: 10.1016/j.celrep.2020.108068

Fridman, J. S., and Lowe, S. W. (2003). Control of apoptosis by p53. *Oncogene* 22, 9030–9040. doi: 10.1038/sj.onc.1207116

Fujise, K., Zhang, D., Liu, J., and Yeh, E. T. (2000). Regulation of apoptosis and cell cycle progression by MCL1. Differential role of proliferating cell nuclear antigen. *J. Biol. Chem.* 275, 39458–39465. doi: 10.1074/jbc.M006626200

Gatti, R. A., Berkel, I., Boder, E., Braedt, G., Charmley, P., Concannon, P., et al. (1988). Localization of an ataxia-telangiectasia gene to chromosome 11q22-23. *Nature* 336, 577–580. doi: 10.1038/336577a0

Georgakilas, A. G., O'Neill, P., and Stewart, R. D. (2013). Induction and repair of clustered DNA lesions: what do we know so far? *Radiat. Res.* 180, 100–109. doi: 10.1667/RR3041.1

Glaser, S. P., Lee, E. F., Trounson, E., Bouillet, P., Wei, A., Fairlie, W. D., et al. (2012). Anti-apoptotic Mcl-1 is essential for the development and sustained growth of acute myeloid leukemia. *Genes Dev.* 26, 120–125. doi: 10.1101/gad.182980.111

Gray, M., Turnbull, A. K., Ward, C., Meehan, J., Martinez-Perez, C., Bonello, M., et al. (2019). Development and characterisation of acquired radioresistant breast cancer cell lines. *Radiat. Oncol.* 14:64. doi: 10.1186/s13014-019-1268-2

Gudmundsdottir, K., and Ashworth, A. (2006). The roles of BRCA1 and BRCA2 and associated proteins in the maintenance of genomic stability. *Oncogene* 25, 5864–5874. doi: 10.1038/sj.onc.1209874

Habraken, Y., Sung, P., Prakash, L., and Prakash, S. (1996). Binding of insertion/deletion DNA mismatches by the heterodimer of yeast mismatch repair proteins MSH2 and MSH3. *Curr. Biol.* 6, 1185–1187. doi: 10.1016/S0960-9822(02)70686-6

Hada, M., and Georgakilas, A. G. (2008). Formation of clustered DNA damage after high-LET irradiation: a review. *J. Radiat. Res.* 49, 203–210. doi: 10.1269/jrr.07123

Han, J., Flemington, C., Houghton, A. B., Gu, Z., Zambetti, G. P., Lutz, R. J., et al. (2001). Expression of bbc3, a pro-apoptotic BH3-only gene, is regulated by diverse cell death and survival signals. *Proc. Natl. Acad. Sci. U. S. A.* 98, 11318–11323. doi: 10.1073/pnas.201208798

Haupt, S., Berger, M., Goldberg, Z., and Haupt, Y. (2003). Apoptosis - the p53 network. *J. Cell Sci.* 116 (Pt 20), 4077–4085. doi: 10.1242/jcs.00739

Holt, S. M., Scemama, J. L., Panayiotidis, M. I., and Georgakilas, A. G. (2009). Compromised repair of clustered DNA damage in the human acute lymphoblastic leukemia MSH2-deficient NALM-6 cells. *Mutat. Res.* 674, 123–130. doi: 10.1016/j.mrgentox.2008.09.014

Hung, C. M., Garcia-Haro, L., Sparks, C. A., and Guertin, A. D. (2012). mTOR-dependent cell survival mechanisms. *Cold Spring Harb. Perspect. Biol.* 4:a008771. doi: 10.1101/cshperspect.a008771

Hyun, S. Y., Hwang, H. I., and Jang, Y. J. (2014). Polo-like kinase-1 in DNA damage response. *BMB Rep.* 47, 249–255. doi: 10.5483/BMBRep.2014.47.5.061

Iliakis, G., Wang, Y., Guan, J., and Wang, H. (2003). DNA damage checkpoint control in cells exposed to ionizing radiation. *Oncogene* 22, 5834–5847. doi: 10.1038/sj.onc.1206682

Iyer, R. R., Pluciennik, A., Burdett, V., and Modrich, P. L. (2006). DNA mismatch repair: functions and mechanisms. *Chem. Rev.* 106, 302–323. doi: 10.1021/cr0404794

Kanakoglou, D. S., Michalettou, T. D., Vasileiou, C., Gioukakis, E., Maneta, D., Kyriakidis, K. V., et al. (2020). Effects of high-dose ionizing radiation in human gene expression: a meta-analysis. *Int. J. Mol. Sci.* 21:1938. doi: 10.3390/ijms21061938

Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360. doi: 10.1038/nmeth.3317

Kim, H., Lee, O. H., Xin, H., Chen, L. Y., Qin, J., Chae, H. K., et al. (2009). TRF2 functions as a protein hub and regulates telomere maintenance by recognizing specific peptide motifs. *Nat. Struct. Mol. Biol.* 16, 372–379. doi: 10.1038/nsmb.1575

Kolas, N. K., Chapman, J. R., Nakada, S., Ylanko, J., Chahwan, R., Sweeney, F. D., et al. (2007). Orchestration of the DNA-damage response by the RNF8 ubiquitin ligase. *Science* 318, 1637–1640. doi: 10.1126/science.1150034

Kontou, P. I., Pavlopoulou, A., and Bagos, P. G. (2018). Methods of analysis and meta-analysis for identifying differentially expressed genes. *Methods Mol. Biol.* 1793, 183–210. doi: 10.1007/978-1-4939-7868-7_12

Kryston, T. B., Georgiev, A. B., Pissis, P., and Georgakilas, A. G. (2011). Role of oxidative stress and DNA damage in human carcinogenesis. *Mutat. Res.* 711, 193–201. doi: 10.1016/j.mrfmmm.2010.12.016

Kutmon, M., Riutta, A., Nunes, N., Hanspers, K., Willighagen, E. L., Bohler, A., et al. (2016). WikiPathways: capturing the full diversity of pathway knowledge. *Nucleic Acids Res.* 44, D488–D494. doi: 10.1093/nar/gkv1024

Lamarche, B. J., Orazio, N. I., and Weitzman, M. D. (2010). The MRN complex in double-strand break repair and telomere maintenance. *FEBS Lett.* 584, 3682–3695. doi: 10.1016/j.febslet.2010.07.029

Larrea, A. A., Lujan, S. A., and Kunkel, T. A. (2010). SnapShot: DNA mismatch repair. *Cell* 141:730.e1. doi: 10.1016/j.cell.2010.05.002

Li, D. (2019). "Statistical methods for rna sequencing data analysis," in *Computational Biology*, ed H. Husi (Brisbane, QLD: Codon Publications).

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

Li, H. S., Zhou, Y. N., Li, L., Li, S. F., Long, D., Chen, X. L., et al. (2019). HIF-1alpha protects against oxidative stress by directly targeting mitochondria. *Redox Biol.* 25:101109. doi: 10.1016/j.redox.2019.101109

Li, Z., Pearlman, A. H., and Hsieh, P. (2016). DNA mismatch repair and the DNA damage response. *DNA Repair* 38, 94–101. doi: 10.1016/j.dnarep.2015.11.019

Liao, Y., Wang, J., Jaehnig, E. J., Shi, Z., and Zhang, B. (2019). WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res.* 47, W199–W205. doi: 10.1093/nar/gkz401

Liu, T., Zhang, L., Joo, D., and Sun, S. C (2017). NF-kappaB signaling in inflammation. *Signal Transduct. Target Ther.* 2:17023. doi: 10.1038/sigtrans.2017.23

Lu, X. X., Cao, L. Y., Chen, X., Xiao, J., Zou, Y., and Chen, Q. (2016). PTEN inhibits cell proliferation, promotes cell apoptosis, and induces cell cycle arrest via downregulating the PI3K/AKT/hTERT pathway in lung adenocarcinoma A549 cells. *Biomed. Res. Int.* 2016:2476842. doi: 10.1155/2016/2476842

Ma, T., Keller, J. A., and Yu, X. (2011). RNF8-dependent histone ubiquitination during DNA damage response and spermatogenesis. *Acta Biochim. Biophys. Sin.* 43, 339–345. doi: 10.1093/abbs/gmr016

Marechal, A., and Zou, L. (2013). DNA damage sensing by the ATM and ATR kinases. *Cold Spring Harb. Perspect. Biol.* 5:a012716. doi: 10.1101/cshperspect.a012716

Marioni, J. C., Mason, C. E., Mane, S. M., Stephens, M., and Gilad, Y. (2008). RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res.* 18, 1509–1517. doi: 10.1101/gr.079558.108

Marston, E., Weston, V., Jesson, J., Maina, E., McConville, C., Agathanggelou, A., et al. (2009). Stratification of pediatric ALL by *in vitro* cellular responses to DNA double-strand breaks provides insight into the molecular mechanisms underlying clinical response. *Blood* 113, 117–126. doi: 10.1182/blood-2008-03-142950

Mavragani, I. V., Nikitaki, Z., Kalospyros, S. A., and Georgakilas, A. G. (2019). Ionizing radiation and complex DNA damage: from prediction to detection challenges and biological significance. *Cancers* 11:1789. doi: 10.3390/cancers11111789

Mavragani, I. V., Nikitaki, Z., Souli, M. P., Aziz, A., Nowsheen, S., Aziz, K., et al. (2017). Complex DNA damage: a route to radiation-induced genomic instability and carcinogenesis. *Cancers* 9:91. doi: 10.3390/cancers9070091

Metzker, M. L. (2010). Sequencing technologies - the next generation. *Nat. Rev. Genet.* 11, 31–46. doi: 10.1038/nrg2626

Mikkelsen, R. B., and Wardman, P. (2003). Biological chemistry of reactive oxygen and nitrogen and radiation-induced signal transduction mechanisms. *Oncogene* 22, 5734–5754. doi: 10.1038/sj.onc.1206663

Modrich, P. (1997). Strand-specific mismatch repair in mammalian cells. *J. Biol. Chem.* 272, 24727–24730. doi: 10.1074/jbc.272.40.24727

Multhoff, G., and Radons, J. (2012). Radiation, inflammation, and immune responses in cancer. *Front. Oncol.* 2:58. doi: 10.3389/fonc.2012.00058

Nakano, K., and Vousden, K. H. (2001). PUMA, a novel proapoptotic gene, is induced by p53. *Mol. Cell* 7, 683–694. doi: 10.1016/S1097-2765(01)00214-3

Nikitaki, Z., Mavragani, I. V., Laskaratou, D. A., Gika, V., Moskvin, V. P., Theofilatos, K., et al. (2016). Systemic mechanisms and effects of ionizing radiation: a new "old" paradigm of how the bystanders and distant can become the players. *Semin. Cancer Biol.* 37–38, 77–95. doi: 10.1016/j.semcancer.2016.02.002

Nikitaki, Z., Michalopoulos, I., and Georgakilas, A. G. (2015). Molecular inhibitors of DNA repair: searching for the ultimate tumor killing weapon. *Future Med. Chem.* 7, 1543–1558. doi: 10.4155/fmc.15.95

Oda, E., Ohki, R., Murasawa, H., Nemoto, J., Shibue, T., Yamashita, T., et al. (2000). Noxa, a BH3-only member of the Bcl-2 family and candidate mediator of p53-induced apoptosis. *Science* 288, 1053–1058. doi: 10.1126/science.288.5468.1053

Patel, M., Nowsheen, S., Maraboyina, S., and Xia, F. (2020). The role of poly(ADP-ribose) polymerase inhibitors in the treatment of cancer and methods to overcome resistance: a review. *Cell Biosci.* 10:35. doi: 10.1186/s13578-020-00390-7

Patil, M., Pabla, N., and Dong, Z. (2013). Checkpoint kinase 1 in DNA damage response and cell cycle regulation. *Cell. Mol. Life Sci.* 70, 4009–4021. doi: 10.1007/s00018-013-1307-3

Pavlopoulou, A., Bagos, P. G., Koutsandrea, V., and Georgakilas, A. G. (2017). Molecular determinants of radiosensitivity in normal and tumor tissue. A bioinformatic approach. *Cancer Lett.* 403, 37–47. doi: 10.1016/j.canlet.2017.05.023

Pavlopoulou, A., Oktay, Y., Vougas, K., Louka, M., Vorgias, C. E., and Georgakilas, A. G. (2016). Determinants of resistance to chemotherapy and ionizing radiation in breast cancer stem cells. *Cancer Lett.* 380, 485–493. doi: 10.1016/j.canlet.2016.07.018

Pawlowski, J., and Kraft, A. S. (2000). Bax-induced apoptotic cell death. *Proc. Natl. Acad. Sci. U. S. A.* 97, 529–531. doi: 10.1073/pnas.97.2.529

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T. C., Mendell, J. T., and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33, 290–295. doi: 10.1038/nbt.3122

Pezuk, J. A., Brassesco, M. S., Morales, A. G., de Oliveira, J. C., de Oliveira, H. F., Scrideli, C. A., et al. (2013). Inhibition of polo-like kinase 1 induces cell cycle arrest and sensitizes glioblastoma cells to ionizing radiation. *Cancer Biother. Radiopharm.* 28, 516–522. doi: 10.1089/cbr.2012.1415

Porta, C., Paglino, C., and Mosca, A. (2014). Targeting PI3K/Akt/mTOR signaling in cancer. *Front. Oncol.* 4:64. doi: 10.3389/fonc.2014.00064

Porter, A. G., and Janicke, R. U. (1999). Emerging roles of caspase-3 in apoptosis. *Cell Death Differ.* 6, 99–104. doi: 10.1038/sj.cdd.4400476

Pouyssegur, J., Dayan, F., and Mazure, N. M. (2006). Hypoxia signalling in cancer and approaches to enforce tumour regression. *Nature* 441, 437–443. doi: 10.1038/nature04871

Ramasamy, A., Mondry, A., Holmes, C. C., and Altman, D. G. (2008). Key issues in conducting a meta-analysis of gene expression microarray datasets. *PLoS Med.* 5:e184. doi: 10.1371/journal.pmed.0050184

Robinson, M. D., McCarthy, D. J., and Smyth, K. G. (2010). edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi: 10.1093/bioinformatics/btp616

Robinson, M. D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 11:R25. doi: 10.1186/gb-2010-11-3-r25

Rouse, J., and Jackson, S. P. (2002). Interfaces between the detection, signaling, and repair of DNA damage. *Science* 297, 547–551. doi: 10.1126/science.1074740

Sakellaropoulos, T., Vougas, K., Narang, S., Koinis, F., Kotsinas, A., Polyzos, A., et al. (2019). A deep learning framework for predicting response to therapy in cancer. *Cell Rep.* 29, 3367–3373.e4. doi: 10.1016/j.celrep.2019.11.017

Sayers, E. W., Beck, J., Brister, J. R., Bolton, E. E., Canese, K., Comeau, D. C., et al. (2020). Database resources of the national center for biotechnology information. *Nucleic Acids Res.* 48, D9–D16. doi: 10.1093/nar/gkz899

Schoenhals, J. E., Seyedin, S. N., Tang, C., Cortez, M. A., Niknam, S., Tsouko, E., et al. (2016). Preclinical rationale and clinical considerations for radiotherapy plus immunotherapy: going beyond local control. *Cancer J.* 22, 130–137. doi: 10.1097/PPO.0000000000000181

Smith, D., Mann, D., and Yong, K. (2016). Cyclin D type does not influence cell cycle response to DNA damage caused by ionizing radiation in multiple myeloma tumours. *Br. J. Haematol.* 173, 693–704. doi: 10.1111/bjh.13982

Sonenshein, G. E. (1997). Rel/NF-kappa B transcription factors and the control of apoptosis. *Semin. Cancer Biol.* 8, 113–119. doi: 10.1006/scbi.1997.0062

Soubeyrand, S., Pope, L., Pakuts, B., and Hache, R. J. (2003). Threonines 2638/2647 in DNA-PK are essential for cellular resistance to ionizing radiation. *Cancer Res.* 63, 1198–1201.

Spratt, D. E., and Speers, C. (2019). RadioGx: a new preclinical tool to model intrinsic radiosensitivity. *Cancer Res.* 79, 6076–6078. doi: 10.1158/0008-5472.CAN-19-3277

StataCorp (2013). *Stata Statistical Software: Release 13.* College Station, TX: StataCorp LP.

Stylianou, S., Clarke, R. B., and Brennan, K. (2006). Aberrant activation of notch signaling in human breast cancer. *Cancer Res.* 66, 1517–1525. doi: 10.1158/0008-5472.CAN-05-3054

Sulai, N. H., and Tan, A. R. (2018). Development of poly(ADP-ribose) polymerase inhibitors in the treatment of BRCA-mutated breast cancer. *Clin. Adv. Hematol. Oncol.* 16, 491–501.

Sutherland, B. M., Bennett, P. V., Sidorkina, O., and Laval, J. (2000). Clustered DNA damages induced in isolated DNA and in human cells by low doses of ionizing radiation. *Proc. Natl. Acad. Sci. U. S. A.* 97, 103–108. doi: 10.1073/pnas.97.1.103

Swanton, E., Savory, P., Cosulich, S., Clarke, P., and Woodman, P. (1999). Bcl-2 regulates a caspase-3/caspase-2 apoptotic cascade in cytosolic extracts. *Oncogene* 18, 1781–1787. doi: 10.1038/sj.onc.1202490

Syed, V. (2016). TGF-beta signaling in cancer. *J. Cell Biochem.* 117, 1279–1287. doi: 10.1002/jcb.25496

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., et al. (2019). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 47, D607–D613. doi: 10.1093/nar/gky1131

Taalman, R. D., Jaspers, N. G., Scheres, J. M., de Wit, J., and Hustinx, T. W. (1983). Hypersensitivity to ionizing radiation, *in vitro*, in a new chromosomal breakage disorder, the Nijmegen breakage syndrome. *Mutat. Res.* 112, 23–32. doi: 10.1016/0167-8817(83)90021-4

Tang, C., Wang, X., Soh, H., Seyedin, S., Cortez, M. A., Krishnan, S., et al. (2014). Combining radiation and immunotherapy: a new systemic therapy for solid tumors? *Cancer Immunol. Res.* 2, 831–838. doi: 10.1158/2326-6066.CIR-14-0069

Tang, Z., Li, C., Kang, B., Gao, G., Li, C., and Zhang, Z. (2017). GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res.* 45, W98–W102. doi: 10.1093/nar/gkx247

Unger, K. (2014). Integrative radiation systems biology. *Radiat. Oncol.* 9:21. doi: 10.1186/1748-717X-9-21

Van den Bossche, J., Domen, A., Peeters, M., Deben, C., De Pauw, I., Jacobs, J., et al. (2019). Radiosensitization of non-small cell lung cancer cells by the Plk1 inhibitor volasertib is dependent on the p53 status. *Cancers* 11:1893. doi: 10.3390/cancers11121893

Vougas, K., Sakellaropoulos, T., Kotsinas, A., Foukas, G. P., Ntargaras, A., Koinis, F., et al. (2019). Machine learning and data mining frameworks for predicting drug response in cancer: an overview and a novel *in silico* screening process based on association rule mining. *Pharmacol. Ther.* 203:107395. doi: 10.1016/j.pharmthera.2019.107395

Wang, M., Chen, X., Chen, H., Zhang, X., Li, J., Gong, H., et al. (2015). RNF8 plays an important role in the radioresistance of human nasopharyngeal cancer cells *in vitro*. *Oncol. Rep.* 34, 341–349. doi: 10.3892/or.2015.3958

Wang, W., Yang, L., Hu, L., Li, F., Ren, L., Yu, H., et al. (2013). Inhibition of UBE2D3 expression attenuates radiosensitivity of MCF-7 human breast cancer cells by increasing hTERT expression and activity. *PLoS ONE* 8:e64660. doi: 10.1371/journal.pone.0064660

Wang, Y., Branicky, R., Noe, A., and Hekimi, S. (2018). Superoxide dismutases: dual roles in controlling ROS damage and regulating ROS signaling. *J. Cell Biol.* 217, 1915–1928. doi: 10.1083/jcb.201708007

Weng, L., Brown, J., and Eng, C. (2001). PTEN induces apoptosis and cell cycle arrest through phosphoinositol-3-kinase/Akt-dependent and -independent pathways. *Hum. Mol. Genet.* 10, 237–242. doi: 10.1093/hmg/10.3.237

Yamamori, T., Yasui, H., Yamazumi, M., Wada, Y., Nakamura, Y., Nakamura, H., et al. (2012). Ionizing radiation induces mitochondrial reactive oxygen species production accompanied by upregulation of mitochondrial electron transport chain function and mitochondrial content under control of the cell cycle checkpoint. *Free Radic. Biol. Med.* 53, 260–270. doi: 10.1016/j.freeradbiomed.2012.04.033

Yang, J., Xu, Z. P., Huang, Y., Hamrick, H. E., Duerksen-Hughes, P. J., and Yu, Y. N. (2004). ATM and ATR: sensing DNA damage. *World J. Gastroenterol.* 10, 155–160. doi: 10.3748/wjg.v10.i2.155

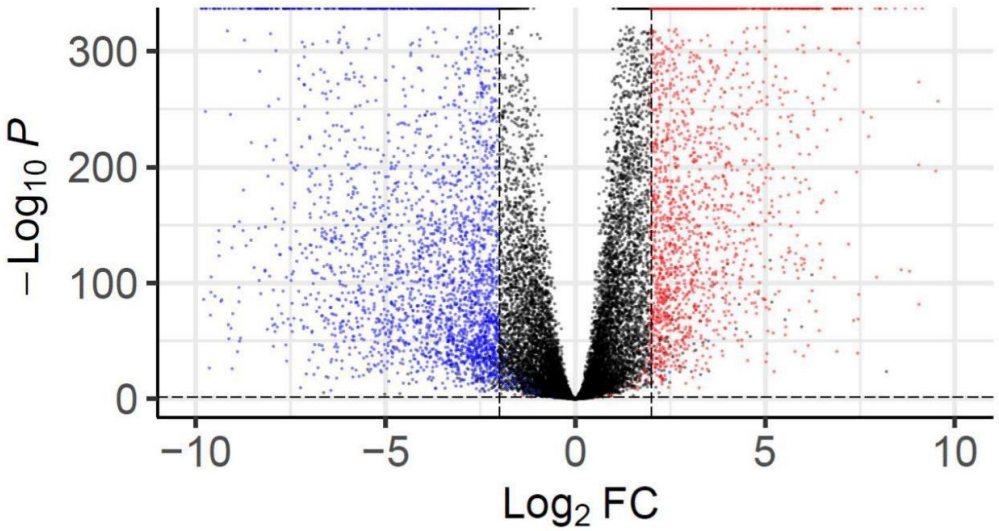
Zhang, B., Kirov, S., and Snoddy, J. (2005). WebGestalt: an integrated system for exploring gene sets in various biological contexts. *Nucleic Acids Res.* 33, W741–W748. doi: 10.1093/nar/gki475

Zhao, S., Weng, Y. C., Yuan, S. S., Lin, Y. T., Hsu, H. C., Lin, S. C., et al. (2000). Functional link between ataxia-telangiectasia and Nijmegen breakage syndrome gene products. *Nature* 405, 473–477. doi: 10.1038/35013083

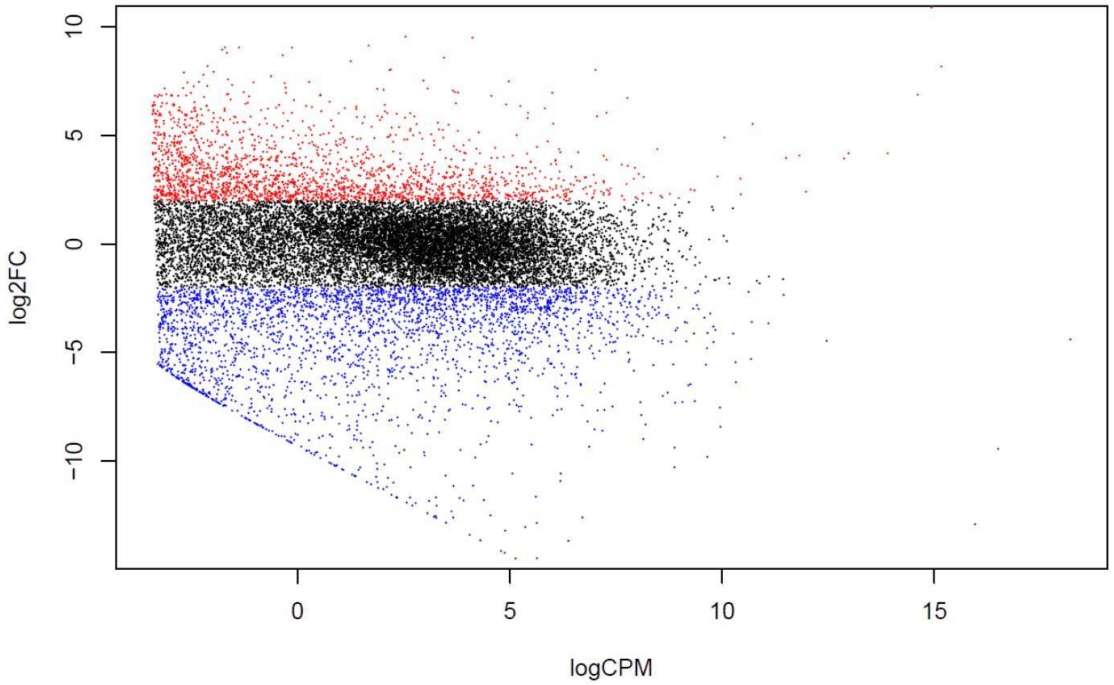
Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Toy, Karakülah, Kontou, Alotaibi, Georgakilas and Pavlopoulou. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

(A)


300
200
100
0
−Log10 P
−10
−5
0
5
10
Log2 FC

(B)


10
5
0
−5
−10
log2FC
0
5
10
15
logCPM

International Journal of
Molecular Sciences

MDPI

Article

In Silico Methods for the Identification of Diagnostic and Favorable Prognostic Markers in Acute Myeloid Leukemia

Hande Yılmaz [1,2,3,†], Halil Ibrahim Toy [1,2,†], Stephan Marquardt [3], Gökhan Karakülah [1,2], Can Küçük [1,2,4], Panagiota I. Kontou [5], Stella Logotheti [3,*] and Athanasia Pavlopoulou [1,2,*]

1 Izmir Biomedicine and Genome Center, Balcova, 35340 Izmir, Turkey; hande.yilmaz@med.uni-rostock.de (H.Y.); ibrahim.toy@msfr.ibg.edu.tr (H.I.T.); gokhan.karakulah@ibg.edu.tr (G.K.); can.kucuk@ibg.edu.tr (C.K.)
2 Izmir International Biomedicine and Genome Institute, Dokuz Eylül University, Balcova, 35340 Izmir, Turkey
3 Institute of Experimental Gene Therapy and Cancer Research, Rostock University Medical Center, 18057 Rostock, Germany; stephan.marquardt@uni-rostock.de
4 Department of Medical Biology, Faculty of Medicine, Dokuz Eylül University, Balcova, 35340 Izmir, Turkey
5 Department of Computer Science and Biomedical Informatics, University of Thessaly, 35131 Lamia, Greece; pankontou@gmail.com
* Correspondence: Styliani.Logotheti@med.uni-rostock.de (S.L.); athanasia.pavlopoulou@ibg.edu.tr (A.P.)
† These authors have contributed equally to this work and share first authorship.

Abstract: Acute myeloid leukemia (AML), the most common type of acute leukemia in adults, is mainly asymptomatic at early stages and progresses/recurs rapidly and frequently. These attributes necessitate the identification of biomarkers for timely diagnosis and accurate prognosis. In this study, differential gene expression analysis was performed on large-scale transcriptomics data of AML patients versus corresponding normal tissue. Weighted gene co-expression network analysis was conducted to construct networks of co-expressed genes, and detect gene modules. Finally, hub genes were identified from selected modules by applying network-based methods. This robust and integrative bioinformatics approach revealed a set of twenty-four genes, mainly related to cell cycle and immune response, the diagnostic significance of which was subsequently compared against two independent gene expression datasets. Furthermore, based on a recent notion suggesting that molecular characteristics of a few, unusual patients with exceptionally favorable survival can provide insights for improving the outcome of individuals with more typical disease trajectories, we defined groups of long-term survivors in AML patient cohorts and compared their transcriptomes versus the general population to infer favorable prognostic signatures. These findings could have potential applications in the clinical setting, in particular, in diagnosis and prognosis of AML.

Keywords: acute myeloid leukemia; transcriptomics; clinical traits; bioinformatics; long-term survivors; minimal residual disease; diagnostics; prognostics

check for updates

Citation: Yılmaz, H.; Toy, H.I.; Marquardt, S.; Karakülah, G.; Küçük, C.; Kontou, P.I.; Logotheti, S.; Pavlopoulou, A. In Silico Methods for the Identification of Diagnostic and Favorable Prognostic Markers in Acute Myeloid Leukemia. *Int. J. Mol. Sci.* **2021**, *22*, 9601. https://doi.org/10.3390/ijms22179601

Academic Editors: Ronald Moura and Sergio Crovella

Received: 16 July 2021
Accepted: 20 August 2021
Published: 5 September 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Acute myeloid leukemia (AML) is a complex and heterogeneous blood condition, and the most prevalent form of acute leukemia in adults, representing 32% of all adult leukemia cases [1]. The average age of AML diagnosis is 67 years [2]. AML is caused by the uncontrolled clonal proliferation of abnormal immature white blood cells or leukemic blasts. Since blasts accumulate in the bone marrow, peripheral blood, and other organs of the body as the disease progresses, the production of normal hematopoietic stem cells is prevented. If AML is not treated within one year of diagnosis, it progresses quickly, leading to excessive bleeding, life-threatening infections, infiltration of lymph nodes with leukemic cells, or metastases to distant organs [3–5].

In the standard practice, AML diagnosis and the decision on the suitable therapeutic strategy rely on the medical history and physical examination of each patient in combination with testing blood samples and bone marrow aspirates and biopsies. Laboratory

tests for AML diagnosis include a series of procedures, such as a complete blood count and differential count, bone marrow biopsy and genetic testing, imaging or lumbar puncture, immunophenotyping, and cytogenetics [6]. AML patients are treated with intensive chemotherapy, targeted therapy, or bone marrow transplantation, which have overall improved survival outcomes, especially in younger patients [7]. However, despite the advances in the understanding of AML pathophysiology, mortality rates remain high. For example, in 2020, there were an estimated 60,530 new leukemia cases and 23,100 people died of this disease in the United States [8]. Long-term survival rates also appear to be low, and most older patients with AML die from disease relapse [9–11]. The poor outcomes reflect two major challenges in AML management. First, most cases arise without any detectable early symptoms and patients usually present with acute complications of bone marrow failure [12]. Second, complete hematologic remission (CR), which is a prerequisite for AML cure, cannot be readily achieved, since a significant percentage of AML patients eventually experience relapse after a seemingly successful treatment. This is because, although the therapeutic interventions can eliminate the majority of malignant cells, a few tumor cells inevitably evade therapy and remain persistently in the patient for years after the initial diagnosis [13]. This condition is termed minimal residual disease (MRD), and is a major cause of AML relapse [14]. These obstacles in AML treatment highlight the urgent need for identification of markers for early diagnosis, accurate prognosis, and estimation of CR.

The rapid progress in '-omics' and high-throughput technologies has promoted the identification of novel genetic abnormalities and holds promise for accelerating the systematic characterization of AML cancer genomes. Gene expression profiling of AML patients has catalyzed the discovery of novel leukemia subgroups and of prognostic signatures [6]. The increasing availability of high-throughput data from blood and bone marrow of AML patients creates unprecedented opportunities for the comprehensive identification of diagnostic biomarkers of AML which could complement and update the existing ones [15–17]. Moreover, despite the poor outcomes of AML, there are a few fortunate cases that manage to survive over exceptionally extended periods of time, perhaps reflecting a remarkable resistance to MRD and disease progression. A novel rationale proposes that patients with exceptionally prolonged survival rates present distinct molecular profiles [18,19] that could be exploited to unveil favorable prognostic factors that differentiate them from patients with a more typical course of disease [20]. Understanding the biological determinants of survival in these extreme outliers may provide a route to improving responses in more typical patients, particularly if these studies identify new biomarkers to guide drug selection or novel pathways that are targetable. Seeking determinants of prolonged survival is particularly important in cancers with generally poor outcomes [20], such as AML. Taking these trends into account, in the present study, we applied integrative bioinformatics approaches to process, analyze, and interpret publicly available AML-relevant transcriptome data towards the identification of biomarkers for diagnosis and favorable prognosis of AML. Our diagnostic pipelines unveiled a signature of 24 differentially expressed genes, mainly involved in the cell cycle and immune response, that discriminate AML patients from normal controls and could be considered as candidate diagnostic biomarkers. It also enabled us to reveal a consensus group of genes that are consistently deregulated in survival outliers.

2. Results

In this study, large-scale transcriptomic data were exploited towards the identification of genes that are differentially expressed between bone marrow AML samples and corresponding normal whole blood specimens. Given that genes with similar expression patterns (i.e., co-expressed genes) are usually implicated in similar biological processes, weighted gene co-expression network analysis (WGCNA) of the differentially expressed genes (DEGs) was performed [21], so as to identify modules (i.e., clusters of densely connected co-expressed genes) in AML with similar expression patterns. The genes contained

in those modules were further correlated to AML clinical traits. Afterwards, different network-based methods were applied on the genes contained in these modules, so as to select the most biologically important genes/gene products. This analysis revealed twenty-four biologically important genes/gene products, which were further compared against two independent gene expression datasets. Furthermore, we set up a pipeline for the characterization of molecular signatures of AML long-term survivors (LTSs) that could enable prediction of MRD and survival probability. To this end, we estimated the median overall survival (OS) of the AML patients, and defined several groups of LTSs whose OS survival largely deviates from the median value. The transcriptomic profiles of the LTSs were compared to those of short-term survivors (STSs) to detect DEGs, and functional enrichment analysis was further performed to functionally interpret those genes. The overall procedure is briefly illustrated in Figure 1.

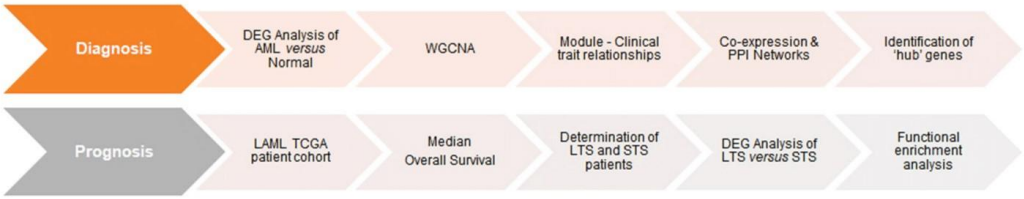


Figure 1. Flowchart diagram of the methodology employed in this study.

2.1. Identification of Diagnostic Markers in AML

2.1.1. Comprehensive Characterization of Consensus AML Transcriptomes Compared to Normal Blood

First, we sought to identify the changes in transcriptomes of leukemic cells versus publicly available normal control samples, such as blood or bone marrow. It should be taken into account that unlike blood sampling, which is a fast and minimally invasive technique, bone marrow biopsy and aspiration are more distressing [6]. This issue poses limitations to obtaining a large biobank from this type of normal tissue, leading to inadequate numbers of relative high-throughput data in public platforms, such as GTEx and GEO. Nevertheless, encouraged by the adequate number (n = 456) of normal whole blood samples with available RNA-seq data in GTEx, we envisaged that comparing the transcriptome of malignant cells to that of normal blood holds a potential to reveal transcripts expressed in circulating leukemic cells. These could be considered as circulatory biochemical markers for blood-based diagnosis via real-time PCR, which is a routine and less painful clinical practice as opposed to other invasive diagnostic techniques. To this end, by using a uniform pipeline (described in the Materials and Methods), gene expression profiles from TCGA and GTEx, 151 AML and 456 normal blood samples, respectively, were compiled. The distribution of TCGA and GTEx samples according to PCA is shown in Figure 2A. As expected, AML and normal samples form discrete clusters, suggesting distinct transcriptomic profiles for the datasets from AML and normal whole blood samples. The statistically significant DEGs obtained by analyzing the RNA-seq data from TCGA-LAML and GTEx include 2056 upregulated, 2959 downregulated genes from edgeR (Figure S1); 1847 upregulated, 3108 downregulated genes from limma (Figure S1); and 1714 upregulated, 2575 downregulated genes from DESEq2 (Figure 2B). There are a total of 3931 common DEGs from edgeR, limma, and DESeq2 (Figure 2C), including 1503 upregulated and 2428 downregulated genes (Table S1). The protein-coding genes constitute 85% of those genes. A Venn diagram was generated using the function *VennDiagram* in the 'VennDiagram' package (v.1.6.20). The bar plot was created with the function *barplot* in the 'graphics' package (v.4.0.4).

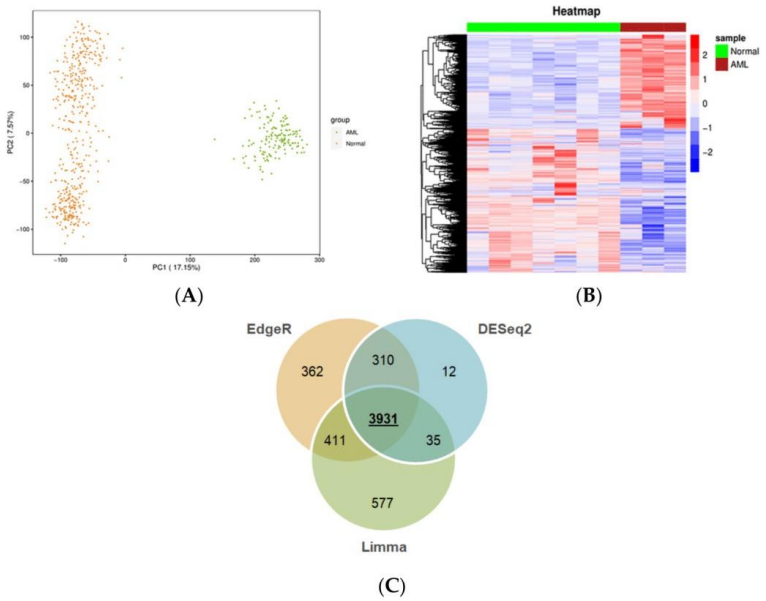
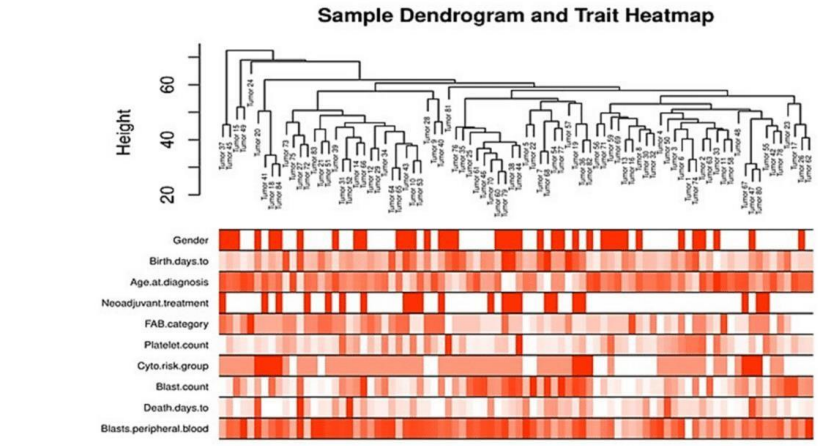


Figure 2. Genes differentially expressed between AML and normal samples. (**A**) Principal component analysis of TCGA and GTEx samples. Orange and green circles represent normal samples from GTEx and AML samples from TCGA, respectively. (**B**) Heatmap of the DEGs generated using DESeq2. Representative AML and normal samples are shown in maroon and green, respectively. Each row corresponds to a gene. The upregulated and the downregulated genes (based on Z-scoring) are shown in red and blue, respectively. (**C**) Venn diagram of the common and unique DEGs across edgeR, limma, and DESeq2.

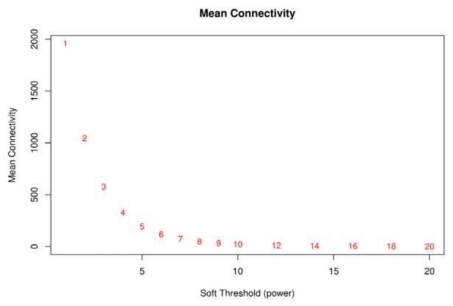
2.1.2. Weighted Gene Co-Expression Network Analysis of DEGs

Sample Clustering for Detecting Outliers

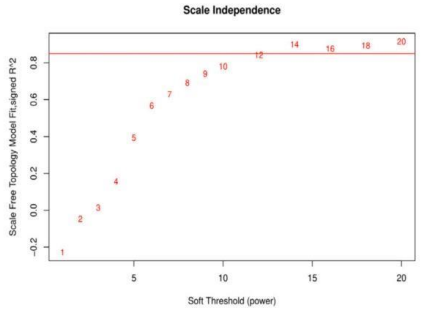
The co-expression network of the DEGs in AML was constructed by WGCNA. First, outlier genes and samples were removed. In this way, a total of 3794 genes and 84 AML samples were used to construct the gene co-expression network. To determine whether all 84 AML samples were suitable for network analysis, the sample dendrogram and corresponding clinical traits were analyzed. All samples were included in the clusters and passed the cutoff thresholds (Figure 3A).



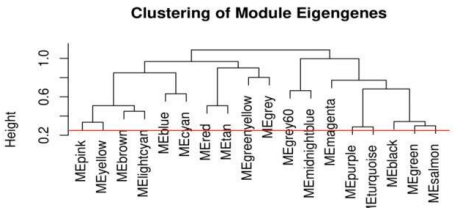
(**A**)



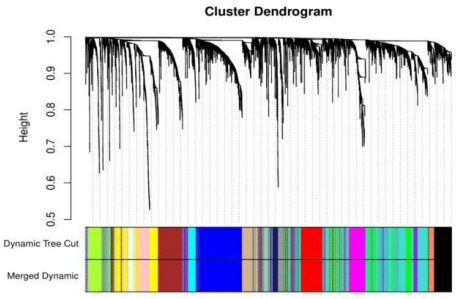
(**B**)



(**C**)

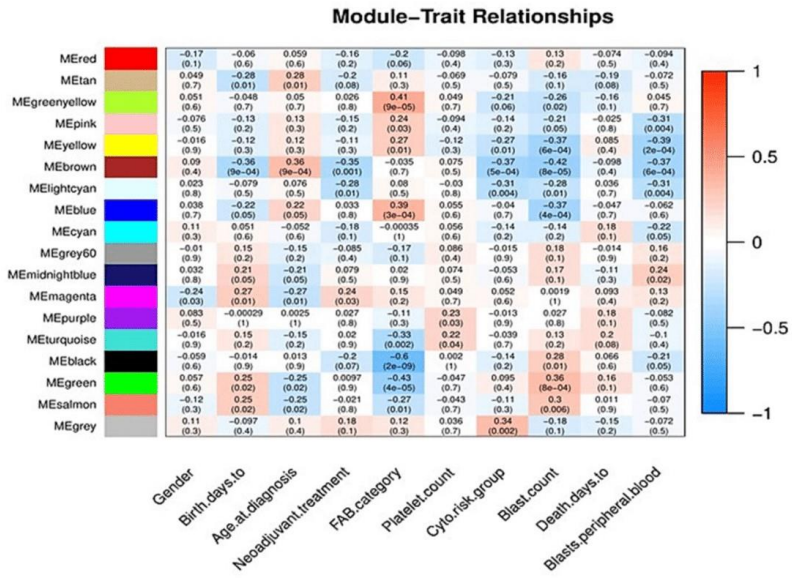


(**D**)



(**E**)

Figure 3. *Cont.*



(**F**)

Figure 3. Sample clustering and module detection in WGCNA. (**A**) Sample dendrogram and trait heatmap for detecting outliers. The clustering was based on the expression data of the 3794 DEGs of the AML samples. All 84 samples are located within clusters. Darker shades of red indicate stronger positive correlation, whereas lighter shades of red denote weaker correlation. (**B**,**C**) Selection of soft-thresholding power (β) based on scale-free fit index and mean connectivity for various soft-thresholding powers. (**D**) Clustering of module eigengenes. The red line indicates cutoff height of 0.25. (**E**) Dendrogram of DEGs produced by hierarchical clustering based on dissTOM (1-TOM). Modules are represented by different colors. The two colored rows below the dendrogram represent the original and merged modules. (**F**) Module–trait correlation heatmap for AML clinical traits and gene module eigengenes. Each cell contains the Pearson correlation coefficient and *p* value. Blue indicates negative correlation, red denotes positive correlation.

A soft-thresholding power (β) of 12 (scale free $R^2 = 0.8$) was chosen by applying the approximate scale-free network topology criterion (Figure 3B,C). A dendrogram of DEGs was generated by average linkage hierarchical clustering using a TOM-based dissimilarity measure, by setting the cut height at 0.25 for module merging and the β value at 12 (Figure 3D,E). Gene modules are represented by different colors. The branches in the dendrogram represent modules, and each leaf, which is a short vertical line, corresponds to a gene. A total of 17 gene co-expression modules were determined, labeled as turquoise (735), blue (512), brown (308), yellow (277), green (258), red (238), black (223), pink (218), magenta (202), purple (186), greenyellow (139), tan (134), salmon (81), cyan (76), midnightblue (70), lightcyan (69), and grey60 (58); the numbers in the parentheses indicate the number of genes per module. The unclustered genes (10) were assigned to the gray module.

Associations of Gene Modules with Clinicopathological Traits of AML

The association between the 17 gene modules and ten important clinicopathological and molecular features of AML patients was investigated. The clinical significance of the selected traits is described below.

- Gender: Men are more prone to leukemia than women [22].
- Chronological age: Elderly people are more prone to AML than young people [23].

- Age at diagnosis: The incidence of AML is almost three times higher in adults aged 50–54 years and thirteen times higher in those aged 70–74 years as compared to adults aged 20–24 years [24,25].
- Neoadjuvant treatment: Whether patients have previously received chemotherapy or not; patients who have previously received chemotherapy are more prone to AML [26].
- FAB category: Based on the type of cell from which leukemia developed and the degree of maturity of these cells. FAB category corresponds to the French–American–British classification scheme which is useful for classifying AML into subtypes from M0 (undifferentiated AML or AML with minimal differentiation) to M7 (acute megakaryoblastic leukemia, AMegL) [27].
- Platelet count: AML is characterized by low platelet count. Normal human blood includes 150,000–450,000 platelets per microliter. Platelet count decreases in AML [28,29].
- Cytogenetic risk group: Cytogenetic analysis is recognized as being the most important prognostic indicator in acute myeloid leukemia. According to the cytogenetic risk group, AML is divided into three groups: favorable risk, intermediate risk, and unfavorable risk. Cytogenetic tests help predict the response of cancer to treatment and allow physicians to design a more effective therapy [30,31].
- Blast count and peripheral blood blast: For the diagnosis and classification of AML, the percentage of peripheral blood (PB) and bone marrow (BM) blasts is especially important. BM blasts normally represent 1% to 5% of marrow cells. Generally, a percentage of 20% blasts is required for AML diagnosis. Most patients with AML have a higher percentage of BM blasts compared to PB blasts [32–34].
- Days to death: Related to overall survival, which is the survival time after initial diagnosis. Measuring the overall survival is required to assess how well a new treatment works in a clinical trial. In AML, the 5-year overall survival is less than 50%; regarding elderly patients, only 20% survive 2 years after diagnosis [35,36].

Pearson correlation coefficients (r) between modules' eigengenes and the clinical traits were calculated to detect important module–trait relationships. In our study, those modules with $|r| > 0.25$ and p value < 0.05 were considered to be statistically significant. As shown in Figure 3F, twelve modules (tan, greenyellow, pink, yellow, brown, lightcyan, blue, magenta, turquoise, black, green, and salmon) are significantly associated with clinical traits. The genes from each of those modules were selected for further analysis.

2.1.3. Reconstruction of AML Molecular Networks

The 3154 genes belonging to the twelve modules (Figure 3E) were used to construct co-expression and PPI networks. In the co-expression network, 90 genes with $\geq$17 connections were identified, and the top 121 genes were selected through cytoHubba by combining the output of eleven different algorithms (Figure 4A). Likewise, in the PPI network, 279 hubs and 153 important protein nodes were identified (Figure 4A). The nodes that correspond to genes or gene products, found to be common between co-expression and PPI networks, based on degree of connectivity and cytoHubba, are shown in Figure 4A. NPY1R was detected by using both approaches (Figure 4A). In this way, a total of 24 genes were found, of which *PPBP*, *PF4*, *CXCL12*, *BUB1B*, *LCK*, *S1PR5*, *EGFR*, *CENPI*, *KNL1*, *RHOJ*, *FGF2*, *EXO1*, *PLK4*, and *MCM10* genes are upregulated, and *NPY1R*, *BRCA1*, *CD3G*, *MYH6*, *KRT4*, *SPARCL1*, *CD3E*, *GALNT15*, *MMP1*, and *SNAP91* genes are downregulated (Figure 4B,C).

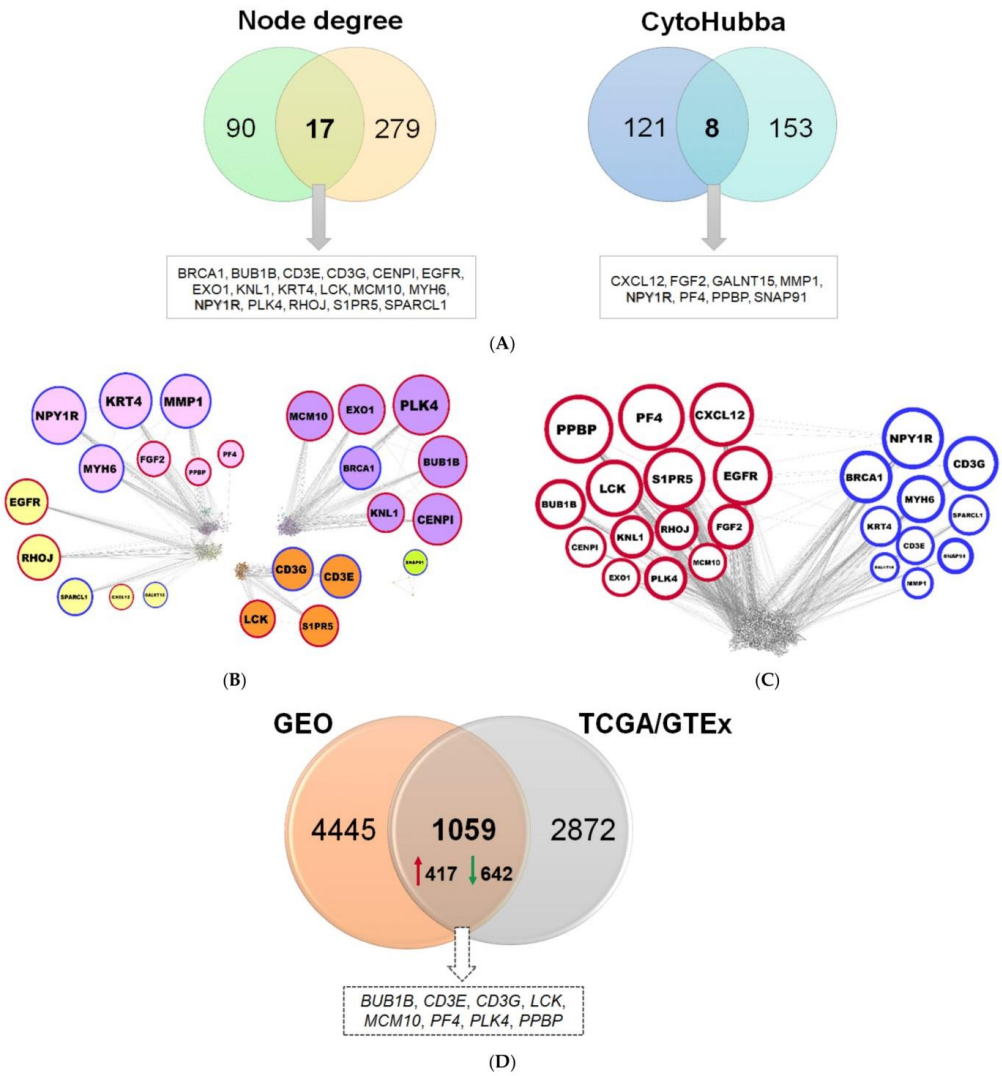


Figure 4. Important nodes identified through gene co-expression and PPI network. (**A**) Venn diagrams depicting the common nodes between the co-expression and PPI networks according to node degree and cytoHubba, respectively. (**B**) Co-expression network showing the 24 common genes, as distributed in pink, magenta, yellow, brown, and greenyellow modules. The nodes correspond to genes and the edges (connecting lines) represent the associations among genes. The size of the nodes is proportional to the number of their immediate connections (degree of connectivity). Red and blue border lines indicate up- and downregulation, respectively. (**C**) PPI network with the 24 common proteins. (**D**) Venn diagram for validation of diagnostic signatures, showing the common DEGs between the TCGA/GTEx and the two merged GEO datasets.

2.1.4. Comparison of Diagnostic Signatures with Independent AML Datasets

Those genes found to be differentially expressed between the TCGA AML and corresponding normal GTEx samples were further compared against two independent datasets obtained from NCBI GEO. The number of DEGs detected between AML and control (i.e., from matched healthy donors) samples for each NCBI GEO dataset is 793 (GSE138702) and 4780 (GSE9476) (Table S1), by applying stringent criteria. The DEGs found in each of the two transcriptome datasets were merged and the duplicates were removed; collectively, 5504 genes were compiled (Table S1). A rather significant overlap was found between the DEGs from TCGA/GTEx and those from the two GEO datasets; in particular, 1059 genes were found in common (representing 27%), eight of which constitute hubs (33%) (Figure 4D).

2.2. Inferring Favorable Prognostic Markers from the Transcriptomic Profiles of Long-Term Survivors of AML

Even after a seemingly successful treatment, a significant percentage of AML patients is at risk of disease recurrence [14], as a consequence of minimal residual disease. It has been estimated that of those patients who will achieve initial remission, approximately 50–70% of them will relapse within three years. Nevertheless, a lower percentage of adult AML patients survive to five years [37] while some infrequent cases with survival over ten years have also been reported [38]. In general, across several cancer types, there are rare cases of patients who manage to remain progression-free for atypically long durations of time, while in some rare cases patients never relapse [20,39]. Intriguingly, tumors of these unusual patients present distinct genetic and gene expression profiles as compared to the patients with a more typical course of disease [18,19]. Patients with exceptionally prolonged survival rates comprise a rare, but distinct, clinical subpopulation which holds the potential to reveal signatures of favorable prognostic factors that could be informative for the effective management of the general population of cancer patients [20]. According to these notions, the patients with unusually prolonged survival (hereafter termed long-term survivors, LTSs) in our study cohort might plausibly better cope with MRD and relapse and, therefore, their distinct transcriptomes, as compared to the general population, can be used to provide prognostic signatures associated with resistance to MRD.

To this end, in the cohort of AML patients of TCGA, we distributed the patients according to their OS and estimated the median OS which would reflect the typical course for the disease for the specific cohort. The median OS for the whole AML population was 366 days (n = 50 patients), while the standard deviation value was estimated as 600.5 days. Then, we defined populations of patients with exceptionally favorable prognosis within the study cohort in a time-dependent manner as the patients who survive i) longer than the median OS + 2 SD (>1567 days, n = 15 patients), or ii) longer than the median OS + 3 SD after initial diagnosis (2168 days, n = 4 patients). Given that in clinical practice the patients who are alive 5+ years past diagnosis are considered LTSs [40], we additionally included the patients who survive past this clinical milestone (1825 days, n = 7 patients) in our analysis (Figure 5A). We sought to identify the distinct tumor gene expression profiles across these LTSs versus those of the patients who are deceased by 366 days (Figure 5B). Our analysis revealed common deregulated genes between all three comparisons (Figure 5C). Subsequent gene set enrichment analysis (GSEA) showed that biological functions such as locomotion, cell migration, and cell junctions are enriched in the DEGs in each LTS subgroup. This is expected, since these traits are linked with a propensity of cancer cells to metastasize. Surprisingly, among the top significantly enriched processes, we also observed deregulations of genes that are involved in developmental processes, in particular neurogenesis, central nervous system development, neuron differentiation, and pattern specification (Figure 5D). We further identified a group of 131 DEGs that remain consistently deregulated in all three groups and, hence, could be important for maintaining survival over extended periods of time (Table S2).

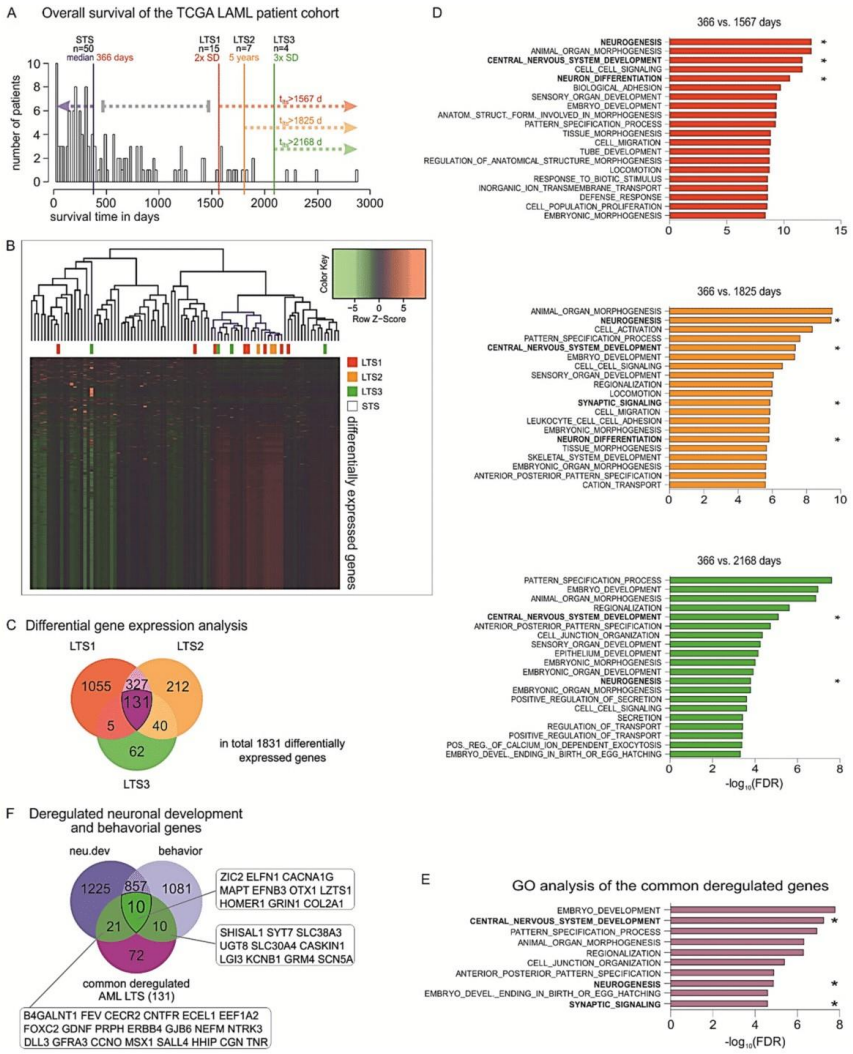


Figure 5. Identification of favorable prognostic factors based on transcriptomic profiles of LTSs. (**A**) Histogram of the overall survival (OS) of the TCGA LAML cohort with median (366 days, n = 50) and cutoffs for the three long-term survivor (LTS) groups: 2 × SD = 1567 days (n = 15), 5-year survival = 1825 days (n = 7), and 3 × SD = 2168 days (n = 4). (**B**) Heatmaps of the DEGs in the three LTS subpopulations versus the STSs. (**C**) Venn diagram of the deregulated genes of all three analyses showing large overlaps between analyses and 131 genes in common. (**D**) GSEA of GO biological functions associated with the DEGs in each LTS subgroup versus the patients with OS < 366 days. (**E**) Top 10 enriched GO biological functions for the 131 genes that are commonly deregulated across all LTS subgroups, ranked by order of significance (−log10 of FDR value). (**F**) Overlapping of the 131 genes with previously identified nervous system-related genes showed that 31.3% of these genes are indispensable for neuronal development, neurological function, or both (gene abbreviations are indicated in the respective boxes). * Nervous system-related pathways.

Strikingly, GSEA showed that the commonly deregulated genes are strongly involved in early and late stages of embryonic development, and particularly in neuronal develop-

ment and neurological function. In detail, the top 10 most significantly enriched biological processes were embryo development, central nervous system (CNS) development, pattern specification process, animal organ morphogenesis, regionalization, cell junction organization, neurogenesis, anterior–posterior pattern specification, synaptic signaling, and embryo development ending in birth or egg hatching (Figure 5E). Among the transcripts found upregulated in LTSs and involved in many of these processes were glial cell-derived neurotrophic factor (GDNF), a neurotrophic molecule that promotes neuronal differentiation and dopaminergic development [41], as well as several products of homeobox genes that are master regulators of development, morphogenesis, and cell differentiation, for example, distal-less homeobox 1 (DLX1), msh homeobox 1 (MSX1), and orthodenticle homeobox 1 (OTX1) [42]. We further considered the cytogenetic background of the patients of the LTS and STS subgroups in regard to the reciprocal translocation t(15;17), which is associated with decreased incidence of relapse and improved survival upon all-trans retinoic acid (ATRA)-containing treatments. In particular, six patients across the LTS subgroups versus six patients in the STS subgroup bear a t(15;17) translocation, leading to fusion of the promyelocytic leukemia gene (PML) to retinoic acid receptor alfa (RARA), which is specifically sensitive to ATRA and arsenical derivatives [43]. To eliminate survival bias in the LTSs caused by this cytogenetic feature, we excluded these patients and performed the comparisons between the transcriptomes of the STSs and LTSs. Strikingly, we found that the DEGs in LTSs remained significantly associated with CNS development, neuronal development, neuronal differentiation, and neurogenesis.

In view of these findings, we sought to further explore the potential link between long-term survival and the nervous system. We have recently performed a comprehensive, phenotype-driven identification of nervous system-related genes. These include all protein-coding genes that are essential for normal nervous system development (neu.dev) and neurological function (behavior), whereby their knockout leads to malformations of the neural system components or behavioral disorders, respectively. Using system-based approaches, we had found 1246 genes that are specific for neuronal development (neu.dev) and 1086 for neurological function (behavior), while 863 genes are involved in both processes [44]. Overlapping of the 'neu.dev' and 'behavior' lists of genes with the 131 commonly deregulated genes in LTSs showed that, overall, 41 (31.3%) of them are nervous system-related genes, whereby 21 are indispensable for neuronal development, 10 for neurological function, and 10 for both processes (Figure 5F). Overall, our analyses indicate a strong association of the nervous system with the exceptionally prolonged survival and patient outcomes in AML.

3. Discussion

The incidence of AML increases with age, and the survival rate is exceptionally poor when diagnosed after age 65. In the present study, we employed integrative bioinformatics approaches on AML-relevant transcriptomic data to identify biomarkers of diagnosis and favorable prognosis.

We identified a set of 24 hub genes which discriminate AML cells from normal controls and could be valuable as a diagnostic signature. As expected, several of these genes (*BRCA1*, *BUB1B*, *CENPI*, *EXO1*, *KNL1*, *MCM10*, and *PLK4*) regulate genomic stability and cell proliferation. DNA damage repair-associated *BRCA1* is a known tumor suppressor gene, essential for DNA synthesis and cellular replication [45]. Mitotic checkpoint serine/threonine kinase B (BUB1B) is a vital component of the mitotic checkpoint and it is necessary for normal mitosis progression [46,47]. Centromere protein I (CENPI) is a structural component of the kinetochore, required for timely progression through G2 phase, mitosis, and chromosome stability [48,49]. Exonuclease I (EXO1) contributes to the regulation of the cell cycle checkpoints, the maintenance of the replication fork, DNA repair, and genomic instability [50,51]. Kinetochore scaffold 1 (KNL1) is essential in spindle assembly checkpoints, chromosome segregation, and kinetochore–microtubule attachments [52,53]. Minichromosome maintenance 10 replication initiation factor (MCM10) is required for

DNA synthesis and entry into or ending of S-phase [54,55]. Polo-like kinase 4 (PLK4) plays a crucial role in controlling centriole duplication, and its deregulation causes centrosome number abnormalities, mitotic defects, chromosomal instability, and, consequently, tumorigenesis [56]. Genomic instability and cell cycle deregulation favor the rapid proliferation and clonal expansion of cells harboring oncogenic DNA alterations in AML [57–59]. The upregulation of hub genes controlling these processes reflects the tendency of hematopoietic precursor cells in bone marrow to accumulate DNA lesions and increase their proliferation rate at the expense of normal hemopoiesis, resulting in bone marrow failure.

Notably, several genes of the diagnostic signature are involved in immune and inflammatory responses. For example, the CD3e molecule (CD3E) and CD3g molecule (CD3G) form a complex with TCR, a T cell receptor that plays an important role in T cell antigen recognition and signal transduction. Decreased expression of *CD3E* and *CD3G*, as in our study, leads to deregulation of T cell function [60]. Moreover, C-X-C motif chemokine ligand 12 (CXCL12) and its receptor CXCR4 are key elements for the adaptive and innate immune response and also for bone marrow organization and maintenance [61]. They are largely responsible for hematopoietic stem cell migration, homing, and survival in bone marrow [61]. CXCL12, found to be upregulated in our study, facilitates trafficking and homing of leukemic cells to the bone marrow microenvironment [62]. Furthermore, we identified the platelet-related genes *platelet factor 4* (*PF4*) and *pro-platelet basic protein* (*PPBP*) as components of the AML diagnostic signature. In general, platelets are anucleate cell fragments known for their central role in coagulation and vascular integrity and their deregulation in AML patients. Beyond their typical roles as fragmentary mediators of hemostasis and thrombosis, platelets orchestrate diverse immunological processes, including fighting against microbial threats, recruitment and promotion of innate effector cell functions, modulating antigen presentation, and enhancement of adaptive immune responses [63,64]. They usually interact with blood cells, as well as circulating tumor cells, via adhering or releasing a variety of soluble mediators, which can modulate the activity of hematopoietic stem and progenitor cells. Such platelet-coated tumor cells not only have an enhanced ability to attach to endothelial cells and endothelial extracellular matrix and thus promote tumor invasion and metastasis, they are also protected from the clearance of immune surveillance [65]. Overall, these findings underscore an implication of the immune system in AML pathophysiology which could be particularly relevant for the prediction of AML patients with favorable responses to immunotherapy. In particular, immunotherapy has catalyzed a paradigm shift in cancer treatment and has become the standard of care for several solid tumors. As our understanding of the genomic landscape of AML moves forward and the arsenal of immunotherapies is being enriched, their extension to AML represents an attractive clinical option, but this endeavor requires the establishment of guidelines for the prediction of groups of patients who would most benefit from immune targeting. A recent study on primary bone marrow samples from independent cohorts of AML patients defined immune-infiltrated and immune-depleted disease classes, revealed critical differences in immune gene expression across age groups and molecular disease subtypes, and provided compelling evidence that immune landscapes predict chemotherapy resistance and immunotherapy response [66]. The immune contexture is an emerging parameter of AML pathophysiology, and further investigation of the tumor microenvironment (TME) of the bone marrow, in combination with identification of immunomodulatory biomarkers in the blood, may allow molecular predictions of immunotherapeutic benefit and guide personalized treatment decisions. Further studies can uncover the molecular mechanisms underlying AML immunobiology and their translation to therapeutic solutions.

The asymptomatic nature of AML in combination with MRD and frequent disease relapse hamper AML prognostics, but a few patients manage to remain progression-free for exceptional durations of time. Despite their sparsity, unusual patients can inform cancer biology [20]. For example, families with rare and high penetrant germline mutations (Li–Fraumeni syndrome) have catalyzed understanding of the function of tumor suppressor genes [67], while populations with diminished cancer incidence (subjects with

Laron syndrome in Ecuador) provided insights on cancer metabolism and the insulin-like growth factor (IGF) pathway in tumor progression [68]. In an analogous manner, cases with exceptionally prolonged survival may bear distinct molecular characteristics, the elucidation of which could ultimately improve the outcome of individuals with more typical disease trajectories [20]. Long-term survivorship has, at least in part, a genetic and immunological basis [18,19,69]. Molecular factors contributing to favorable survival in a patient, who may otherwise have had a more typical course of disease, may include germline genetics, endogenous antitumor immune responses, tumor mutational load, and characteristics of the microenvironment [20]. Putting this novel rationale under testing, we defined subgroups of LTSs in our study population and investigated their distinct transcriptomic profiles as compared to patients below the median OS in order to infer favorable prognostic signatures which could be associated with resistance to disease progression. Surprisingly, this approach revealed that malignant cells of LTS AML patients exhibit prominent deregulations of genes involved in neuronal processes, thereby providing hints for associations between exceptionally prolonged survival and neurogenic cell traits.

The effect of the nervous system on tumor progression has recently come into the spotlight as an enigmatic phenomenon with a paramount effect in disease outcome. We recently demonstrated that all cancer types can activate neuronal networks 'off-context' and this ability largely influences the acquisition of aggressive characteristics. This trend seems to have a universal character across several cancers, since nervous system-related gene transcripts are upregulated in invasive stages and predict clinical outcomes in all cancer types, including AML. The prognostic potential of each nervous system-related gene depends on the cancer type [44]. Depending on the nervous system-related genes that are ectopically activated in a cancer cell context, patient survival can be affected either positively or negatively. Furthermore, different cancer types co-opt genes essential for the development of distinct anatomical parts of the nervous system. In the case of AML, genes indispensable for the somatic sensory system, the cerebellum fissum, the sympathetic nervous system, and the dentate gyrus granule cell layer are associated with AML patient outcomes [44]. Herein, we show that deregulations of nervous system-related genes are strongly associated with exceptionally long survival periods in AML. The relevant underlying mechanism awaits to be unveiled. On the one hand, it has been recently described that leukemia stem cells can evade the arachnoid space, migrate, proliferate in the cerebrospinal fluid (CSF), and consequently infiltrate perivascular spaces and brain parenchyma. The CNS is an immune-privileged site that can protect leukemic cells from chemotherapy. Overexpression of the neural cell adhesion molecule NCAM/CD56 in the surface of leukemic stem cells increases their ability to infiltrate the CNS [70]. In this regard, deterministic deregulation of selected neuronal signaling pathways may hypothetically enable leukemic cells to 'mingle' within the neuronal tissue microenvironment and rewire with the electrically active neurons [70]. On the other hand, Hanoun et al. (2014) demonstrated that adrenergic signals regulate AML and that when leukemic cell injections in mice are followed by sympathetic denervation, the course of disease is accelerated. They further provided compelling evidence that sympathetic neuropathy represents a mechanism by which AML co-opts the microenvironment to its own advantage in order to deplete niche cells that maintain healthy hematopoietic stem cells and expand leukemia-supportive mesenchymal progenitors [71]. The interactions between leukemic and neuronal cells in the context of AML recurrence, MDR, and/or metastasis remain a *terra incognita*. Future research can unveil how 'hijacking' of networks of the nervous system might offer several selective advantages in leukemic cells, especially those that persist in the circulation after initial AML management, and represent the 'seeds' of MRD. Conversely, it would be interesting to investigate whether activation of other neuronal signaling pathways can be protective against MRD in AML patients. Deciphering the neuronal signaling pathways that are beneficial for leukemic cell survival over those with a detrimental effect is anticipated to shed more light on AML pathophysiology and might have important translational implications in disease prevention and management. Such studies could open up novel

avenues for therapeutic interventions, for example, by repurposing already approved neurological drugs to target the cancer–neuronal crosstalk and prevent progression to aggressive stages [72].

Last, but not least, the profiles of the rare LTS patients who either never regress or manage to remain relapse-free for extended periods of time might be particularly meaningful in the context of AML prognosis. Their alleged resistance to disease progression or relapse could be an indicator of better abilities to cope both with MRD, which may be associated not only with epidemiological, environmental, or lifestyle factors, but also with their inherent molecular traits [20]. Hints supporting the influence of the hereditable genetic characteristics on the resistance to recurrence arise from studies demonstrating that ethnicities with inherently different genetic backgrounds, such as Caucasians and African Americans, exhibit differences in the AML prognosis, whereby Caucasian males have better CR and overall survival than African American males [73]. It would be interesting to investigate whether patients with unusually long survival periods share favorable intrinsic characteristics that overall predispose them to better detect and self-eliminate residual cancer cells during or after treatment. Checking for additional associations between the inherent characteristics of the LTSs and several clinical traits that are important for treatment personalization, such as cytogenetic characteristics and therapy regimens, could also be highly informative for the management of AML. This endeavor would require appropriate study designs that are specifically focused on LTS patient subgroups, where patients surviving over 5, 10, or more years will be oversampled and compared to those with more conventional disease trajectories [20]. Unveiling the cancer genomes and intrinsic genetic traits of this novel clinical population as opposed to relapsing cancer patients or even the general non-cancer population could reveal predisposing factors for resistance to recurrence. Future clinical trials could aim at the characterization of the molecular factors that are associated with LTSs. Such clinical study endpoints would be invaluable for the prognosis and management of AML and could enable not only the identification of patients with unfavorable prognosis who will need special treatments, but also the design of appropriate targeted strategies that prolong survival.

4. Materials and Methods

All analyses were performed in the R statistical computing environment v.4.0.3 (https://www.r-project.org, accessed on 1 September 2020) [74], unless otherwise stated.

4.1. Data Retrieval, Processing, and Analysis

Non-normalized, raw RNA-seq read count data of AML tissue samples were downloaded from The Cancer Genome Atlas (TCGA) using the 'TCGAbiolinks' R/Bioconductor package [75]. Since there are no available matched normal tissue samples for AML in TCGA, raw RNA-seq read counts per gene for whole blood normal tissue samples were downloaded from the Genotype-Tissue Expression (GTEx) Portal (https://gtexportal.org/home/, accessed on 14 September 2020) using the 'TCGAbiolinks' package. The RNA-seq data were generated from two different technological platforms (TCGA and GTEx), and due to their inherent variability, they had to be processed in a uniform way in order to be comparable. To ensure comparability across TCGA-derived AML and GTEx-derived whole blood normal tissue samples, all raw RNA-seq read counts were normalized and filtered, as suggested by the original TCGAbiolinks workflow [76,77] The overall procedure is illustrated in (Figure 6).

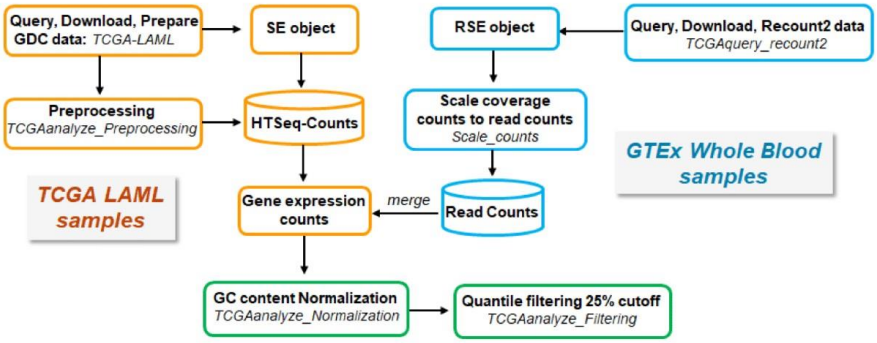


Figure 6. Schematic workflow of retrieving, processing, and merging TCGA and GTEx data.

4.1.1. RNA-seq Data Acquisition and Processing

AML Samples from TCGA

Acute myeloid leukemia (LAML) harmonized data (hg38) (Illumina HiSeq platform) were obtained from the Genomic Data Common (GDC) Portal (https://portal.gdc.cancer.gov/, accessed on 14 September 2020) of the TCGA database. GDC data were queried, downloaded, and prepared using the functions *GDCquery*, *GDCdownload*, and *GDCprepare*, respectively. A total of 151 samples including expression data of primary blood-derived cancer—peripheral blood were analyzed. Samples lacking clinical information were excluded from our study, and therefore 84 tumor samples were included. Pre-processing steps were applied according to the TCGA's workflow as recommended in the 'TCGAbiolinks' package in order to find possible outliers in transcriptome data using the *TCGAanalyze_Preprocessing* function. This function generates a count matrix ready to be provided as input in the downstream analysis pipeline.

Normal Tissue Samples from GTEx

Collectively, 456 of normal whole blood GTEx samples were downloaded through the Recount2 project using the function *TCGAquery_recount2* in the 'TCGAbiolinks' package, as ranged summarized experiment (RSE) objects. The raw counts were scaled with the '*scale_counts*' function in the 'Recount' package (v.1.16.1) [78,79].

Processing and Merging of TCGA- and GTEx-Derived Data

HTSeq counts from TCGA and read counts from GTEx were merged and gene expression counts were acquired, which were normalized for 'GC content' using the function *TCGAanalyze_Normalization* of the EDASeq protocol for exploratory data analysis (EDA) [80] in the 'TCGAbiolinks' package. Quantile filtering was applied with a cutoff of 25%.

Principal Component Analysis

Principal component analysis (PCA) of the TCGA and GTEx datasets was employed to detect batch effects. Raw RNA-seq read counts were converted to counts per million reads mapped (CPM) values with the function *cpm* in the 'edgeR' package (v.3.32.1) [81]. The functions *prcomp* and *ggplot* in the 'ggplot2' package (v.3.3.3) [82] were used to create a PCA plot.

4.2. Identification of Differentially Expressed Genes

To identify those genes that are differentially expressed between AML and normal samples, with the same direction (i.e., up- or downregulated), three different methods implemented in the R packages edgeR (v 3.32.1) [81], limma (v 3.44.3) [83], and DESeq2 (v.1.28.1) [84] were used.

Normalization by trimmed mean of M values (TMM) was performed with the *calcNormFactors* function in the 'edgeR' package (v.3.32.1). The negative binomial (NB) distribution was used to model the RNA-seq read counts per gene per sample in edgeR. Then, estimating dispersion was calculated with the *estimateDisp* function. DEG analysis between the two RNA-seq groups (i.e., AML versus normal) was performed using likelihood ratio tests (LRTs) including the *glmFit* and *glmLRT* functions in edgeR.

The *Voom* function in the 'limma' package was utilized to transform RNA-seq count data to log2 counts per million (logCPM). The *lmFit* and *contrasts.fit* functions in the 'limma' package were utilized for linear modeling. Next, the *eBayes* function in the 'limma' package was used to carry out empirical Bayes moderation by borrowing information across all genes, so as to obtain more accurate estimates of variability across genes. The DEGs were extracted by using the *decideTests* function.

The package 'DESeq' uses NB generalized linear models including the estimates of dispersion from the edgeR method. The standard differential expression analysis was performed using the function *DESeq*. Result tables were generated with the function *results*.

For detecting statistically significant differentially expressed genes, the threshold for the absolute $\log_2$ fold change (FC) was set at two ($|\log_2 FC \geq 2|$), and the adjusted p value by the Benjamini and Hochberg false discovery rate (FDR) method [85] less than 0.05.

To visualize DEGs, heatmap, MA and volcano plots were created. In particular, the *EnhancedVolcano* function in the R package 'EnhancedVolcano' (v.1.8.0) was utilized to generate volcano plots of the DEGs. Heatmap plots of logCPM values were created with the *pheatmap* function in the 'pheatmap' package in R (v.1.0.12). The *plotMD* function in the R package 'limma' was used to generate MA plots of the DEGs.

4.3. Construction of Weighted Gene Co-Expression Network

The common DEGs from the 84 AML samples were used to construct a gene co-expression network with the package WGCNA [21] (v.1.69) implemented in R. To this end, the FPKM values of the common genes were downloaded from the GDC Portal (https://portal.gdc.cancer.gov/; accessed on 10 December 2020) of the TCGA database and processed as described earlier (Section 4.1.1); analysis was conducted using the log2 (fpkm + 1) value. The *goodSamplesGenes* function was used to iteratively filter genes and samples.

The soft-threshold power was applied to construct pairwise Pearson correlation matrices so as to measure gene co-expression similarities. The function *pickSoftThreshold* was used to analyze network topology and select a suitable soft-thresholding power. The adjacency matrices were calculated with the function *adjacency*, which converts the Pearson correlation coefficients into gene connection strengths. The adjacency matrices were transformed into a topological overlap matrix (TOM) with the function *TOMsimilarity* in order to minimize the effects of noise. To group genes into modules, the average linkage hierarchical clustering method was applied by using the function *hclust* to cluster module eigengenes (i.e., first principal component) with a TOM-based dissimilarity measure; a relatively high minimum module size of thirty (minModuleSize = 30) was selected. The *dynamicTreeCut* algorithm was applied to detect modules with an eigengene correlation coefficient above 0.75, by setting the cut height threshold (MEDissThress) at 0.25.

4.4. Identification of Clinically Important Modules

Significant clinical data including gender, days to birth, age at diagnosis, history neoadjuvant treatment, French–American–British (FAB) category, platelet count preresection, cytogenetic risk group, blast count, days to death, and percentage of blasts in peripheral blood were acquired from the GDC Portal (https://portal.gdc.cancer.gov/, accessed on 14 September 2020). GDC data were processed as described earlier (Section 4.1.1).

To determine whether all AML samples were suitable for network analysis or to detect any outliers in these samples, the clustering dendrogram of the samples and the

corresponding clinical traits was generated using the *hclust* function in the 'WGCNA' package based on their Euclidean distance.

4.5. Co-Expression and Protein–Protein Interaction (PPI) Network Analysis

Given that the modules detected by WGCNA are usually large, consisting of many genes, it is important to determine which genes are more representative of each module, as well as of the entire network. In the biological networks, the most highly connected genes, 'hub' genes, are generally considered to be of higher biological significance and more relevant to the overall function of the network [86,87]. The intra-modular hubs are central to a given module, whereas inter-modular hubs are intermediate between two or more modules.

To this end, the genes contained in the modules associated with AML clinical traits were used to construct a gene co-expression network; genes with a Pearson correlation coefficient more than 0.3 were selected. Furthermore, those genes were provided as input to the Search Tool for the Retrieval of Interacting Genes (STRING) (v.11.0) (https://string-db.org/, accessed on 1 February 2021) database [88] in order to investigate the functional and/or physical associations among their corresponding gene products. A highest confidence score of $\geq$0.9 was set as a cutoff.

The topological properties of the gene co-expression and protein–protein interaction (PPI) networks were further investigated through Cytoscape (v.3.8.2) (https://cytoscape.org/, accessed on 12 February 2021) [89], an open-source platform for network visualization and manipulation. In both networks, those nodes with high connectivity [90,91], that is, connected to more than 17 immediate neighboring nodes, were selected. Moreover, the Cytoscape plugin cytoHubba [92], which provides eleven local and global methods for network topological analysis, was utilized to select the top thirty nodes (for each of the 11 methods).

4.6. NCBI GEO Gene Expression Datasets

The publicly accessible repository NCBI GEO DataSets (https://www.ncbi.nlm.nih.gov/gds/, accessed on 20 February 2021) [93,94] was searched thoroughly for transcriptomic datasets using the relevant keywords: ('acute myeloid leukemia' OR 'AML') AND ('homo sapiens' OR 'human'). The selection criteria were the following: (i) gene expression data derived from human AML patient samples and corresponding normal tissue samples, (ii) inclusion of more than 5000 genes in the dataset, and (iii) data from animal models or cell lines were excluded. In this way, two eligible datasets were selected.

The GEO series GSE9476 [95] (Table S1) contains genome-wide gene expression by microarray of blast cells from AML patients and (a) unselected bone marrow (BM), and (b) unselected peripheral blood (PB) samples from healthy donors. The Affymetrix Human Genome U133A Array [HG-U133A] platform was employed.

In GSE138702 [96] (Table S1), global gene expression analysis was performed using RNA-seq of total RNA isolated from the bone marrow of AML patients and matched healthy donors. The Illumina HiSeq 3000 (Homo sapiens) GPL21290 platform was used.

4.6.1. Microarray-Based Transcriptomic Data Analysis

For microarray data analysis, the two-sample *t*-test was employed to identify genes differentially expressed in two groups: (i) leukemic blast cells from AML (n = 26) versus BM (n = 10), and (ii) AML (n = 26) versus PB (n = 10) (Table S1), by applying a methodology described in detail in previous studies [97,98]. The Benjamini–Hochberg method [85] for controlling FDR was applied; an FDR-adjusted p value less than or equal to 0.01 was considered as statistically significant in this study. For all statistical analyses, the Stata 13 statistical software package (https://www.stata.com/, accessed on 10 March 2021) [99] was used.

4.6.2. RNA-seq-Based Transcriptomic Data Analysis

The gene expression profiles of the bone marrow of 9 AML patients versus 10 healthy control donors were compared (Table S1). For the RNA sequencing analysis, the following pipeline was performed. FASTQ files that contained raw 2 × 50 bp paired-end RNA-seq reads were downloaded from the respective Sequence Read Archive (SRA) by using the SRA Tool Kit v.2.9.0 [100]. Raw RNA-seq reads were mapped to the human reference genome GRCh38 (Ensembl version 97) by employing the splice junction aligner HISAT2 v.2.1.0 [101]. Transcriptome normalization, reconstruction, and quantification were conducted by using StringTie v.1.3.5 [102]. Differentially expressed genes (DEGs) were detected using the edgeR package version 3.32.0 [81], by setting $|\log_2 FC \geq 2|$ and the FDR-corrected p value less than 0.05.

4.7. Identification of DEGs Correlated with Long-Term Survival

Patient data from the TCGA LAML cohort were retrieved via the 'TCGAbiolinks' package (data.type: Gene Expression Quantification, workflow.type: HTSeq—Counts). Only the genes that were expressed in at least half of the patients were retained and the outliers that were >6 standard deviations away from the median were removed. We defined the groups of LTS patients (2 × SD = 1567 days, 3 × SD = 2168 days, and 5-year survival = 1825 days) and calculated the respective differentially expressed genes (DEGs) in the population that died within 365 days using edgeR [81], with $\log_2 FC$ cutoff > 1. Afterwards, we performed gene set enrichment analysis (GSEA) on the DEGs for each comparison, as well as for the common 131 genes that were deregulated in all three analyses. For GSEA, the GSEA-P 2.0 software (Broad Institute, Cambridge, MA, USA) [103] was used. Enriched Gene Ontology (GO) terms were plotted against the negative log10 of their individual FDR value (<0.05).

5. Conclusions

In conclusion, we applied in silico workflows on publicly available AML-relevant transcriptomic data and identified 24 genes mainly related to cell cycle and immune response, several of which were subsequently verified in two independent datasets. These genes could be extrapolated for the establishment of powerful AML diagnostic biomarkers and also be considered for defining patients that would most benefit from immunotherapeutics. Moreover, taking advantage of the fact that transcriptomic data of AML cases with unusually extended survival periods already exist in well-established databases, we tested a novel notion which proposes that rare patients at the extremes can be informative for prognosis and management of the general cancer population. Comparisons of the transcriptomes of these exceptional cases with those of more typical disease trajectories in our study cohort revealed an association of deregulations of nervous system-related genes with favorable AML prognosis. Future experimental and clinical studies can provide further insights on the underlying molecular mechanisms, inform clinical decisions, and catalyze the translation of these findings to diagnostic and therapeutic solutions.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/ijms22179601/s1.

Author Contributions: Conceptualization, A.P. and S.L.; methodology, A.P., H.Y., H.I.T., S.M. and S.L; software, H.Y., H.I.T., S.M. and P.I.K.; validation, H.Y., H.I.T., S.M., S.L. and A.P.; formal analysis, H.Y., H.I.T., S.M., S.L. and A.P.; investigation, H.Y., H.I.T., S.M., G.K., C.K., S.L. and A.P.; data curation, H.Y., H.I.T., S.M., P.I.K. and A.P.; writing—original draft preparation, H.Y., H.I.T., S.M., G.K., C.K., P.I.K., S.L. and A.P.; writing—review and editing, H.Y., H.I.T., S.M., G.K., C.K., P.I.K., S.L. and A.P.; visualization, H.Y., H.I.T. and S.M.; supervision, A.P.; project administration, A.P. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Deutsche Forschungsgemeinschaft (DFG grant PU 188/17-1) and FORUN program of Rostock University Medical School (889006 to SL). H.T. acknowledges Yükseköğretim Kurulu (YÖK) scholarship and TÜBİTAK 2211-C scholarship. P.I.K. is co-financed by

Greece and the European Union (European Social Fund- ESF) through the Operational Programme "Human Resources Development, Education and Lifelong Learning" in the context of the project "Reinforcement of Postdoctoral Researchers—2nd Cycle" (MIS-5033021), implemented by the State Scholarships Foundation (IKY).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were generated or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Leukemia-acute Myeloid-aml: Statistics. Available online: https://www.cancer.net/cancer-types/leukemia-acute-myeloid-aml/statistics (accessed on 7 June 2021).
2. Cancer Stat Facts: Leukemia—Acute Myeloid Leukemia (aml). Available online: https://seer.cancer.gov/statfacts/html/amyl.html (accessed on 7 June 2021).
3. Estey, E.; Dohner, H. Acute myeloid leukaemia. *Lancet* **2006**, *368*, 1894–1907. [CrossRef]
4. Lowenberg, B.; Downing, J.R.; Burnett, A. Acute myeloid leukemia. *N. Engl. J. Med.* **1999**, *341*, 1051–1062. [CrossRef]
5. Shipley, J.L.; Butera, J.N. Acute myelogenous leukemia. *Exp. Hematol.* **2009**, *37*, 649–658. [CrossRef]
6. Dohner, H.; Estey, E.H.; Amadori, S.; Appelbaum, F.R.; Buchner, T.; Burnett, A.K.; Dombret, H.; Fenaux, P.; Grimwade, D.; Larson, R.A.; et al. Diagnosis and management of acute myeloid leukemia in adults: Recommendations from an international expert panel, on behalf of the european leukemianet. *Blood* **2010**, *115*, 453–474. [CrossRef] [PubMed]
7. Kantarjian, H.; Kadia, T.; DiNardo, C.; Daver, N.; Borthakur, G.; Jabbour, E.; Garcia-Manero, G.; Konopleva, M.; Ravandi, F. Acute myeloid leukemia: Current progress and future directions. *Blood Cancer J.* **2021**, *11*, 41. [CrossRef] [PubMed]
8. Available online: https://www.cancer.gov/news-events/press-releases/2020 (accessed on 7 June 2021).
9. Pulte, D.; Gondos, A.; Brenner, H. Improvement in survival in younger patients with acute lymphoblastic leukemia from the 1980s to the early 21st century. *Blood* **2009**, *113*, 1408–1411. [CrossRef] [PubMed]
10. Farag, S.S.; Ruppert, A.S.; Mrozek, K.; Mayer, R.J.; Stone, R.M.; Carroll, A.J.; Powell, B.L.; Moore, J.O.; Pettenati, M.J.; Koduru, P.R.; et al. Outcome of induction and postremission therapy in younger adults with acute myeloid leukemia with normal karyotype: A cancer and leukemia group b study. *J. Clin. Oncol.* **2005**, *23*, 482–493. [CrossRef] [PubMed]
11. Klepin, H.D.; Balducci, L. Acute myelogenous leukemia in older adults. *Oncology* **2009**, *14*, 222–232. [CrossRef] [PubMed]
12. Abelson, S.; Collord, G.; Ng, S.W.K.; Weissbrod, O.; Mendelson Cohen, N.; Niemeyer, E.; Barda, N.; Zuzarte, P.C.; Heisler, L.; Sundaravadanam, Y.; et al. Prediction of acute myeloid leukaemia risk in healthy individuals. *Nature* **2018**, *559*, 400–404. [CrossRef] [PubMed]
13. Paez, D.; Labonte, M.J.; Bohanes, P.; Zhang, W.; Benhanim, L.; Ning, Y.; Wakatsuki, T.; Loupakis, F.; Lenz, H.J. Cancer dormancy: A model of early dissemination and late cancer recurrence. *Clin. Cancer Res.* **2012**, *18*, 645–653. [CrossRef]
14. Benton, C.B.; Ravandi, F. A mind map for managing minimal residual disease in acute myeloid leukemia. *Clin. Adv. Hematol. Oncol. HO* **2017**, *15*, 859–867.
15. Lauber, C.; Correia, N.; Trumpp, A.; Rieger, M.A.; Dolnik, A.; Bullinger, L.; Roeder, I.; Seifert, M. Survival differences and associated molecular signatures of dnmt3a-mutant acute myeloid leukemia patients. *Sci. Rep.* **2020**, *10*, 12761. [CrossRef]
16. Guo, C.; Gao, Y.Y.; Ju, Q.Q.; Zhang, C.X.; Gong, M.; Li, Z.L. The landscape of gene co-expression modules correlating with prognostic genetic abnormalities in aml. *J. Transl. Med.* **2021**, *19*, 228. [CrossRef]
17. Pan, Y.; Zhang, Q.; Deng, X.; An, N.; Du, X.; Liu, J. Gene coexpression network analysis revealed biomarkers correlated with blast cells and survival in acute myeloid leukemia. *Mol. Clin. Oncol.* **2020**, *12*, 475–484. [CrossRef] [PubMed]
18. Reifenberger, G.; Weber, R.G.; Riehmer, V.; Kaulich, K.; Willscher, E.; Wirth, H.; Gietzelt, J.; Hentschel, B.; Westphal, M.; Simon, M.; et al. Molecular characterization of long-term survivors of glioblastoma using genome- and transcriptome-wide profiling. *Int. J. Cancer* **2014**, *135*, 1822–1831. [CrossRef] [PubMed]
19. Shen, J.; Hu, J.; Wu, J.; Luo, X.; Li, Y.; Li, J. Molecular characterization of long-term survivors of hepatocellular carcinoma. *Aging* **2021**, *13*, 7517–7537. [CrossRef] [PubMed]
20. Saner, F.A.M.; Herschtal, A.; Nelson, B.H.; deFazio, A.; Goode, E.L.; Ramus, S.J.; Pandey, A.; Beach, J.A.; Fereday, S.; Berchuck, A.; et al. Going to extremes: Determinants of extraordinary response and survival in patients with cancer. *Nat. Rev. Cancer* **2019**, *19*, 339–348. [CrossRef]
21. Langfelder, P.; Horvath, S. Wgcna: An r package for weighted correlation network analysis. *BMC Bioinform.* **2008**, *9*, 559. [CrossRef]
22. Jackson, N.; Menon, B.S.; Zarina, W.; Zawawi, N.; Naing, N.N. Why is acute leukemia more common in males? A possible sex-determined risk linked to the abo blood group genes. *Ann. Hematol.* **1999**, *78*, 233–236. [CrossRef]
23. Almeida, A.M.; Ramos, F. Acute myeloid leukemia in the older adults. *Leuk. Res. Rep.* **2016**, *6*, 1–7. [CrossRef] [PubMed]

24. Seer Cancer Statistics Review, 1975–2000. Available online: http://seer.cancer.gov/csr/1975_2000/sections.html (accessed on 7 June 2021).
25. De Kouchkovsky, I.; Abdul-Hay, M. Acute myeloid leukemia: A comprehensive review and 2016 update. *Blood Cancer J.* **2016**, *6*, e441. [CrossRef]
26. Morton, L.M.; Dores, G.M.; Schonfeld, S.J.; Linet, M.S.; Sigel, B.S.; Lam, C.J.K.; Tucker, M.A.; Curtis, R.E. Association of chemotherapy for solid tumors with development of therapy-related myelodysplastic syndrome or acute myeloid leukemia in the modern era. *JAMA Oncol.* **2019**, *5*, 318–325. [CrossRef]
27. Bennett, J.M.; Catovsky, D.; Daniel, M.T.; Flandrin, G.; Galton, D.A.; Gralnick, H.R.; Sultan, C. Proposals for the classification of the acute leukaemias. French-american-british (fab) co-operative group. *Br. J. Haematol.* **1976**, *33*, 451–458. [CrossRef] [PubMed]
28. Psaila, B.; Bussel, J.B.; Frelinger, A.L.; Babula, B.; Linden, M.D.; Li, Y.; Barnard, M.R.; Tate, C.; Feldman, E.J.; Michelson, A.D. Differences in platelet function in patients with acute myeloid leukemia and myelodysplasia compared to equally thrombocytopenic patients with immune thrombocytopenia. *J. Thromb. Haemost. JTH* **2011**, *9*, 2302–2310. [CrossRef]
29. Saultz, J.N.; Garzon, R. Acute myeloid leukemia: A concise review. *J. Clin. Med.* **2016**, *5*, 33. [CrossRef]
30. Grimwade, D.; Hills, R.K.; Moorman, A.V.; Walker, H.; Chatters, S.; Goldstone, A.H.; Wheatley, K.; Harrison, C.J.; Burnett, A.K.; National Cancer Research Institute Adult Leukaemia Working, G. Refinement of cytogenetic classification in acute myeloid leukemia: Determination of prognostic significance of rare recurring chromosomal abnormalities among 5876 younger adult patients treated in the united kingdom medical research council trials. *Blood* **2010**, *116*, 354–365. [CrossRef] [PubMed]
31. Gupta, M.; Mahapatra, M.; Saxena, R. Cytogenetics' impact on the prognosis of acute myeloid leukemia. *J. Lab. Physicians* **2019**, *11*, 133–137. [CrossRef] [PubMed]
32. DiNardo, C.D.; Garcia-Manero, G.; Pierce, S.; Nazha, A.; Bueso-Ramos, C.; Jabbour, E.; Ravandi, F.; Cortes, J.; Kantarjian, H. Interactions and relevance of blast percentage and treatment strategy among younger and older patients with acute myeloid leukemia (aml) and myelodysplastic syndrome (mds). *Am. J. Hematol.* **2016**, *91*, 227–232. [CrossRef]
33. Holtzman, N.G.; El Chaer, F.; Baer, M.R.; Ali, O.; Patel, A.; Duong, V.H.; Sausville, E.A.; Singh, Z.N.; Koka, R.; Zou, Y.S.; et al. Peripheral blood blast rate of clearance is an independent predictor of clinical response and outcomes in acute myeloid leukaemia. *Br. J. Haematol.* **2020**, *188*, 881–887. [CrossRef]
34. Weinkauff, R.; Estey, E.H.; Starostik, P.; Hayes, K.; Huh, Y.O.; Hirsch-Ginsberg, C.; Andreeff, M.; Keating, M.; Kantarjian, H.M.; Freireich, E.J.; et al. Use of peripheral blood blasts vs. bone marrow blasts for diagnosis of acute leukemia. *Am. J. Clin. Pathol.* **1999**, *111*, 733–740. [CrossRef]
35. Oran, B.; Weisdorf, D.J. Survival for older patients with acute myeloid leukemia: A population-based study. *Haematologica* **2012**, *97*, 1916–1924. [CrossRef] [PubMed]
36. Acute Myeloid Leukaemia (Aml)—Survival. Available online: https://www.cancerresearchuk.org/about-cancer/acute-myeloid-leukaemia-aml/survival (accessed on 7 June 2021).
37. Cheng, M.J.; Hourigan, C.S.; Smith, T.J. Adult acute myeloid leukemia long-term survivors. *J. Leuk.* **2014**, *2*, 1–8. [CrossRef]
38. Vasu, S.; Kohlschmidt, J.; Mrozek, K.; Eisfeld, A.K.; Nicolet, D.; Sterling, L.J.; Becker, H.; Metzeler, K.H.; Papaioannou, D.; Powell, B.L.; et al. Ten-year outcome of patients with acute myeloid leukemia not treated with allogeneic transplantation in first complete remission. *Blood Adv.* **2018**, *2*, 1645–1650. [CrossRef] [PubMed]
39. Bloomfield, C.D.; Shuma, C.; Regal, L.; Philip, P.P.; Hossfeld, D.K.; Hagemeijer, A.M.; Garson, O.M.; Peterson, B.A.; Sakurai, M.; Alimena, G.; et al. Long-term survival of patients with acute myeloid leukemia: A third follow-up of the fourth international workshop on chromosomes in leukemia. *Cancer* **1997**, *80*, 2191–2198. [CrossRef]
40. Arndt, V.; Koch-Gallenkamp, L.; Jansen, L.; Bertram, H.; Eberle, A.; Holleczek, B.; Schmid-Hopfner, S.; Waldmann, A.; Zeissig, S.R.; Brenner, H. Quality of life in long-term and very long-term cancer survivors versus population controls in germany. *Acta Oncol.* **2017**, *56*, 190–197. [CrossRef]
41. Roussa, E.; Krieglstein, K. Gdnf promotes neuronal differentiation and dopaminergic development of mouse mesencephalic neurospheres. *Neurosci. Lett.* **2004**, *361*, 52–55. [CrossRef]
42. Holland, P.W.; Booth, H.A.; Bruford, E.A. Classification and nomenclature of all human homeobox genes. *BMC Biol.* **2007**, *5*, 47. [CrossRef]
43. Adès, L.; Guerci, A.; Raffoux, E.; Sanz, M.; Chevallier, P.; Lapusan, S.; Recher, C.; Thomas, X.; Rayon, C.; Castaigne, S.; et al. Very long-term outcome of acute promyelocytic leukemia after treatment with all-trans retinoic acid and chemotherapy: The European APL Group experience. *Blood J. Am. Soc. Hematol.* **2010**, *115*, 1690–1696. [CrossRef]
44. Logotheti, S.; Marquardt, S.; Richter, C.; Sophie Hain, R.; Murr, N.; Takan, I.; Pavlopoulou, A.; Putzer, B.M. Neural networks recapitulation by cancer cells promotes disease progression: A novel role of p73 isoforms in cancer-neuronal crosstalk. *Cancers* **2020**, *12*, 3789. [CrossRef]
45. Friedenson, B. The brca1/2 pathway prevents hematologic cancers in addition to breast and ovarian cancers. *BMC Cancer* **2007**, *7*, 152. [CrossRef]
46. Fu, X.; Chen, G.; Cai, Z.D.; Wang, C.; Liu, Z.Z.; Lin, Z.Y.; Wu, Y.D.; Liang, Y.X.; Han, Z.D.; Liu, J.C.; et al. Overexpression of bub1b contributes to progression of prostate cancer and predicts poor outcome in patients with prostate cancer. *Onco Targets Ther.* **2016**, *9*, 2211–2220. [PubMed]
47. Zhu, L.J.; Pan, Y.; Chen, X.Y.; Hou, P.F. Bub1 promotes proliferation of liver cancer cells by activating smad2 phosphorylation. *Oncol. Lett.* **2020**, *19*, 3506–3512. [CrossRef]

48. Ding, N.; Li, R.; Shi, W.; He, C. Cenpi is overexpressed in colorectal cancer and regulates cell migration and invasion. *Gene* **2018**, *674*, 80–86. [CrossRef]
49. Thangavelu, P.U.; Lin, C.Y.; Vaidyanathan, S.; Nguyen, T.H.M.; Dray, E.; Duijf, P.H.G. Overexpression of the e2f target gene cenpi promotes chromosome instability and predicts poor prognosis in estrogen receptor-positive breast cancer. *Oncotarget* **2017**, *8*, 62167–62182. [CrossRef] [PubMed]
50. He, D.; Li, T.; Sheng, M.; Yang, B. Exonuclease 1 (exo1) participates in mammalian non-homologous end joining and contributes to drug resistance in ovarian cancer. *Med Sci. Monit.* **2020**, *26*, e918751. [CrossRef]
51. Keijzers, G.; Bakula, D.; Petr, M.A.; Madsen, N.G.K.; Teklu, A.; Mkrtchyan, G.; Osborne, B.; Scheibye-Knudsen, M. Human exonuclease 1 (exo1) regulatory functions in DNA replication with putative roles in cancer. *Int. J. Mol. Sci.* **2018**, *20*, 74. [CrossRef]
52. Bai, T.; Zhao, Y.; Liu, Y.; Cai, B.; Dong, N.; Li, B. Effect of knl1 on the proliferation and apoptosis of colorectal cancer cells. *Technol. Cancer Res. Treat.* **2019**, *18*. [CrossRef] [PubMed]
53. Tromer, E.; Snel, B.; Kops, G.J. Widespread recurrent patterns of rapid repeat evolution in the kinetochore scaffold knl1. *Genome Biol. Evol.* **2015**, *7*, 2383–2393. [CrossRef] [PubMed]
54. Du, W.; Stauffer, M.E.; Eichman, B.F. Structural biology of replication initiation factor mcm10. *Sub-Cell. Biochem.* **2012**, *62*, 197–216.
55. Yang, W.D.; Wang, L. Mcm10 facilitates the invaded/migrated potentials of breast cancer cells via wnt/beta-catenin signaling and is positively interlinked with poor prognosis in breast carcinoma. *J. Biochem. Mol. Toxicol.* **2019**, *33*, e22330. [CrossRef] [PubMed]
56. Zhang, X.; Wei, C.; Liang, H.; Han, L. Polo-like kinase 4's critical role in cancer development and strategies for plk4-targeted therapy. *Front. Oncol.* **2021**, *11*, 587554.
57. de Oliveira Lisboa, M.; Brofman, P.R.S.; Schmid-Braz, A.T.; Rangel-Pozzo, A.; Mai, S. Chromosomal instability in acute myeloid leukemia. *Cancers* **2021**, *13*, 2655. [CrossRef]
58. Papaemmanuil, E.; Gerstung, M.; Bullinger, L.; Gaidzik, V.I.; Paschka, P.; Roberts, N.D.; Potter, N.E.; Heuser, M.; Thol, F.; Bolli, N.; et al. Genomic classification and prognosis in acute myeloid leukemia. *N. Engl. J. Med.* **2016**, *374*, 2209–2221. [CrossRef]
59. Schnerch, D.; Yalcintepe, J.; Schmidts, A.; Becker, H.; Follo, M.; Engelhardt, M.; Wasch, R. Cell cycle control in acute myeloid leukemia. *Am. J. Cancer Res.* **2012**, *2*, 508–528.
60. Yang, Y.; Zang, Y.; Zheng, C.; Li, Z.; Gu, X.; Zhou, M.; Wang, Z.; Xiang, J.; Chen, Z.; Zhou, Y. Cd3d is associated with immune checkpoints and predicts favorable clinical outcome in colon cancer. *Immunotherapy* **2020**, *12*, 25–35. [CrossRef]
61. Garcia-Cuesta, E.M.; Santiago, C.A.; Vallejo-Diaz, J.; Juarranz, Y.; Rodriguez-Frade, J.M.; Mellado, M. The role of the cxcl12/cxcr4/ackr3 axis in autoimmune diseases. *Front. Endocrinol.* **2019**, *10*, 585. [CrossRef] [PubMed]
62. Cho, B.S.; Kim, H.J.; Konopleva, M. Targeting the cxcl12/cxcr4 axis in acute myeloid leukemia: From bench to bedside. *Korean J. Intern. Med.* **2017**, *32*, 248–257. [CrossRef]
63. Ali, R.A.; Wuescher, L.M.; Worth, R.G. Platelets: Essential components of the immune system. *Curr. Trends Immunol.* **2015**, *16*, 65–78. [PubMed]
64. Morrell, C.N.; Aggrey, A.A.; Chapman, L.M.; Modjeski, K.L. Emerging roles for platelets as immune and inflammatory cells. *Blood* **2014**, *123*, 2759–2767. [CrossRef] [PubMed]
65. Qian, X.; Wen-jun, L. Platelet changes in acute leukemia. *Cell Biochem. Biophys.* **2013**, *67*, 1473–1479. [CrossRef] [PubMed]
66. Vadakekolathu, J.; Minden, M.D.; Hood, T.; Church, S.E.; Reeder, S.; Altmann, H.; Sullivan, A.H.; Viboch, E.J.; Patel, T.; Ibrahimova, N.; et al. Immune landscapes predict chemotherapy resistance and immunotherapy response in acute myeloid leukemia. *Sci. Transl. Med.* **2020**, *12*, 546. [CrossRef] [PubMed]
67. Malkin, D.; Li, F.P.; Strong, L.C.; Fraumeni, J.F., Jr.; Nelson, C.E.; Kim, D.H.; Kassel, J.; Gryka, M.A.; Bischoff, F.Z.; Tainsky, M.A.; et al. Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science* **1990**, *250*, 1233–1238. [CrossRef]
68. Guevara-Aguirre, J.; Bautista, C.; Torres, C.; Pena, G.; Guevara, C.; Palacios, C.; Guevara, A.; Gavilanes, A.W.D. Insights from the clinical phenotype of subjects with laron syndrome in ecuador. *Rev. Endocr. Metab. Disord.* **2021**, *22*, 59–70. [CrossRef]
69. Balachandran, V.P.; Luksza, M.; Zhao, J.N.; Makarov, V.; Moral, J.A.; Remark, R.; Herbst, B.; Askan, G.; Bhanot, U.; Senbabaoglu, Y.; et al. Identification of unique neoantigen qualities in long-term survivors of pancreatic cancer. *Nature* **2017**, *551*, 512–516. [CrossRef] [PubMed]
70. Deak, D.; Gorcea-Andronic, N.; Sas, V.; Teodorescu, P.; Constantinescu, C.; Iluta, S.; Pasca, S.; Hotea, I.; Turcas, C.; Moisoiu, V.; et al. A narrative review of central nervous system involvement in acute leukemias. *Ann. Transl. Med.* **2021**, *9*, 68. [CrossRef]
71. Hanoun, M.; Zhang, D.; Mizoguchi, T.; Pinho, S.; Pierce, H.; Kunisaki, Y.; Lacombe, J.; Armstrong, S.A.; Duhrsen, U.; Frenette, P.S. Acute myelogenous leukemia-induced sympathetic neuropathy promotes malignancy in an altered hematopoietic stem cell niche. *Cell Stem Cell* **2014**, *15*, 365–375. [CrossRef]
72. Dolgin, E. Cancer-neuronal crosstalk and the startups working to silence it. *Nat. Biotechnol.* **2020**, *38*, 115–117. [CrossRef]
73. Sekeres, M.A.; Peterson, B.; Dodge, R.K.; Mayer, R.J.; Moore, J.O.; Lee, E.J.; Kolitz, J.; Baer, M.R.; Schiffer, C.A.; Carroll, A.J.; et al. Differences in prognostic factors and outcomes in african americans and whites with acute myeloid leukemia. *Blood* **2004**, *103*, 4036–4042. [CrossRef]
74. R core team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2017.
75. Colaprico, A.; Silva, T.C.; Olsen, C.; Garofano, L.; Cava, C.; Garolini, D.; Sabedot, T.S.; Malta, T.M.; Pagnotta, S.M.; Castiglioni, I.; et al. Tcgabiolinks: An r/bioconductor package for integrative analysis of tcga data. *Nucleic Acids Res.* **2016**, *44*, e71. [CrossRef]

76. Mounir, M.; Lucchetta, M.; Silva, T.C.; Olsen, C.; Bontempi, G.; Chen, X.; Noushmehr, H.; Colaprico, A.; Papaleo, E. New functionalities in the tcgabiolinks package for the study and integration of cancer data from gdc and gtex. *PLoS Comput. Biol.* **2019**, *15*, e1006701. [CrossRef]
77. Silva, T.C.; Colaprico, A.; Olsen, C.; D'Angelo, F.; Bontempi, G.; Ceccarelli, M.; Noushmehr, H. Tcga workflow: Analyze cancer genomics and epigenomics data using bioconductor packages. *F1000Research* **2016**, *5*, 1542. [CrossRef] [PubMed]
78. Collado-Torres, L.; Nellore, A.; Jaffe, A.E. Recount workflow: Accessing over 70,000 human rna-seq samples with bioconductor. *F1000Research* **2017**, *6*, 1558. [CrossRef] [PubMed]
79. Collado-Torres, L.; Nellore, A.; Kammers, K.; Ellis, S.E.; Taub, M.A.; Hansen, K.D.; Jaffe, A.E.; Langmead, B.; Leek, J.T. Reproducible rna-seq analysis using recount2. *Nat. Biotechnol.* **2017**, *35*, 319–321. [CrossRef]
80. Risso, D.; Schwartz, K.; Sherlock, G.; Dudoit, S. Gc-content normalization for rna-seq data. *BMC Bioinform.* **2011**, *12*, 480. [CrossRef] [PubMed]
81. Robinson, M.D.; McCarthy, D.J.; Smyth, G.K. Edger: A bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **2010**, *26*, 139–140. [CrossRef]
82. Wickham, H. *Ggplot2: Elegant Graphics for Data Analysis*; Springer: New York, NY, USA, 2016.
83. Ritchie, M.E.; Phipson, B.; Wu, D.; Hu, Y.; Law, C.W.; Shi, W.; Smyth, G.K. Limma powers differential expression analyses for rna-sequencing and microarray studies. *Nucleic Acids Res.* **2015**, *43*, e47. [CrossRef]
84. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for rna-seq data with deseq2. *Genome Biol.* **2014**, *15*, 550. [CrossRef]
85. Benjamini, Y.; Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Society. Ser. B* **1995**, *57*, 289–300. [CrossRef]
86. Barabasi, A.L.; Gulbahce, N.; Loscalzo, J. Network medicine: A network-based approach to human disease. *Nat. Rev. Genet.* **2011**, *12*, 56–68. [CrossRef]
87. Kontou, P.I.; Pavlopoulou, A.; Dimou, N.L.; Pavlopoulos, G.A.; Bagos, P.G. Network analysis of genes and their association with diseases. *Gene* **2016**, *590*, 68–78. [CrossRef]
88. Szklarczyk, D.; Gable, A.L.; Nastou, K.C.; Lyon, D.; Kirsch, R.; Pyysalo, S.; Doncheva, N.T.; Legeay, M.; Fang, T.; Bork, P.; et al. The string database in 2021: Customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.* **2021**, *49*, D605–D612. [CrossRef] [PubMed]
89. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **2003**, *13*, 2498–2504. [CrossRef]
90. Albert, R.; Jeong, H.; Barabasi, A.L. Error and attack tolerance of complex networks. *Nature* **2000**, *406*, 378–382. [CrossRef] [PubMed]
91. Horvath, S.; Dong, J. Geometric interpretation of gene coexpression network analysis. *PLoS Comput. Biol.* **2008**, *4*, e1000117. [CrossRef] [PubMed]
92. Chin, C.H.; Chen, S.H.; Wu, H.H.; Ho, C.W.; Ko, M.T.; Lin, C.Y. Cytohubba: Identifying hub objects and sub-networks from complex interactome. *BMC Syst. Biol.* **2014**, *8* (Suppl. 4), S11. [CrossRef]
93. Barrett, T.; Wilhite, S.E.; Ledoux, P.; Evangelista, C.; Kim, I.F.; Tomashevsky, M.; Marshall, K.A.; Phillippy, K.H.; Sherman, P.M.; Holko, M.; et al. Ncbi geo: Archive for functional genomics data sets—Update. *Nucleic Acids Res.* **2013**, *41*, D991–D995. [CrossRef]
94. Clough, E.; Barrett, T. The gene expression omnibus database. *Methods Mol. Biol.* **2016**, *1418*, 93–110.
95. Stirewalt, D.L.; Meshinchi, S.; Kopecky, K.J.; Fan, W.; Pogosova-Agadjanyan, E.L.; Engel, J.H.; Cronk, M.R.; Dorcy, K.S.; McQuary, A.R.; Hockenbery, D.; et al. Identification of genes with abnormal expression changes in acute myeloid leukemia. *Genes Chromosomes Cancer* **2008**, *47*, 8–20. [CrossRef]
96. Celik, H.; Lindblad, K.E.; Popescu, B.; Gui, G.; Goswami, M.; Valdez, J.; DeStefano, C.; Lai, C.; Thompson, J.; Ghannam, J.Y.; et al. Highly multiplexed proteomic assessment of human bone marrow in acute myeloid leukemia. *Blood Adv.* **2020**, *4*, 367–379. [CrossRef]
97. Toy, H.I.; Karakulah, G.; Kontou, P.I.; Alotaibi, H.; Georgakilas, A.G.; Pavlopoulou, A. Investigating molecular determinants of cancer cell resistance to ionizing radiation through an integrative bioinformatics approach. *Front. Cell Dev. Biol.* **2021**, *9*, 620248. [CrossRef]
98. Kontou, P.; Pavlopoulou, A.; Braliou, G.; Bogiatzi, S.; Dimou, N.; Bangalore, S.; Bagos, P. Identification of gene expression profiles in myocardial infarction: A systematic review and meta-analysis. *BMC Med. Genom.* **2018**, *11*, 109. [CrossRef]
99. *StataCorp, Stata Statistical Software: Release 13*; StataCorp LP: College Station, TX, USA, 2013.
100. Alnasir, J.; Shanahan, H.P. Investigation into the annotation of protocol sequencing steps in the sequence read archive. *GigaScience* **2015**, *4*, 23. [CrossRef] [PubMed]
101. Kim, D.; Langmead, B.; Salzberg, S.L. Hisat: A fast spliced aligner with low memory requirements. *Nat. Methods* **2015**, *12*, 357–360. [CrossRef] [PubMed]
102. Pertea, M.; Pertea, G.M.; Antonescu, C.M.; Chang, T.C.; Mendell, J.T.; Salzberg, S.L. Stringtie enables improved reconstruction of a transcriptome from rna-seq reads. *Nat. Biotechnol.* **2015**, *33*, 290–295. [CrossRef] [PubMed]
103. Subramanian, A.; Tamayo, P.; Mootha, V.K.; Mukherjee, S.; Ebert, B.L.; Gillette, M.A.; Paulovich, A.; Pomeroy, S.L.; Golub, T.R.; Lander, E.S.; et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 15545–15550. [CrossRef] [PubMed]