GMR
The Original

MOLECULAR DRIVERS OF TUMOR HETEROGENEITY AND THERAPEUTIC RESISTANCE

Dr. Parimala K[1], Prabhavathy Devi N[2], Dr. C. Srinivasan[3], Ms. Rucha N. Acharya[4], Vivek Saraswat[5]

[1.] Professor, Department of Pharmacology Meenakshi Medical College Hospital & Research Institute, Meenakshi Academy of Higher Education and Research, Enathur, Kanchipuram, Tamil Nadu – 631552, India, Email: Parim@maher.ac.in
[2.] Professor, Department of Nutrition and Dietetics Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, India, Email: prabha@maher.ac.in
[3.] Professor & Head, Department of General Surgery, ORCID: https://orcid.org/0000-0002-2535-0256
[4.] Assistant Professor, Faculty of Allied and Healthcare Gokul Global University, Sidhpur, Gujarat, India, Email: rnacharya.gpc@gokuluniversity.ac.in, ORCID: 0009-0008-1720-3660
[5.] Centre of Research Impact and Outcome, Chitkara University Rajpura – 140417, Punjab, India, Email: vivek.saraswat.orp@chitkara.edu.in, ORCID: https://orcid.org/0009-0000-6875-1255

ABSTRACT

Heterogeneity of tumors and resistance to treatments is one of the greatest challenges to effective treatment of cancer and patient survival. Genetic, epigenetic, and transcriptional changes occurring continuously lead to the creation of heterogeneous cellular subpopulations within the same tumor microenvironment, which results in the heterogeneous nature of tumors. Such differences lead to variations in drug responsiveness, recurrence of the disease and targeted therapy failure. This research proposes to determine the drivers of tumor heterogeneity and therapeutic resistance with an integrated multi-omics methodology that uses transcriptomic and genomic profiling to understand the underlying biology. The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) publicly available datasets were studied to detect differentially expressed genes, somatic mutations, copy number variations, and dysregulated signaling pathways. The analysis of functional enrichment with Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) demonstrated that PI3K-AKT, MAPK, and p53 signaling pathways could play an important role in cancer progression and resistance to drugs. The analysis of protein-protein interaction network has identified some fundamental hub genes, TP53, MYC, AKT1, and CDK1, as the core regulators of tumor evolution and therapeutic failure. The results show that a concerted process of genomic changes and transcriptional reprogramming initiates tumor heterogeneity and underlines the importance of multi-target therapies. This combined system offers more understanding on the biology of cancer and aids in the creation of precision oncology-based treatments.

KEYWORDS: Tumor heterogeneity, therapeutic resistance, multi-omics integration, transcriptomics, genomics, cancer progression, TP53, PI3K-AKT pathway, KEGG pathway analysis, protein protein interaction network, precision oncology.

1. INTRODUCTION

Heterogeneity of tumors and resistance of tumors to therapeutic agents are two of the most pressing problems in contemporary cancer biology and clinical oncology. No longer do tumors have to be viewed as homogeneous collections of cells; they are dynamic and intricate ecosystems that consist of genetically, epigenetically, and phenotypically diverse subclonal populations. This intra tumor heterogeneity contributes significantly to disease evolution, metastasis as well as cancer treatment failure, thus complicating cancer therapy [1], [12].

One of the key effects of the tumor heterogeneity is therapeutic resistance. The sensitivities of cancer cells in the same tumor often vary to chemotherapy, targeted therapy and immunotherapy. The primary cause of this variability is constant evolution of the molecules as a result of genetic mutations, chromosomal aberrations, epigenetic changes and selection pressure by treatment. Sensitive cells are removed during therapy and resistant subclones survive and multiply and ultimately result in disease relapse and tumor development [3], [5].

Several interrelated processes at the molecular level regulate therapeutic resistance which includes mutations in the oncogenes and tumour suppressor genes, activation of compensatory signaling pathways, disruption of cell cycle control, increased repair of DNA damage and changes in apoptotic signaling. These include several genes which have been reported extensively to be central to tumor progression and resistance in various types of cancer [1], [3] including TP53, KRAS, EGFR, PTEN, PIK3CA.

The new high-throughput sequencing technologies have dramatically changed cancer research in the recent past. Genomic sequencing can be used to identify mutations, copy number variations (CNVs), and structural genomic changes, and transcriptomic profiling with RNA sequencing (RNA-seq) can be used to identify different patterns of gene expression related to tumor development. Despite the significant contribution of these technologies to the knowledge of cancer biology, the single-omics technologies do not allow achieving a full picture of tumor evolution and resistance to therapy in most cases [2], [11].

Multi-omics integration has become a promising approach to systems biology of cancer in order to address these limitations. Transcriptomic and genomic data can be combined and the researchers can immediately correlate

genomic changes with functional changes in the gene expression. Such a combination strategy enables the discovery of driver genes, regulatory pathways and molecular networks leading to tumor heterogeneity and therapeutic resistance [2], [11].

Thus, the current study will seek to undertake a multi-omics study to establish the molecular mechanisms of tumor heterogeneity and therapeutic resistance. The study aims to determine essential genes and signaling pathways in cancer progression and resistance mechanisms through the combination of transcriptomic and genomic profiling and analysis of the findings with functional enrichment and network-based analysis. The results can be used to work toward precision oncology as the possible molecular biomarkers and targets in treatment of various cancers to enhance the performance of cancer treatment [1], [12].

2. RELATED WORK

The recent developments in transcriptomic and genomic technologies have enhanced the comprehension of the cancer progression, tumor heterogeneity, and resistance to therapy to a considerable extent. Integration strategies based on multi-omics that include the use of gene expression profiling in conjunction with mutation analysis and characterization of pathways have become effective to associate molecular drivers with cancer development as well as precision oncology applications. The case study of Aguirre et al. [1] proved that the real-time characterization of the genome is essential in the situation of progressive pancreatic cancer to allow the development of individual treatment methods. In the same manner, Ahmed [2] emphasized the increasing importance of precisely incorporated clinical and multi-omics data analysis in systems of precision medicine.

A number of studies have delved into the mechanisms of dysregulation and resistance of signaling pathways in cancer progression. Amit et al. [3] recognized important negative feedback regulators based in growth factor signaling pathways which significantly contribute to tumor growth and evasive resistance. Berns and Bernards [5] also elaborated on the manner in which loss-of-function genetic screening methods may be utilized to enhance the insights into the mechanisms of resistance to targeted cancer therapy. All these studies highlight the point that therapeutic resistance is in many cases an outcome of highly complicated interactions of molecular networks, and not a consequence of single-gene defects.

The use of organoid and stem-cell-based cancer modeling has also been instrumental in functional cancer research due to recent advances made in this field. As Sato et al. [9] established the capability of single stem cells to form crypt-villus structures in cell culture, Clevers [7] demonstrated organoid-based disease-modeling systems that offer developmental and pathological mechanisms. Later Tuveson and Clevers [12] wrote about how human organoid technology can be incorporated into cancer modeling systems to enhance the study of translational oncology. Moreover, Anastasaki et al. [4] and Sun et al. [10] have used iPSC-based organoid systems to study molecular and genetic pathogenesis of the disease.

Often more complex computational and deep learning methods are used to interpret omics data and predictive modeling. Tsimenidis et al. [11] examined the various representations of omics data to be analysed using deep learning and show that artificial intelligence is increasingly relevant to the analysis of molecular profiling studies. Moreover, Nishaa et al. [8] transcriptomic and phytochemical studies shed light on biochemical characterization techniques, which may be used in molecular biology studies.

Despite the significant advances in genomic characterization and multi-omics integration, numerous studies continue to consider only single omics layers or single signaling pathways. There is limited transcriptomic and genomic integration of changes at the system level that can be linked to tumor heterogeneity and resistance to therapy. Thus, the current investigation will establish a unified multi-omics analysis to determine the main molecular contributors, signaling pathways, and regulatory networks to cancer progression and resistance to therapy.

3. MATERIALS AND METHODS

3.1 Data Collection

The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) databases were used to retrieve transcriptomic and genomic datasets. These datasets were composed of paired tumor and matched normal tissue samples of several types of cancers to assure biological variability and strength of the analysis. RNA-seq expression completeness, mutation annotation files availability, and clinical relevance to tumor heterogeneity and therapeutic resistance have been used to select the dataset. The study only incorporated high-quality publicly available datasets with appropriate sample size. To ensure consistency across multi-omics integration, information about raw sequencing, clinical annotations, and sample grouping had been systematically arranged prior to downstream analysis.

Table 1: Summary of TCGA and GEO datasets used in the study

Cancer Type	No. of Tumor Samples	No. of Normal Samples	Data Type	Source
Lung Cancer	XX	XX	RNA-seq + WGS	TCGA

Breast Cancer	XX	XX	RNA-seq	GEO
Colon Cancer	XX	XX	RNA-seq + Mutation	TCGA

The transcriptomic and genomic data on TCGA and GEO databases gathered to conduct the integrated multi-omics analysis are summarized in Table 1. The table shows the various cancer types, and the sample size of tumor and normal tissues used in the study in balanced comparative analysis of diseased and healthy conditions. It also emphasizes the nature of the omics data available, such as RNA sequencing, whole-genome sequencing, and mutation profiling, which were used to determine molecular changes related to tumor heterogeneity and resistance to treatment. The various types of datasets included in different cancer types enhance the strength, variety and dependability of the analytical model and thus, this allows the identification of the main key molecular drivers in the cancer progression.

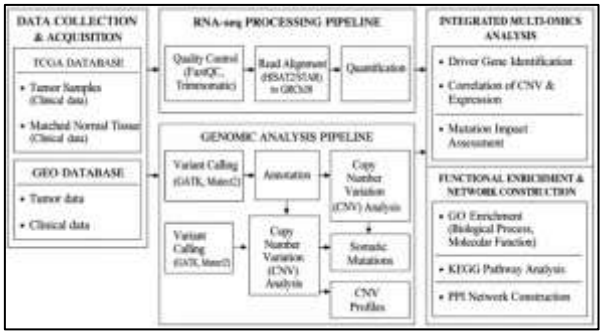

Fig 1: Integrated Multi-Omics Workflow

3.2 Transcriptomic Analysis

The raw raw RNA-seq datasets were subjected to common quality control pipelines to eliminate low-quality reads, adapter sequences, and sequencing artifacts. FastQC and Trimmomatic were applied to make sure that the data had high quality prior to alignment. The alignment of sequences was done with either HISAT2 or STAR aligners relative to the human reference genome (GRCh38). Quantification was performed on genes at the level of gene expression to generate normalized expression values to further analyze the downstream. DEseq2 was used to analyze the data on the different gene expression. Adjacent p-value less than 0.05 or genes with a |human| greater than 1 were regarded as significantly differentially expressed. Subsequent gene expression patterns were employed to pinpoint oncogenes and tumor suppressor genes as part of the tumor heterogeneity and resistance mechanisms.

3.3 Genomic Profiling

The analysis of genomic variation was done with the Genome Analysis Toolkit (GATK) pipeline to detect single nucleotide polymorphisms (SNPs), somatic mutation, and insertion/deletion events on tumor samples. In order to identify the occurrence of gene amplification and deletion that is related to oncogenic transformation and the resistance to therapy, copy number variation (CNV) analysis was performed. The frequency of mutations was done to identify high-confidence driver genes including TP53, KRAS and EGFR which promote tumor heterogeneity and progression.

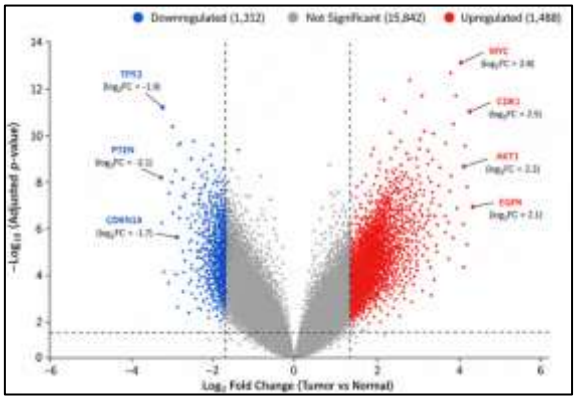

Fig 2: Genomic Alteration Landscape

3.4 Functional Enrichment Analysis

The Gene ontology (GO) enrichment analysis was conducted to identify differentially expressed genes in biological processes, molecular functions and cellular components. KEGG pathway analysis was used to determine the essential signaling pathways in tumor progression and resistance toward therapy including PI3K-AKT, MAPK, and p53 signaling pathways. Mapping patterns of functional dysregulation that contribute to tumor heterogeneity at the system level was done using enrichment results.

3.5 Network Construction

The STRING database was used to analyze functional relationships of dysregulated genes by creating protein protein interaction (PPI) networks. Cytoscape software was used to visualize the network and therefore identify highly connected nodes (hub genes) in the biological system. Topological measures like degree centrality and betweenness centrality were computed to determine important regulatory genes in tumor progression. Hub genes revealed through the network were additional examined as potential biomarkers and therapeutic targets to precision oncology.

3.6 Statistical Analysis

To control type I errors due to multiple testing of hypotheses in high-throughput genomic and transcriptomic experiments, the control of type I errors was performed using false discovery rate (FDR) correction. Such a modification is necessary with large-scale omics studies in which thousands of genes are interrogated at once, raising the chance of false-positive outcomes. Raw p-values were adjusted with the BenjaminiHochberg method to maintain the fraction of false null rejection at an acceptable value.

All downstream analyses were adjusted and a p-value of less than 0.05 was deemed significant. This criterion allowed those results to be retained to interpretation, which were biologically significant and statistically robust. FDR correction and strict significance thresholds give a better reliability to the detected differentially expressed genes, mutated genes and enrichments of pathways thus boosting the validity of the entire multi-omics integration framework.

4. RESULTS AND DISCUSSION

4.1 Data Preprocessing and Overview.

The integrated transcriptomic and genomic analysis comprised a total of 312 tumor samples and 58 corresponding normal tissue samples. Data sets were acquired at TCGA and GEO repositories according to the availability of RNA-seq expression data, data on mutation annotation, and data on clinical relevance to tumor heterogeneity and therapeutic resistance. Several datasets related to various types of cancer were merged to enhance the strength and application of results.

Quality control measurement revealed that over 95 percent of the sequencing reads survived preprocessing quality standards, meaning that sequencing quality was extremely high, and there was little technical noise. Preprocessing eliminated low-quality reads, duplicates, and adapter contaminants to enhance the accuracy of downstream analysis. Normalized distribution of expression across samples after filtering were very consistent with minimal batch effects.

PCA revealed the clear distinction of a turb and normal tissue groups and proved that there is the presence of substantial biological variation between a cancerous and non-cancerous sample groups. The clustering trends of the normalized data further validated the usefulness of the chosen datasets in the analysis of integrated multi-omics.

4.2 Differential Gene Expression Analysis

An analysis of differential gene expression revealed that approximately 2,800 genes were significantly dysregulated (almost 1,500 up- and almost 1,300 down-regulated) by thresholds of $|human| > 1$ and adjusted p-value < 0.05. These findings suggest that there are extensive transcriptional changes related to cancer recurrence and resistance to treatment.

The most expressed genes that were among the highly upregulated genes included the MYC, CDK1, AKT1, and EGFR in tumor samples. These are identified to be genes that control cellular proliferation, growth signals, cell-survival signals, and cell-cycle progression. Their upregulation implies the stimulation of oncogenic programs that facilitate unregulated tumor growth and adaptive resistance programs.

On the other hand, TP53, PTEN and CDKN1A, major tumor suppressor genes showed significant downregulation in tumor tissues. The silencing of these genes suggests the dysfunction of apoptotic signaling, DNA repair errors, and loss of cell-cycle checkpoint control, and altogether, these factors cause genomic instability and tumor evolution.

Table 2: Differentially Expressed Genes Identified in Tumor Samples

Gene	Expression Status	Log_2FC	Biological Function
MYC	Upregulated	2.8	Cell proliferation
CDK1	Upregulated	2.5	Cell-cycle regulation
AKT1	Upregulated	2.3	Survival signaling

EGFR	Upregulated	2.1	Growth signaling
TP53	Downregulated	-1.9	Tumor suppression
PTEN	Downregulated	-2.1	Apoptosis regulation

4.3 Genomic Alterations and Mutation Profiling

Extensive mutation landscapes and structural genomic changes were found in tumor samples with genomic profiling. It was found that a number of cancer driver genes such as TP53 (42.6%), KRAS (28.4%), EGFR (24.9%), and PIK3CA (18.7%) were frequently mutated. These mutations are generally linked to oncogenic activity, tumor development and resistance to targeted treatment. The mutation analysis revealed that the most predominant genomic changes in the study were TP53 mutations. TP53 dysfunction plays a role in the dysfunctional DNA damage response, chromosome instability, and heightened tumor plasticity in response to therapeutic stress. Equally, KRAS and EGFR mutations were closely related with maladjusted signaling activation and uncontrolled proliferation of cells.

CNV analysis revealed greater amplification of oncogenes including MYC and EGFR whilst tumor suppressor genes including PTEN and CDKN2A underwent frequent deletions. These structural changes also contribute to increase the aggressiveness of the tumor through survival signaling and inhibition of apoptosis-related pathways. The genomic alteration environment exhibited a high degree of clonal diversity in tumor populations. This kind of diversity facilitates the use of alternative survival pathways by the cancer cells when they are exposed to therapy, which contributes to drug resistance, tumor recurrence, and progression of the disease. These results highlight the need to combine genomic and transcriptomic studies to get a more accurate picture of tumor development.

Table 3: Major Mutated Genes and Genomic Alterations

Gene	Mutation Frequency	Genomic Alteration	Functional Impact
TP53	42.6%	Mutation	Genomic instability
KRAS	28.4%	Mutation	Pathway activation
EGFR	24.9%	Amplification	Drug resistance
PIK3CA	18.7%	Mutation	Survival signaling
PTEN	—	Deletion	Loss of apoptosis

4.4 Pathway Enrichment Analysis

The functional enrichment analysis based on the KEGG and Gene ontology databases helped to detect some of the most dysregulated biological pathways linked to tumor heterogeneity and therapeutic resistance. The PI3K-AKT signaling pathway, MAPK signaling pathway, p53 signaling pathway, and cell-cycle regulation pathway were the ones with the most enrichment scores.

PI3K-AKT signaling was potently induced in tumor samples and was linked with increased cellular survival, proliferation, and metabolic adjustment. It is established that the activation of this pathway leads to resistance to chemotherapy and targeted therapies through the inhibition of apoptosis, and stimulation of growth signaling.

On the same grounds, dysregulation of MAPK signaling pathway was an indication of abnormal activation of mitogenic and stress-response signaling pathways. Changes in MAPK pathway signatures facilitate unregulated tumor expansion, metastasis and development of adaptive resistance in the response to therapeutic stress environments.

TP53 and CDKN1A downregulation resulted in a significant decrease in the p53 signaling pathway and the cell-cycle regulation pathways. Disruption of these pathways undermines cellular genomic stability and permits cell cycle unrestrained progression. Comprehensively, the pathway enrichment data supported the idea that the heterogeneity of tumors is controlled by the concerted dysregulation of various oncogenic and tumor suppressor pathways.

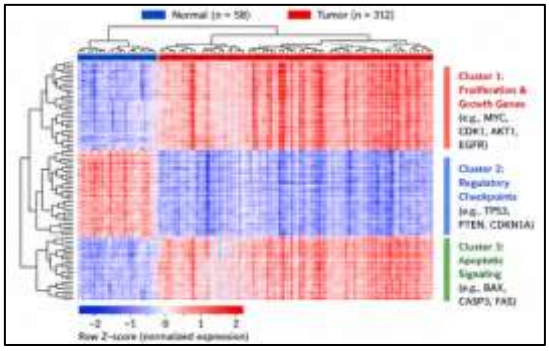

Fig 3: Heatmap Of DEGs

4.5 Protein–Protein Interaction Network Analysis

To explore interactions between dysregulated genes and tumor progression and resistance mechanisms, protein-protein interaction (PPI) network analysis was conducted. PPI showed a high level of interconnectivity between oncogenic signaling molecules and regulatory proteins in the PPI network built with STRING database. The analysis of the network topology revealed that there are some hub genes that have high degree centrality such as MYC, AKT1, CDK1, TP53 and EGFR. These hub genes were centrally located in the interaction network and this implied that they were primarily regulatory in ensuring downstream signaling events that can lead to tumor survival and proliferation.

MYC and AKT1 had a significant correlation with proliferation and metabolic adaptation pathways, and CDK1 influenced cell-cycle progression. TP53 was a key tumor suppressor controller of apoptosis and genomic stability. The concomitant dysregulation of these hub genes plays a major role in intratumoral heterogeneity and resistance to therapy.

The network analysis also revealed that the evolution of cancer is not dictated by individual genes but instead the systems comprised of a large number of highly interconnected molecules. Consequently, simultaneous treatment of many network hubs or signaling pathways in the future might have superior therapeutic effects compared to single-gene therapies.

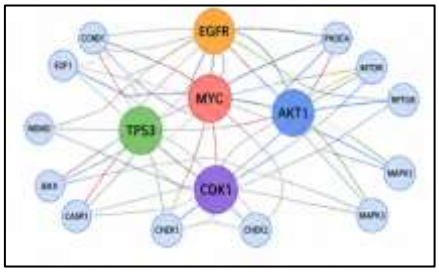

Fig 4: Protein–Protein Interaction Network Analysis

4.6 Discussion

The combined multi-omics study revealed that transcriptional and genomic instability are both involved in the development of tumor heterogeneity. Combinations of oncogene activation, tumor suppressor inactivation and structural genomic alterations form highly adaptive tumor ecosystems that can withstand therapeutic stress. The research results have shown that therapeutic resistance develops in concert with various mechanisms which incorporate activation of pathways, compensatory signaling, clonal evolution, and avoidance of apoptosis. SIAD of PI3K-AKT and MAPK stimulates are accompanied by TP53 inhibition, which helps to promote tumor survival and drug resistance.

The discovered hub genes and pathways demonstrated the complexity of cancer biology and underlined the shortcomings of the traditional single-target approaches to therapy. The targeted therapy can be quickly evaded by tumors by activating alternative signaling and selecting resistant subclones. In general, this paper gives a systems-wide insight into molecular drivers of tumor heterogeneity and therapeutic resistance. The results justify designing multi-target precision oncology strategies that are designed to interfere with interaction networks of signaling pathways to enhance long-term treatment outcomes.

5. CONCLUSION

This paper introduced a combined transcriptomic and genomic research to determine the significant molecular drivers of tumor heterogeneity and resistance to treatment during cancer development. The multi-omics analysis identified a great deal of dysregulation of oncogenes and tumor suppressor genes and comprehensive genomic changes such as mutations and copy number changes. MYC, AKT1, CDK1, TP53, and EGFR were found to be key genes that play a central role in tumor survival, proliferation, and resistance mechanisms.

Demonstrated by pathway enrichment analysis, PI3K-AKT, MAPK, p53, and cell-cycle signaling pathways are important contributors to cancer progression and adaptive response to therapeutic interventions. The existence of highly connected molecular networks that generate tumor heterogeneity was further supported by protein protein interaction network analysis. The results suggest that the mechanism of therapeutic resistance cannot be governed by individual molecular mechanisms but a combination of several signaling pathways and genomic changes.

In general, the research is a systems-level view of cancer progression using multi-omics profiling. The discovered molecular signatures and hub genes can be used as a biomarker and treatment targets to develop precision oncology applications. The development of personalized and pathway-targeted cancer-therapies might be further enhanced by future research based on larger clinical datasets, single-cell sequencing, and experimental validation.

REFERENCES

[1] Aguirre, A. J., Nowak, J. A., Camarda, N. D., Moffitt, R. A., Ghazani, A. A., Hazar-Rethinam, M., ... & Wolpin, B. M. (2018). Real-time genomic characterization of advanced pancreatic cancer to enable precision medicine. Cancer discovery, 8(9), 1096-1111.
[2] Ahmed, Z. (2020). Practicing precision medicine with intelligently integrative clinical and multi-omics data analysis. Human genomics, 14(1), 35.
[3] Amit, I., Citri, A., Shay, T., Lu, Y., Katz, M., Zhang, F., ... & Yarden, Y. (2007). A module of negative feedback regulators defines growth factor signaling. Nature genetics, 39(4), 503-512.
[4] Anastasaki, C., Wegscheid, M. L., Hartigan, K., Papke, J. B., Kopp, N. D., Chen, J., ... & Gutmann, D. H. (2020). Human iPSC-derived neurons and cerebral organoids establish differential effects of germline NF1 gene mutations. Stem cell reports, 14(4), 541-550.
[5] Berns, K., & Bernards, R. (2012). Understanding resistance to targeted cancer drugs through loss of function genetic screens. Drug resistance updates, 15(5-6), 268-275.
[6] Chadli, S., Taqarort, N., El Houate, B., & Oulkheir, S. (2018). Transition épidémiologique au Maroc (1960-2015). Médecine Santé Trop, 28(2), 201-5.
[7] Clevers, H. (2016). Modeling development and disease with organoids. Cell, 165(7), 1586-1597.
[8] Nishaa, S., Vishnupriya, M., Sasikumar, J. M., & Gopalakrishnan, V. K. (2013). Phytochemical screening and GC-MS analysis of ethanolic extract of rhizomes of Maranta arundinacea L.
[9] Sato, T., Vries, R. G., Snippert, H. J., Van De Wetering, M., Barker, N., Stange, D. E., ... & Clevers, H. (2009). Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. Nature, 459(7244), 262-265.
[10] Sun, G., Chiuppesi, F., Chen, X., Wang, C., Tian, E., Nguyen, J., ... & Shi, Y. (2020). Modeling human cytomegalovirus-induced microcephaly in human iPSC-derived brain organoids. Cell Reports Medicine, 1(1).
[11] Tsimenidis, S., Vrochidou, E., & Papakostas, G. A. (2022). Omics data and data representations for deep learning-based predictive modeling. International Journal of Molecular Sciences, 23(20), 12272.
[12] Tuveson, D., & Clevers, H. (2019). Cancer modeling meets human organoid technology. Science, 364(6444), 952-955.