UNIVERSITÀ
degli STUDI
di CATANIA

Dipartimento
di Fisica
e Astronomia
"Ettore Majorana"

PhD PROGRAMME IN COMPLEX SYSTEMS FOR
PHYSICAL, SOCIO-ECONOMIC AND LIFE SCIENCES
XXXVIII CYCLE

GIULIA CHISARI

Phenotypic Dynamics of Rectal Cancer under Neoadjuvant Therapy: A Multi-omic Single-cell Perspective from Patient-Derived Organoids

---

PhD THESIS

SUPERVISOR:
CHIAR.MO PROF. A. PULVIRENTI

CO- SUPERVISOR:
DOTT. S. FORTE

ACADEMIC YEAR 2022/2025

Abstract

This work investigates the mechanisms of phenotypic adaptation and intratumor heterogeneity (ITH) in colorectal cancer, with a particular focus on locally advanced rectal cancer. Therapy resistance remains one of the major challenges in oncology, and understanding the adaptive trajectories available to tumor cells is crucial to identify molecular dependencies that may be exploited therapeutically. To address this, we established patient-derived organoids (PDOs) as high-fidelity models of rectal cancer and used them to recapitulate the therapeutic regimens administered to the donor patients.

By coupling longitudinal sampling with multi-resolution transcriptomic profiling, including bulk, single-cell, and long-read sequencing approaches, we traced the evolution of tumor cells under treatment pressure. This strategy enabled the identification of transcriptional programs, phenotypic states, and splicing isoforms that emerge during therapy-induced selective bottlenecks. The analysis revealed that clonal evolution and phenotypic plasticity play pivotal roles in shaping resistance, highlighting how treatment can enrich resistant subpopulations that ultimately drive disease recurrence.

Overall, this work underscores the importance of investigating phenotypic heterogeneity alongside genomic variability to fully comprehend tumor adaptation. The integration of patient-derived organoid models with advanced sequencing technologies provides new insights into the dynamics of therapeutic resistance, identifies critical windows of tumor plasticity, and highlights potential vulnerabilities that could inform the development of novel treatment strategies in colorectal cancer.

Contents

List of figures

21. Gene set enrichment analysis (GSEA) of Hallmark pathways in patient-derived organoids (PDOs) treated at T1 versus untreated controls (T0).
22. Vulcano Plot of Differentially Expressed Genes in PDOs T2 relapse vs T0 control
23. Gene ontology enrichment analysis of differentially expressed genes in PDOs T2 relapse vs T0 control
24. Analysis of alternative splicing events in selected genes in PDOs. Comparison between the relapse condition (T2) and control (T0)
25. Gene set enrichment analysis (GSEA) of Hallmark pathways in patient-derived organoids (PDOs) at relapse (T2) compared to untreated controls (T0).
26. Vulcano Plot of Differentially Expressed Genes in PDOs T2 relapse vs T1 treated
27. Gene ontology enrichment analysis of differentially expressed genes in PDOs T2 relapse vs T1 treated
28. Analysis of Alternative splicing events in selected genes in PDOs. Comparison between the relapse condition (T2) vs T1 treated
29. Gene set enrichment analysis (GSEA) of Hallmark pathways in patient-derived organoids (PDOs) at relapse (T2) compared treated (T1).
30. UMAP projection of annotated cell clusters.
31. Marker-based annotation of cell clusters.
32. UMAP visualization of cell populations across timepoints and treatment conditions.
33. Dynamics of cell population proportions across timepoints and treatment conditions

List of Tables

Chapter 1. Introduction

1.1 LARC (LOCALLY ADVANCED RECTAL CANCER)

Cancer represents a major societal, public health, and economic challenge in the 21st century, accounting for nearly one in six deaths (16.8%) and one in four deaths (22.8%) from noncommunicable diseases (NCDs) worldwide.

Colorectal cancer (CRC) comprises a heterogenous group of neoplastic disorders arising from the malignant transformation of epithelial cells lining the external surface of the large intestine (colon and rectum).

According to GLOBOCAN 2022 estimates by the International Agency for Research on Cancer (IARC), colorectal cancer (CRC) accounts for 9.6 % of all cancers diagnosed worldwide [1].

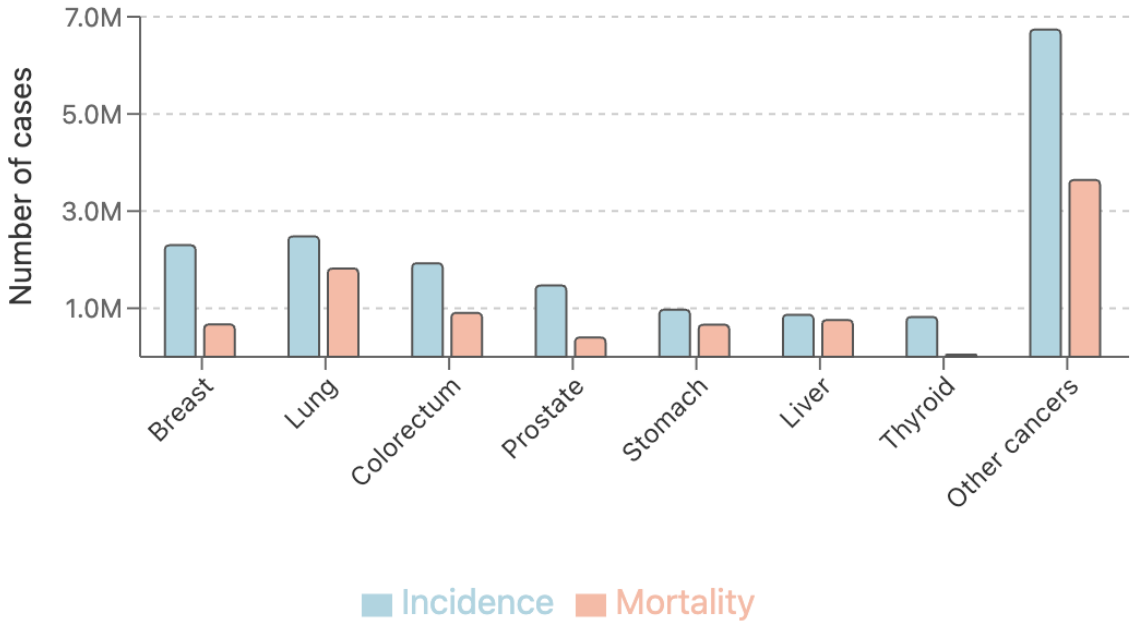


Figure 1. Incidence and Mortality numbers for the most prevalent cancers. The figure shows the mortality and incidence numbers of the most prevalent cancers worldwide in the year 2022 (GLOBOCAN 2022). Worldwide colorectal cancer is the third most prevalent and accounts for the second most cancer-related deaths.

In Italy, CRC is the most frequently diagnosed malignancy, representing 13% of all new cancer cases annually in both sexes. It ranks second among women, after breast cancer, and third among men, after lung and prostate cancer.

Moreover, it is the second leading cause of cancer-related mortality in both sexes, accounting for 11% of all cancer deaths recorded in the country [2].

CRC is a complex, multifactorial disease, influenced by genetic and environmental risk factors [3], including positive family history, hereditary CRC syndrome, inflammatory bowel disease, a high-fat diet, lacking physical exercise, and age [4].

Due to its initially asymptomatic disease, CRC is often diagnosed at an advanced stage. As a result, many patients present with locally or metastatic disease at diagnosis [5].

Diagnosis is typically based on a digital rectal examination (DRE) and endoscopy evaluation with biopsy for histopathological confirmation [6]. Staging follows the TNM System, based on tumor invasion depth (T stage), lymph node involved (N stage) and distal metastasis (M stage) in four stages.

Stage II (cT3-T4, N0) and stage III (cT1-T4, N1-N3) rectal cancers are commonly categorized as Locally Advanced rectal cancer (LARC), a group of highly heterogeneous tumors [7].

Since the late 1980s, the management of LARC has witnessed gradual progress, due to precise staging and risk stratification of magnetic resonance imaging (MRI), the performance of neo-chemoradiotherapy (neo-CRT) and total mesorectal excision surgery (TME). These measures have reduced five-year local recurrence rates from over 25% to approximately 5-10% and have improved sphincter preservation in patients [8,9,10].

Standard treatment for LARC involves long-course neoadjuvant concurrent chemoradiation therapy or short-course radiation without chemotherapy, followed by TME and optional adjuvant chemotherapy. However, this standard approach has several limitations, including low pathological complete response (pCR) (10–25%), high metastasis rate (30–35%), and highly inconsistent compliance with adjuvant chemotherapy (25–75%).

To address these issues, in 2021, the intensification of neoadjuvant treatment started to be implemented, defining the Total Neoadjuvant Treatment (TNT) strategy. The results obtained in terms of local control and survival supported the TNT as a new standard of care for LARC [11].

TNT incorporates both chemotherapy and radiotherapy before surgery to enhance tumor response and reduce systemic failure. TNT can be administered in two main strategies: induction chemotherapy followed by CRT or CRT followed by consolidation chemotherapy, both of which have shown an improving treatment outcome.

1.2 Intratumor Heterogeneity (ITH)

Despite advances that cancer therapies have made in recent years, resistance to systemic therapies for advanced cancer is still a challenge to overcome. This resistance is not limited to targeted therapies but is also a significant problem in conventional chemotherapy and radiotherapy, which remain the backbone of cancer treatment in many settings. Tumor cells often develop or possess intrinsic mechanisms to evade the cytotoxic effects of chemotherapeutic agents and ionizing radiation, leading to therapeutic failure. Such resistance is largely driven by the phenotypic heterogeneity within tumors, where distinct subpopulations of cancer cells differ in their sensitivity to treatments due to variations in DNA repair capacity, drug metabolism, cell cycle status, and microenvironmental interactions [12]. Unfortunately, most advanced cancers, including those with robust initial responses, eventually acquire resistance to therapies and relapse.

This therapeutic failure is largely attributable to traditional approaches about therapeutic decisions that are carried out at the whole-population level. Standard treatment strategies, in fact, target a single receptor or pathway, treating cancer as a homogenous disease.

In the last decades, attention has been focused on a fundamental property of many solid tumors: intra-tumor heterogeneity (ITH).

Tumor heterogeneity refers to the existence of subpopulation of cells, with distinct genotypes and phenotypes that may harbor divergent biological behaviors within a primary tumor [13].

Intratumor heterogeneity arises through various mechanisms.

In clonal evolution models, stochastic accumulation of mutations through genomic instability results in increasing genetic diversity, with the tumor acquiring subclones with distinct genotypes over time.

This model, initially introduced by Peter Nowell in 1976, is based on the hypothesis that tumor initiations occur in a stochastic manner, beginning with an induced change in a previously non-malignant cell that confers a selective growth advantage and leads to neoplastic proliferation. Subsequently, the genomic instability of the expanding tumor population creates additional genetic diversity that is subjected to evolutionary selection pressures, resulting in the sequential emergence of increasingly genetically abnormal and heterogeneous subpopulations. In this model of tumor evolution, linear evolution describes evolution owing to the successive acquisition of mutations that confer a growth and/or survival advantage, with sequential clones harboring these advantageous mutations outcompeting ancestral clones [14].

Alternatively, heterogeneity is also generated through cellular differentiation. Cancer stem cell (CSC) models, cancers are hierarchically organized with a stem cell-like population, sustaining tumor growth through self-renewal and differentiation. This evolution denotes the emergence and divergent propagation of multiple subclonal tumor cell populations that share a common ancestor. Branched evolution enables a greater level of opportunity to create a more heterogeneous tumor. It should be noted however, that this hierarchy leading is not a one-way route, but can be reversible or plastic whereby the terminally differentiated cells can also dedifferentiate and gain CSCs properties under specific conditions [15].

These models are not mutually exclusive and act together to create a complex system with multiple layers of heterogeneity established by the distinct genetic, epigenetic, transcriptomic, proteomic and functional properties of different cells [16].

Cellular phenotypic heterogeneity within a tumor is a complex, multifactorial phenomenon, which integrates genetic, epigenetic and environmental inputs [17].

While ITH has traditionally been explored from a genomic perspective, recent evidence highlights the importance of phenotypic heterogeneity, particularly transcriptomic variation, in shaping tumor behavior.

Because the transcriptome is a feature of the cancer cell phenotype, it is natural to view changes in expression, and the pattern of transcriptomic intratumor heterogeneity as "functional" and the substrate for the tumor evolution. Potentially tITH could be driven entirely by underlying heritable (epi)genetic variation that evolves during tumor growth. However, the observation that local invasion is polyclonal in both CRC and early breast cancer challenges the notion that cancer cell phenotype (here, the ability to invade) is driven solely by the accrual of genetic mutations. Furthermore, observations of rapid transcriptional shifts following treatment (for example, in melanoma) and, in CRC, variation in subclone proliferation rates through serial retransplantation despite largely stable patterns of genetic alterations, discount the notion that transcriptomic phenotypes are determined solely by clonal replacement. It has previously been determined that most driver mutations are clonal in metastatic CRC, meaning that intratumoral transcriptional variation often happens in the absence of the acquisition of new key driver mutations.

The majority of intratumor variation in gene expression in fact is not strongly heritable, but rather "plastic".

Knowledge of the genotype-phenotype map of cancer cells is limited and thus, while genomics offers us a window into determination of which clones are selected, the methodology provides limited information on precisely why they are selected [18].

So, studies have primarily focused on cancer cells genomes even though selection forces act on phenotypes rather than genotypes. This insight underscores the rationale of this study: investigating phenotypic adaptation in response to pharmacological perturbations using models that recapitulate this complexity.

1.3 Patient-Derived Organoids (PDOs)

The heterogeneity of clinical response to cancer therapies highlights the need to develop more complex, biologically relevant and predictive cellular assays.

The current landscape to understand the genomic, metabolomic and pharmacological bases of tumor relies heavily on traditional models, such as two-dimensional cell lines (2D) and patient-derived tumor xenografts (PDXs). However, these models have inherent limitations that hinder their effectiveness.

Two-dimensional cell lines (2D) are derived from primary patient tumors, but their generation is very inefficient and involves extensive adaptation and selection to in vitro 2D culture conditions. As only rare clones are able to expand and can be maintained over many passages, the derived cell lines may have undergone substantial changes and no longer recapitulate the heterogeneity of the original tumors. PDTXs have the advantage of mimicking the biological characteristics of the human tumor much better than in vitro culture models. PDTXs are generated by transplanting freshly derived patient material subcutaneously or orthotopically into immunodeficient mice. The ability to serially transplant tumor tissues into increasing numbers of animals allows for preclinical testing of novel therapies for cancer treatment. In addition, PDTXs frequently undergo mouse-specific tumor evolution, characterized by the rapid accumulation of copy number alterations (CNAs), often due to the selection of pre-existing minor subclones. This genomic divergence from the parental tumor increases with each passage and can result in the loss of recurrent driver alterations that are under selection in patients but not in the murine host. Consequently, the selection pressures in mice may lead to clonal dynamics and tumor trajectories that diverge substantially from those in human patients, thus potentially affecting the relevance of drug response and biomarker studies conducted in PDTX models [19]. However, the limitations of PDTXs include the use of animals and limited engraftment efficiencies for subsets of patient tumors. Moreover, the approach is expensive, time and resource consuming, and influenced by species-specific stromal interactions that may distort tumor behavior [20].

So, in order to reduce in vivo experimentation by national and community legislation (Modernization Act 2.0) introduced in 2023, it is urgent to establish a model that may accurately reflect the heterogeneity and specificity of cancer to better aid in the clinical diagnosis and treatment of colon cancer.

For these reasons, patient-derived organoids (PDOs) have come to our aid, emerging as a highly promising tool in the field of cancer research.

Patient-derived organoids (PDOs) are three-dimensional (3D), self-organized multicellular structures generated by cancer stem cells (CSC), derived from patient-specific tissue samples, embedded in an extracellular matrix (ECM). They have long-term proliferation and differentiation capacities and mimic key histopathological, genetic and phenotypic features of the parent tumor [21].

The first organoid was successfully established by Hans Clevers in 2009 [22] .

Since then, organoid has increasingly gained in popularity. In recent years, organoid has become a research hotspot, showing significant promise in the biological analysis of the tumor.

More recently, precision medicine studies have been extended to several types of cancer and treatments, demonstrating the feasibility of PDOs for treatment testing, and as a promising tool for predicting patient responses.

Organoids can be used to predict clinical outcomes in CRC [23].

The results of a study illustrate the potential of PDOs in precision treatment of LARC. In fact, was generated a patient-derived organoid (PDOs) biobank from patients with locally advanced rectal cancer (LARC), reported that organoids had similar molecular profiles to those of the patient tumor and that PDOs can predict LARC patient chemoradiation responses, reinforcing their value as a companion diagnostic tool in LARC treatment. In fact, chemoradiation responses in patients are highly matched to PDOs responses, with 84.43% accuracy, 78.01% sensitivity, and 91.97% specificity [24].

So, PDOs can be a biomarker tool to guide and tailor treatment of LARC.

1.4 RNA Sequencing

Patient-derived organoids offer a powerful preclinical model to study tumor biology while preserving the cellular and molecular heterogeneity of the original tissue. However, to accurately dissect the complex transcriptional programs involved in tumor heterogeneity,

therapy resistance, and plasticity, advanced transcriptomic technologies are required. In this context, RNA sequencing (RNA-seq) has become an essential tool to investigate the roles of all RNA species forming the tumor transcriptome.

The elucidation of DNA's double helix structure by Watson and Crick in 1953 laid the foundation for modern molecular biology, highlighting the central role of nucleic acids in the regulation of life processes [25] .

In the following decades, RNA sequencing technologies have undergone rapid evolution, progressively transforming our ability to interrogate the transcriptome with increasing resolution, throughput and cost-effectiveness. The timeline of RNA sequencing technologies is shown in Figure 2 [26].

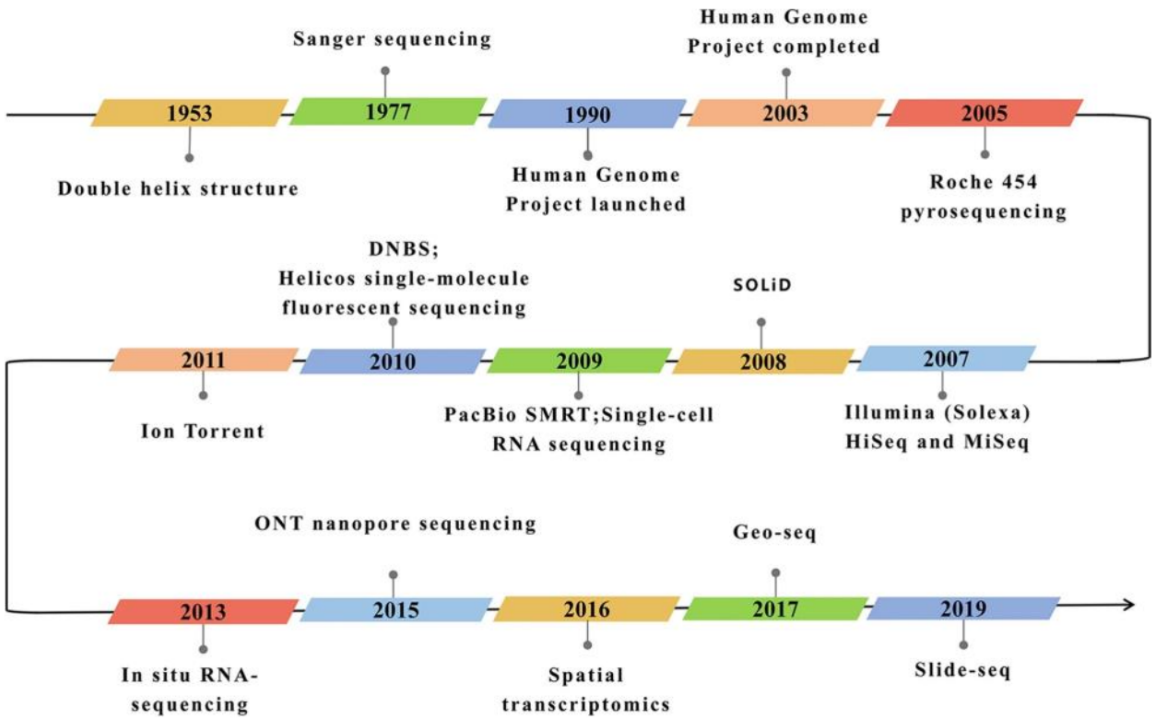


Figure 2. The development timeline of RNA sequencing technologies from Hong et al (2020)

Early transcriptomics profiling relied on hybridization-based microarray technologies and offered a limited ability to fully catalogue and quantify the diverse RNA molecules that are expressed from genomes over wide ranges of levels [27] .

Microarray consists of systematic arrays of cDNAs or oligonucleotides with known sequences, printed or synthesized at specific spots on a glass or silicon surface [28, 29].

This technology enables a comprehensive experimental strategy, allowing the simultaneous examination of transcript levels across entire genomes in response to a variety of stimuli.

This genome- wide transcriptional analysis, or "transcriptional profiling," offers comparative data on the relative expression levels of individual transcript within an organism and correlates these changes to defined cellular stimuli. This broad approach has introduced challenges in data management and has required sophisticated analytical methods to extract biologically meaningful information from the large volumes of primary data [30].

While microarray technology has significantly advanced our understanding of gene expression, the advent of RNA sequencing (RNA-seq) has further revolutionized the field by providing even greater resolution and accuracy in transcriptome analysis.

The first-generation sequencing technology, also referred to as Sanger sequencing, was pioneered by Sanger in 1997 [31], introducing the chain termination method, followed by the chemical degradation method developed by Maxam and Gilbert [32] .

Since the first appearance of high-throughput sequencing in 2005 [33], it has become possible to understand life activities at the molecular level and to conduct detailed research to elucidate the genome and transcriptome and multiple next-generation sequencing platforms have been introduced.

The advent of second-generation sequencing, also known as next-generation sequencing (NGS), marked a revolution in throughput and cost-effectiveness. Platforms such as Illumina and Ion Torrent perform massively parallel sequencing, generating millions of short reads with high accuracy and relatively low cost per base [34, 35].

Ion Torrent platform relies on detecting the hydrogen ion released during the sequencing process. When a nucleotide is incorporated into the DNA chain, while it is duplicated on the bead, a hydrogen ion is released, altering the pH of the solution. This change is detected by a sensor and converted into a voltage signal. By avoiding optical scanning to differentiate nucleotides, the process is faster.

Illumina-based sequencing technology utilizes a technique known as "reversible terminator sequencing." This method involves a single step where samples are randomly fragmented into sequences, to which adapters are ligated. These adapters then bind to complementary adapters on a solid plate. Each fragment is amplified through bridge PCR, creating clusters of identical sequences. Modified nucleotides, specifically fluorescently labeled 3’-O-azidomethyldNTPs, are used to identify each base. In contrast to microarrays, NGS offers the capability to identify previously unknown gene expression patterns [36].

Second-generation approaches, such as on the Illumina or Ion Torrent platforms, generally start with DNA fragmentation, DNA end-repair, adapter ligation, surface attachment, and in-situ amplification. These “short-read” sequencing technologies involve the massively parallel sequencing of short reads (50-200 bp), whereby millions of individual sequencing

reactions occur in parallel [37]. For the past 20 years, short-read sequencing (srRNA-seq) has been the gold standard for transcriptomics. However, sr-RNAseq falls short of fully capturing the complexity and dynamics of transcriptomics [38]. Despite its widespread utility short-read RNAseq is constrained by the resulting 100-150 bp read length making difficult to resolve the connectivity between distant exons, as they are never represented on the same sequenced fragment [39]. Moreover, srRNAseq cannot directly detect epitranscriptome modifications and struggles with mapping repetitive regions, which limits its capacity to distinguish between similar transcripts [40].

RNA sequencing has thus opened a new chapter in the molecular understanding of cancer. It enables the study of differential gene expression, the identification of cancer-specific biomarkers, the investigation of intra-tumor heterogeneity and tumor evolution, the characterization of drug resistance mechanisms, the profiling of the tumor microenvironment and immune landscape, and even the discovery of neoantigens for personalized immunotherapy [41].

Features	ONT (Third- Generation Sequencing)	Illumina (Next-Generation Sequencing)
Sequencing Chemistry	Nanopore-based	Sequencing by synthesis
Average Reads Length	10kb- 4Mb	2 x 300 bp
Base-calling accuracy	99.9%	99.9%
Maximum throughput/run	290 Gb [2]	8 Tb [4]
Complex Genomic region analysis	Yes	No
Pros	Long reads Allows Direct RNA analysis Allows direct methylation Availability of portable sequencer	High Accuracy High Sensitivity High multiplexing capacity High versatility in several application fields
Cons	Sequencing costs are still higher than NGS Bioinformatic requirements	No long read Requires PCR amplification Does not allow direct RNA analysis Low accuracy in complex genomic analysis Time-consuming workflows

Table I. ONT and Illumina comparison to highlight TGS differences with respect to the widely used NGS technology.

1.5 The Third Generation Sequencing: Oxford Nanopore Technologies

To address the limitations of short-read sequencing, third-generation technologies (TGS) emerged, offering a new level of transcriptomic resolution, enabling long-read sequencing of full-length RNA molecules. These approaches allow direct observation of transcript isoforms and splicing events without assembly, significantly enhancing transcriptomic resolution [42].

Currently, two different TGS platforms such as PacBio's Single Molecule Real-Time (SMRT) sequencing [43] and Oxford Nanopore Technologies' nanopore-based sequencing [44] have shown their robustness to profiling transcriptome complexity.

The first TGS technology that accomplished the long-reads goal appeared in 2011 when Pacific Biosciences (PacBio) released its PacBio RS sequencer, still characterized by high error rates (~13%) and relatively long read lengths (~1.5 kb). PacBio or Smart Sequencing technology is based on the use of a silicon chip called a SMRT (Single Molecule, Real-Time) cell. Circularized DNA fragments are loaded in the SMRT cell and flow within the ZMWs, being immobilized at their bottom. A single DNA molecule is immobilized in each ZMW, so when the labeled nucleotides are added, the DNA polymerase, attached to the DNA during the library preparation procedure, starts to replicate them. Whenever a nucleotide is incorporated in the newly synthesized strand, the fluorescent light is registered and associated with a specific base to ensure base-calling accuracy [45].

In addition to Pac Bio's Technology, TGS saw the development of another approach based on a totally different sequencing method: nanopore sequencing.

The concept of nanopore sequencing, where single-stranded nucleic acids pass through a nanopore in a membrane under an electric field, was first proposed by David Deamer in the 1980s [46]. Despite initial skepticism, technological advances eventually made nanopore sequencing a reality [47]. Proteins such as α-hemolysin from Staphylococcus aureus [48, 49, 50] and Mycobacterium smegmatis porin A (MspA) [51, 52] were shown to distinguish the four bases on single-stranded nucleotide molecules. The use of phi29 DNA polymerase slowed down the translocation of nucleic acid molecules through the nanopore, improving the signal-to-noise ratio [53, 54]. Oxford Nanopore Technologies (ONT), founded in 2005 by Oxford professor Bayley and colleagues [55], facilitated the commercialization of nanopore sequencing with the release of their MinION sequencer in 2014 [47].

When single-stranded DNA or RNA molecules pass through nanopores, the nucleic acid molecule and the nucleobases, obstruct the pore and limit the electrical ionic current. As different nucleobases have different geometries, they obstruct the pore to a different extent, and thus, the change of electrical ionic current is sequence dependent.

The change in electrical current in the nanopore is recorded by the ASIC and transferred to software, which uses a base-calling algorithm to infer the DNA or RNA sequence. This is performed by applying a machine learning approach, such as a hidden Markov model or recurrent neural network [56]

Ideally, the nanopore should accommodate a single nucleobase, but currently used nanopores accommodate several nucleobases. As a result, the electrical current depends on an oligonucleotide rather than a single nucleobase, which further complicates the base-calling process [57, 58, 59].

Importantly, each nucleobase(s) should reside in the nanopore for a minimum time needed for robust measurement, and this requires the use of dedicated mechanisms to impede and control the nucleic acid translocation through the pore. Currently, for this purpose, motor proteins –helicases are used. They cannot pass through the pore so their movement on single-stranded DNA or RNA molecules limits and determines the speed of translocation [59].

Helicases are preloaded on adapter oligonucleotides called “leader” or “hairpin,” which are ligated to the nucleic acid molecule during the process of library preparation for sequencing (Figure 3) [60].

The sequencing process will begin after the leader motor protein unzips the dsDNA, and the template strand could move through the pore. The hairpin motor protein, which is located at the end of the template strand, guides the translocation of the complement strand through the nanopore. After the translocation of the DNA through the nanopore is completed, the motor proteins are separated, and the whole process can be repeated until the sequencing is completed successfully [56].

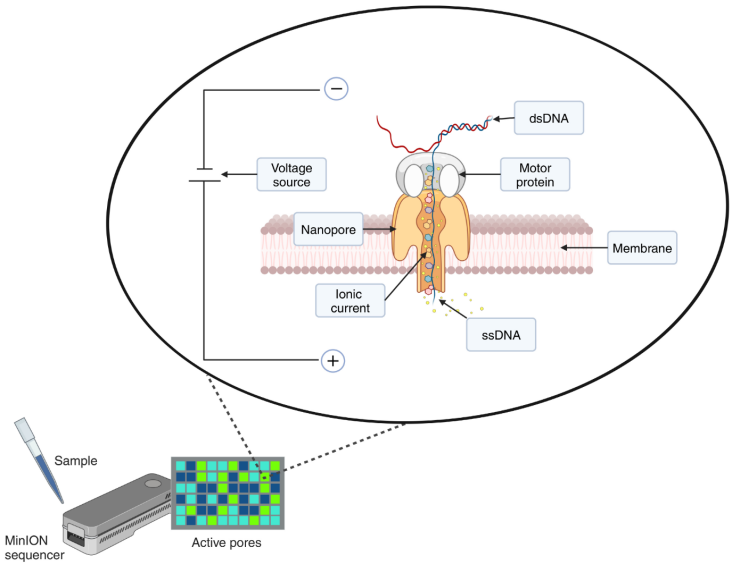


Figure 3. Structure of nanopore and principle of Nanopore sequencing from Dyshlovoy SA et al, Int J Cancer. 2024

A major strength of ONT sequencing technology is the ability to determine the sequence of full-length RNA molecules, either via cDNA sequencing [61, 62] or via native RNA sequencing [63, 64] .

ONT released its first prototype of nanopore sequencer in 2014, the pocket-sized and portable MinION, while the updated platform PromethION was released the following year with an improved throughput. The reaction system for nanopore sequencing is carried out in a flow cell, in which two ionic solution-filled

Compartments are separated by membranes containing either 2048 (MinION, generating up to 50 Gb/sequencing run) or 12,000 (PromethION, generating up to 290 Gb/sequencing run) nanopores. Specifically, two versions of PromethION, namely 24 and 48, integrate 24 and 48 independent flow cells, respectively, to increase productivity by simultaneously using several flow cells. In addition, the PromethION 2 (P2) is a small device allowing for the analysis of up to two PromethION flow cells and has been just released for laboratories with a smaller sample processing number. All these systems are based on the same sequencing technology but with a scalable throughput. While the PromethION system could output up to 7.6 Tb data (with a theoretical maximum of 15 Tb), the MinION could only generate 50 Gb within 72 h. This feature makes the PromethION a serious competitor for Illumina's HiSeq X Ten, which generates a theoretical maximum output of 16–18 Tb/run. In early 2017, ONT also released the GridION X5, which can hold up to five MinION flow cells and can

generate up to 240 Gb of data per run. ONT also announced the release of the SmidgION, which is even smaller than a MinION and can be controlled by a smartphone [65, 66, 44].

Despite its advantages in portability, real-time analysis, and ultra-long reads, Oxford Nanopore Technology (ONT) still presents several limitations. Notably, its relatively high error rate—especially in homopolymeric and high-GC regions—requires extensive error correction, often combining ONT with short-read data. Read truncation, particularly at the 5' end in RNA sequencing, and variable base-calling accuracy further complicate isoform identification. Additionally, ONT performance is highly sensitive to nucleic acid quality, and its throughput is lower compared to short-read platforms, making large-scale projects more costly. Biases in coverage and the lack of standardized workflows limit its clinical application. Continuous improvements in chemistry and software are narrowing these gaps, but ONT still demands sophisticated bioinformatic pipelines and validation for high-resolution or diagnostic use [67, 68].

1.5.1 Data Analysis of Long reads and isoform identification

A characteristic feature of ONT is the huge amount of data generated through massive parallel sequencing technologies.

This substantial data output posed significant analytical challenges, especially considering that ONT's first commercial device, the MinION, was launched without accompanying bioinformatics tools.

Only after introducing this technology, ONT launched the MinION Access Programme (MAP), a beta-testing program for a developer community of more than 1000 laboratories that allowed research groups to evaluate the base throughput, read quality, and performance of the platform [69]. This community has developed several computational approaches and algorithms to manage and analyze ONT data, including base-calling, read mapping, de novo assembly and variant detection and discovery. The operating software of ONT devices is MinKNOW [70], which carries out data acquisition, real-time analysis, and local base-calling to produce raw signal data (FAST5 file) that can be used for base-calling.

From 2021, ONT's teams develop Epi2me Labs [71], an open-analysis platform, with several predefined workflows for data analysis.

Among these, the wf-transcriptomes pipeline (Figure 4) developed by Epi2me Labs provides a robust, scalable, and modular solution for transcriptomic data analysis using Nanopore reads

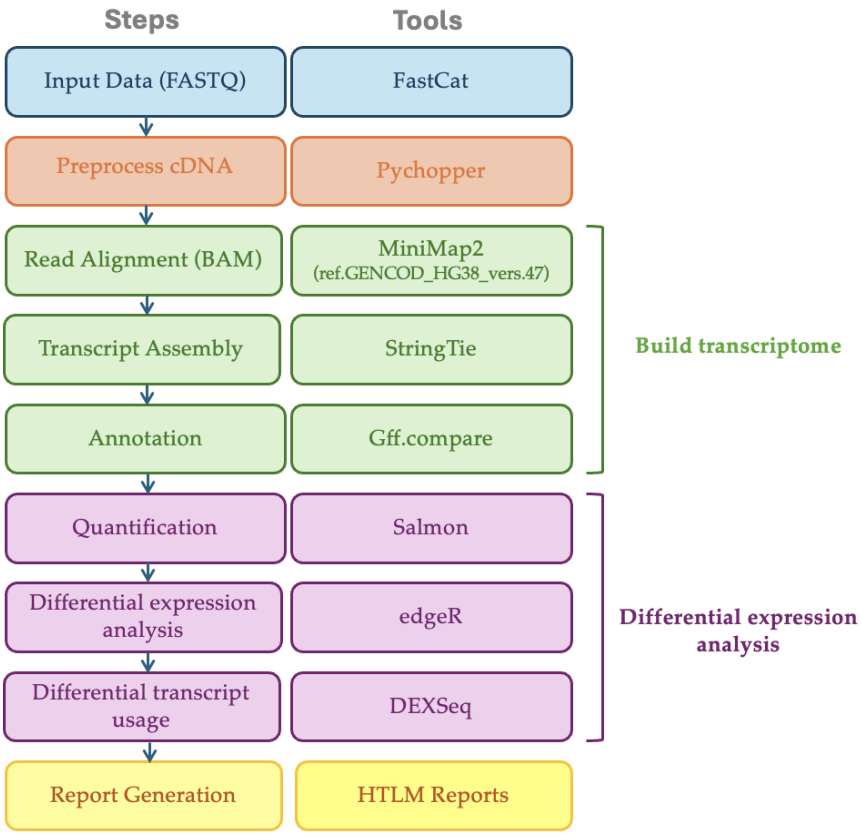


Figure 4. Pipeline data analysis wf-transcriptomes Epi2Me

The wf-transcriptomes workflow is implemented in Nextflow [72], a widely used workflow management system that supports reproducible and scalable analysis across diverse computational environments. This pipeline is optimized for transcriptome-wide studies using either direct RNA or cDNA sequencing reads generated by ONT devices. It supports two main analytical modes: (1) Reference-guided transcriptome assembly, where long reads are aligned to a reference genome using Minimap2 [73], followed by transcript reconstruction using StringTie2 [74] and annotation with GFFcompare [75] ; and (2) Precomputed mode, which leverages an existing transcriptome annotation for downstream quantification. The workflow accepts single-sample or multiplexed datasets in FASTQ or BAM format and includes a customizable sample sheet for batch processing. Containerized environments (Docker [76] or Singularity [77]) ensure portability and reproducibility. Recommended computational resources include a minimum of 8 CPUs and 32 GB RAM, although 16 CPUs and 64 GB RAM are advised for optimal performance. Typical execution times are approximately 15 minutes per million reads, depending on read length and complexity. Downstream analysis includes differential gene expression (DGE) and differential transcript usage (DTU). DGE

is performed using statistical packages such as edgeR [78], while DTU is assessed through a combination of DEXSeq [79], DRIMSeq [80], and stageR [81], enabling a comprehensive evaluation of both gene-level and isoform-level changes across conditions.

In addition to outputs suitable for integration with further bioinformatics workflows, the pipeline generates interactive HTML reports (Figure 5) summarizing quality control metrics, alignment statistics, transcript annotations, and results from DGE and DTU analyses. Genome browser-compatible files (e.g., bigWig and GTF) are also produced, facilitating visualization of transcript structures and expression patterns in tools such as IGV.

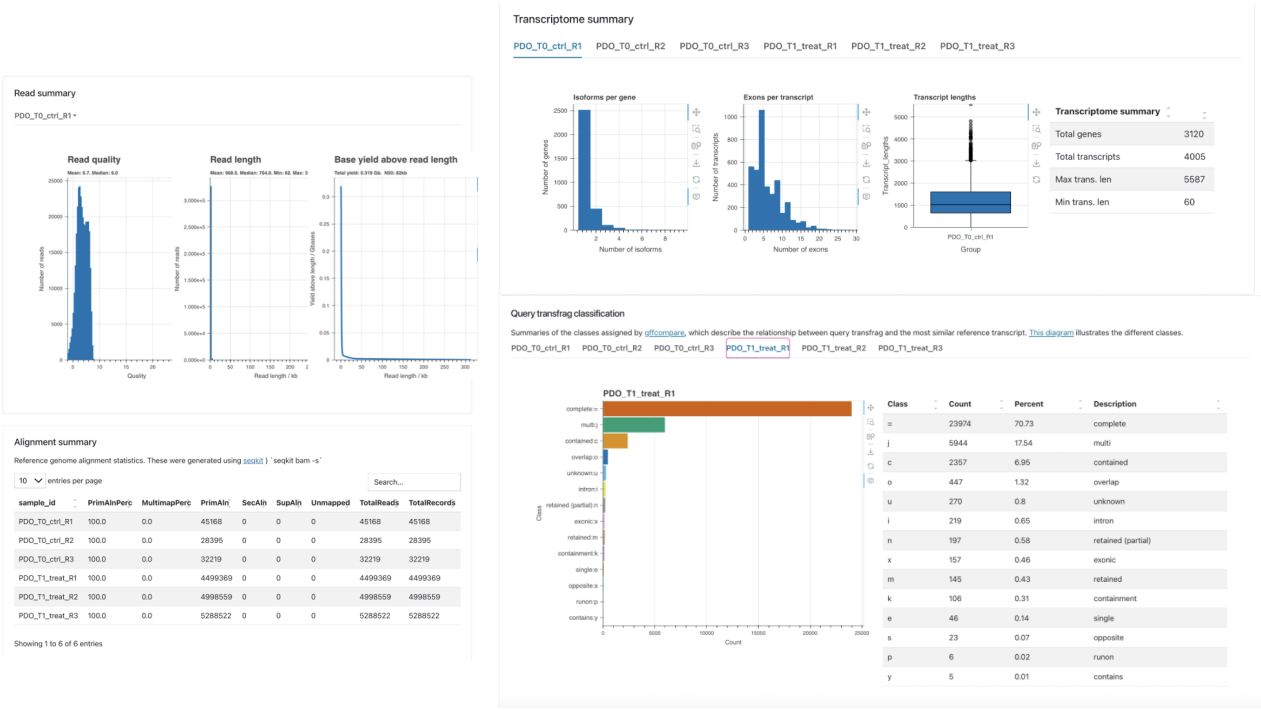


Figure 5. Example of results generated through the wf-transcriptomes workflow provided by Oxford Nanopore Technologies.

Overall, the wf-transcriptomes workflow provides a complete, automated solution for the analysis of ONT transcriptomic data, supporting key applications such as isoform discovery, splicing analysis, and expression quantification. Its flexibility and integration with established statistical tools make it particularly suitable for studying complex transcriptional landscapes, including those found in cancer and other heterogeneous biological systems.

It is important to emphasize that this platform and workflow enable a detailed characterization of isoform splicing, a process that has gained increasing relevance in recent

years for its pivotal role in cancer biology. Understanding alternative splicing patterns can uncover novel biomarkers and therapeutic targets, offering promising strategies for cancer diagnosis and treatment

1.6 Alternative Splicing and colorectal cancer

Alternative splicing (AS) is a tightly regulated process that generates multiple mRNA and protein isoforms from a single gene [82], affecting up to 95% of human genes with multiple exons [83]. This mechanism leads to the production of various splicing isoforms, including canonical and alternative variants. Canonical isoforms, as defined by UniProtKB [84], are typically the most prevalent, conserved, highly expressed and longest, while alternative isoforms encompass all other variants that differ from canonical form.

Pre-mRNA splicing was first discovered in 1977 and has since been recognized as a crucial role in post-transcriptional regulation of gene expression [85].

The proceed is carried out by spliceosome, a large multi-protein machinery consisting of five small nuclear ribonucleoprotein particles (snRNP U1, U2, U3, U5 and U4/U6) and more than 150 accessory proteins which bind the pre-mRNA in a stepwise fashion and form highly dynamic complexes [86]. Introns are excised via two transesterification reactions, resulting in joining the exons together and the release of the introns in a form of a lariat [87, 88].

Five types of Alternative Splicing have been defined, based on the exon configurations in mature mRNAs (Figure 6) [89]:

(1) Exon skipping: exons are 'skipped over' and not included in the mRNA. A process that affects numerous transcripts and that can occur via multiple mechanisms, for example via mutation or transcriptional alterations in the proteins that regulate splicing or via mutations in splice donor or acceptor sites on the pre-mRNA [90].

(2) Mutually exclusive exons: the generation of alternative transcripts in a process in which only a single exon among multiple neighbouring ones is retained in the mature transcript [91].

(3) Alternative 5′ splice site: an alternate donor splice site that changes the 3′ position of the upstream exon.

(4) Alternative 3′ splice site: an alternate acceptor splice site changing the 5′ position of the downstream exon.

(5) Intron retention: the non-coding intronic regions are retained in the mature mRNA.

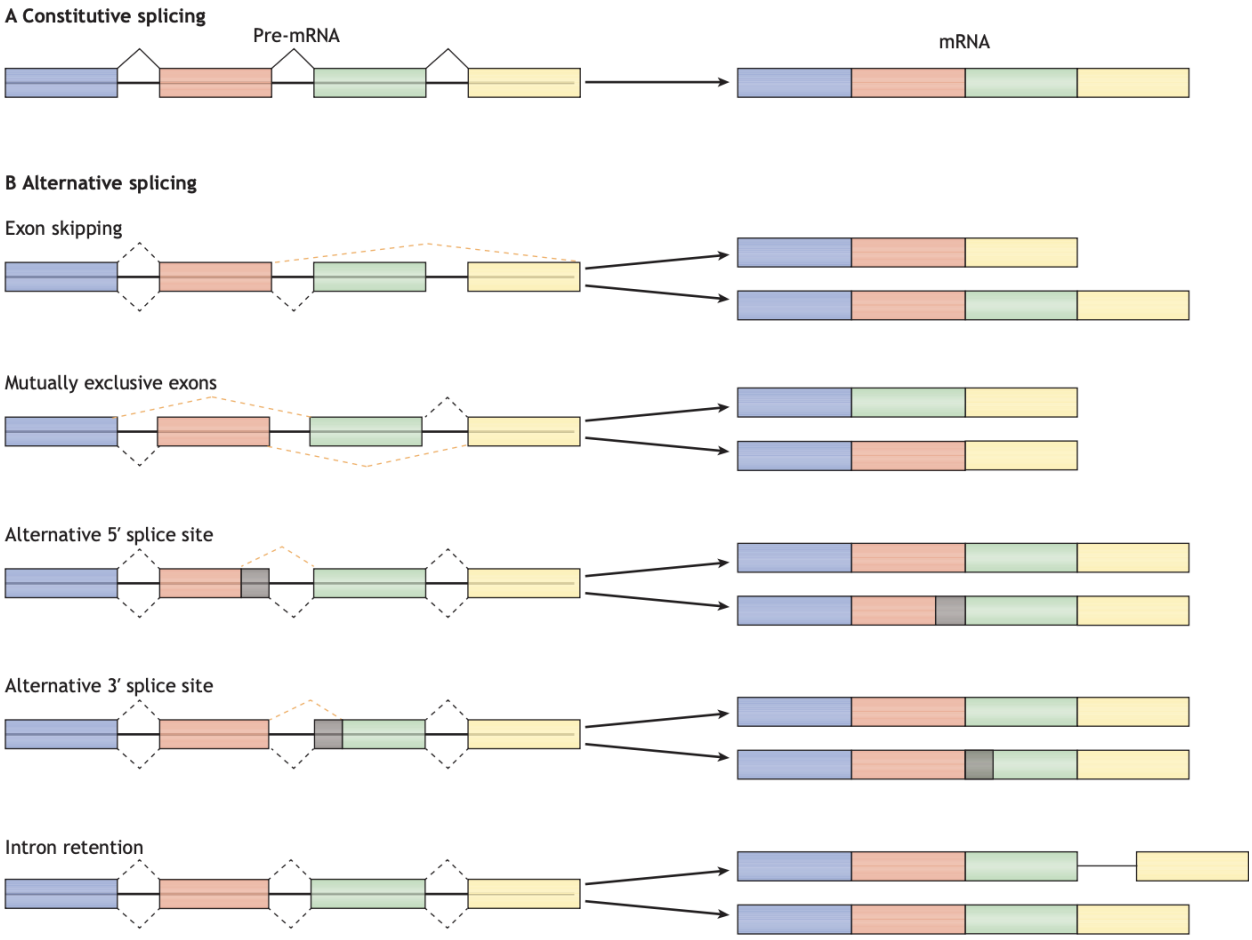


Figure 6. Mechanisms of Alternative Splicing from Öther-Gee Pohl S et al (2022)

These splicing isoforms often encode different protein structures, which can result in different functions, playing crucial roles in a wide range of biological processes and disease, in particular in cancer [92, 93].

Aberrant splicing in tumours can arise from several causes, including altered expression of key splicing regulatory proteins or RNAs, which themselves can function as proto-oncoproteins or tumour suppressors; cis-acting somatic mutations that alter splicing of the genes bearing those lesions; and trans-acting somatic mutations that cause gain-of-function or loss-of-function alterations affecting splicing regulators, driving pervasive splicing changes across the transcriptome [94, 95].

Changes in alternative splicing have been linked to cancer through post-transcriptional regulation of components of many of the cellular processes considered as “hallmarks” of cancer, including cell proliferation, apoptosis, metabolism, invasion, and angiogenesis.

In solid tumors, particularly colorectal cancer (CRC), several advances suggest how specific spliced isoforms contribute to tumorigenesis have paved the way for novel therapeutic strategies targeting these events [96].

Capon et al. [97] were the first to discover that in CRC cell lines, c-Ki-ras (KRAS) mutates at different points within the same codon, resulting in the production of two transcript variants. So, investigating the influence of splicing variants in CRC is of paramount importance for the diagnosis and treatment of CRC.
The occurrence of cancer involves a complex process that has to do with the interaction of multiple genes and molecular pathways. Anomalies in alternative splicing have been identified as a significant contributor to the development of CRC, and studying this phenomenon has the potential to shed light on the mechanisms of tumor formation.

Several studies focus on the major splicing patterns—including exon skipping (ES), intron retention (IR), mutually exclusive exons (MXE), and alternative 5' and 3' splice sites (A5SS, A3SS)—that alter the transcriptome and proteome of cancer cells. These aberrant splicing events frequently arise from mutations or dysregulation of core spliceosome components and splicing regulators such as SRSF1 [98] , hnRNPA1 [99], or SF3B1 [100] , resulting in tumor-specific isoforms with functional consequences.

There are also several key oncogenic splice variants in CRC, including: *KRAS4A* and *KRAS4B*, derived from alternative exon 4 usage [101] ; *APC* isoforms generated by exon skipping (e.g., exon 14) and intron retention (e.g., intron 11) [102]; *EIF4H* isoform 1 (via inclusion of exon 5), which enhances cyclin D1 expression [103]; and *BARD1* isoforms (e.g., BARD1k, BARD1b), which exhibit oncogenic properties [104] . Conversely, some isoforms exert tumor-suppressive effects, such *VEGF165b*, an anti-angiogenic splice variant formed via distal 3' splice site selection in exon 8 [105].

In last decades it is also emerging strategies targeting specific splicing isoforms, including antisense oligonucleotides (ASOs) to modulate splice site selection, and small molecules that interfere with splicing factor activity or spliceosome assembly (e.g., Spliceostatin A, H3B-8800). Tumor-specific AS isoforms such as *c-FLIPL* [106], *CCND1b* [107] and *RONΔ165* [108] are proposed as promising diagnostic markers and potential therapeutic targets. Moreover, CRC cells show an increased PKM2/PKM1 ratio driven by hnRNPA1 and PTBP1, which promotes aerobic glycolysis and tumor growth [109, 110].

As one of the primary functions of Alternative Splicing is to generate proteomic and transcriptomic diversity, its link to both cellular and molecular tumor heterogeneity and the therapeutic consequences of this starts to emerge. These findings underscore the importance of AS and how it could predict patient survival and potential answers to therapies.

1.7 Single-cell RNA Sequencing

Bulk RNA sequencing detects all RNA transcripts in each tissue or group of cells, capturing average gene expression signals while ignoring differences between individual cells. In contrast, single-cell RNA sequencing addresses this limitation by providing transcriptome profiling at single-cell resolution [111].

A typical scRNA-seq protocol includes several steps: sample acquisition, single-cell isolation, lysis, reverse transcription (RT), complementary DNA (cDNA) amplification, library construction, sequencing, and data analysis [112].

Currently, several methods are utilized to isolate single-cell, including manual cell selection [113], limiting dilution [114], laser-capture microdissection (LCM) [115], fluorescence-activated cell sorting (FACS) [116], magnetic activated cell sorting (MACS)[117], and microfluidics[118].

Currently, the most popular high-throughput platform is based on droplet-based microfluidics (microdroplets). In 2017, the commercial sequencing platform 10x Genomics was successfully developed based on the above techniques, enabling a significant increase in cell throughput and a considerable reduction in single-cell sequencing costs [119, 120].

The basic chemistry that drives 10x Genomics Chromium X Series is related to the foundational Drop-seq method published by Harvard researchers in 2015 [121].

10x Genomics Chromium X Series does so by partitioning thousands of cells into nanoliter-scale Gel Bead-In-Emulsions (GEMs).

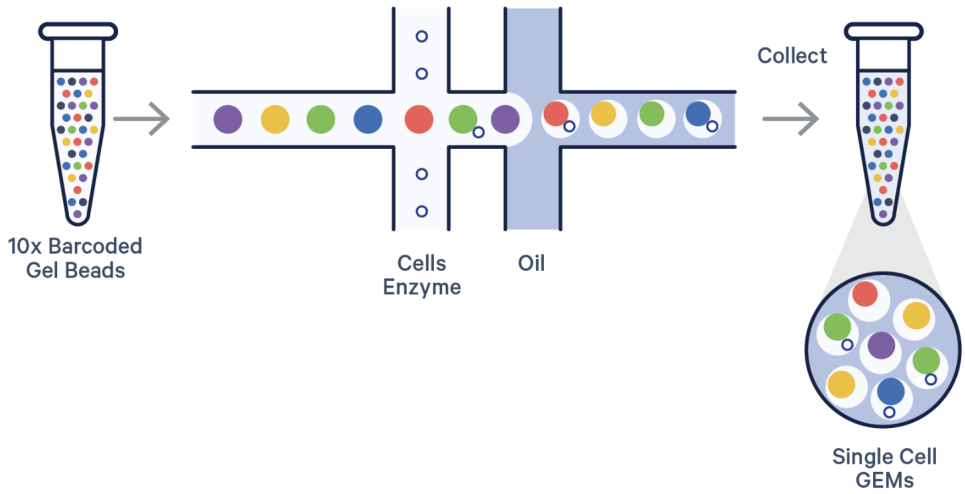


Figure 7. A schematic overview of GEMs generation and barcoding with the GEM-X chip workflow.

The mixture is composed of a cell suspension with lysis buffer, Gel Beads coated with oligonucleotides—which include a 10x Barcode that marks each RNA molecule's cell of origin and a unique molecular identifier (UMI) that gives each transcript a unique fingerprint—a poly(dT) sequence for capturing mRNA at the 3' or 5' end, and adapter sequences for downstream next-generation sequencing. Once the chip is placed in our [104], cells move through a microfluidic labyrinth at a limiting dilution to generate GEMs, which appear as nanometer-sized rain drops, each containing a single cell. The cell is lysed and the Gel Bead dissolves. The 3' end of the released RNA binds to the poly-dT sequences that coated the Gel Bead. The primed RNA is reverse transcribed, forming complementary DNA (cDNA).

All cDNAs from a single cell, originating from a single Gel Bead, will have the same 10x Barcode, enabling the user to map each transcript back to its cell of origin. The individual oligos on a single Gel Bead, however, have different UMIs, which distinctly mark each cDNA molecule with a unique, random 12-base sequence. This allows the user to eliminate any PCR duplicates during downstream processing. The barcoded cDNA is then used to prepare a sequencing library, which includes the addition of sequencing adapters and PCR amplification. Finally, the library is sequenced with NGS [120]. During reverse transcription, the cell barcodes are labeled, then the tiny reaction chambers are broken, and the pooled fractions are recovered.

As a more mature commercialized system, 10x Genomics generally requires less time, has higher molecular sensitivity and precision, and is accompanied by less technical noise. By rule of thumb, 10x Genomics is currently a safe choice for most applications and has been used in cancer research most widely. However, these techniques can only identify the 3′ or 5′ end sequence of transcripts and have a limited depth of sequencing [122].

In human cancer research, scRNA-seq has been widely used to study heterogeneity, the TME, gene expression profiles, transcriptome profiles, and cell–cell interactions, and other biology related to cancer research.

Because our research is focusing on ITH, it is important to underline the application of scRNA-seq to dissecting Intratumor heterogeneity in cancer research. Almost all single-cell studies of cancer have focused on tumor heterogeneity. Dissecting tumor heterogeneity with scRNA-seq has been used to facilitate cancer diagnosis [123] and prognosis prediction [124],

increase the understanding of disease progression and cancer metastasis [125, 126,] and guide therapy [127].

Solid tumors account for the vast majority of cancer incidence and death. scRNA-seq has been used in the study and application of various solid tumors.

In the context of CRC, scRNA-seq has proven instrumental in dissecting intratumoral heterogeneity and uncovering transcriptionally distinct tumor cell subpopulations. Studies have revealed the coexistence of epithelial subclones with varying degrees of stemness, proliferation, and differentiation potential [128]. Such heterogeneity has been linked to variable responses to chemoradiotherapy and to divergent evolutionary trajectories within the same tumor.

For example, Li et al. [129] identified functionally distinct CRC subpopulations expressing differential levels of LGR5, a well-established intestinal stem cell marker. These subpopulations were found to exhibit differential proliferative capacity and drug sensitivity, suggesting a hierarchical organization within the tumor epithelium. Moreover, the use of scRNA-seq in patient-derived organoids has shown that these subclones can persist and dynamically re-establish heterogeneity in vitro, reinforcing their potential role in tumor relapse.

scRNA-seq has also been employed to explore mechanisms of resistance to therapy. In preclinical models, CRC cells exhibiting a quiescent or dormant transcriptional profile prior to treatment were shown to evade apoptosis and drive tumor regrowth. This emphasizes the importance of capturing functional heterogeneity at baseline, which would be masked in population-level bulk transcriptomics [130].

Despite its advantages, scRNA-seq presents several limitations. First, although scRNA-seq has been the most mature and widely used single-cell omics in cancer research, it can't completely reflect cancer biology, such as the level of genomics, proteomics, epigenomics, etc. Second, scRNA-seq has high requirements for specimens, and can only use live cells as detection objects; specimens analyzed at different times will have associated batch effects. Third, enzymatically digesting cells to obtain single-cell suspensions is likely to kill cells and degrade RNA. This will also result in the loss of spatial and morphologic information. Though single-cell data enable single-cell transcriptomics, it may lose spatial information during single-cell isolation. To solve this problem, spatial transcriptomics has emerged. Spatial transcriptomics employs unique positional barcodes to visualize RNA distributions in RNA sequencing of tissue sections. Fourth, not all nucleated cells are detected, the

sequencing depth may be limited, and scRNA-seq data do not fully reflect the entire transcriptional landscape and genetic information of cancer samples. The heterogeneity of patient tumors and differences in detection platforms also increase the difficulty of result interpretation and decrease data uniformity. In addition, scRNA-seq is expensive and time-consuming, and data processing is complex.

Nevertheless, as the technology matures, scRNA-seq continues to advance our understanding of cancer biology. Its integration with clinical research holds the potential to support precision oncology through the identification of predictive biomarkers and actionable targets.

1.7.1 Data Analysis: Epi2me and wf-single cell

Single-cell RNA sequencing (scRNA-seq) has emerged as a powerful tool for characterizing cellular heterogeneity, identifying rare but biologically relevant cell types, and exploring cell–cell communication and dynamic interactions within tissues. To effectively analyze transcriptomic data at single-cell resolution, numerous computational methods and pipelines have been developed to address challenges such as technical noise, sparse gene detection, and high biological variability.

Oxford Nanopore Technologies (ONT) has developed Epi2me Labs (https://github.com/epi2me-labs), an open-access analysis platform that includes several pre-configured workflows.

Among these, the wf-single-cell pipeline is specifically designed for the analysis of single-cell libraries generated using ONT sequencing, particularly those prepared with 10x Genomics protocols. The pipeline was originally implemented as a Nextflow-based adaptation of the Sockeye framework [131, 132] , enabling scalable and reproducible execution across different computing environments.

The `wf-single-cell` pipeline (Figure 8) encompasses all major steps of scRNA-seq data analysis, including preprocessing, quality control, normalization, dimensionality reduction and clustering. It accepts multiplexed single-cell cDNA libraries in FASTQ format and supports upstream barcoding strategies used in ONT's single-cell protocols.

The initial stage of the pipeline involves barcode demultiplexing and UMI (Unique Molecular Identifier) extraction, which enable accurate assignment of reads to individual cells. This process includes the identification and optional error correction of cell barcodes,

filtering of low-quality reads, and parsing of UMIs. The filtered reads are then aligned to the reference genome using **minimap2** [133], a spliced aligner optimized for long-read RNA sequencing, which ensures high-accuracy mapping of full-length transcripts.

Following alignment, the pipeline performs barcode correction. At this stage, each read is assigned to a gene or transcript. First, a transcriptome is assembled from the aligned reads using **StringTie2** [74], followed by transcript annotation with **Gff.compare** [75]**,** which integrates metadata from the supplied reference annotation. Transcript assignment is conducted using a strategy inspired by FLAMES, wherein reads are mapped to annotated transcript models. UMI correction is then carried out using **UMI-tools**, which clusters reads by gene and collapses PCR duplicates.

The pipeline further includes an isoform detection module that reconstructs full-length transcript isoforms on a per-cell basis by clustering aligned reads. These isoforms are annotated and quantified, generating both gene-level and isoform-level count matrices. Predicted transcript structures are also output in GTF format, allowing for the analysis of alternative splicing events and transcript diversity across single cells.

Throughout the workflow, detailed quality control metrics are generated and compiled into interactive HTML reports. These reports provide insights into read quality, barcode distributions, alignment statistics, and gene coverage, ensuring transparency and reproducibility. Final outputs are compatible with widely used downstream analysis tools such as Seurat [134], Scanpy [135] , and genome visualization platforms like IGV [136].

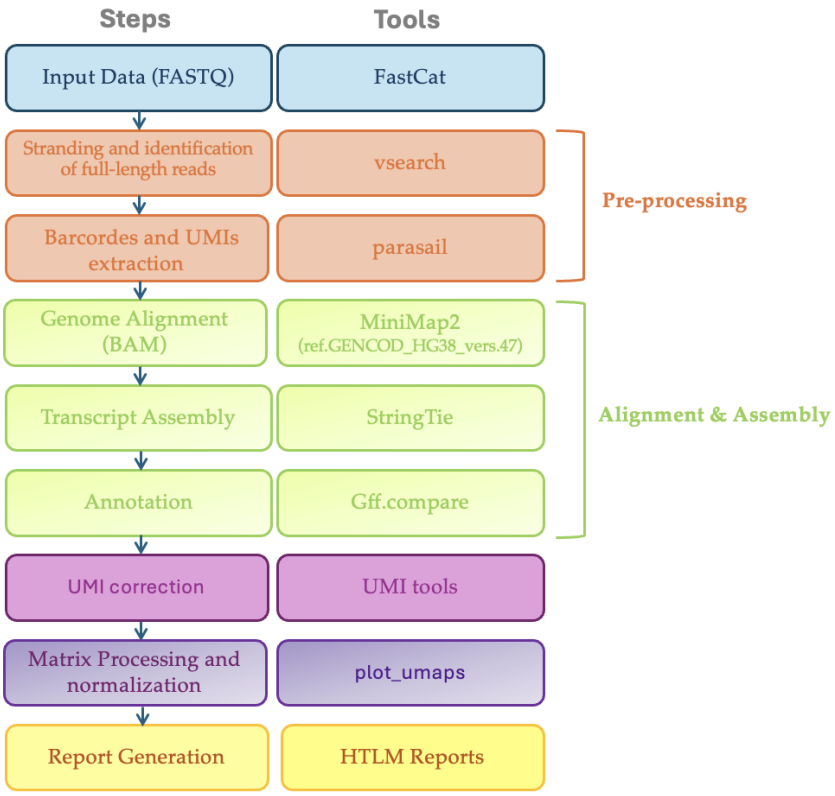


Figure 8. Pipeline data analysis wf-single-cell Epi2Me

1.8 Adaptive trajectories in CRC cancer: insights from longitudinal single-cell transcriptomics

Cancer progression and therapeutic resistance are governed by the dynamic reorganization of tumor cell populations in response to internal and external cues. In colorectal cancer (CRC), as in other solid tumors, intra-tumoral heterogeneity and cellular plasticity enable subsets of cells to survive, adapt, and reprogram their transcriptional states under treatment pressure [137]. These dynamic transitions are best conceptualized as adaptive trajectories, in which tumor cells follow non-linear evolutionary paths through phenotypic space—paths that are often undetectable in bulk analyses but can be reconstructed with longitudinal single-cell RNA sequencing (scRNA-seq).

Longitudinal scRNA-seq allows the temporal resolution of transcriptional changes at the single-cell level across multiple stages of treatment. By integrating these datasets using algorithms such as pseudotime inference, RNA velocity, and cell fate mapping, it becomes possible to model the directionality and branching of cell state transitions over time [138, 139]. These methods go beyond snapshot clustering to predict future transcriptional states by

estimating the rate and direction of gene expression change from unspliced and spliced mRNA ratios. In doing so, they provide critical insight into which cellular states give rise to others under therapeutic pressure—whether towards apoptosis, quiescence, resistance, or trans differentiation.

The RNA velocity, particularly in its dynamical formulation, has proven instrumental in identifying such transitions, though it remains sensitive to assumptions about kinetic parameters and data resolution [139]. Notably, current implementations may fail to capture complex or nonlinear behaviors across heterogeneous subpopulations. Despite these limitations, RNA velocity has revealed that drug treatment does not merely select pre-existing resistant clones but can actively induce new cell states—many of which exhibit distinct transcriptomic programs and altered vulnerabilities.

Recent pan-cancer single-cell analyses reinforce this view, showing that expression heterogeneity is not random, but frequently structured into recurrent transcriptional programs across tumors and cell lines [140]. These programs, including stress response, epithelial–mesenchymal transition (EMT), interferon signaling, and senescence, often emerge as adaptive endpoints of dynamic trajectories. In organoid models of CRC, these programs can be observed to intensify, rewire, or resolve during longitudinal exposure to drugs, revealing branching paths of adaptation. Such transitions often culminate in the expansion of drug-tolerant persister populations, characterized by metabolic rewiring, transcriptional infidelity, or epigenetic reprogramming.

Importantly, mapping adaptive trajectories through time enables the identification of early biomarkers of resistance and transient vulnerabilities that would otherwise be obscured. For example, trajectory-based modeling can inform therapeutic interventions aimed at intercepting cell states before they become fixed or combining drugs to target multiple fates simultaneously[137]. These insights are foundational for the development of adaptive therapy strategies, which seek not to eliminate all tumor cells, but to steer their evolution toward less aggressive or more treatable phenotypes.

In summary, adaptive trajectories represent the evolving architecture of tumor plasticity under treatment, and longitudinal single-cell approaches—powered by dynamic models such as RNA velocity—are essential to their characterization. As tools mature and integrate multimodal data (e.g., epigenomics, proteomics), the ability to predict and redirect cancer evolution in real time becomes increasingly feasible.

Chapter 2. AIM OF THE PROJECT

This project aims to advance the understanding of the mechanisms underlying phenotypic adaptation in response to pharmacological perturbations, utilizing high-fidelity models of colorectal cancer (CRC). Specifically, it seeks to elucidate the adaptive trajectories available to tumor cells, how these trajectories contribute to therapy resistance, and which molecular dependencies render them pharmacologically vulnerable. Gaining such insights represents a significant step forward in addressing the urgent need to expand and enhance the therapeutic options available in oncological medicine.

The experimental model, based on patient-derived rectal cancer organoids, has been designed to investigate the phenotypic trajectories that tumor cells undergo during Total Neoadjuvant Therapy (TNT). By mimicking the therapeutic regimen administered to the donor patient, the model enables a detailed analysis of tumor cell evolution under treatment pressure. This study aims to clarify how therapeutic bottlenecks influence clonal selection and tumor progression, with the hypothesis that clonal evolution plays a pivotal role in shaping resistance mechanisms.

A central focus of the project is intra-tumor heterogeneity (ITH), which refers to the diversity among cancer cells within the same tumor. ITH has profound implications for treatment efficacy, as it enables evolutionary mechanisms that can lead to secondary resistance and treatment failure. While traditionally investigated from a genomic standpoint, emerging evidence suggests that phenotypic heterogeneity—observable at the level of gene expression, morphology, and behavior—can significantly influence therapeutic response, even among cells with identical genomic backgrounds. This underlines the necessity of investigating phenotypic ITH to fully comprehend resistance dynamics.

Moreover, therapy-induced selective pressures can enrich resistant cellular subpopulations, which may ultimately drive recurrence. In this context, a pilot study has been developed to trace and characterize the phenotypic adaptation of rectal cancer organoids throughout the course of neoadjuvant therapy. By coupling longitudinal sampling with transcriptomic profiling, this approach aims to identify critical windows of phenotypic plasticity and potential vulnerabilities exploitable for therapeutic intervention.

Chapter 3. STUDY DESIGN

3.1 Graphical Abstract

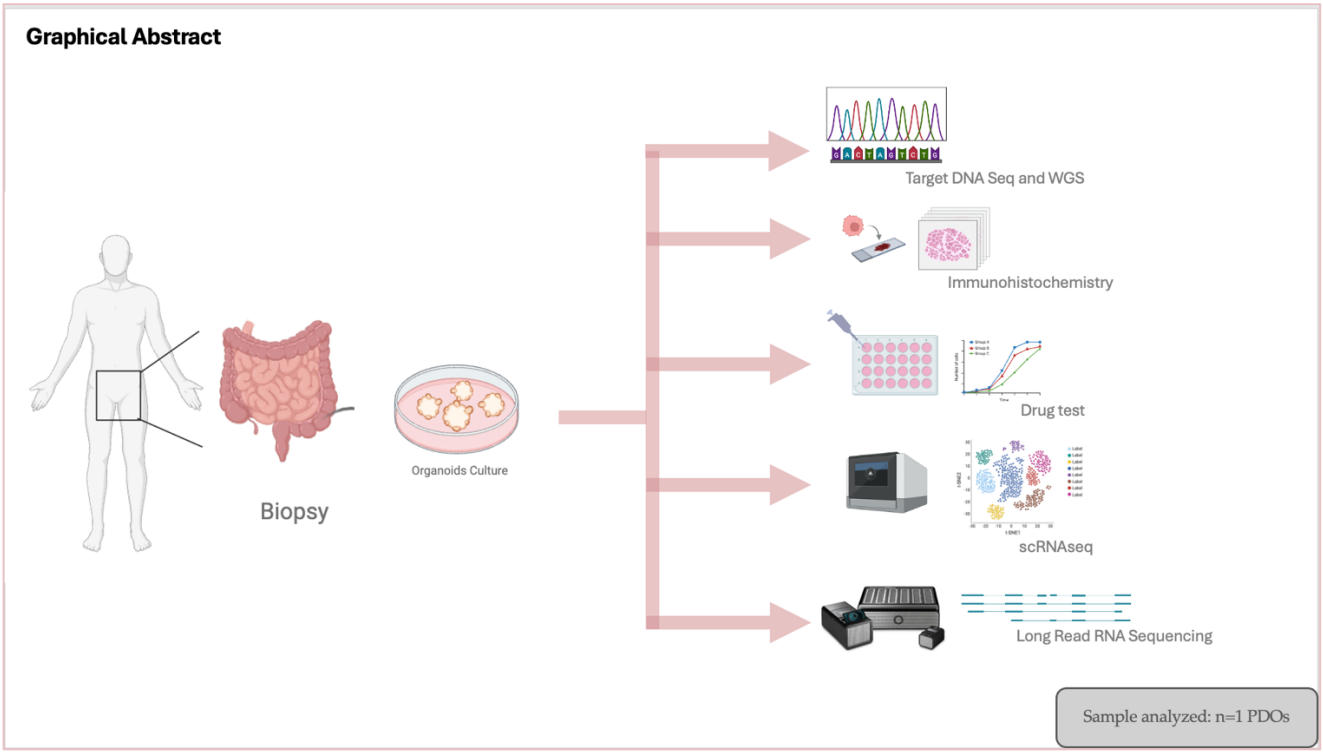


Figure 9. Graphical Abstract. Patient-derived rectal cancer organoids enable multi-level analyses, including targeted DNA/WGS, immunohistochemistry, drug testing, scRNA-seq, and long-read RNA sequencing.

3.2 Experimental workflow

The experimental design involves the use of a pilot line, PDOs CA-04, of patient derived organoids (PDOs) generated from a treatment - naïve LARC patient.

Organoid line will undergo comprehensive multi-omics profiling, including targeted DNA sequencing, whole genome sequencing (WGS), histological (H&E), immunohistochemical and immunofluorescence staining, as well as both bulk and single-cell RNA sequencing, in order to define their baseline molecular and phenotypic landscape.

Following initial characterization, PDO line will be subjected to a pharmacological bottleneck, they will be exposed to a neoadjuvant chemoradiotherapy regimen (radiation + FOLFOX) for 24 days. In parallel, a matched control (untreated) PDOs culture will be maintained to serve as a negative control.

Throughout the 24-day treatment period, cell viability will be assessed at multiple time points using fluorescence-based assays. These measurements will provide insight into the sensitivity of each PDO line to treatment.

At the end of the treatment period (Day 24), biological material from each PDO line will be collected and processed for bulk and single-cell RNA sequencing. These data will be integrated with previously acquired information to enable comparative analyses of pre- and post-treatment phenotypes. Single-cell RNA sequencing will be instrumental in identifying qualitative and quantitative drug-induced changes affecting intratumoral phenotypic architecture.

Following treatment, surviving PDOs will be cultured drug-free for an additional 12 weeks to model treatment relapse dynamics. At the end of this period, phenotypic and transcriptomic analyses will be repeated to assess long-term adaptation. This experimental design will ultimately allow for a longitudinal and mechanistic dissection of phenotypic adaptation trajectories in response to pharmacological pressure.

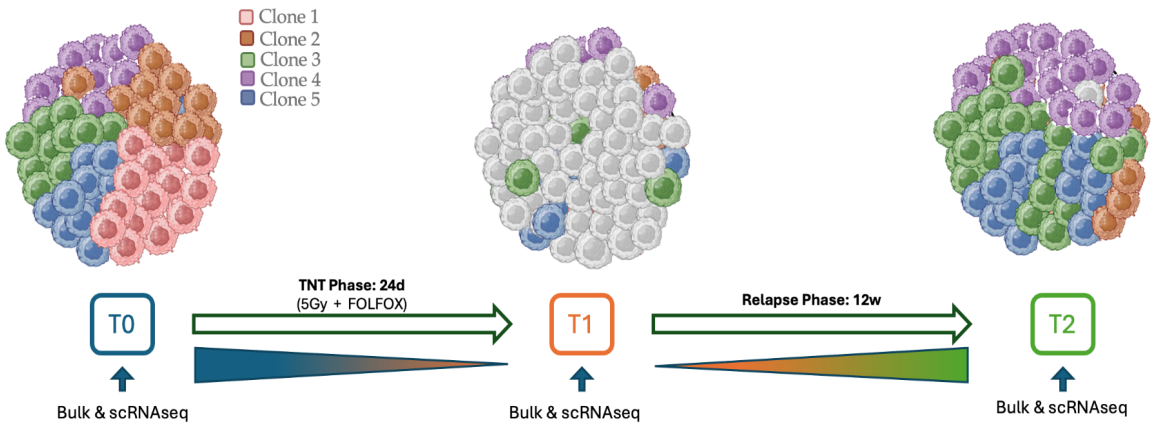


Figure 10. Experimental design. Experimental design of patient-derived rectal cancer organoids. Samples were collected at baseline (T0), following total neoadjuvant therapy (T1;

24 days, 5 Gy + FOLFOX), and at relapse (T2; 12 weeks). Bulk and single-cell RNA sequencing were performed at each stage to track clonal evolution.

Chapter 4. Materials and Methods

4.1 Patient Enrolment and Primary Human Tumor Collection

Biopsy samples from primary rectal cancer tumor were obtained from patients with LARC enrolled in the non-randomized clinical study, STAR (Stem Cells and NeoAdjuvant Treatment in K-Rectum) (authorization document protocol 116/CE of February 22, 2022) conducted at the Department of Oncology, Mediterranean Oncology Institute, Viagrande (CT), Italy.

The enrolled cohort consisted of patients diagnosed with operable rectal adenocarcinoma (stage II and III), eligible for neoadjuvant therapy.

Biopsies were collected prior to the administration of neoadjuvant radiotherapy.

Twelve weeks after treatment, therapeutic responses were assessed by CT scan and magnetic resonance imaging, and patients were classified accordingly.

Table II summarizes the clinical characteristics of the involved patients. No clear correlation emerged between clinical features and the ability to establish viable organoid cultures. The overall organoid derivation rate (63%) was lower than previously reported in the literature [141], potentially due to biopsy heterogeneity, including low tumor cellularity, presence of necrosis, stromal contamination, or bacterial contamination.

PDO	Age	Sex	Diagnosis	T	N	M	Stage	TRG	Mutation Status			
									KRAS	NRAS	BRAF	MSI/MSS
CA-01	77	M	ADC	3	0	X	II A	1	-	-	-	90%
CA-02	71	M	ADC	3	0		IIA	2	-	-	-	90%
CA-03	80	F	ADC	3	0		IIA		-	-	-	-
CA-04	74	F	ADC	3	2b	X	IIIC	2	-	-	-	90%
CA-05	73	M	ADC	3	0	X	IIA		-	-	-	90%
CA-06	73	M	ADC	3	1c	X	IIIB	3	-	-	-	90%
CA-07	70	M	ADC	4	0		IIB		-	-	-	-
CA-08	68	M	ADC	3	0	X	IIA		-	-	-	-
CA-09	62	F	ADC	4b	2		IIIC		Q61X	-	-	-

Table II. Clinical characteristics of the patients enrolled. Patients for whom PDOs lines were not obtained are reported with grey background, pilot patient is reported with green background.

The studies were conducted in accordance with recognized ethical guidelines (Declaration of Helsinki). Informed consent was obtained from all of the participants.

4.2 Biopsy tissue processing

Primary rectal cancer biopsy (6 mm) was performed by experienced clinical surgeons. Biopsies were obtained in compliance with the standards of the ethics committee on human experimentation of the IOM. Biopsy samples were collected before the administration of neo-adjuvant radiotherapy.

Tissue samples were quickly washed two to three times in cold saline buffer (PBS), then transferred in Dulbecco's modified Eagle's medium (Gibco, Thermo Fisher Scientific, Carlsbad, CA, USA, 12491023) containing 5% penicillin-streptomycin-amphotericin B solution (Sial, Rome, Italy, SIAL-AAS-B) and kept in this medium at 4°C until processing within 24 h.

4.3 PDOs isolation and culture

Within 24 hours of collection, tissues were first extensively washed in PBS (Gibco, Thermo Fisher Scientific, Carlsbad, CA, USA, D8537) and then subjected to mechanical (2-3 mm in size) and enzymatic digestion with Collagenase type II (Gibco, Thermo Fisher Scientific, Carlsbad, CA, USA, 17101015) at 37°C for 1 h. Cell suspension was then filtered through a 100-µm nylon to remove large undigested tissue fragments.

The cell suspension was centrifuged at 1200 rpm for 5 minutes, washed with PBS, and centrifuged again under the same conditions. The resulting cell pellet was kept on ice and resuspended in 200–400 µL of cold Basement Membrane Extract (BME) (R&D SYSTEMS, Biotechne, Minneapolis, USA, BME001-05).

Aliquots (200 µL) of the cell-BME mixture were seeded into the center of pre-chilled 6-well plates (Corning Costar, Cambridge, MA, USA, Minneapolis, USA, 3516), allowed to polymerize at 37°C and 5% CO_2 for approximately 30 minutes, and overlaid with 2 mL of Tumorsphere Medium XF (Tumorsphere Medium XF, PromoCell, Heidelberg, Germany, C28070), pipetted along the well wall to avoid disrupting the dome.

Cultures were maintained at 37°C in a humidified atmosphere of 5% CO^2. Every 2–3 days, half of the culture medium was replaced with fresh medium.

PDOs became evident after a variable length of time, ranging from 5 to 7 days to 3 weeks. Regular culture splitting (1:2) was usually needed after 3–6 weeks from isolation.

4.4 Histological and Molecular Validation of Organoid Models

4.4.1 HE and Immunohistochemistry Staining

Histological sections were prepared from both patient-derived tumor tissues and corresponding organoid cultures. Tumor samples were obtained from formalin-fixed, paraffin-embedded (FFPE) diagnostic biopsies. Tissue blocks were sectioned at 4 μm thickness and mounted on TOMO adhesive glass slides (Matsunami Glass, Osaka, Japan, 10748-174) for hematoxylin and eosin (H&E) staining and immunohistochemical (IHC) analysis.

Organoids were processed for paraffin embedding following a protocol optimized for three-dimensional cultures [142]. Briefly, organoids were collected by centrifugation at 400×g for 5 minutes, washed in ice-cold phosphate-buffered saline (PBS), and fixed in 10% neutral buffered formalin for 30 minutes at room temperature. The samples were dehydrated in a graded ethanol series (70%, 80%, 90%, 100%; 30 minutes each), cleared in xylene (2 × 10 minutes), and infiltrated with molten paraffin at 60 °C for 2 hours. Paraffin-embedded organoids were transferred into embedding molds, allowed to solidify, sectioned at 5 μm, and mounted on adhesive glass slides for H&E and IHC staining.

Immunohistochemical analyses were conducted to evaluate the expression of CDX2, β-catenin, and CK20 in both tumor tissues and corresponding organoids. Four-micron-thick sections were used for the assessment of CDX2, β-catenin and CK20 expression using the fully automated BenchMark Ultra system (Ventana Medical Systems, Tucson, AZ, USA). Slides underwent deparaffinization at 72 °C, followed by heat-induced epitope retrieval with CC1 buffer at 95 °C for 36 minutes. The CDX2 rabbit monoclonal antibody (EPR2764Y, Roche Diagnostics) was incubated for 36 minutes, the β-catenin mouse monoclonal antibody (clone 14, Roche Diagnostics) was incubated for 24 minutes, while the CK20 rabbit monoclonal antibody (SP33, Roche Diagnostics) was incubated for 32 minutes. Signal detection was performed using Ventana UltraView Universal DAB Detection Kit (Ventana Medical Systems). Images were taken with a Ventana DP200 digital scanner. H&E staining was performed according to standard protocols.

4.4.2 3D-Immunofluorescence Staining

Organoids embedded in BME were mechanically dissociated using a P1000 pipette directly in the culture well and transferred to a 15 ml tube.

After centrifugation (1200 rpm, 10 min) and PBS wash (Gibco, Thermo Fisher Scientific, Carlsbad, CA, USA, D8537), organoids were transferred to FBS-coated 1.5 mL tubes to prevent adhesion. Fixation was carried out with 4% paraformaldehyde (PFA) in PBS for 16 hours at 4°C (or 2 hours at room temperature), with constant rotation to ensure uniform fixation.

Post-fixation, organoids were washed three times with PBS for 5 minutes each and permeabilized in PBS with 0.4% Triton X-100 (Sigma-Aldrich, Merck, Germany, 92046-34-9) for 30 minutes at room temperature. Afterward, blocking was performed in 1 ml of blocking solution (PBS with 4% bovine serum albumin, BSA) for 30 minutes at room temperature.

Primary antibody incubation was performed overnight at 4 °C using 200 µl in a 1:1 dilution of blocking solution and PBS, containing the following antibodies: goat polyclonal anti-Ki-67 (Santa Cruz Biotechnology, Texas, USA, sc-7846, dil.1:50) and mouse monoclonal anti-β-actin (Santa Cruz Biotechnology, Texas, USA, sc-69879, dil.1:50). For co-staining, both antibodies were applied simultaneously. After incubation, organoids were washed three times with PBS for 10 minutes each.

After washing, secondary antibody incubation was carried out for 1 hour at room temperature using 200 µl of blocking solution diluted 1:1 with PBS, containing Alexa Fluor-conjugated secondary antibodies: anti-goat 568 (1:250) (Invitrogen™, ThermoFisher, Carlsbad, CA, USA, A11057) and anti-mouse 488 (1:250) (Invitrogen™, ThermoFisher, Carlsbad, CA, USA, A21202). All steps were performed under constant rotation, and washes following secondary incubation were done in the dark (3 × 10 minutes in PBS).

DAPI nuclear staining was performed (1:20,000 dilution in PBS, 10 minutes), followed by PBS (2 × 10 min) and Milli-Q water washes to prevent salt crystal formation. Organoids were resuspended using FBS-coated tips, to minimize adhesion loss.

For imaging, organoids were transferred to 96-well plates in 200 µl PBS Imaging was conducted using the Olympus IXplore™ IX85 fluorescence microscope.

4.4.3 Targeted Genomic Profiling of FFPE and PDOs

The patient-derived tumor organoids and primary tumor sample from donor patient were molecularly characterized using next-generation sequencing (NGS) technology.

DNA from Tumor sample (FFPE) and PDOs were isolated using QIAamp DNA FFPE Advanced Kit (Qiagen, Germantown, MD, USA, 56604) and QiaDNeasy Blood and Tissue kit (Qiagen, Germantown, MD, USA, 69504) respectively, according to the manufacturer's instructions.

Mutational analysis was performed using AmpliSeq for Illumina Comprehensive Panel v3 (Illumina, California, 20019109), following the manufacturer's protocol, which offers coverage of 161 cancer-associated genes, including kinases and genes involved in DNA repair. DNA libraries have been employed to assess for single nucleotide, short indel, and CN variants.

The library preparation workflow for DNA using the AmpliSeq for Illumina Comprehensive Panel v3 involves several key steps that transform genomic DNA into sequence-ready libraries for high-throughput sequencing. The process begins by quantifying the input DNA using Qubit DNA HS Assay Kit (Invitrogen™, ThermoFisher, Carlsbad, CA, USA, Q32851). DNA input required for the assay was 180 ng DNA.
Once the DNA is prepared, target regions are amplified using PCR. This step requires mixing the DNA with AmpliSeq HiFi Mix and primer pools specific to the comprehensive panel. The PCR amplification focuses on the regions of interest, generating many copies of the target sequences. After amplification, the amplicons undergo partial digestion with FuPa reagent, which helps reduce unwanted primer dimers and other artifacts from the PCR reaction.
Following this, index adapters are ligated to the amplicons. These adapters are crucial as they allow multiple libraries to be pooled together for sequencing, enabling efficient high-throughput workflows. Each sample receives unique index combinations, ensuring that data from different samples can be accurately distinguished during sequencing.

The next step involves cleaning up the amplified libraries using magnetic beads, such as Agencourt AMPure XP beads (Beckman Coulter, California, USA, A63881). This step removes any remaining contaminants or unwanted products from the previous reactions,

leaving behind a purified library that contains the target DNA sequences ready for sequencing.

To ensure there is enough material for sequencing, the libraries are amplified again in a second round of PCR. This final amplification step boosts the quantity of the DNA library to meet the input requirements of Illumina sequencing platforms. After amplification, a second cleanup is performed to purify the libraries further and remove any remaining primers or byproducts.

At this stage, the quality of the libraries is checked using a Bioanalyzer (Agilent, DNA 1000 kit, 5067-1504). These instruments provide a profile of the DNA fragments, helping to confirm that the library has the expected size distribution and concentration. Once the libraries are validated, they are diluted to the appropriate starting concentration for Illumina sequencing system.

NGS libraries have been sequenced on an Illumina MiSeq sequencer (2x300bp) and raw data have been processed using the embedded primary and secondary analysis pipeline "Amplicon Studio".

Variant call format (VCF) files generated from this step were subsequently intersected with a BED file containing the reference genome coordinates, to accurately map and annotate the detected variants. The resulting datasets were analyzed in RStudio using the *mutationalPatterns* package within the *PlotR* framework, enabling the visualization of mutational profiles. Finally, comparative analyses were performed between FFPE-derived samples and patient-derived organoids (PDOs) to assess concordance and potential discrepancies in mutational landscapes.

4.4.4 Whole genome sequencing, mutation and copy number analysis

Whole-genome sequencing (WGS) was performed on DNA isolated from patient-derived organoids. 1 µg of high molecular weight DNA was used to prepare sequencing libraries using the SQK-LSK114 ligation sequencing kit (Oxford Nanopore Technologies, UK), following the manufacturer's protocol. Briefly, the protocol involves initial DNA repair using NEBNext FFPE DNA Repair Mix and end-preparation with NEBNext Ultra II End Repair/dA-Tailing Module (New England Biolabs, Massachusetts, USA, E7645). Sequencing adapters provided in the SQK-LSK114 kit were then ligated using the NEBNext Quick Ligation Module, and libraries were purified with AMPure XP beads (Beckman

Coulter, California, USA A63881). Final library quantification was performed using the Qubit DNA HS Assay Kit (Invitrogen™, ThermoFisher, Carlsbad, CA, USA, Q32851).

All sequencing data were analyzed using the wf-humanvariation pipeline provided through the EPI2ME platform by Oxford Nanopore Technologies, which accepts a reference genome (hg19 or hg38) and BAM files as input, performs quality control (coverage estimation and alignment statistics), and carries out variant calling with Clair3, followed by functional annotation with SnpEff and ClinVar.

4.5 Organoids preparation for irradiation response and drug tests

Organoids were seeded, after counting, in 35 × 10 mm dishes (Corning, NY, USA) approximately 3 days before start of treatment.

4.5.1 Treatment Response Evaluation in PDOs: Irradiation and Drug Exposure

The treated PDO line was subjected to the following protocol. The in vitro irradiation experiments were conducted using a dedicated system and a custom-designed irradiation setup, as previously described by Puglisi et al. [143]. This system allows cell cultures to be treated through the same equipment used for patients (Varian Novalis-TrueBeamSTx linear accelerator). To simulate the flow of radiation beams through human tissues before reaching the cells, each dish containing PDOs was inserted inside a niche of a custom-built phantom made of plexiglass to position the cells at the radiation isocenter. From a physical point of view, this material resembles water, which is the main component of human tissues, when interacting with radiation during treatment. PDOs were irradiated through a Varian Novalis-TrueBeamSTx linear accelerator, a radiotherapy equipment able to perform stereotactic treatments with very high precision, using the high-dose rate FFF technique and the high-definition multilamellar collimator (MLC), with minimum leaves size at the isocenter of 2.5 mm, specifically designed to treat small lesions.

Two different radiation doses were tested: 8 Gy, as reported by Yao et al [144] and a lower dose of 5 Gy.

Immediately following irradiation, the PDOs were re-seeded into 6-well plates (Corning Costar, Cambridge, MA, USA, Minneapolis, USA, 3516) and was added drug-containing culture medium with 5-fluorouracil (Merck, USA, F6627-10G) and oxaliplatin (Merck, USA, 09512-5Mg) at 1M final concentration. A second dose (drug boost) was administered three days later, after which the culture medium was replaced every three days with fresh drug-free medium.

To monitor treatment effects, organoids were imaged every six days over a 24-day period. Cell viability was assessed on days 6, 12, 18, and 24 using the CellTiter-Glo® 3D Cell Viability Assay (Promega, G9683), according to the manufacturer's protocol.

A second PDOs line served as the untreated control. It was seeded and maintained under standard culture conditions without irradiation or chemotherapy. This control was followed over the same 24-day period, with medium changes every seven days, to establish baseline growth and viability parameters.

4.6 Transcriptomic Profiling: Bulk RNAseq Analysis

Patient-derived organoids (PDOs) established from rectal cancer biopsies were subjected to bulk RNA sequencing at multiple time points: baseline (T0), post-treatment (T1), and relapse (T2). All experiments were performed in biological triplicates to ensure reproducibility and enable robust statistical analysis.

At T0, total RNA was extracted using the RNeasy Mini Kit (Qiagen, Germantown, MD, USA, 74104). RNA integrity was assessed using the Agilent 2100 Bioanalyzer with the Eukaryote Total RNA Nano 6000 Assay (Agilent, Santa Clara, USA, 5067-1511), and quantification was performed using the Qubit RNA Broad Range Assay Kit (Invitrogen™, ThermoFisher, Carlsbad, CA, USA Q32851), with 500 ng of total RNA used as input for library preparation.

Full-length cDNA libraries were prepared using the Oxford Nanopore Technologies (ONT) cDNA-PCR Barcoding Kit V14 (Oxford Nanopore Technologies, Oxford, UK, SQK-PCB114.24). The protocol included a reverse transcription step using strand-switching primers, which enabled the capture of full-length polyadenylated transcripts and the incorporation of unique molecular identifiers (UMIs) to enhance isoform-level resolution. Reverse transcription was followed by PCR amplification using rapid attachment barcode

primers, allowing multiplexing of up to 24 samples without the need for ligation-based adapter attachment. After approximately 40 minutes of amplification, a bead-based clean-up was performed to remove short cDNA fragments. The barcoded PCR products were then pooled and ligated with sequencing adapters using the Rapid Adapter Mix in a 5-minute reaction. The final libraries were loaded onto PromethION flow cells (Oxford Nanopore Technologies, Oxford, UK, FLO-PROMIN114) and sequenced using the MinKNOW software. Sequencing was conducted on R10.4.1 flow cells using Kit 14 chemistry, optimized for high accuracy (Q20+) and efficient isoform detection.

On day 24 (T1), organoids from each experimental condition (control and treated) were processed for bulk RNA-seq following the same workflow. The data generated at T1 were integrated with those from T0 to assess transcriptomic changes associated with treatment-induced phenotypic plasticity.

After 12 weeks of culture post-treatment, organoids that had undergone various bottleneck conditions were sampled again (T2). RNA extraction, library preparation, and long-read sequencing were repeated as described above to enable a longitudinal analysis of transcriptional changes associated with adaptive relapse dynamics.

All sequencing data were analyzed using the wf-transcriptomics pipeline provided through the EPI2ME platform by Oxford Nanopore Technologies.

Long-read bulk sequencing data were analyzed using R (version 4.3.0) [145], with the R packages DEXSeq [146] and stageR [147], to identify differential transcriptional isoforms between experimental conditions. Additionally, we investigated isoform structural differences to assess whether isoform switching within the same gene between conditions was associated with the isoform's biotype (e.g., a shift from a non-coding to a protein-coding transcript). To explore the functional and biological implications of the observed changes, we performed pathway enrichment analysis using Gene Ontology (GO) [148], KEGG database [149] and Gene Set Enrichment Analysis (GSEA)[150].

4.7 Transcriptomic Profiling: Single- cell RNA sequencing analysis

Patient-derived organoids (PDOs) established were also subjected to single cell-RNA sequencing at multiple time points: baseline (T0), post-treatment (T1), and relapse (T2), across all experimental conditions: untreated and treated.

For each condition and timepoint, organoids were dissociated into single-cell suspensions following an optimized protocol involving enzymatic digestion with Trypsin-EDTA, mechanical disruption, filtration through 40 µm cell strainers (Corning, 352340) to remove debris and clumps, and dead cell depletion using the Miltenyi Dead Cell Removal Kit (Miltenyi Biotec, Deutschland, 130-090-101). Cell viability and concentration were assessed using an automatic cell counter EVE™ (Avantor, Pennsylvania, USA, 734-2675). The resulting viable single-cell suspensions were resuspended in PBS (Gibco, Thermo Fisher Scientific, Carlsbad, CA, USA, D8537) containing 0.04% BSA at a concentration of approximately 700–1500 cells/µL, aiming to recover ~10,000 cells per sample. Single-cell RNA-seq libraries were prepared using the Chromium GEM-X Single Cell 3' Gene Expression v4 platform (10x Genomics, Pleasanton, CA, USA). Cell suspensions were loaded into a Chromium chip along with barcoded gel beads, partitioning oil, and RT reagents. Inside the microfluidic system, single cells were encapsulated into Gel Beads-in-Emulsion (GEMs), where cell lysis and reverse transcription occurred. During this step, polyadenylated RNA molecules were captured by barcoded oligo-dT primers, resulting in the synthesis of first-strand cDNA tagged with a cell-specific barcode and a unique molecular identifier (UMI). After reverse transcription, the emulsion was broken and barcoded cDNA was recovered and purified using silane magnetic beads, followed by cDNA amplification via PCR.

The amplified cDNA was then used as input for long-read library preparation following the Oxford Nanopore Technologies protocol for single-cell transcriptomics using 10x-derived cDNA (Oxford Nanopore Technologies, UK, SQK-LSK 114). A second PCR was performed using biotinylated primers to enrich full-length cDNA, which was subsequently purified using streptavidin-coated magnetic beads to remove unbarcoded fragments. The resulting material underwent end-repair and dA-tailing using NEBNext Ultra II reagents, followed by ligation of sequencing adapters provided in the SQK-LSK114 kit. Final libraries were purified using AMPure XP beads and quantified using the Qubit DNA HS Assay Kit (Invitrogen™, ThermoFisher, Carlsbad, CA, USA, Q32851). Libraries were sequenced on a PromethION device (Oxford Nanopore Technologies) using FLO-PRO114M (R10.4.1) flow cells. Basecalling and demultiplexing were performed using MinKNOW and the EPI2ME

wf-single-cell pipeline, enabling the assignment of reads to individual cells based on their 10x barcodes and UMIs.

Single-cell RNA sequencing (scRNA-seq) data from human colon organoids were processed using the Seurat R package [151]. Raw count matrices were imported with the Read10X function. To assess cell quality, mitochondrial gene expression was calculated, and cells exhibiting more than 20% mitochondrial content were excluded. The filtered dataset was then normalized and log-transformed.

Subsequently, the data were scaled, and principal component analysis (PCA) was conducted. The first 10 principal components were used for clustering and dimensionality reduction using UMAP and t-SNE. Cell type annotation was performed using CellMarker [152], the Cell Annotation Platform, and curated markers from the literature.

To explore transcriptional differences, we identified the most upregulated genes in each cluster across conditions. These gene sets were subjected to Gene Ontology (GO) [148] enrichment analysis to investigate Biological Processes, Molecular Functions, and Cellular Components [153]. Additionally, we applied Gene Set Enrichment Analysis (GSEA) [154].

Chapter 5. Results

5.1 Evaluation of Model Fidelity: Immunohistochemical and Genomic Comparison between Patient-Derived Organoids and Original FFPE Tumor Samples

Hematoxylin and eosin (H&E) staining demonstrated that patient-derived organoids (PDOs) recapitulated the key architectural and cytological features of the matched rectal adenocarcinoma. Primary tumor tissue exhibited irregular, infiltrative glandular structures with nuclear atypia and desmoplastic stroma, consistent with moderately to poorly differentiated colorectal adenocarcinoma. Correspondingly, PDOs displayed dense epithelial architecture with hyperchromatic nuclei and partial glandular organization, reflecting the neoplastic phenotype in a simplified, stroma-free 3D context.

Subsequent marker expression analysis revealed that organoids and paired primary tumors exhibited similar staining patterns for CDX2, CK20 and β-catenin (Figure 11).

Collectively, these findings indicate that organoids preserve both the histological features and marker expression of the original tumor tissues, as reported previously [155,156].

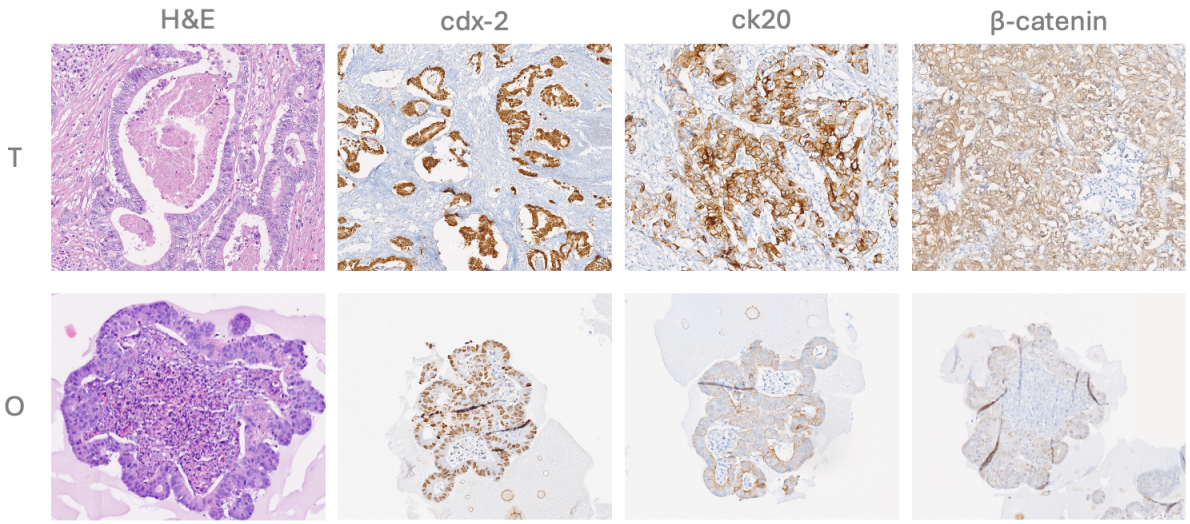


Figure 11. H&E staining compares PDOs with corresponding primary tumor (T, primary tumor; O, PDOs). Immunohistochemistry staining of cdx-2, ck20,b-catenin on PDOs and corresponding primary tumor.

Previous studies have revealed that organoids derived from cancer patients recapitulate the genomic profiles of corresponding tumors, including DNA mutation and DNA copy number variations (CNVs) [157, 158, 159, 160].

To validate this, targeted next-generation sequencing (NGS) was performed on PDOs and matched tumor biopsies. Mutation profiles exhibited high concordance between models and tumors (Cosine similarity = 0.993), further supporting their genomic fidelity (Figure 12).

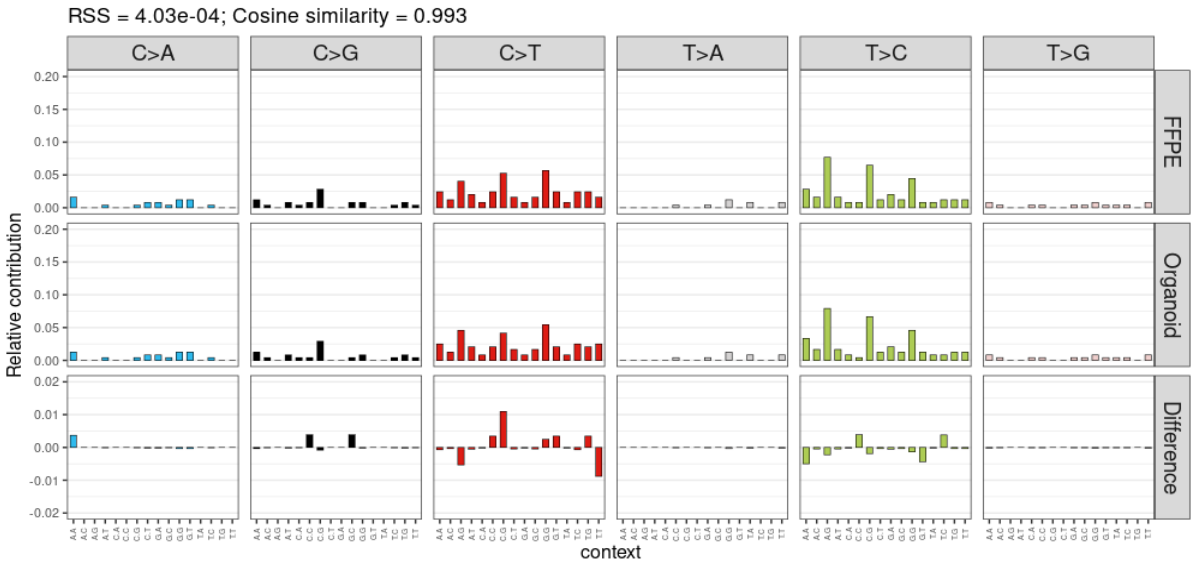


Figure 12. 96-bar mutational plot displaying the different contribution of point mutation type.

5.2 Immunofluorescence and Genomic Characterization of PDOs

Whole-genome sequencing (WGS) was performed to further characterize the genomic features of PDOs.

The top 50 genes with clinically relevant variants, as annotated in ClinVar, are shown in Figure X. Among these, several are well-known players in colorectal cancer (CRC) biology. The APC gene, frequently mutated in CRC, acts as a gatekeeper in the Wnt/β-catenin signaling pathway; loss-of-function mutations in APC drive early tumorigenic events by deregulating cell proliferation. NOTCH1, a key regulator of cell fate and stem cell

maintenance, has been implicated in CRC progression, particularly in the transition from adenoma to carcinoma. BRCA2 and FANCA, both involved in homologous recombination repair, suggest a potential defect in DNA double-strand break repair, which may contribute to genomic instability in the tumor. SYNE1, a large spectrin-repeat protein, has been recurrently found mutated in CRC, possibly affecting nuclear structure and chromosomal segregation. Other genes, such as COL4A2 and COL11A1, are related to extracellular matrix composition and may influence tumor invasion and microenvironment interactions.

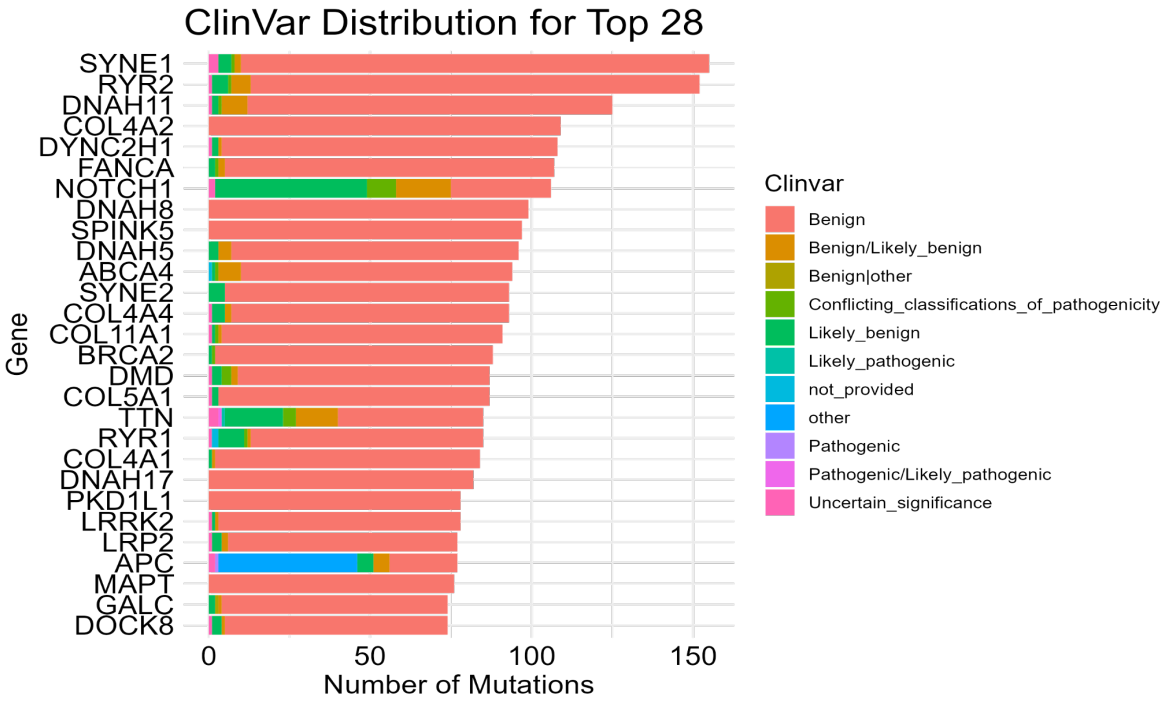


Figure 13. Top 50 ClinVar genes with clinically relevant variants in CRC organoids.
Bar plot showing the 50 genes with the highest number of ClinVar-reported variants identified by WGS in patient-derived CRC organoids.

The mutational spectrum of single nucleotide substitutions (Figure 14) revealed a predominance of C>T and G>A transitions, which together represented the most frequent substitution types. These changes are often associated with the spontaneous deamination of 5-methylcytosine at CpG sites, a process commonly observed in colorectal carcinogenesis. The elevated frequency of A>G transitions suggest additional mutational processes, potentially linked to oxidative damage or transcription-associated mutagenesis. Interestingly, a notable proportion of T>C substitutions was detected, which might be indicative of treatment-related mutational signatures or specific enzymatic activities, such

as APOBEC-mediated deamination. Less frequent substitutions, including C>A, T>G, and G>T, could reflect exogenous damage from reactive oxygen species or replication stress. The overall profile aligns with mutational patterns described in CRC, supporting the clinical relevance of the genomic alterations identified in the organoid model.

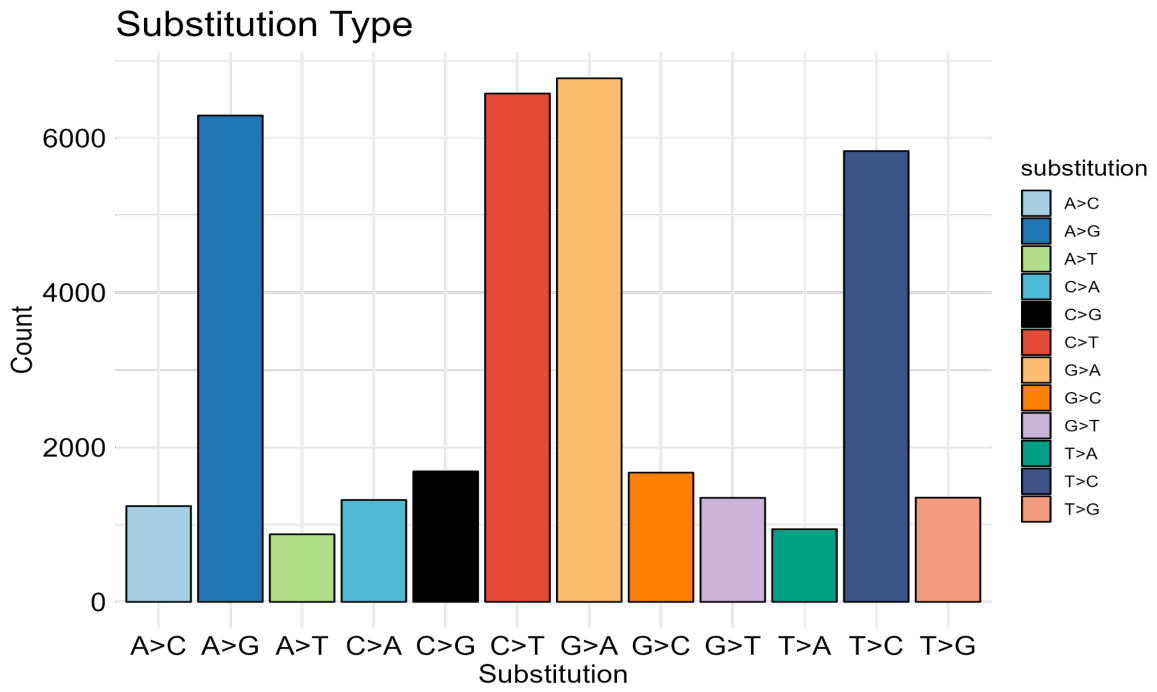


Figure 14. Single nucleotide substitution profile of CRC organoids. Distribution of substitution types detected in WGS analysis of CRC organoids. C>T and G>A transitions are the most frequent, consistent with mutational patterns typically observed in colorectal cancer, while a notable proportion of T>C substitutions may reflect treatment-related or enzymatic mutational processes.

Further morphological analyses were conducted on organoids through immunofluorescent imaging. A triple-staining immunofluorescent protocol was successfully tested on floating CRC organoids using **DAPI** (blue) to visualize nuclei, **β-actin** (green) to highlight cytoskeletal organization, and **Ki-67** (red) to detect proliferating cells.

In the deconvoluted images, the DAPI channel reveals a dense nuclear distribution within the organoid, indicating high cellularity typical of CRC-derived structures. The β-actin signal outlines the cytoplasmic borders and structural framework, emphasizing the spherical, lobulated morphology and the integrity of the epithelial organization. The Ki-67 signal shows punctate and nuclear-localized staining in a subset of cells, reflecting active proliferation within specific regions of the organoid.

Merged images demonstrate the spatial relationship between these markers: DAPI-positive nuclei are surrounded by β-actin–rich cytoskeleton, while Ki-67–positive nuclei are interspersed throughout, suggesting heterogeneous proliferative activity. The final composite image clearly shows that proliferating cells are not uniformly distributed but rather concentrated in defined clusters,consistent with patterns observed in colorectal cancer tissue, where proliferation often occurs at the invasive front or in niche-like microenvironments, consistent with proliferative heterogeneity observed in colorectal tumors.

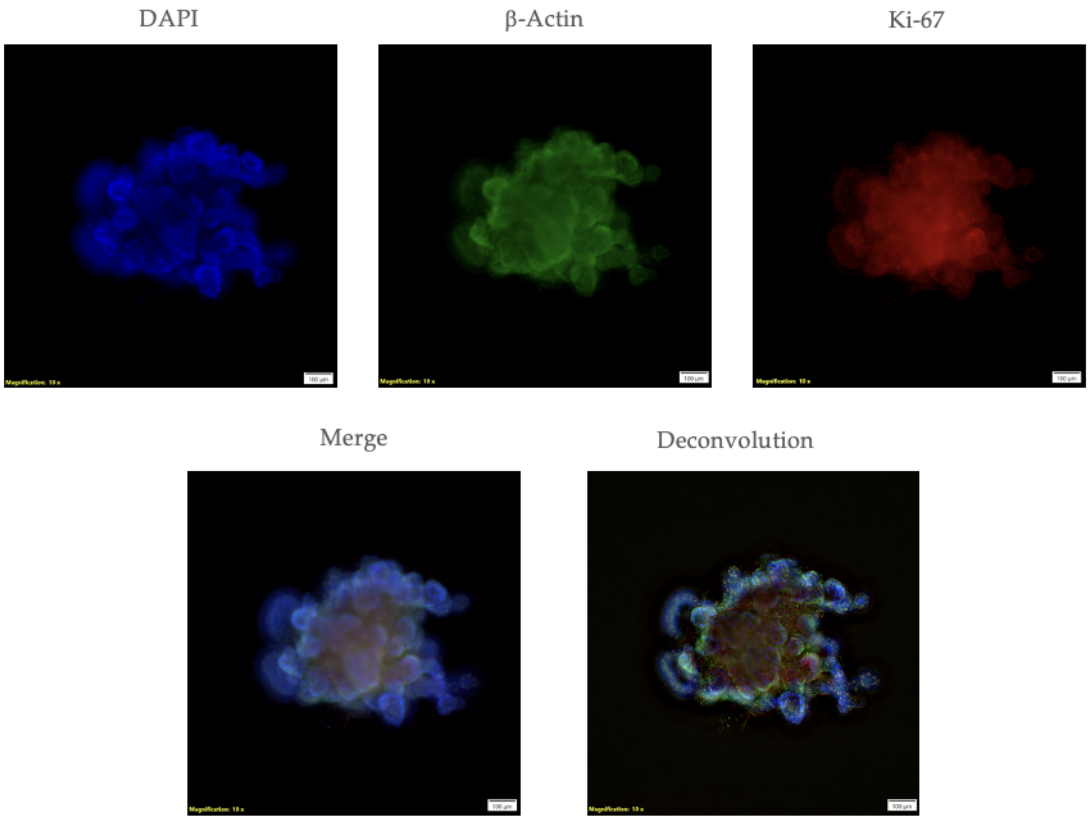


Figure 15. Triple-staining immunofluorescence imaging of CRC organoids.
Representatives deconvoluted and merged confocal images of floating CRC organoids labeled with DAPI (blue, nuclei), β-actin (green, cytoskeleton), and Ki-67 (red, proliferating cells). Merged images reveal heterogeneous distribution of Ki-67–positive cells, often clustered in specific regions, consistent with proliferation patterns observed in colorectal tumors. Scale bars: 100 µm.

5.3 Functional Response to Chemotherapy and Irradiation

Cell viability of patient-derived colorectal cancer organoids was evaluated by CellTiter-Glo® following combined FOLFOX chemotherapy and irradiation at 5 Gy or 8 Gy, and monitored at days 6, 12, 18, and 24. At 5 Gy, viability decreased from ~50% at day 6 to ~35% at day 18, with partial recovery to ~55% by day 24 (Figure 16A). At 8 Gy, viability declined more sharply, from ~40% at day 6 to <15% by day 18, and remained suppressed thereafter. Untreated controls retained >95% viability across all timepoints (Figure 16B).

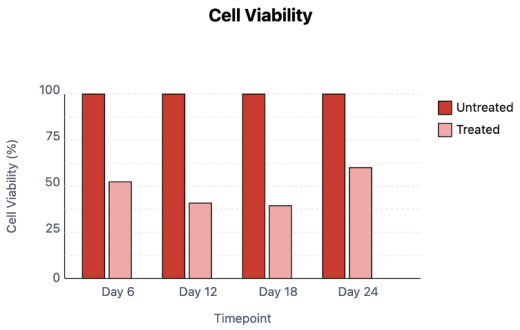


Figure A. Bar-plot cell viability Radiotherapy 5 Gy + Folfox

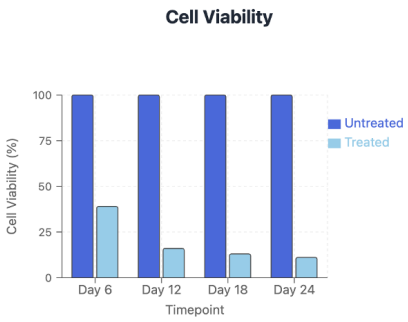


Figure B. Bar-plot cell viability Radiotherapy 8 Gy + Folfox

Figure 16 A-B. Bar plots showing cell viability of colorectal cancer organoids at different timepoints (Day 6, 12, 18, and 24) following combined treatment with radiotherapy and FOLFOX. (A) Radiotherapy at 5 Gy + FOLFOX. (B) Radiotherapy at 8 Gy + FOLFOX. Untreated controls maintained consistently high viability, while treated organoids exhibited a marked reduction in viability, more pronounced at higher radiation dose and later timepoints.

Sequential brightfield imaging of the same cultures confirmed the quantitative results. Untreated organoids maintained compact, well-defined spheroids with preserved density and morphology. In contrast, treated organoids exhibited progressive structural deterioration, including size reduction, irregular borders, and loss of cellular compaction, which was more pronounced and irreversible at 8 Gy. The 5 Gy–treated organoids showed visible damage but retained partial growth capacity, consistent with the modest recovery in viability measurements. Based on these findings, the 5 Gy condition was selected for downstream

functional and molecular analyses to preserve a viable population while maintaining a measurable treatment effect (Figure 17).

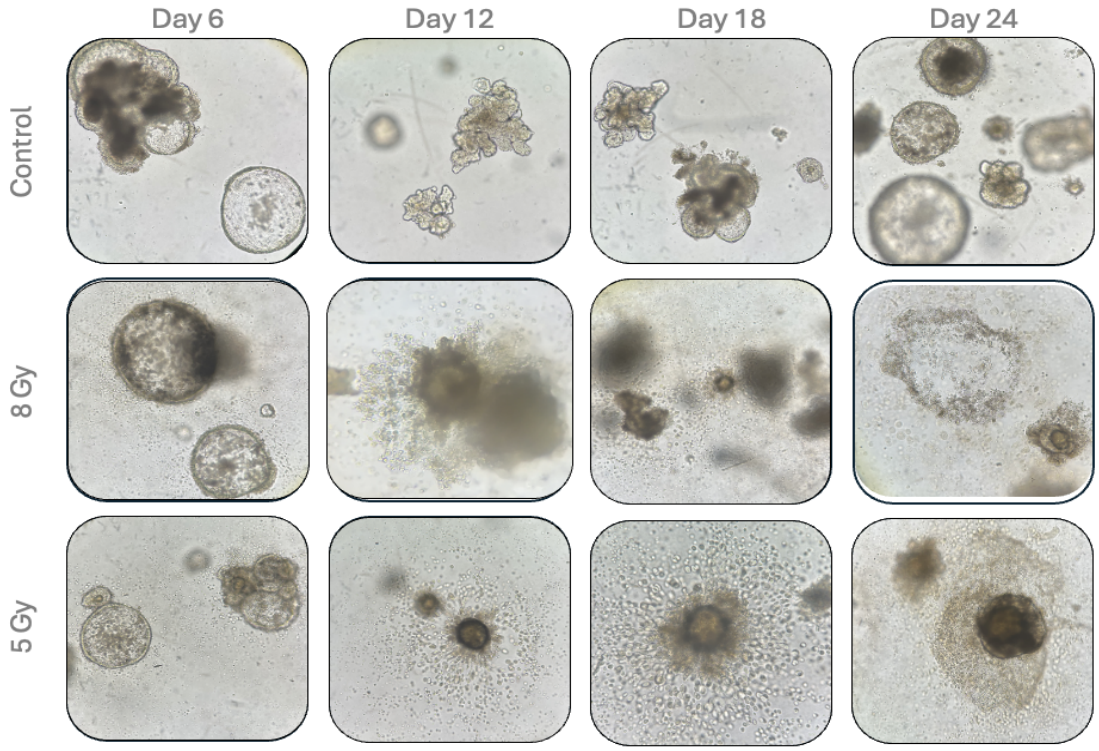


Figure 17. Morphological changes in patient-derived colorectal cancer organoids following chemoradiotherapy vs control.

5.4 Transcriptomic Profiling: Bulk RNA-Sequencing Analysis

The transcriptional landscape of PDOs was investigated using bulk RNA sequencing. Analyses included differential expression (DE), Gene Ontology (GO) enrichment, Gene Set Enrichment Analysis (GSEA), and long-read nanopore sequencing for isoform-level characterization. Collectively, these analyses provide a multi-layered perspective on the molecular responses of the organoids.

5.4.1 T1 treated vs T0 control

Differential expression analysis between treated organoids (T1) and untreated controls (T0) identified a large number of transcripts significantly deregulated at $|logFC| \geq 2$ and FDR < 0.05. Among the most upregulated, we observed genes involved in RNA processing and

splicing (*HNRNPAB, HNRNPD, DDX3X, SNRNP25*), highlighting extensive post-transcriptional reprogramming. The transcriptional regulator ID1 emerged as one of the most significantly upregulated factors, consistent with its role in promoting cell proliferation, epithelial–mesenchymal transition (EMT), and stress adaptation. Additional upregulated genes included those related to DNA replication and nuclear integrity (*MCM4, LMNA*), mitochondrial metabolism and oxidative phosphorylation (*ACO2, CYB5R3*), and extracellular matrix remodeling (*TGFBI, CEMIP*), reflecting combined transcriptional, metabolic, and microenvironmental adaptations to treatment.

Conversely, the most strongly downregulated transcripts included several ribosomal protein genes and ribosomal pseudogenes (RPL21P family, RPS15AP1, RPL39L). While pseudogenes do not directly contribute to ribosome biogenesis, their co-regulation with ribosomal protein genes is consistent with a global suppression of ribosome-related transcriptional programs. Together, these findings point to a dual adaptive response in treated organoids: activation of stress-related transcriptional and metabolic programs on one hand, and concomitant repression of biosynthetic capacity on the other. These findings are further contextualized by Gene Ontology (GO) enrichment analysis.

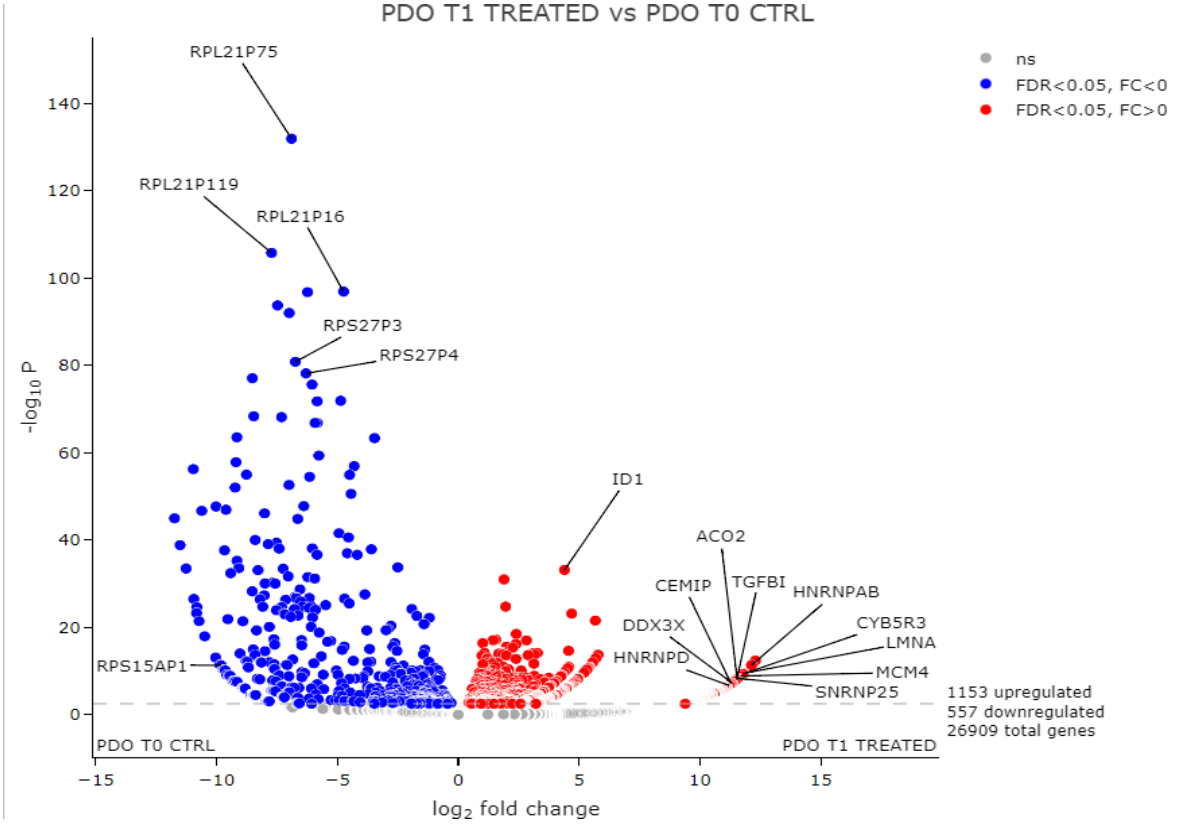


Figure 18. Volcano plot of differential gene expression analysis between patient-derived organoids (PDOs) after treatment (T1, treated) and at baseline (T0, control).

Gene Ontology (GO) enrichment analysis revealed a significant upregulation of pathways associated with the cellular response to therapeutic stress, most prominently RNA splicing and mRNA processing, consistent with extensive transcriptional and post-transcriptional reprogramming.

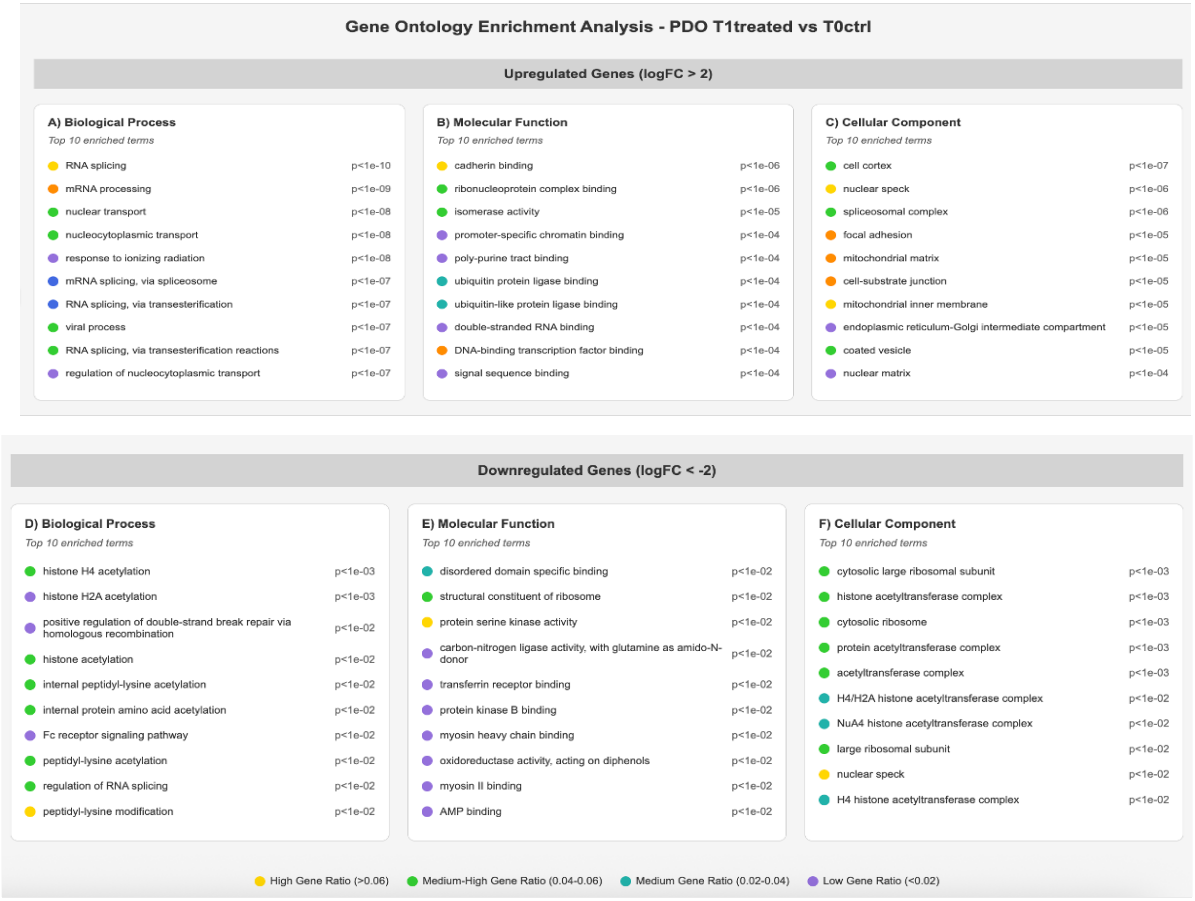


Figure 19. Gene ontology enrichment analysis of differentially expressed genes in PDOs T1 treated vs T0 control. Dot color represents gene ratio (number of genes in category/total genes in gene set). Only terms with p-value < 0.05 and top 10 terms per category are shown.

Long-read nanopore sequencing enabled a detailed characterization of alternative splicing, highlighting five key genes with treatment-relevant isoform changes.

Two main splicing trends emerged: (i) a shift toward non-functional isoforms in TPM3, IMPDH2, and DDX5, with increased retained introns and the production of transcripts subject to nonsense-mediated decay (NMD), potentially reducing protein function, cell

proliferation, and cytoskeletal stability; (ii) qualitative modulation of functional isoforms in CTSB and EPCAM, leading to alteration in protein structure or expression levels without abolishing coding potential, a pattern that could influence cellular invasiveness, adhesion, and epithelial identity.

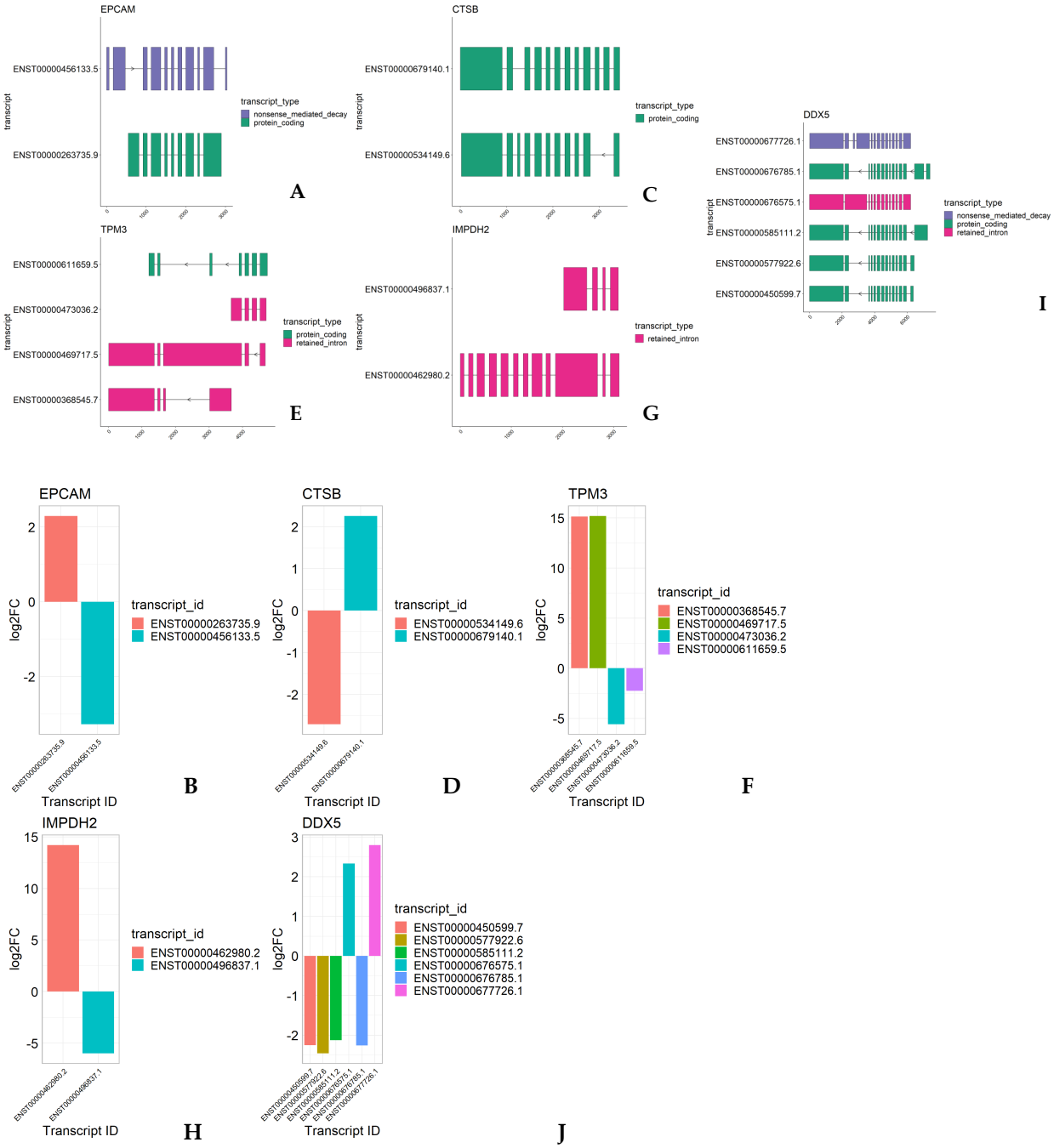


Figure 20. Analysis of alternative splicing events in selected genes in PDOs. Comparison between the treated condition (T1) and control (T0) highlights differences in isoform usage, suggesting a potential role of the treatment in modulating splicing patterns.

(A) ***EPCAM***: Isoform ENST00000263735.9 (orange) is upregulated, while ENST00000456133.5 (light blue) is downregulated in treated samples compared to controls.
(B) ENST00000263735.9 is protein-coding, whereas ENST00000456133.5 undergoes nonsense-mediated decay.
(C) ***CTSB***: Isoform ENST00000534149.6 (orange) is significantly downregulated, while ENST00000679140.1 (light blue) is upregulated in treated samples.
(D) Both isoforms are protein-coding; however, the exon–intron architecture differs, as illustrated in the panel.
(E) ***TPM3***: Four isoforms are detected, two upregulated (ENST00000368545.7, orange; ENST00000469717.5, green) and two downregulated (ENST00000473036.2, light blue; ENST00000611659.5, purple).
(F) ENST00000611659.5 is protein-coding, while the other three isoforms (ENST00000368545.7, ENST00000469717.5, ENST00000473036.2) are retained-intron, potentially producing truncated proteins.
(G) ***IMPDH2***: Isoform ENST00000462980.2 (orange) is upregulated, whereas ENST00000496837.1 (light blue) is downregulated in treated samples.
(H) Both transcripts are retained-intron but differ structurally: ENST00000462980.2 is longer, with multiple exons and a centrally retained intron, while ENST00000496837.1 is shorter, with few terminal exons and retained introns.
(I) ***DDX5***: Six isoforms are detected, two upregulated (ENST00000676575.1, light blue; ENST00000677726.1, pink) and four downregulated (ENST00000450599.7, orange; ENST00000577922.6, gold; ENST00000585111.2, green; ENST00000676785.1, blue).
(J) Transcript types include nonsense-mediated decay (ENST00000677726.1), retained-intron (ENST00000676575.1), and protein-coding with structural differences (ENST00000450599.7, ENST00000577922.6, ENST00000585111.2, ENST00000676785.1).

Beyond splicing-related processes, additional upregulated pathways included nucleocytoplasmic transport and DNA damage response programs, reflecting enhanced protein turnover and the acquisition of features associated with stress adaptation or radio resistance. Conversely, downregulated were enriched in pathways linked to protein synthesis, energy metabolism, and specialized functions, indicating a general reduction of biosynthetic activity and a shift toward a transient quiescent or adaptive silencing state.
Hallmark gene set analyses supported these findings, showing robust activation of p53 pathway, modulation of MYC targets, and induction of apoptotic signaling.

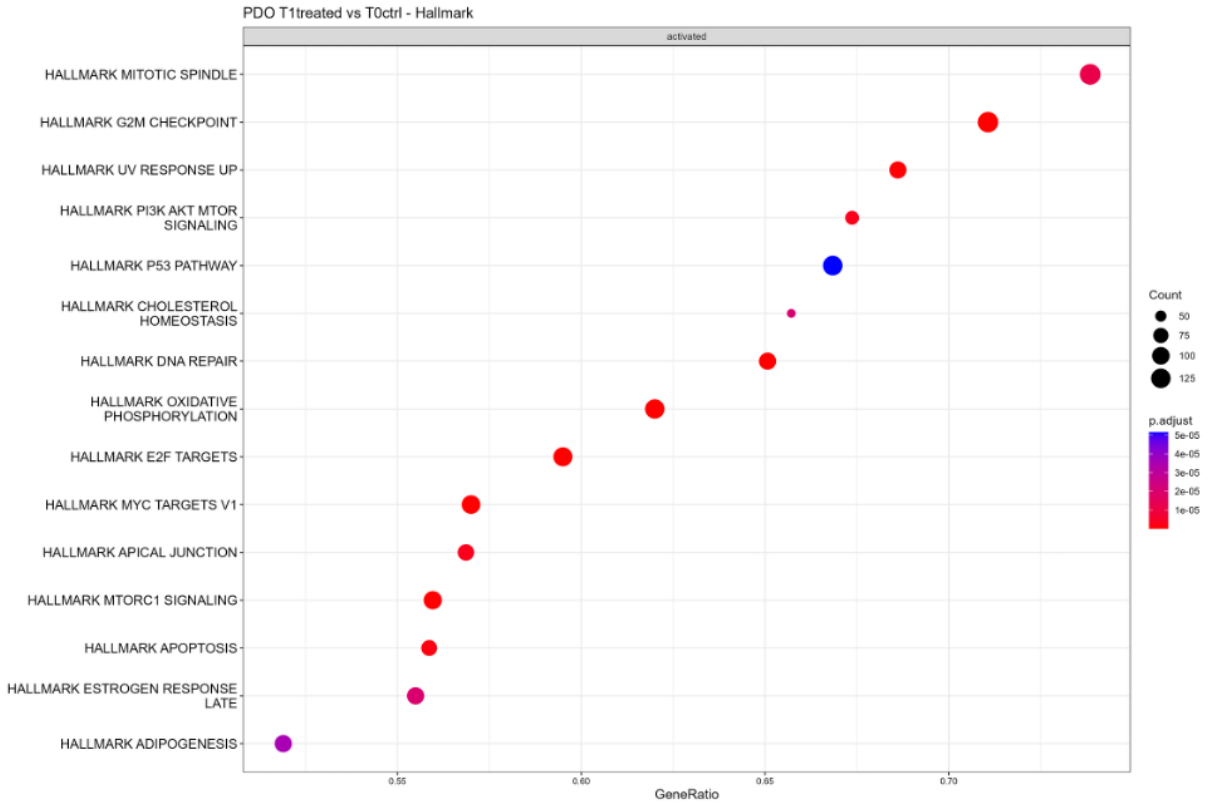


Figure 21. Gene set enrichment analysis (GSEA) of Hallmark pathways in patient-derived organoids (PDOs) treated at T1 versus untreated controls (T0). Dot plot shows significantly enriched pathways, with dot size indicating gene count and color representing adjusted p-value.

Taken together, these patterns indicate that treatment does not lead to complete eradication of tumor organoids but instead drives a phenotypic switch. Surviving cells appear to adopt a quiescent, stem-like state, which may underlie the emergence of therapeutic resistance.

5.4.2 T2 treated vs T0 control

Differential expression analysis between PDOs at T2 post-treatment versus baseline control (T0) identified 199 significantly deregulated transcripts ($|\log FC| \geq 2$, FDR < 0.05), comprising 63 upregulated and 136 downregulated genes. Among the most upregulated, we observed LCN2, a neutrophil-associated lipocalin implicated in stress adaptation and inflammatory responses, and several immune mediators such as HLA-DQB1 and RNF144A, together with long noncoding RNAs (*LINC01127*, *GAPLINC*) that have been linked to tumor progression and chemoresistance. These findings suggest the activation of inflammatory and stress-related programs in the surviving tumor cells. Conversely, the most downregulated

genes included NRG1, a trophic factor relevant for epithelial growth and survival, SMARCA2, a chromatin remodeler associated with differentiation, and the oncofetal lncRNA H19, alongside regulators of the hematopoietic/lymphoid niche such as IL7 and MME. This profile indicates a marked loss of differentiation cues and epithelial identity, with suppression of signaling pathways critical for tissue homeostasis. These results have been confirmed with the GO enrichment analysis.

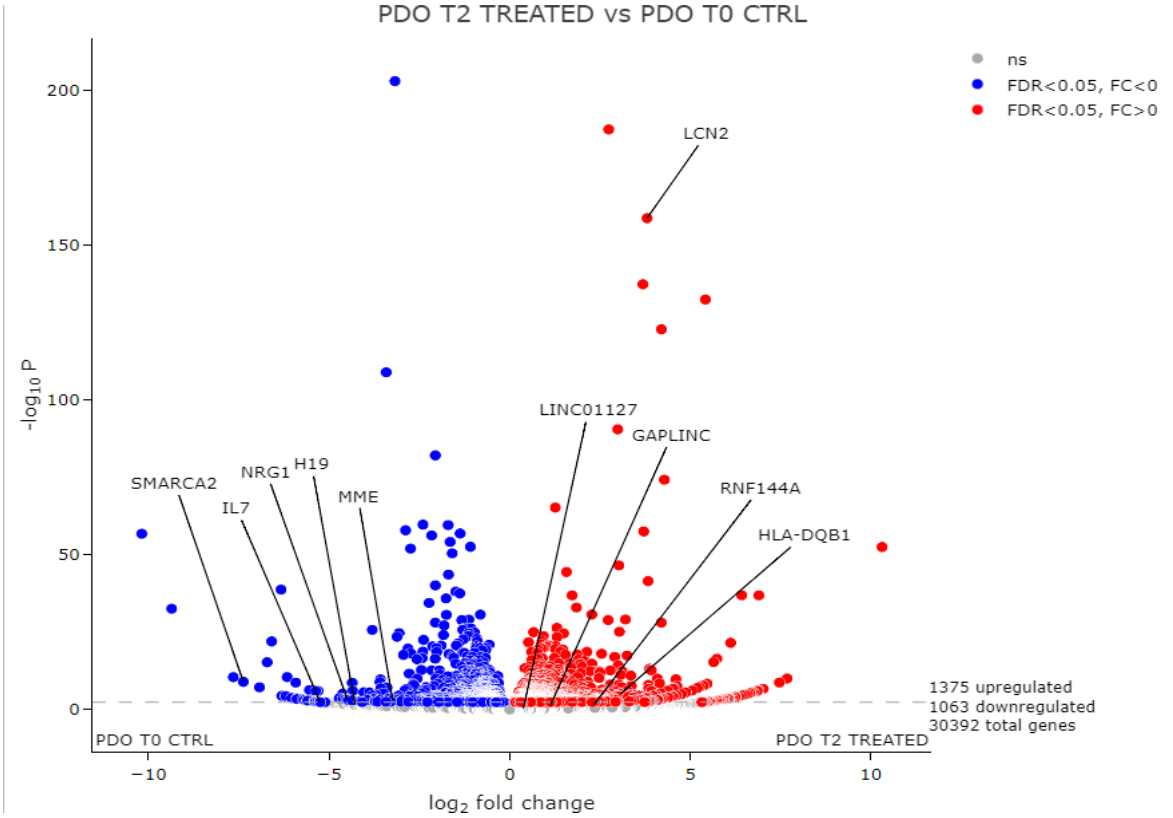


Figure 22. Volcano plot of differential gene expression analysis between patient-derived organoids (PDOs) at relapse time (T2, relapse) and baseline (T0, control).

Gene Ontology Enrichment Analysis revealed a persistent upregulation of genes associated with cellular responses to heavy metals and oxidative stress, consistent with the presence of a transcriptional "memory" of chemoradiotherapy. In parallel, chronic inflammatory pathways and cytoskeletal remodeling were activated, suggesting enhanced motility and structural reorganization at the nucleolar and cytoskeletal levels. Conversely, downregulated pathways indicated a pronounced loss of epithelial functions, disruption of the extracellular matrix organization, and alterations in acid and iron metabolism, pointing toward the selection of a quiescent, inflammatory "survivor" cell population with increased

malignant potential.

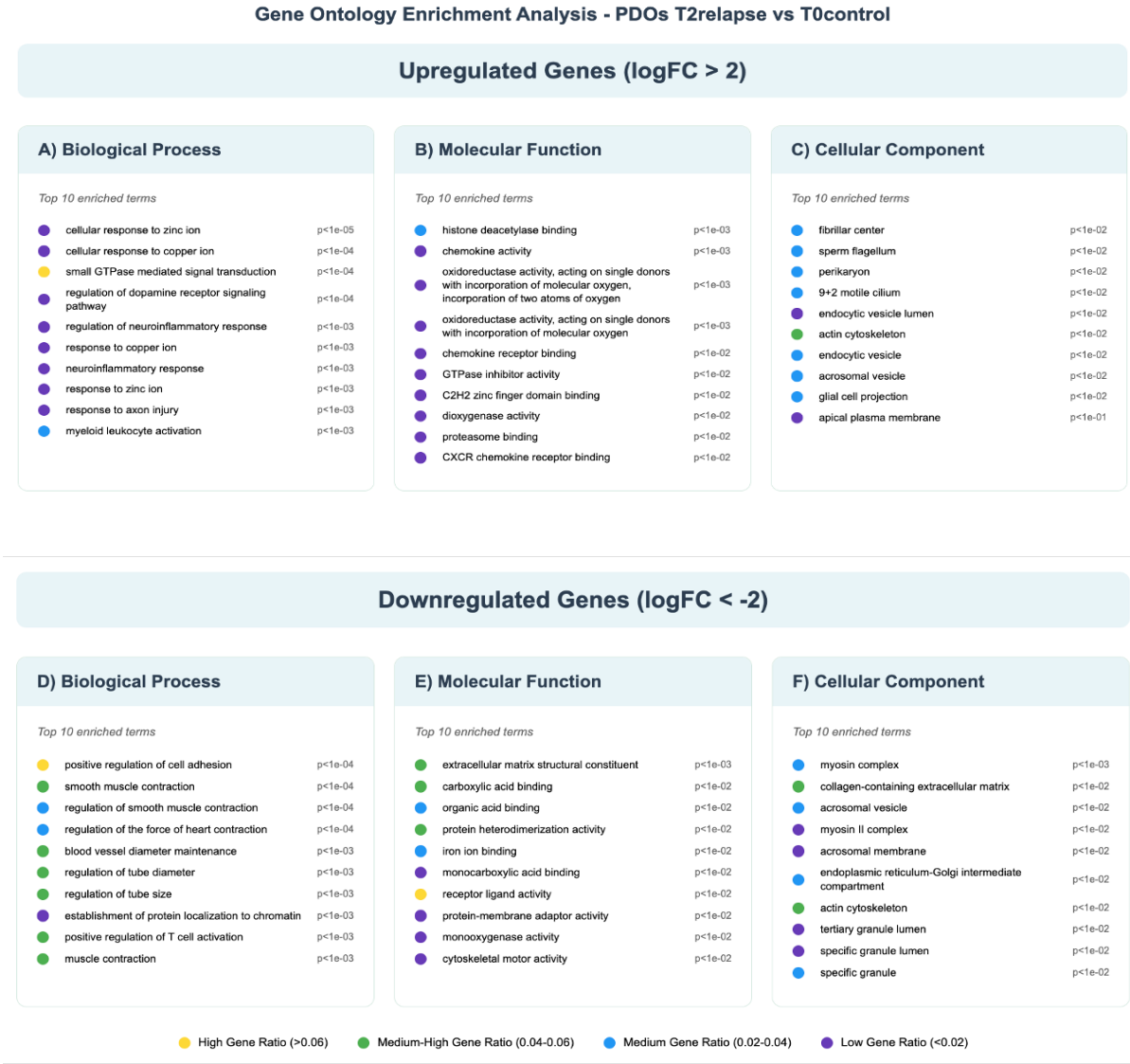
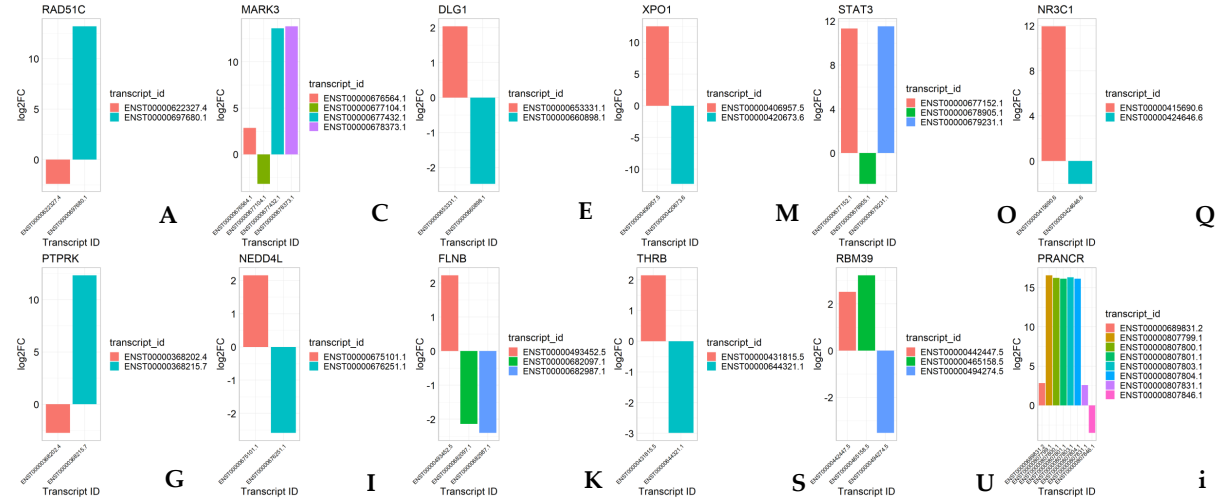


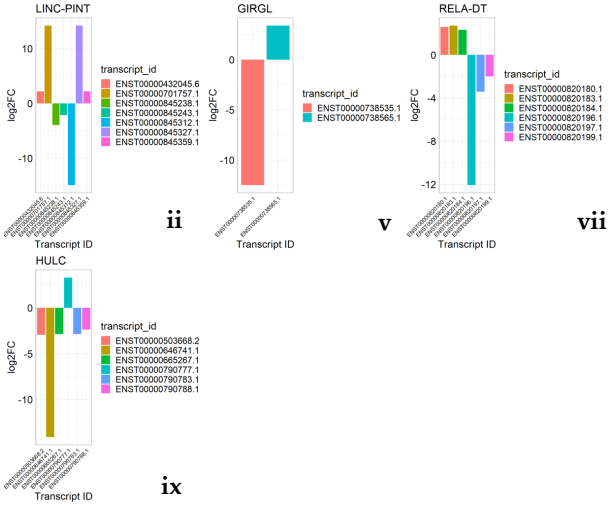
Figure 23. Gene ontology enrichment analysis of differentially expressed genes in PDOs T2 relapse vs T0 control. Dot color represents gene ratio (number of genes in category/total genes in gene set). Only terms with p-value < 0.05 and top 10 terms per category are shown.

Alternative splicing analysis between T2 and T0 revealed widespread transcriptomic remodeling. Key genes involved in DNA repair, cell cycle control, and epithelial integrity such as RAD51C, MARK3, DLG1, and PTPRK exhibited a shift toward nonfunctional isoforms, potentially impairing homologous recombination and promoting EMT. In contrast, NEDD4L, FLNB, and XPO1 showed upregulation of active isoforms supporting protein turnover, cytoskeletal dynamics, and nuclear export of tumor suppressors. Transcription factors like STAT3 displayed isoform redistribution likely modulating its pro-survival

activity, while nuclear receptors NR3C1 and THRB shifted to shorter isoforms that could alter hormonal and metabolic responses.

The analysis also highlighted altered expression of oncogenic and tumor-suppressive lncRNA isoforms. HULC and PRANCR were upregulated in stable, pro-tumor forms, while LINC-PINT showed splicing changes that may reduce its suppressive function. Additional changes in GIRGL and RELA-DT pointed to metabolic adaptation and NF-κB pathway activation, and RBM39 showed intron-retained variants, suggesting an impact on global splicing regulation.





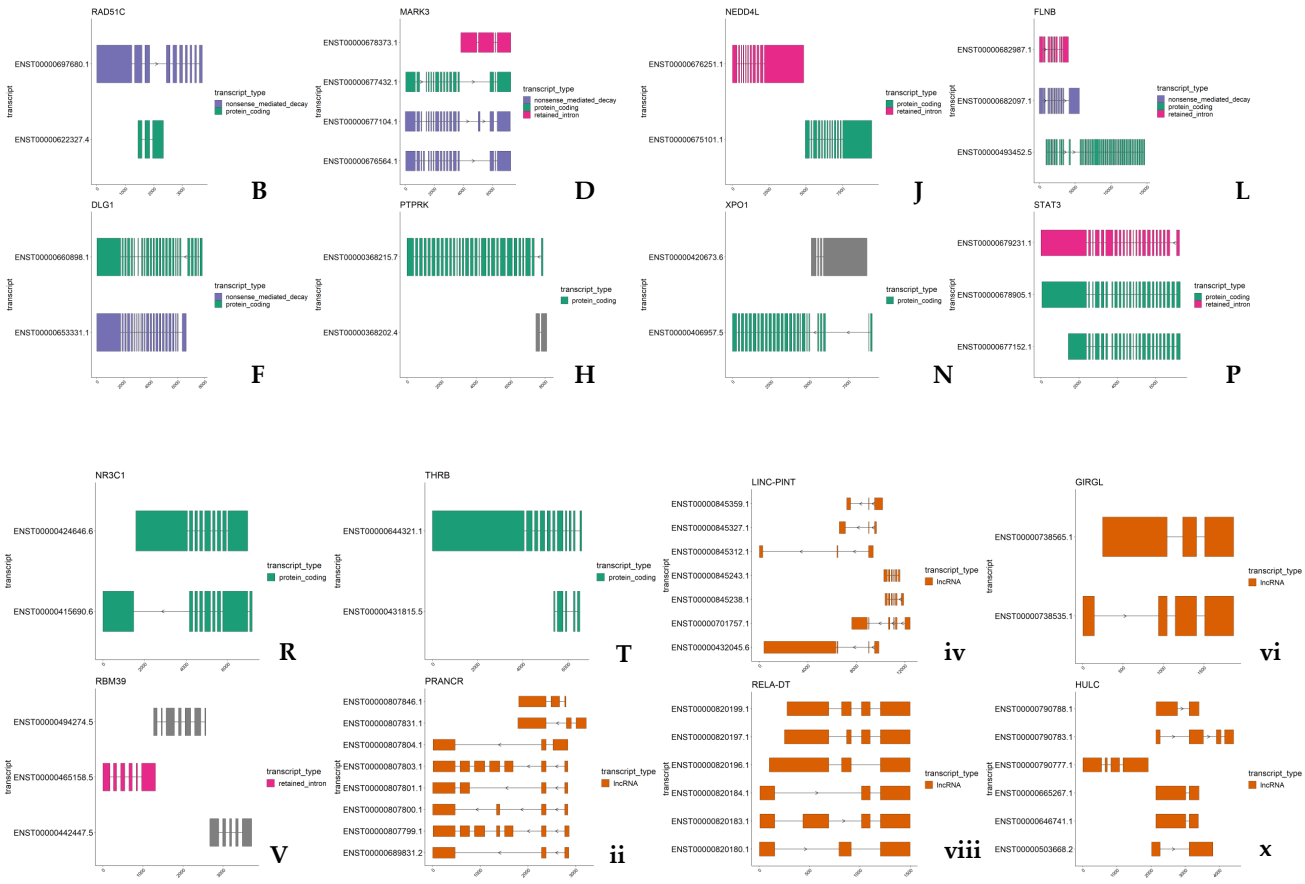


Figure 24. Analysis of alternative splicing events in selected genes in PDOs. Comparison between the relapse condition (T2) and control (T0) highlights differences in isoform usage, suggesting a potential role of the treatment in modulating splicing patterns.

(A) **RAD51C**: Isoform ENST00000697680.1 (light blue) is upregulated, while ENST00000622327.4 (orange) is downregulated in relapsed samples compared to controls.

(B) ENST00000622327.4 is protein-coding, whereas ENST00000697680.1 undergoes nonsense-mediated decay.

(C) **MARK3**, four isoforms are detected, three upregulated (ENST00000676564.1, orange; ENST00000677432.1, light blue, ENST00000678373.1, purple) and one downregulated (ENST00000677104.1, green)

(D) Transcript types include nonsense-mediated decay (ENST00000677104.1; ENST00000676564.1), retained-intron (ENST00000678373.1), and protein-coding (ENST00000677432.1)

(E) **DLG1**: Isoform ENST00000653331.1 (orange) is upregulated, while ENST00000660898.1 (light blue) is downregulated in relapsed samples compared to controls.

(F) ENST00000660898.1 is protein-coding, whereas ENST00000653331.1 undergoes nonsense-mediated decay

(G) **PTPRK**: Isoform ENST00000368215.7 (orange) is downregulated, while ENST00000368215.7 (light blue) is upregulated in treated samples compared to controls.

(H) Both isoforms are protein-coding; however, the exon–intron architecture differs, as illustrated in the panel.

(I) **NEDD4L**: Isoform ENST00000675101.1 (orange) is upregulated, while ENST00000676251.1 (light blue) is downregulated in relapsed samples compared to controls.

(J) ENST00000675101.1 is protein-coding while ENST00000676251.1 is retained intron

(K) **FLNB**: three isoforms are detected, one upregulated ENST00000493452.5 (orange), while the other two are downregulated in relapsed samples compared to controls ENST00000682097.1 (green), ENST00000682987.1 (blue).

(L) Transcript types include nonsense-mediated decay (ENST00000682097.1), retained-intron (ENST00000682987.1), and protein-coding (ENST00000493452.5)

(M) **XPO1**: Isoform ENST00000406957.5 (orange) is upregulated, while ENST00000420673.6 (light blue) is downregulated in relapsed samples compared to controls.

(N) Both isoforms are protein-coding; however, the exon–intron architecture differs, as illustrated in the panel.

(O) **STAT3**: three isoforms are detected, two upregulated ENST00000677152.1 (orange) and ENST00000679231.1 (blue), and one is downregulated ENST00000678905.1 (green) in relapsed samples compared to controls.

(P) Two isoforms are protein coding (ENST00000678905.1, ENST00000677152.1) while ENST00000679231.1 is retained intron.

(Q) **NR3C1**: Isoform ENST00000415690.6 (orange) is upregulated, while ENST00000424646.6 (light blue) is downregulated in relapsed samples compared to controls

(R) Both isoforms are protein-coding; however, the exon–intron architecture differs, as illustrated in the panel.

(S) **THRB**: Isoform ENST00000431815.5 (orange) is upregulated, while ENST00000644321.1 (light blue) is downregulated in relapsed samples compared to controls.

(T) Both isoforms are protein-coding; however, the exon–intron architecture differs, as illustrated in the panel.

(U) **RBM39**: three isoforms are detected, two upregulated ENST00000442447.5 (orange), ENST00000465158.5 (green) and one is downregulated ENST00000494274.5 (blue) in relapsed samples compared to controls
(V) Two isoforms are protein coding with different architecture (ENST00000494274.5 and ENST00000442447.5) while ENST00000465158.5 is retained intron.
(i-x) Different lcRNA isoforms with different transcript types.

Overall, the data indicate that relapse after radio+FOLFOX treatment is associated with selective alternative splicing reprogramming, favoring loss of tumor suppressor function, enhancement of pro-survival and pro-metastatic pathways, and modulation of regulatory lncRNAs, supporting the emergence of a therapy-resistant, transcriptionally plastic phenotype.

Gene Set Enrichment Analysis (GSEA) of cancer hallmark pathways corroborated these findings, showing a metabolic reprogramming characterized by robust induction of glycolysis—typical of resistant tumor cells—and moderate activation of the p53 pathway, consistent with adaptive resistance mechanisms. The concomitant activation of oncogenic signaling together with suppression of metabolic homeostasis may represent a key driver of therapy-resistant recurrence.

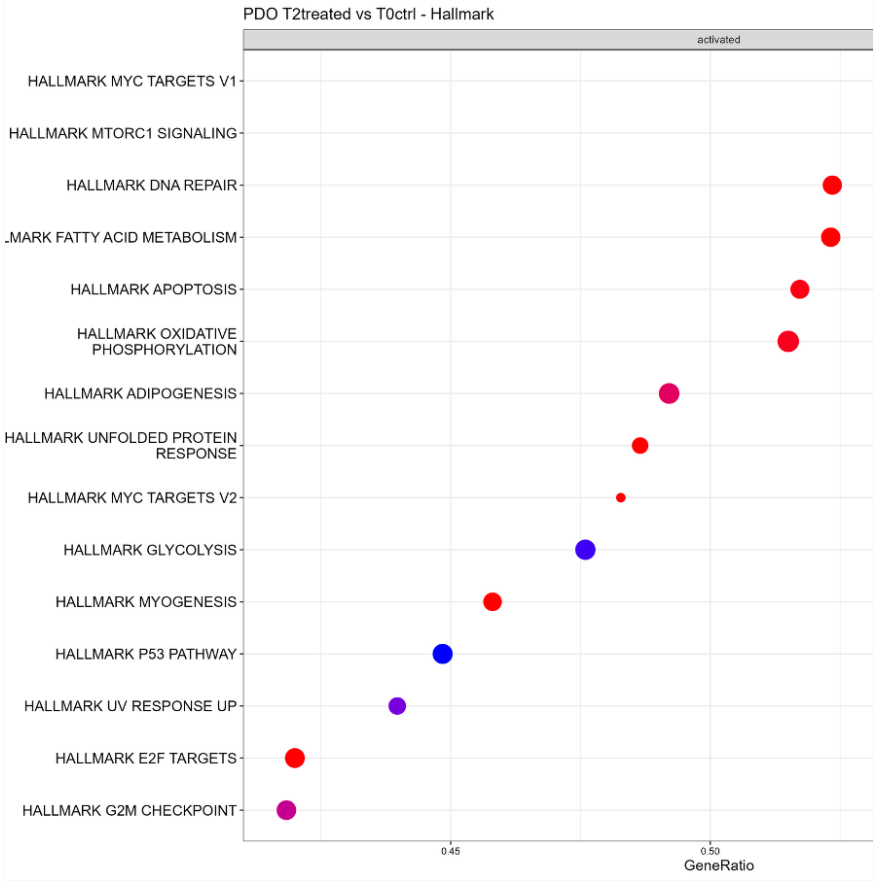


Figure 25. Gene set enrichment analysis (GSEA) of Hallmark pathways in patient-derived organoids (PDOs) at relapse (T2) compared to untreated controls (T0). Dot plot indicates significantly enriched pathways, with dot size representing gene count and color denoting adjusted p-value.

5.4.3 T2 treated vs T1 treated

Differential expression analysis between T2- and T1-treated colorectal cancer organoids identified 293 significantly deregulated transcripts ($|\log FC| \geq 2$, FDR < 0.05), including 113 upregulated and 180 downregulated. Prominent induced genes comprised inflammatory and stress mediators (LCN2, NLRP1), regulators of proliferation and immune signaling (CDCA7, PIGR, RERG), as well as transcripts linked to neuronal-like functions. Conversely, strongly repressed genes included chromatin remodelers (SMARCA2), growth factors (NRG1), epithelial defense molecules (TFF2), and cancer-testis antigens (XAGE1A).

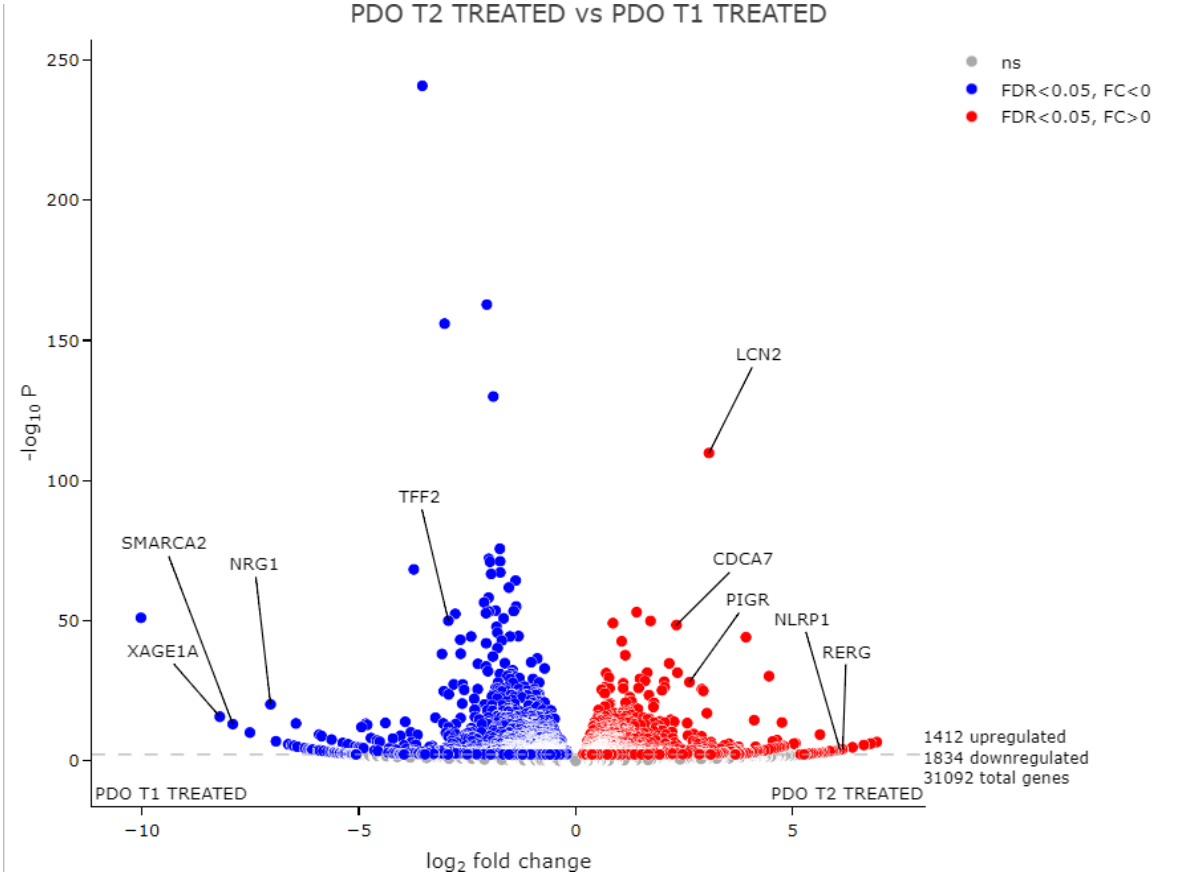


Figure 26. Volcano plot of differential gene expression analysis between patient-derived organoids (PDOs) at relapse time (T2, relapse) and treated (T1, treated).

Gene Ontology Enrichment Analysis revealed a pronounced upregulation of pathways associated with axonal growth, neuronal projection regeneration, and palmitoylation-driven protein modification, indicating a shift toward neurodevelopmental mimicry and post-translational adaptation. Enhanced activity of ion channel regulators and mechanosensory behavior further suggested increased cellular plasticity, environmental sensing, and potential motility. Metabolic rewiring and redox-related functions, including oxidoreductase activity and ubiquitin ligase adaptor activity, reflected an adaptive, stress-resistant phenotype emerging after chemoradiotherapy.

In contrast, downregulated pathways revealed a broad suppression of intestinal epithelial identity, marked by the loss of apical membrane polarity, lipoprotein metabolism, and amino acid transport. Extracellular matrix organization and cell adhesion were markedly reduced, supporting a transition toward a mesenchymal, invasive state. Concurrent downregulation of TGF-β, GPCR, and steroid binding pathways, along with decreased responsiveness to oxidative and toxic stimuli, suggests a shift toward transcriptional autonomy and resistance.

Collectively, these changes highlight the emergence of a dedifferentiated, motile, and therapy-resistant cell population driving tumor relapse.

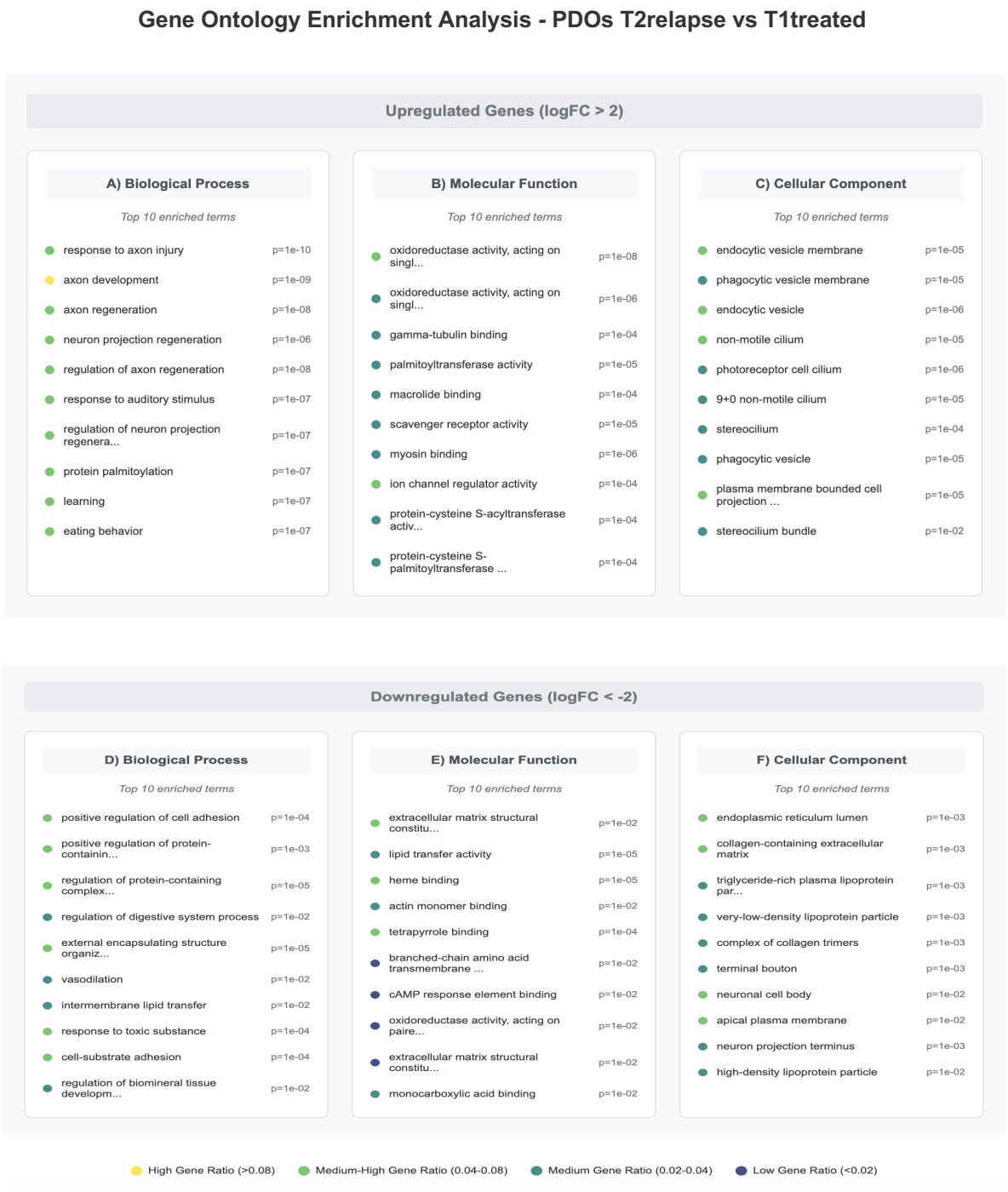
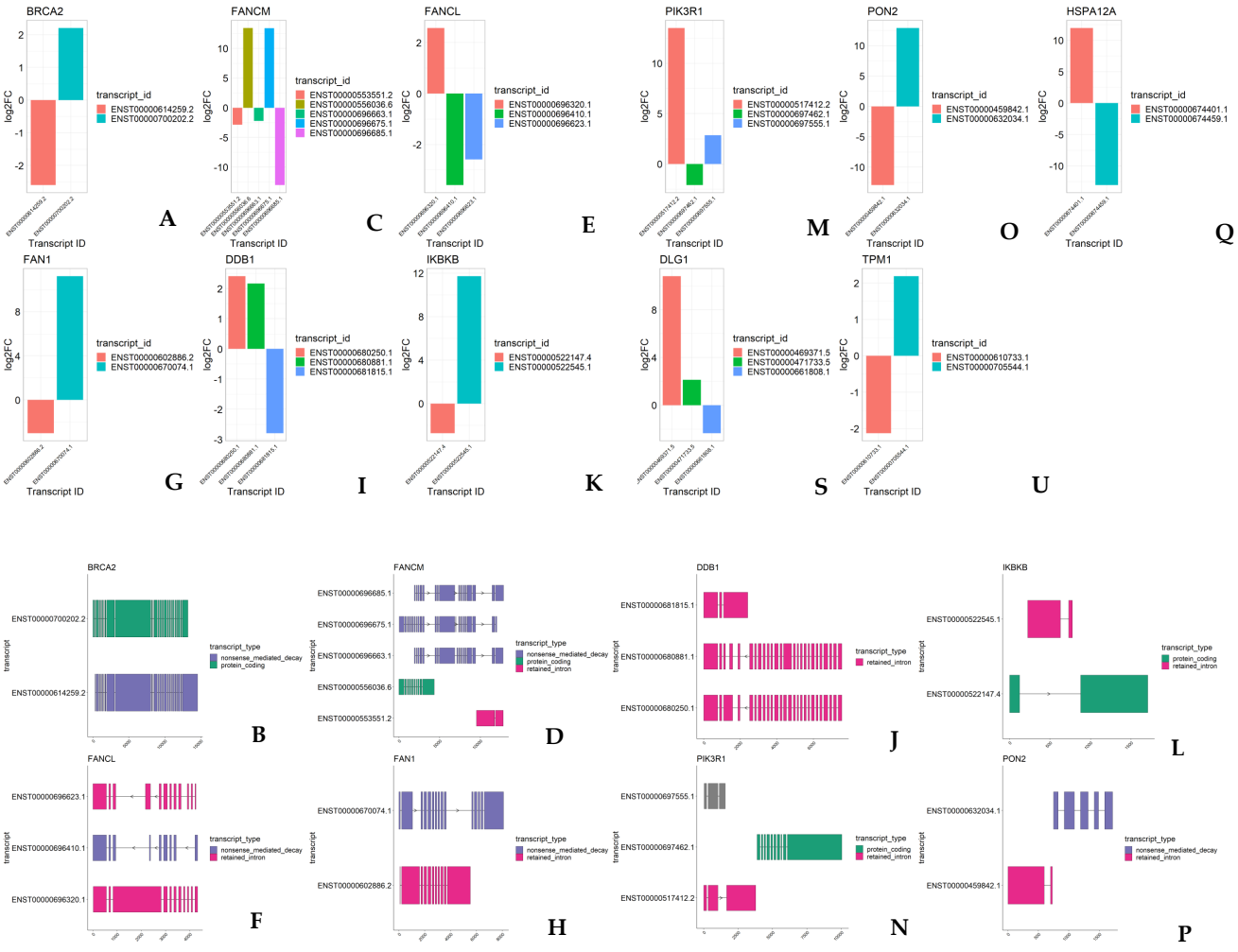


Figure 27. Gene ontology enrichment analysis of differentially expressed genes in PDOs T2 relapse vs T1 treated. Dot color represents gene ratio (number of genes in category/total genes in gene set). Only terms with p-value < 0.05 and top 10 terms per category are shown.

At the level of alternative splicing, T2 was characterized by a selective isoform switch that reinforced these adaptive traits. Functional isoforms of key homologous recombination (HR) genes such as *BRCA2* and *FANCM* were enriched, indicating a reactivation of HR as the dominant mechanism of tolerance to crosslinking and double-strand break damage induced by therapy. In contrast, components of the Fanconi pathway (*FANCL*, *FAN1*) and NER (*DDB1*) were predominantly expressed as non-functional isoforms, reflecting the abandonment of canonical repair routes in favor of HR.

Classical survival pathways such as NF-κB (*IKBKB*) and PI3K/AKT (*PIK3R1*) were silenced through shifts toward defective isoforms, while oxidative stress defense was no longer supported by *PON2* but instead by a more protective isoform of *HSPA12A*. Moreover, tumor suppressor structural genes such as *TPM1* and *DLG1* lost their coding isoforms in favor of non-functional transcripts, promoting cytoskeletal instability, loss of polarity, and an invasive/metastatic phenotype.



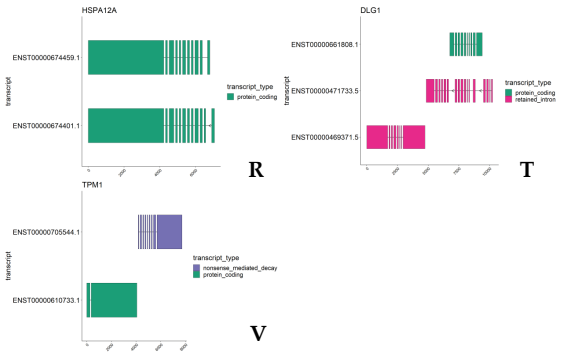


Figure 28. Analysis of alternative splicing events in selected genes in PDOs. Comparison between the relapse condition (T2) and control (T0) highlights differences in isoform usage, suggesting a potential role of the treatment in modulating splicing patterns.

(A) **BRCA**: Isoform ENST00000700202.2 (light blue) is upregulated, while ENST00000614259.2 (orange) is downregulated in relapsed samples compared to treated.

(B) ENST00000700202.2 is protein-coding, whereas ENST00000614259.1 undergoes nonsense-mediated decay.

(C) **FANCM**: Five isoforms are detected, two upregulated ENST00000556036.6 (gold) and ENST00000696675.1 (blue) while three are down regulated (ENST00000553551.2, orange; ENST00000696663.1, green; ENST00000696685.1, pink) in relapsed samples compared to treated.

(D) Transcript types include nonsense-mediated decay (ENST00000696685.1; ENST00000696675.1; ENST00000696663.1), retained-intron (ENST00000553551.2), and protein-coding (ENST00000556036.6).

(E) **FANCL**: three isoforms are detected, one upregulated ENST00000696320.1 (orange) while two are down regulated (ENST00000696410.1, green; ENST00000696623.1, blue) in relapsed samples compared to treated.

(F) ENST00000696410.1 is nonsense-mediated decay, whereas ENST00000696623.1 and ENST00000696320.1 are retained intron

(G) **FAN1**: Isoform ENST00000670074.1 (light blue) is upregulated, while ENST00000602886.2 (orange) is downregulated in relapsed samples compared to treated.

(H) ENST00000670074.1 is nonsense-mediated decay, while ENST00000602886.2 is retained intron

(I) **DDB1**: three isoforms are detected, two upregulated ENST00000680250.1 (orange) and ENST00000680881.1 (green) while one is down regulated (ENST00000681815.1, blue).

(J) All isoforms are retained intron.

(K) **IKBKB**: Isoform ENST00000522545.1 (light blue) is upregulated, while ENST00000522147.4 (orange) is downregulated in relapsed samples compared to treated.

(L) ENST00000522545.1 is retained intron while ENST00000522147.4 is protein coding.

(M) **PIK3R1**: three isoforms are detected, two upregulated ENST00000517412.2 (orange) and ENST00000697555.1 (blue) while one is down regulated (ENST00000697462.1, green) in relapsed samples compared to treated

(N) Two isoforms are protein-coding; however, the exon–intron architecture differs (ENST00000697555.1, ENST00000697462.1) while one ENST00000517412.2 is retained intron.

(O) **PON2**: Isoform ENST00000459842.1 (orange) is down regulated while ENST00000632034.1 (light blue) is upregulated in relapsed samples compared to treated.

(P) ENST00000459842.1 is retained intron, whereas ENST00000632034.1 undergoes nonsense-mediated decay.

(Q) **HSPA12A**: Isoform ENST00000674401.1 (orange) is upregulated, while ENST00000674459.1 (light blue) is downregulated in relapsed samples compared to treated.

(R) Both isoforms are protein-coding; however, the exon–intron architecture differs, as illustrated in the panel

(S) **DLG1**: three isoforms are detected, two upregulated ENST00000469371.5 (orange) and ENST00000471733.5 (green) while one is down regulated (ENST00000661808.1, blue) in relapsed samples compared to treated.

(T) Two isoforms are retained intron (ENST00000471733.5; ENST00000469371.5) while ENST00000661808.1 is protein coding.

(U) **TPM1**: Isoform ENST00000705544.1 (light blue) is upregulated, while ENST00000610733.1 (orange) is downregulated in relapsed samples compared to treated.

(V) ENST00000610733.1 is protein-coding while ENST00000705544.1 is nonsense mediated decay

The Gene Set Enrichment Analysis (GSEA) of cancer Hallmark pathway comparing PDOs at T2 versus T1 revealed a significant activation of transcriptional programs

associated with proliferation and cellular metabolism. Specifically, enrichment was observed for E2F, MYC, and mTORC1 signaling, together with metabolic pathways including oxidative phosphorylation, fatty acid metabolism, and glycolysis. In parallel, activation of pathways related to hypoxia, apoptosis, protein secretion, and unfolded protein response was detected, indicating cellular adaptation to stress conditions. Overall, the enrichment profile suggests that treatment induces a profound functional reprogramming, characterized by enhanced proliferative programs and remodeling of metabolic and stress-response pathways.

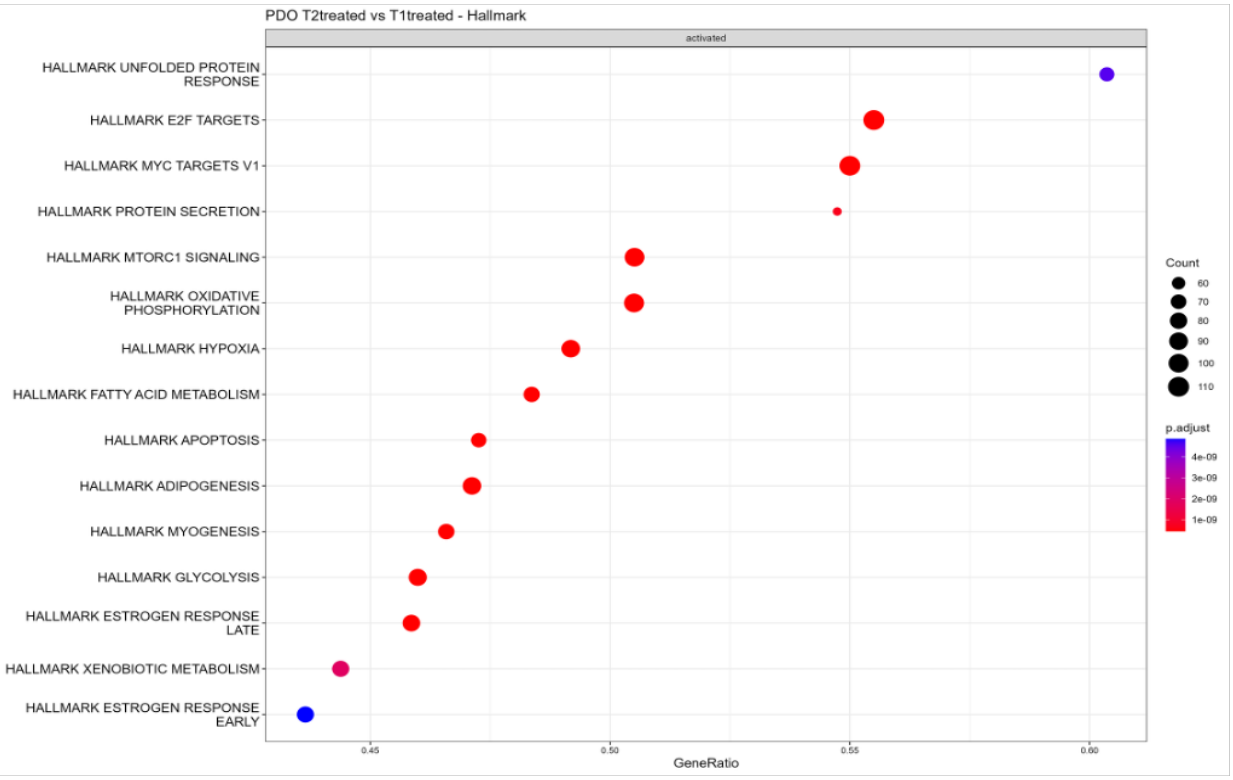


Figure 29. Gene set enrichment analysis (GSEA) of Hallmark pathways in patient-derived organoids (PDOs) at relapse (T2) compared to treated (T1). Bubble plot indicates significantly enriched pathways, with dot size representing gene count and color denoting adjusted p-value.

The three expression analyses were interpreted within a temporal trajectory (T0 → T1 → T2), each representing a distinct layer of PDOs adaptation to pharmacological bottleneck. The T1 vs T0 comparison captures the acute transcriptional response to chemoradiotherapy, characterized by a massive induction of stress- responsiveness, activation of RNA processing and splicing, and repression of protein synthesis and metabolic functions. The T2

vs T0 comparison, instead, represents the cumulative divergence from baseline, highlighting the long-term effects of treatment and adaptation during culture. This condition is marked by persistent activation of inflammatory and DNA damage response programs, cytoskeletal and extracellular matrix remodeling, and concomitant suppression of epithelial adhesion, differentiation cues, and metabolic homeostasis. So, T2 vs T0 does not simply recapitulate temporal drift in culture, but rather reveals durable transcriptional reprogramming driven by therapy, consolidating resistant and plastic phenotypes. Finally, the T2 vs T1 comparison isolates the late transition from the immediate treatment response to relapse, uncovering adaptive traits such as dedifferentiation, neuro-mimetic signaling, and enhanced motility. Altogether, this temporal framing delineates a functional trajectory in which stress response, splicing plasticity, and glycolytic metabolism progressively increase, while epithelial identity, oxidative metabolism, and protein synthesis steadily decline, driving the emergence of a transcriptionally plastic and therapy-resistant state.

5.5 Single-cell analysis

To identify the clusters of individual cell populations, has been integrated the datasets from the different time points (T0, T1, and T2) and conditions (treated and control) using the Seurat pipeline. Based on Gene Ontology (GO) and KEGG enrichment analyses, as well as the expression of established cellular markers, we identified seven distinct clusters.
Cluster 0 (Metabolically inactive) was characterized by the absence of positive discriminative markers; instead, it displayed downregulation of ribosomal and mitochondrial genes. This profile is consistent with a metabolically inactive or quiescent population, in which reduced protein synthesis suggests a state of low metabolic demand. In contrast, Cluster 1 (Proliferating cells) exhibited a strong enrichment of genes involved in cell cycle regulation, DNA replication, and mitotic division, accompanied by ribosomal activation, which clearly defined this cluster as composed of actively proliferating cells. Two clusters were associated with differentiated epithelial phenotypes. Cluster 2 (Enterocytes / Goblet cells) expressed canonical markers of Enterocytes and Goblet cells, with pathway signatures including brush border and apical membrane organization, intestinal absorption, tight junctions, and PPAR signaling, consistent with fully mature and functionally active epithelial cells. Cluster 3 (High mitochondrial metabolism/ stressed cells) was instead distinguished by the overexpression of mitochondrial genes and enrichment in oxidative phosphorylation and electron transport pathways, together with signatures often associated

with cellular stress and neurodegeneration. These features suggest that this population may correspond to cells with hyperactive mitochondrial metabolism and an altered energetic state, potentially reflecting stress-adapted subpopulations. Cluster 4 represented mature intestinal epithelial cells, displaying enrichment in metabolic processes such as lipid digestion, arginine and proline metabolism, and ion homeostasis, further supporting their role as absorptive and differentiated epithelial populations.

Stemness-associated features emerged in Cluster 5 (Stem cell), which displayed activation of Wnt/β-catenin signaling, together with developmental pathways typically linked to progenitor function. This profile strongly supports the interpretation of Cluster 5 as intestinal stem or progenitor epithelial cells, capable of sustaining self-renewal and differentiation. Finally, Cluster 6 showed expression of genes involved in DNA repair, ErbB/EGFR signaling, focal adhesion, and cytoskeletal remodeling, suggesting a population characterized by proliferative signaling combined with active genome maintenance and cellular adhesion dynamics.

Taken together, these data delineate a heterogeneous cellular landscape in which proliferative, differentiated, metabolically stressed, progenitor, and DNA repair–competent populations coexist, reflecting the complexity and plasticity of colorectal cancer organoids.

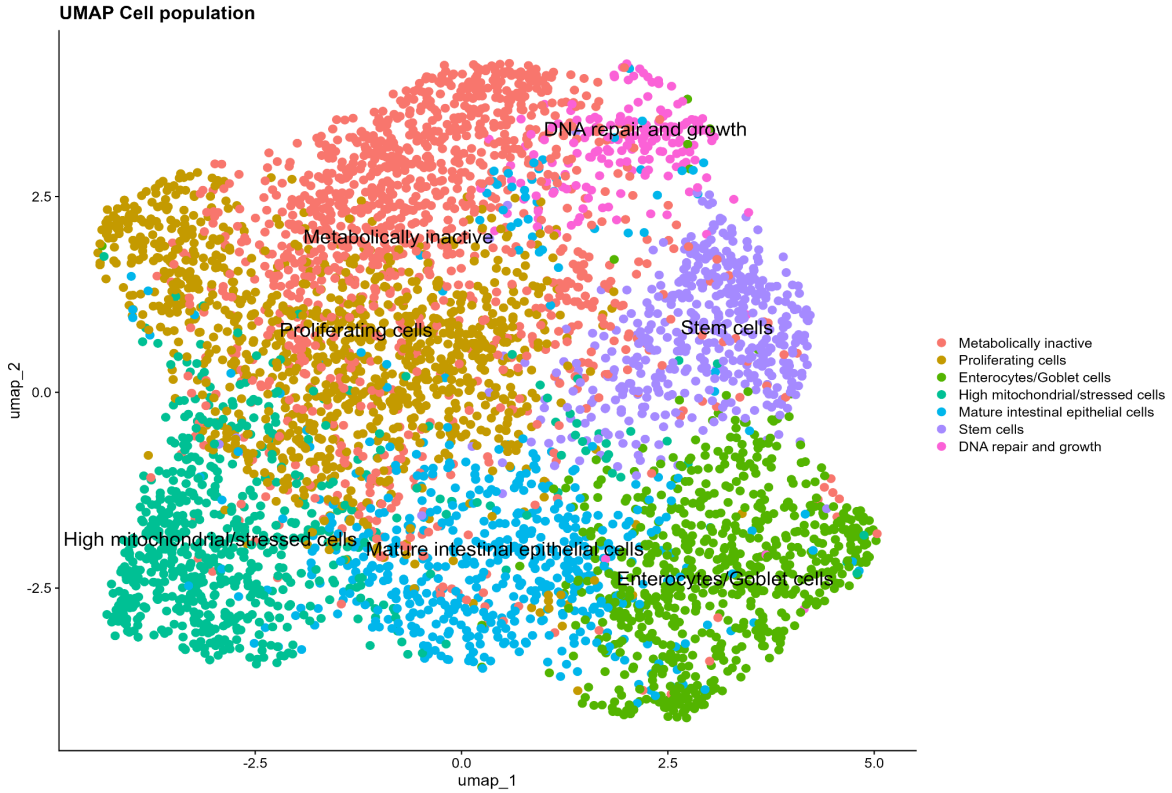
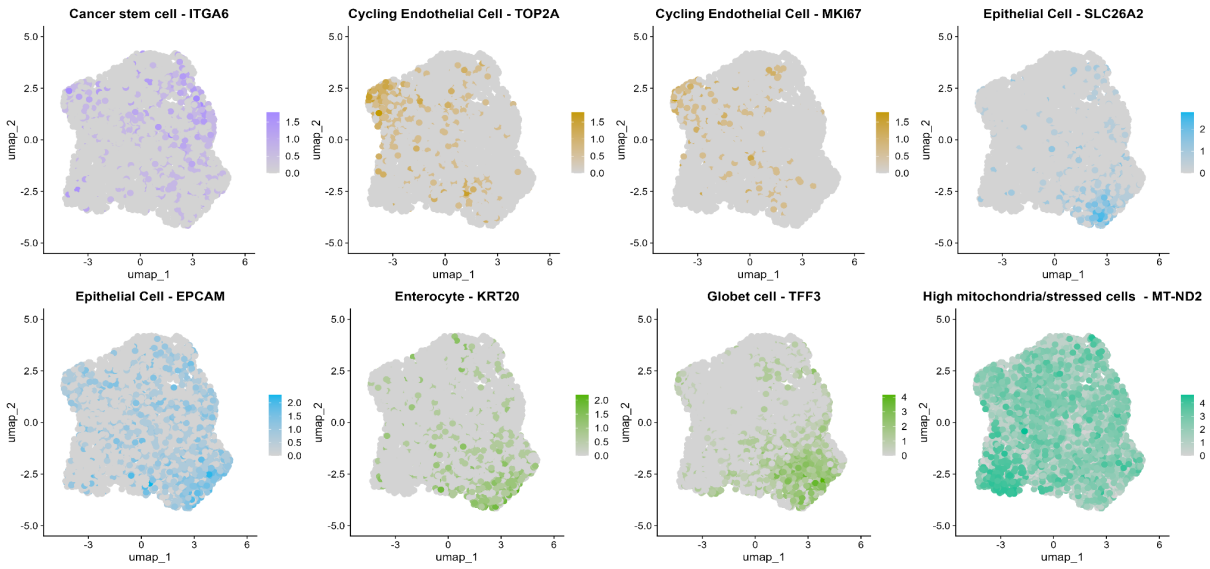


Figure 30. UMAP projection of annotated cell clusters.

Uniform Manifold Approximation and Projection (UMAP) plot displaying transcriptionally distinct clusters, identified as Cluster 0 (Proliferating cells), Cluster 1 (Metabolically inactive cells), Cluster 2 (Enterocytes/Goblet cells), Cluster 3 (High mitochondrial/stressed cells), Cluster 4 (Mature intestinal epithelial cells), Cluster 5 (Stem cells), and Cluster 6 (DNA repair and growth).



Cluster	Markers	Pathway analysis
Cluster 0 - Metabolically inactive	No positive markers, downregulation of ribosomal proteins and mitochondrial genes.	Electron transport chain
Cluster 1 - Proliferating Cells	CENPF; CCNB1; CDKN3; TOP2A; NUSAP1	Cell cycle, DNA replication, Mitotic division
Cluster 2 - Enterocytes/Goblet cells	CEACAM7; SLC11A2; PTPRR; TFF3; SLC26A2	Intestinal absorption and microvillus organization
Cluster 3 - High mitochondrial metabolism/stressed cells	MT-ND2; MT-CYB	Oxidative phosphorylation, mitochondrial transport
Cluster 4 - Mature Intestinal epithelial cells	FABP1; CKB; FXYD3; TFF3; PHGR1	Digestion and absorption
Cluster 5 - Stem Cells	APCDD1; NKD1; DACH1; NKD1	Wnt signaling, β-catenin binding
Cluster 6 - DNA repair active cells	SPIDR; LINC02163; WWOX	ATP- dependent activity acting on DNA and damaged-DNA binding

Figure 31. Marker-based annotation of cell clusters. UMAP feature plots and summary table showing representative marker genes and pathway enrichments defining the major cell clusters: metabolically inactive, proliferating, enterocytes/goblet cells, high mitochondrial/stressed cells, mature intestinal epithelial cells, stem cells, and DNA repair–active cells.

To investigate how the distribution of clusters varied across time points (T0, T1, and T2) and between experimental conditions (control and treated), it has been examined the relative abundance of the seven cellular populations identified by single-cell RNA sequencing.

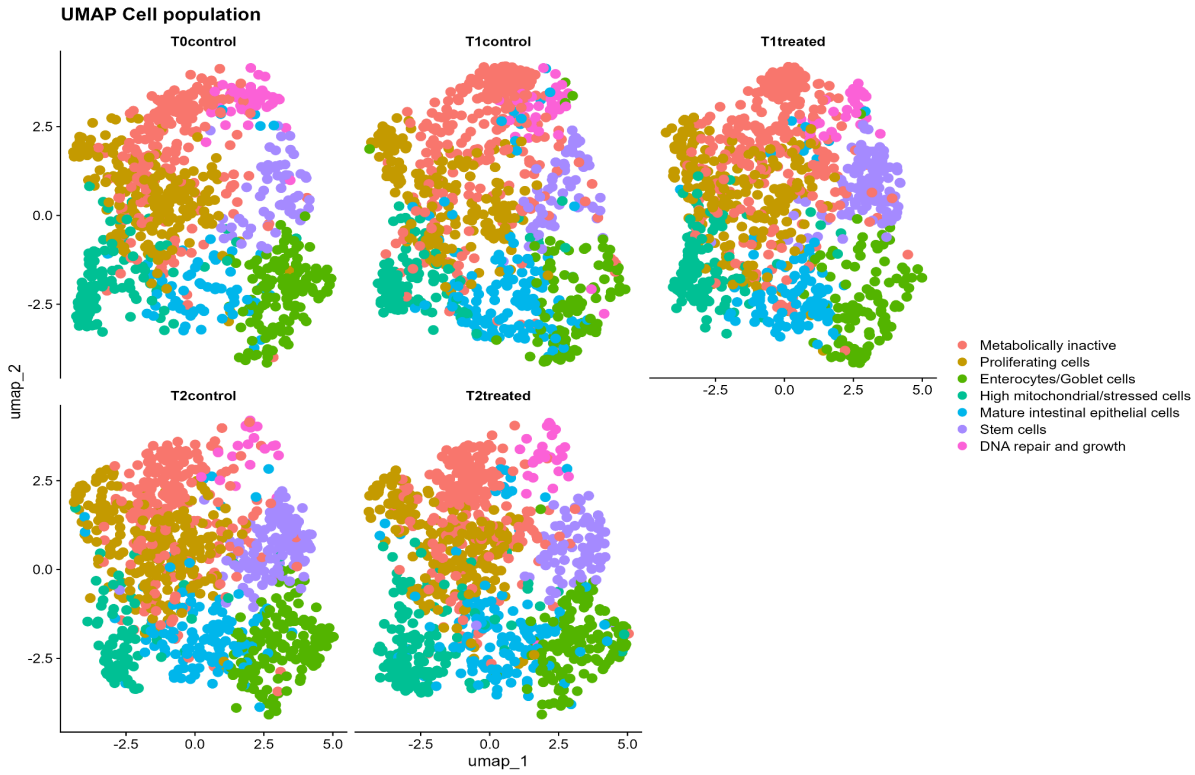


Figure 32. UMAP visualization of cell populations across timepoints and treatment conditions.

UMAP plots showing the distribution of transcriptionally defined cell populations at three timepoints (T0, T1, T2) under control and treated conditions, annotated as metabolically inactive, proliferating, enterocytes/goblet cells, high mitochondrial/stressed cells, mature intestinal epithelial cells, stem cells, and DNA repair–active cells.

To investigate temporal dynamics, we performed a preliminary analysis quantifying the relative abundance of each cellular cluster across the three timepoints (T0, T1, and T2) in both control and treated PDOs.

In the control samples, clusters displayed oscillatory changes across timepoints without a consistent temporal trend, likely reflecting the physiological evolution of organoids in culture.

In contrast, in treated PDOs some clusters did not differ substantially from controls, whereas two notable patterns emerged.

First, between T0 and T1, therapy induced a transient expansion of Cluster 5, corresponding to the stem cell pool. This phenomenon may represent a compensatory response to acute cellular damage, consistent with the early regenerative pressure triggered by combined

radiotherapy and FOLFOX treatment. Second, from T1 to T2, we observed an increase in Cluster 3 (stressed/mitochondrial cells), compatible with persistent oxidative stress and mitochondrial dysfunction, a plausible consequence of DNA damage and ROS generation elicited by both ionizing radiation and oxaliplatin-based chemotherapy.

These observations are still under active investigation, and further analyses will be required to fully elucidate the biological significance and mechanistic basis of these temporal changes.

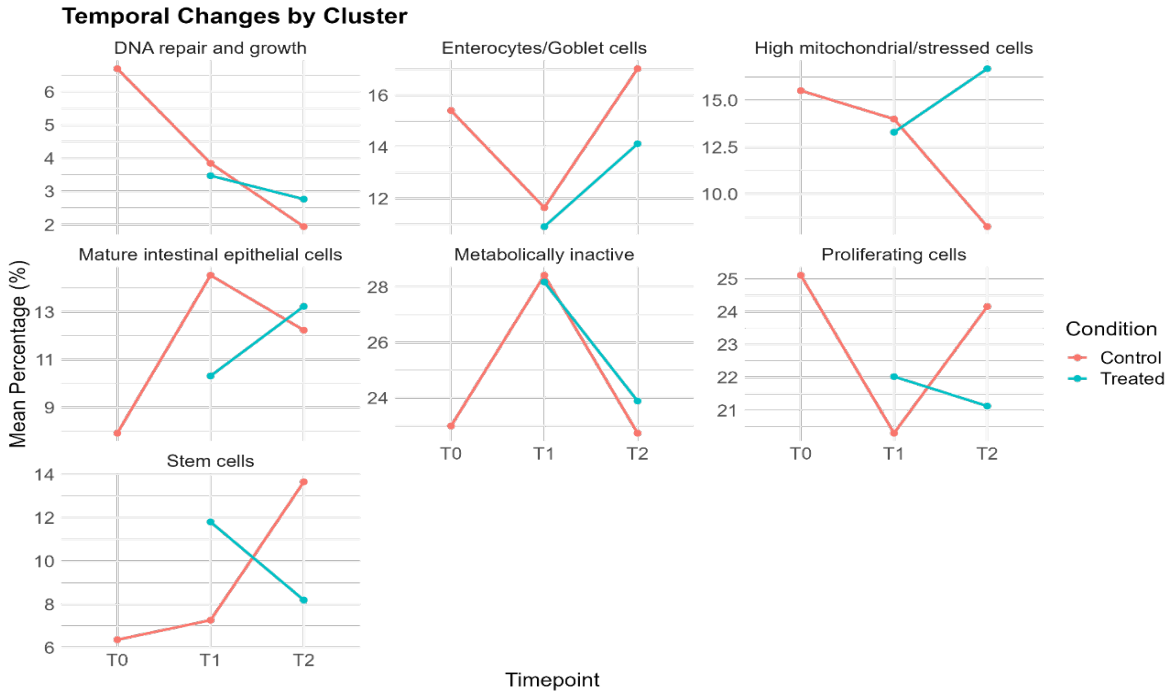


Figure 33. Dynamics of cell population proportions across timepoints and treatment conditions. Line plots showing temporal changes in the relative abundance of transcriptionally defined clusters (DNA repair and growth, enterocytes/goblet cells, high mitochondrial/stressed cells, mature intestinal epithelial cells, metabolically inactive cells, proliferating cells, and stem cells) across three timepoints (T0, T1, T2) under control (red) and treated (blue) conditions.

Chapter 6. Discussion and Conclusion

In this study, patient-derived rectal cancer organoids were employed as a high-fidelity model to investigate the adaptive trajectories of colorectal cancer cells under therapeutic pressure. The model provided a unique opportunity to capture phenotypic transitions during Total Neoadjuvant Therapy (TNT) and to dissect the cellular programs that shape therapy resistance. The results highlight the central role of phenotypic intra-tumor heterogeneity (ITH) in mediating adaptation, demonstrating that resistance emerges not as a static property but through dynamic and multi-layered reprogramming.

Bulk transcriptomic profiling revealed that therapy-exposed organoids undergo profound transcriptional remodeling. After 24 days of treatment (T1), resistant populations activated stress-responsive pathways, including RNA splicing, nucleocytoplasmic transport, and DNA damage repair, while simultaneously downregulating ribosome biogenesis and biosynthetic programs. This dual response suggests that tumor cells cope with treatment-induced bottlenecks by transiently reducing anabolic demand while enhancing transcriptional plasticity and stress tolerance. Isoform-level analyses further supported this interpretation, uncovering widespread alternative splicing events that favored non-functional variants in genes involved in cytoskeletal stability and DNA repair, while modulating functional isoforms of adhesion and signaling genes. Such splicing-driven plasticity likely facilitates survival under selective pressure by diversifying transcriptomic outputs beyond what is captured by canonical expression changes alone.

At relapse stage (T2), the transcriptional landscape no longer resembled baseline but instead consolidated into a therapy-shaped state. Persistent activation of inflammatory signaling, metabolic rewiring with glycolytic enrichment, and loss of epithelial identity defined this condition, together with enrichment of stress-resistance pathways. Isoform switching reinforced these traits, silencing DNA repair and tumor suppressor functions while favoring pro-survival and pro-invasive variants, including stable isoforms of oncogenic lncRNAs. These findings indicate that relapse does not reflect a passive return to pre-treatment equilibrium, but rather the emergence of a transcriptionally plastic and stress-adapted ecosystem, able to sustain long-term resistance.

The temporal framing (T0→T1→T2) delineates a functional trajectory: an acute stress-adaptive phase characterized by transcriptional shutdown of anabolic programs and splicing-driven diversification; a consolidation phase with chronic inflammatory signaling and metabolic reprogramming; and a relapse phase marked by dedifferentiation and enhanced motility. This progression underscores the dynamic nature of adaptation, in which transcriptional memory and isoform remodeling progressively reshape cellular identity.

Single-cell transcriptomic analysis provided a complementary layer, directly resolving the cellular heterogeneity underlying these adaptations. In control PDOs, the relative abundance of clusters fluctuated across timepoints without converging toward a coherent trajectory, reflecting the physiological evolution of organoids in culture. In contrast, therapy-exposed PDOs exhibited distinct stage-specific dynamics. Between T0 and T1, treatment induced a transient expansion of the stem cell cluster, consistent with an early regenerative response to acute therapy-induced damage. From T1 to T2, this stem-like compartment persisted, while a progressive accumulation of metabolically stressed/mitochondrial cells became evident, in line with sustained oxidative stress and therapy-related DNA damage. Taken together, these preliminary findings suggest that therapy reshapes the cellular landscape by transiently mobilizing progenitor pools and progressively enriching stress-adapted subpopulations, thereby maintaining a reservoir of cells with regenerative and survival potential.

Overall rectal cancer organoids exposed to TNT recapitulate key features of clinical resistance, revealing that therapy acts as an evolutionary bottleneck: it reduces cellular diversity but enriches for quiescent, stem-like, and stress-adapted populations that sustain recurrence.

By integrating long-reads bulk and single-cell transcriptomic approaches, this work highlights resistance as a multi-scale adaptive process, where transcriptional reprogramming and alternative splicing converge to remodel phenotypic landscapes. The integration of multiple omics technologies, including single-cell, RNA sequencing in long read and whole genome sequencing, represents a key strength of this study, providing unprecedented resolution and comprehensive insights into the phenotypic dynamics of rectal cancer from patient derived organoids.

These results underscore the critical role of phenotypic ITH in shaping therapeutic outcomes. Even in the absence of new genomic alterations, transcriptional plasticity and isoform diversity enable tumor cells to explore alternative adaptive trajectories, ultimately fostering resistance. This supports a paradigm shift in which therapy failure in CRC is understood not only as the product of genetic selection but as the outcome of dynamic phenotypic reprogramming. Clinically, this suggests that targeting adaptive programs—such as splicing machinery, metabolic rewiring, or inflammatory signaling—may provide novel opportunities to disrupt resistant trajectories and improve the efficacy of standard treatments.

Nevertheless, several limitations of the organoid model must be acknowledged. Organoids lack the native tumor microenvironment, including stromal, immune, and vascular compartments that critically influence therapy response. The absence of such interactions may underestimate immune clearance or stromal-driven protection. Furthermore, long-term culture may impose artificial selective pressures, favoring subsets adapted to in vitro conditions. Finally, single-cell analysis was restricted to epithelial components, limiting the assessment of non-epithelial contributors to ITH.

In conclusion, this study provides a comprehensive view of how colorectal cancer organoids adapt to TNT, revealing a trajectory of stress-driven, splicing-mediated, and population-level reprogramming that fuels recurrence. By uncovering the central role of phenotypic heterogeneity and plasticity, these findings highlight potential therapeutic vulnerabilities and emphasize the need for integrative approaches to intercept adaptive resistance.

Bibliography

1. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2024; 74(3): 229-263. doi:10.3322/caac.21834.
2. Linee Guida Aiom- Neoplasie dei Retto e Ano 2021.
3. Yang, T., Li, X., Montazeri, Z., Little, J., Farrington, S. M., Ioannidis, J. P. A., et al. (2019). Gene-environment Interactions and Colorectal Cancer Risk: An Umbrella Review of Systematic Reviews and Meta-analyses of Observational Studies. Int. J. Cancer 145, 2315–2329. doi:10.1002/ijc.32057.
4. Nguyen, L. H., Liu, P.-H., Zheng, X., Keum, N., Zong, X., Li, X., et al. (2018). Sedentary Behaviors, Tv Viewing Time, and Risk of Young-Onset Colorectal Cancer. JNCI Cancer Spectr. 2, pky073. doi:10.1093/jncics/pky073.
5. Yang C., Xiao W, Wang R, Hu Y, Yi K, Sun X, Wang G, Xu X. Tumor organoid model of colorectal cancer (Review). Oncol Lett. 2023 Jun 15;26(2):328. doi: 10.3892/ol.2023.13914.
6. Rectal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up Glynne-Jones, R. et al. Annals of Oncology, Volume 28, iv22 - iv40.
7. Hu, X.; Xue, Z.; He, K.; Tian, Y.; Chen, Y.; Zhao, M.; Yu, J.; Yue, J. Strategies to Optimize Treatment for Locally Advanced Rectal Cancer. Cancers 2023, 15, 219. doi:10.3390/cancers15010219.
8. Gollins S, Sebag-Montefiore D. Neoadjuvant Treatment Strategies for Locally Advanced Rectal Cancer. Clin Oncol (R Coll Radiol). 2016 Feb;28(2):146-151. doi: 10.1016/j.clon.2015.11.003.

9. Kaur H, Choi H, You YN, Rauch GM, Jensen CT, Hou P, Chang GJ, Skibber JM, Ernst RD. MR imaging for preoperative evaluation of primary rectal cancer: practical considerations. Radiographics. 2012 Mar-Apr;32(2):389-409. doi: 10.1148/rg.322115122.

10. Cedermark, B.; Dahlberg, M.; Glimelius, B.; Påhlman, L.; Rutqvist, L.E.; Wilking, N. Improved survival with preoperative radiotherapy in resectable rectal cancer. New Engl. J. Med. 1997, 336, 980–987. doi: 10.1056/NEJM199704033361402.

11. Gandini, A.; Sciallero, S.; Martelli, V.; Pirrone, C.; Puglisi, S.; Cremante, M.; Grassi, M.; Andretta, V.; Fornarini, G.; Caprioni, F.; et al. A Comprehensive Approach to Neoadjuvant Treatment of Locally Advanced Rectal Cancer. Cancers 2025, 17, 330. doi:10.3390/cancers17020330.

12. Barker HE, Paget JT, Khan AA, Harrington KJ. The tumour microenvironment after radiotherapy: mechanisms of resistance and recurrence. Nat Rev Cancer. 2015 Aug;15(8):509-25. doi: 10.1038/nrc3958.

13. Fisher, R., Pusztai, L. & Swanton, C. Cancer heterogeneity: implications for targeted therapeutics. Br J Cancer 108, 479–485 (2013). doi: 10.1038/bjc.2012.581.

14. Dagogo-Jack I, Shaw AT. Tumour heterogeneity and resistance to cancer therapies. Nat Rev Clin Oncol. 2018 Feb;15(2):81-94. doi: 10.1038/nrclinonc.

15. Meacham, C., Morrison, S. Tumour heterogeneity and cancer cell plasticity. Nature 501, 328–337 (2013). doi: 10.1038/nature12624.

16. Lawson, D.A., Kessenbrock, K., Davis, R.T. et al. Tumour heterogeneity and metastasis at single-cell resolution. Nat Cell Biol 20, 1349–1360 (2018). doi: 10.1038/s41556-018-0236-7.

17. Marusyk A, Janiszewska M, Polyak K. Intratumor Heterogeneity: The Rosetta Stone of Therapy Resistance. Cancer Cell. 2020 Apr 13;37(4): 471-484.doi: 10.1016/j.ccell.2020.03.007.

18. Househam, J., Heide, T., Cresswell, G.D. et al. Phenotypic plasticity and genetic control in colorectal cancer evolution. Nature 611, 744–753 (2022). doi: 10.1038/s41586-022-05311-x.

19. Ben-David, U., Ha, G., Tseng, YY. et al. Patient-derived xenografts undergo mouse-specific tumor evolution. Nat Genet 49, 1567–1575 (2017). doi:10.1038/ng.3967.

20. Drost, J., Clevers, H. Organoids in cancer research. Nat Rev Cancer 18, 407–418 (2018). doi:10.1038/s41568-018-0007-6.

21. Barbáchano A, Fernández-Barral A, Bustamante-Madrid P, Prieto I, Rodríguez-Salas N, Larriba MJ, Muñoz A. Organoids and Colorectal Cancer. Cancers (Basel). 2021 May 28;13(11):2657. doi: 10.3390/cancers13112657.

22. Sato T, Vries RG, Snippert HJ, van de Wetering M, Barker N, Stange DE, van Es JH, Abo A, Kujala P, Peters PJ, Clevers H. Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. Nature. 2009 May 14;459(7244):262-5. doi: 10.1038/nature07935.

23. DeHaan R.K., Sarvestani S.K., Huang E.H. Organoid Models of Colorectal Pathology: Do They Hold the Key to Personalized Medicine? A Systematic Review. Dis. Colon Rectum. 2020;63:1559–1569. doi: 10.1097/DCR.0000000000001806.

24. Ganesh K., Wu C., O'Rourke K.P., Szeglin B.C., Zheng Y., Sauve C.G., Adileh M., Wasserman I., Marco M.R., Kim A.S., et al. A rectal cancer organoid platform to study individual responses to chemoradiation. Nat. Med. 2019;25:1607–1614. doi: 10.1038/s41591-019-0584-2.

25. WATSON, J., CRICK, F. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. Nature 171, 737–738 (1953). doi:10.1038/171737a0.

26. Hong, M., Tao, S., Zhang, L. et al. RNA sequencing: new technologies and applications in cancer research. J Hematol Oncol 13, 166 (2020). doi: 10.1186/s13045-020-01005-x.

27. Ozsolak, F., Milos, P. RNA sequencing: advances, challenges and opportunities. Nat Rev Genet 12, 87–98 (2011). doi:10.1038/nrg2934.

28. Lockhart DJ, Dong H, Byrne MC, Follettie MT, Gallo MV, Chee MS, Mittmann M, Wang C, Kobayashi M, Horton H, Brown EL. Expression monitoring by hybridization to high-density oligonucleotide arrays. Nat Biotechnol. 1996 Dec;14(13):1675-80. doi: 10.1038/nbt1296-1675.

29. Wodicka L, Dong H, Mittmann M, Ho MH, Lockhart DJ. Genome-wide expression monitoring in Saccharomyces cerevisiae. Nat Biotechnol. 1997 Dec;15(13):1359-67. doi: 10.1038/nbt1297-1359.

30. Harkin DP. Uncovering functionally relevant signaling pathways using microarray-based expression profiling. Oncologist. 2000;5(6):501-7. doi: 10.1634/theoncologist.5-6-501.

31. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A. 1977 Dec;74(12):5463-7. doi: 10.1073/pnas.74.12.5463.

32. Maxam AM, Gilbert W. A new method for sequencing DNA. Proc Natl Acad Sci U S A. 1977 Feb;74(2):560-4. doi: 10.1073/pnas.74.2.560.

33. Margulies, M., Egholm, M., Altman, W. et al. Genome sequencing in microfabricated high-density picolitre reactors. Nature 437, 376–380 (2005). doi: 10.1038/nature03959.

34. Bentley, D., Balasubramanian, S., Swerdlow, H. et al. Accurate whole human genome sequencing using reversible terminator chemistry. Nature 456, 53–59 (2008). doi: 10.1038/nature07517.

35. Rothberg, J., Hinz, W., Rearick, T. et al. An integrated semiconductor device enabling non-optical genome sequencing. Nature 475, 348–352 (2011). doi: 10.1038/nature10242.

36. Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y. RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. Genome Res. 2008 Sep;18(9):1509-17. doi: 10.1101/gr.079558.108.

37. Hu T, Chitnis N, Monos D, Dinh A. Next-generation sequencing technologies: An overview. Hum Immunol. 2021 Nov;82(11):801-811. doi: 10.1016/j.humimm.2021.02.012.

38. Tilgner H, Jahanbani F, Blauwkamp T, Moshrefi A, Jaeger E, Chen F, Harel I, Bustamante CD, Rasmussen M, Snyder MP. Comprehensive transcriptome analysis using synthetic long-read sequencing reveals molecular co-association of distant splicing events. Nat Biotechnol. 2015 Jul;33(7):736-42. doi: 10.1038/nbt.3242. Epub 2015 May 18.

39. Park E, Pan Z, Zhang Z, Lin L, Xing Y. The Expanding Landscape of Alternative Splicing Variation in Human Populations. Am J Hum Genet. 2018 Jan 4;102(1):11-26. doi: 10.1016/j.ajhg.2017.11.002.

40. Monzó, C., Liu, T. & Conesa, A. Transcriptomics in the era of long-read sequencing. Nat Rev Genet (2025). doi: 10.1038/s41576-025-00828-z.

41. Supplitt S, Karpinski P, Sasiadek M, Laczmanska I. Current Achievements and Applications of Transcriptomics in Personalized Cancer Medicine. Int J Mol Sci. 2021 Jan 31;22(3):1422. doi: 10.3390/ijms22031422.

42. Schadt EE, Turner S, Kasarskis A. A window into third-generation sequencing. Hum Mol Genet. 2010 Oct 15;19(R2):R227-40. doi: 10.1093/hmg/ddq416.

43. Eid J, Fehr A, Gray J, et al. Real-time DNA sequencing from single polymerase molecules. Science. 2009 Jan 2;323(5910):133-8. doi: 10.1126/science.1162986.

44. Jain, M., Olsen, H.E., Paten, B. et al. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. Genome Biol 17, 239 (2016). doi: 10.1186/s13059-016-1103-0.

45. Ardui S, Ameur A, Vermeesch JR, Hestand MS. Single molecule real-time (SMRT) sequencing comes of age: applications and utilities for medical diagnostics. Nucleic Acids Res. 2018 Mar 16;46(5):2159-2168. doi: 10.1093/nar/gky066.

46. Deamer D, Akeson M, Branton D. Three decades of nanopore sequencing. Nat Biotechnol. 2016 May 6;34(5):518-24. doi: 10.1038/nbt.3423.

47. Wang Y, Zhao Y, Bollas A, Wang Y, Au KF. Nanopore sequencing technology, bioinformatics and applications. Nat Biotechnol. 2021 Nov;39(11):1348-1365. doi: 10.1038/s41587-021-01108-x.

48. Bezrukov SM, Kasianowicz JJ. Current noise reveals protonation kinetics and number of ionizable sites in an open protein ion channel. Phys Rev Lett. 1993 Apr 12;70(15):2352-2355. doi: 10.1103/PhysRevLett.70.2352.

49. Walker B, Kasianowicz J, Krishnasastry M, Bayley H. A pore-forming protein with a metal-actuated switch. Protein Eng. 1994 May;7(5):655-62. doi: 10.1093/protein/7.5.655.

50. Stoddart D, Heron AJ, Mikhailova E, Maglia G, Bayley H. Single-nucleotide discrimination in immobilized DNA oligonucleotides with a biological nanopore. Proc Natl Acad Sci U S A. 2009 May 12;106(19):7702-7. doi: 10.1073/pnas.0901054106.

51. Butler TZ, Pavlenok M, Derrington IM, Niederweis M, Gundlach JH. Single-molecule DNA detection with an engineered MspA protein nanopore. Proc Natl Acad Sci U S A. 2008 Dec 30;105(52):20647-52. doi: 10.1073/pnas.0807514106.

52. Derrington IM, Butler TZ, Collins MD, Manrao E, Pavlenok M, Niederweis M, Gundlach JH. Nanopore DNA sequencing with MspA. Proc Natl Acad Sci U S A. 2010 Sep 14;107(37):16060-5. doi: 10.1073/pnas.1001831107.

53. Lieberman KR, Cherf GM, Doody MJ, Olasagasti F, Kolodji Y, Akeson M. Processive replication of single DNA molecules in a nanopore catalyzed by phi29 DNA polymerase. J Am Chem Soc. 2010 Dec 22;132(50):17961-72. doi: 10.1021/ja1087612.

54. Cherf GM, Lieberman KR, Rashid H, Lam CE, Karplus K, Akeson M. Automated forward and reverse ratcheting of DNA in a nanopore at 5-Å precision. Nat Biotechnol. 2012 Feb 14;30(4):344-8. doi: 10.1038/nbt.2147.

55. Chen P, Sun Z, Wang J, Liu X, Bai Y, Chen J, Liu A, Qiao F, Chen Y, Yuan C, Sha J, Zhang J, Xu LQ, Li J. Portable nanopore-sequencing technology: Trends in development and applications. Front Microbiol. 2023 Feb 1;14:1043967. doi: 10.3389/fmicb.2023.1043967.

56. Magi A., Semeraro R., Mingrino A., Giusti B., D'Aurizio R. Nanopore sequencing data analysis: State of the art, applications and challenges. Brief. Bioinform. 2018;19(6):1256–1272. doi: 10.1093/bib/bbx062.

57. Wei G., Hu R., Li Q., Lu W., Liang H., Nan H., Lu J., Li J., Zhao Q. Oligonucleotide discrimination enabled by tannic acid- coordinated film-coated solid-state nanopores. Langmuir. 2022;38(20):6443–6453. doi: 10.1021/acs.langmuir.2c00638.

58. Asandei A., Mereuta L., Bucataru I.C., Park Y., Luchian T. A single-molecule insight into the ionic strength-dependent, cationic peptide nucleic acids-oligonucleotides interactions. Chem. Asian J. 2022;17(12):e202200261. doi: 10.1002/asia.202200261.

59. de Lannoy C., de Ridder D., Risse J. The long reads ahead: De novo genome assembly using the MinION. F1000 Res. 2017;6:1083. doi: 10.12688/f1000research.12012.2.

60. Levkova M, Chervenkov T, Angelova L, Dzenkov D. Oxford Nanopore Technology and its Application in Liquid Biopsies. Curr Genomics. 2023 Dec 28;24(6):337-344. doi: 10.2174/0113892029286632231127055733.

61. Sharon, D., Tilgner, H., Grubert, F. et al. A single-molecule long-read survey of the human transcriptome. Nat Biotechnol 31, 1009–1014 (2013). doi:10.1038/nbt.2705.

62. Au KF, Sebastiano V, Afshar PT, Durruthy JD, Lee L, Williams BA, van Bakel H, Schadt EE, Reijo-Pera RA, Underwood JG, Wong WH. Characterization of the human ESC transcriptome by hybrid sequencing. Proc Natl Acad Sci U S A. 2013 Dec 10;110(50):E4821-30. doi: 10.1073/pnas.1320101110.

63. Garalde, D., Snell, E., Jachimowicz, D. et al. Highly parallel direct RNA sequencing on an array of nanopores. Nat Methods 15, 201–206 (2018). doi: 10.1038/nmeth.4577.

64. Soneson C, Yao Y, Bratus-Neuenschwander A, Patrignani A, Robinson MD, Hussain S. A comprehensive examination of Nanopore native RNA sequencing for characterization of complex transcriptomes. Nat Commun. 2019 Jul 31;10(1):3359. doi: 10.1038/s41467-019-11272-z.

65. van Dijk EL, Jaszczyszyn Y, Naquin D, Thermes C. The Third Revolution in Sequencing Technology. Trends Genet. 2018 Sep;34(9):666-681. doi: 10.1016/j.tig.2018.05.008.

66. Athanasopoulou K, Boti MA, Adamopoulos PG, Skourou PC, Scorilas A. Third-Generation Sequencing: The Spearhead towards the Radical Transformation of Modern Genomics. Life (Basel). 2021 Dec 26;12(1):30. doi: 10.3390/life12010030.

67. Rang FJ, Kloosterman WP, de Ridder J. From squiggle to basepair: computational approaches for improving nanopore sequencing read accuracy. Genome Biol. 2018 Jul 13;19(1):90. doi: 10.1186/s13059-018-1462-9.

68. Workman, R.E., Tang, A.D., Tang, P.S. et al. Nanopore native RNA sequencing of a human poly(A) transcriptome. Nat Methods 16, 1297–1305 (2019). doi: 10.1038/s41592-019-0617-2.

69. Alberto Magi, Roberto Semeraro, Alessandra Mingrino, Betti Giusti, Romina D'Aurizio, Nanopore sequencing data analysis: state of the art, applications and challenges, Briefings in Bioinformatics, Volume 19, Issue 6, November 2018, Pages 1256–1272, doi: 10.1093/bib/bbx062.

70. https://nanoporetech.com/document/experiment-companion-minknow.

71. https://github.com/epi2me-labs.

72. Di Tommaso, P., Chatzou, M., Floden, E. et al. Nextflow enables reproducible computational workflows. Nat Biotechnol 35, 316–319 (2017). doi: 10.1038/nbt.3820.

73. Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics. 2018 Sep 15;34(18):3094-3100. doi: 10.1093/bioinformatics/bty191.

74. Kovaka S, Zimin AV, Pertea GM, Razaghi R, Salzberg SL, Pertea M. Transcriptome assembly from long-read RNA-seq alignments with StringTie2. Genome Biol. 2019 Dec 16;20(1):278. doi: 10.1186/s13059-019-1910-1.

75. Pertea G, Pertea M. GFF Utilities: GffRead and GffCompare. F1000Res. 2020 Apr 28;9:ISCB Comm J-304. doi: 10.12688/f1000research.23297.2.

76. Docker, Inc. (2024). Docker Documentation. https://docs.docker.com/.

77. https://docs.sylabs.io/guides/3.9/user-guide/environment_and_metadata.html.

78. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010 Jan 1;26(1):139-40. doi: 10.1093/bioinformatics/btp616.

79. Anders S, Reyes A, Huber W. Detecting differential usage of exons from RNA-seq data. Genome Res. 2012 Oct;22(10):2008-17. doi: 10.1101/gr.133744.111.

80. Nowicka M, Robinson MD. DRIMSeq: a Dirichlet-multinomial framework for multivariate count outcomes in genomics. F1000Res. 2016 Jun 13;5:1356. doi: 10.12688/f1000research.8900.2.

81. Van den Berge K, Soneson C, Robinson MD, Clement L. stageR: a general stage-wise method for controlling the gene-level false discovery rate in differential expression and differential transcript usage. Genome Biol. 2017 Aug 7;18(1):151. doi: 10.1186/s13059-017-1277-0.

82. Zheng CL, Fu XD, Gribskov M. Characteristics and regulatory elements defining constitutive splicing and different modes of alternative splicing in human and mouse. RNA. 2005 Dec;11(12):1777-87. doi: 10.1261/rna.2660805.

83. Pan, Q., Shai, O., Lee, L. et al. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat Genet 40, 1413–1415 (2008). doi: 10.1038/ng.259.

84. UniProt Consortium. UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Res. 2021 Jan 8;49(D1):D480-D489. doi: 10.1093/nar/gkaa1100.

85. Berget SM, Moore C, Sharp PA. Spliced segments at the 5’ terminus of adenovirus 2 late mRNA. Proc Natl Acad Sci U S A. 1977 Aug;74(8):3171-5. doi: 10.1073/pnas.74.8.3171.

86. Wahl MC, Will CL, Lührmann R. The spliceosome: design principles of a dynamic RNP machine. Cell. 2009 Feb 20;136(4):701-18. doi: 10.1016/j.cell.2009.02.009.

87. Bonnal SC, López-Oreja I, Valcárcel J. Roles and mechanisms of alternative splicing in cancer - implications for care. Nat Rev Clin Oncol. 2020 Aug;17(8):457-474. doi: 10.1038/s41571-020-0350-x.

88. Matera AG, Wang Z. A day in the life of the spliceosome. Nat Rev Mol Cell Biol.2014 Apr;15(4):294. doi: 10.1038/nrm3742.

89. Sciarrillo R, Wojtuszkiewicz A, Assaraf YG, Jansen G, Kaspers GJL, Giovannetti E, Cloos J. The role of alternative splicing in cancer: From oncogenesis to drug resistance. Drug Resist Updat. 2020 Dec;53:100728. doi: 10.1016/j.drup.2020.100728.

90. Anna A, Monika G. Splicing mutations in human genetic disorders: examples, detection, and confirmation.J Appl Genet. 2019 May;60(2):231. doi: 10.1007/s13353-019-00493-z.

91. Hatje K, Rahman RU, Vidal RO, Simm D, Hammesfahr B, Bansal V, Rajput A, Mickael ME, Sun T, Bonn S, Kollmar M. The landscape of human mutually exclusive splicing. Mol Syst Biol. 2017 Dec 14;13(12):959. doi: 10.15252/msb.20177728.

92. Gallego-Paez LM, Bordone MC, Leote AC, Saraiva-Agostinho N, Ascensão-Ferreira M, Barbosa-Morais NL. Alternative splicing: the pledge, the turn, and the prestige : The key role of alternative splicing in human biological systems. Hum Genet. 2017 Sep;136(9):1015-1042. doi: 10.1007/s00439-017-1790-y.

93. Wang R, Helbig I, Edmondson AC, Lin L, Xing Y. Splicing defects in rare diseases: transcriptomics and machine learning strategies towards genetic diagnosis. Brief Bioinform. 2023 Sep 20;24(5):bbad284. doi: 10.1093/bib/bbad284.

94. Urbanski LM, Leclair N, Anczuków O. Alternative-splicing defects in cancer: Splicing regulators and their downstream targets, guiding the way to novel cancer therapeutics. Wiley Interdiscip Rev RNA. 2018 Jul;9(4):e1476. doi: 10.1002/wrna.1476.

95. Dvinge H, Kim E, Abdel-Wahab O, Bradley RK. RNA splicing factors as oncoproteins and tumour suppressors. Nat Rev Cancer. 2016 Jul;16(7):413-30. doi: 10.1038/nrc.2016.51.

96. Stanley RF, Abdel-Wahab O. Dysregulation and therapeutic targeting of RNA splicing in cancer. Nat Cancer. 2022 May;3(5):536-546. doi: 10.1038/s43018-022-00384-z.

97. Capon DJ, Seeburg PH, McGrath JP, Hayflick JS, Edman U, Levinson AD, Goeddel DV. Activation of Ki-ras2 gene in human colon and lung carcinomas by two different point mutations. Nature. 1983 Aug 11-17;304(5926):507-13. doi: 10.1038/304507a0.

98. Gonçalves V, Henriques AF, Pereira JF, Neves Costa A, Moyer MP, Moita LF, Gama-Carvalho M, Matos P, Jordan P. Phosphorylation of SRSF1 by SRPK1 regulates alternative splicing of tumor-related Rac1b in colorectal cells. RNA. 2024 Mar 18;30(4):463. doi: 10.1261/rna.079955.123.

99. Flodrops M, Dujardin G, Busson A, Trouvé P, Ka C, Simon B, Arzur D, Le Jossic-Corcos C, Corcos L. TIMP1 intron 3 retention is a marker of colon cancer progression controlled by hnRNPA1. Mol Biol Rep. 2020 Apr;47(4):3031-3040. doi: 10.1007/s11033-020-05375-w.

100. Larrayoz M, Blakemore SJ, Dobson RC, Blunt MD, Rose-Zerilli MJ, Walewska R, Duncombe A, Oscier D, Koide K, Forconi F, Packham G, Yoshida M, Cragg MS, Strefford JC, Steele AJ. The SF3B1 inhibitor spliceostatin A (SSA) elicits apoptosis in chronic lymphocytic leukaemia cells through downregulation of Mcl-1. Leukemia. 2016 Feb;30(2):351-60. doi: 10.1038/leu.2015.286.

101. Ahearn IM, Haigis K, Bar-Sagi D, Philips MR. Regulating the regulator: post-translational modification of RAS. Nat Rev Mol Cell Biol. 2011 Dec 22;13(1):39-51. doi: 10.1038/nrm3255.

102. Schwarzová L, Štekrová J, Florianová M, Novotný A, Schneiderová M, Lněnička P, Kebrdlová V, Kotlas J, Veselá K, Kohoutová M. Novel mutations of the APC gene and genetic consequences of splicing mutations in the Czech FAP families. Fam Cancer. 2013 Mar;12(1):35-42. doi: 10.1007/s10689-012-9569-8.

103. Wu D, Matsushita K, Matsubara H, Nomura F, Tomonaga T. An alternative splicing isoform of eukaryotic initiation factor 4H promotes tumorigenesis in vivo and is a potential therapeutic target for human cancer. Int J Cancer. 2011 Mar 1;128(5):1018-30. doi: 10.1002/ijc.25419.

104. Zhang YQ, Pilyugin M, Kuester D, Leoni VP, Li L, Casula G, Zorcolo L, Schneider-Stock R, Atzori L, Irminger-Finger I. Expression of oncogenic BARD1

isoforms affects colon cancer progression and correlates with clinical outcome. Br J Cancer. 2012 Aug 7;107(4):675-83. doi: 10.1038/bjc.2012.297.

105. Díaz R, Peña C, Silva J, Lorenzo Y, García V, García JM, Sánchez A, Espinosa P, Yuste R, Bonilla F, Domínguez G. p73 Isoforms affect VEGF, VEGF165b and PEDF expression in human colorectal tumors: VEGF165b downregulation as a marker of poor prognosis. Int J Cancer. 2008 Sep 1;123(5):1060-7. doi: 10.1002/ijc.23619.

106. Wilson TR, McLaughlin KM, McEwan M, Sakai H, Rogers KM, Redmond KM, Johnston PG, Longley DB. c-FLIP: a key regulator of colorectal cancer cell death. Cancer Res. 2007 Jun 15;67(12):5754-62. doi: 10.1158/0008-5472.CAN-06-3585.

107. Kim CJ, Tambe Y, Mukaisho KI, Sugihara H, Kawauchi A, Inoue H. Akt-dependent activation of Erk by cyclin D1b contributes to cell invasiveness and tumorigenicity. Oncol Lett. 2016 Dec;12(6):4850-4856. doi: 10.3892/ol.2016.5286.

108. Wang MH, Lao WF, Wang D, Luo YL, Yao HP. Blocking tumorigenic activities of colorectal cancer cells by a splicing RON receptor variant defective in the tyrosine kinase domain. Cancer Biol Ther. 2007 Jul;6(7):1121-9. doi: 10.4161/cbt.6.7.4337.

109. Wan L, Yu W, Shen E, Sun W, Liu Y, Kong J, Wu Y, Han F, Zhang L, Yu T, Zhou Y, Xie S, Xu E, Zhang H, Lai M. SRSF6-regulated alternative splicing that promotes tumour progression offers a therapy target for colorectal cancer. Gut. 2019 Jan;68(1):118-129. doi: 10.1136/gutjnl-2017-314983.

110. Sun Y, Luo M, Chang G, Ren W, Wu K, Li X, Shen J, Zhao X, Hu Y. Phosphorylation of Ser6 in hnRNPA1 by S6K2 regulates glucose metabolism and cell growth in colorectal cancer. Oncol Lett. 2017 Dec;14(6):7323-7331. doi: 10.3892/ol.2017.7085.

111. Suvà ML, Tirosh I. Single-Cell RNA Sequencing in Cancer: Lessons Learned and Emerging Challenges. Mol Cell. 2019 Jul 11;75(1):7-12. doi: 10.1016/j.molcel.2019.05.003.

112. Yeo AT, Rawal S, Delcuze B, Christofides A, Atayde A, Strauss L, Balaj L, Rogers VA, Uhlmann EJ, Varma H, Carter BS, Boussiotis VA, Charest A. Single-cell RNA sequencing reveals evolution of immune landscape during glioblastoma progression. Nat Immunol. 2022 Jun;23(6):971-984. doi: 10.1038/s41590-022-01215-0.

113. Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, Wang X, Bodeau J, Tuch BB, Siddiqui A, Lao K, Surani MA. mRNA-Seq whole-transcriptome analysis of a single cell. Nat Methods. 2009 May;6(5):377-82. doi: 10.1038/nmeth.1315.

114. Gross A, Schoendube J, Zimmermann S, Steeb M, Zengerle R, Koltay P. Technologies for Single-Cell Isolation. Int J Mol Sci. 2015 Jul 24;16(8):16897-919. doi: 10.3390/ijms160816897.

115. Datta S, Malhotra L, Dickerson R, Chaffee S, Sen CK, Roy S. Laser capture microdissection: Big data from small samples. Histol Histopathol. 2015 Nov;30(11):1255-69. doi: 10.14670/HH-11-622.

116. Jackson K, Milner RJ, Doty A, Hutchison S, Cortes-Hinojosa G, Riva A, Sahay B, Lejeune A, Bechtel S. Analysis of canine myeloid-derived suppressor cells (MDSCs) utilizing fluorescence-activated cell sorting, RNA protection mediums to yield quality RNA for single-cell RNA sequencing. Vet Immunol Immunopathol. 2021 Jan;231:110144. doi: 10.1016/j.vetimm.2020.110144.

117. Welzel G, Seitz D, Schuster S. Magnetic-activated cell sorting (MACS) can be used as a large-scale method for establishing zebrafish neuronal cell cultures. Sci Rep. 2015 Jan 22;5:7959. doi: 10.1038/srep07959.

118. Lecault V, White AK, Singhal A, Hansen CL. Microfluidic single cell analysis: from promise to practice. Curr Opin Chem Biol. 2012 Aug;16(3-4):381-90. doi: 10.1016/j.cbpa.2012.03.022.

119. Li PH, Kong XY, He YZ, Liu Y, Peng X, Li ZH, Xu H, Luo H, Park J. Recent developments in application of single-cell RNA sequencing in the tumour immune

microenvironment and cancer therapy. Mil Med Res. 2022 Sep 26;9(1):52. doi: 10.1186/s40779-022-00414-y.

120. Zheng GX, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, Ziraldo SB, Wheeler TD, McDermott GP, Zhu J, Gregory MT, Shuga J, Montesclaros L, Underwood JG, Masquelier DA, Nishimura SY, Schnall-Levin M, Wyatt PW, Hindson CM, Bharadwaj R, Wong A, Ness KD, Beppu LW, Deeg HJ, McFarland C, Loeb KR, Valente WJ, Ericson NG, Stevens EA, Radich JP, Mikkelsen TS, Hindson BJ, Bielas JH. Massively parallel digital transcriptional profiling of single cells. Nat Commun. 2017 Jan 16;8:14049. doi: 10.1038/ncomms14049.

121. Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, Tirosh I, Bialas AR, Kamitaki N, Martersteck EM, Trombetta JJ, Weitz DA, Sanes JR, Shalek AK, Regev A, McCarroll SA. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell. 2015 May 21;161(5):1202-1214. doi: 10.1016/j.cell.2015.05.002.

122. Huang D, Ma N, Li X, Gou Y, Duan Y, Liu B, Xia J, Zhao X, Wang X, Li Q, Rao J, Zhang X. Advances in single-cell RNA sequencing and its applications in cancer research. J Hematol Oncol. 2023 Aug 24;16(1):98. doi: 10.1186/s13045-023-01494-6.

123. Ruan H, Wang Z, Zhai Y, Xu Y, Pi L, Zheng J, Zhou Y, Zhang C, Huang R, Chen K, Li X, Ma W, Wu Z, Shen J, Deng X, Zhang C, Guan M. Single-cell transcriptome analysis of diffuse large B cells in cerebrospinal fluid of central nervous system lymphoma. iScience. 2021 Aug 11;24(9):102972. doi: 10.1016/j.isci.2021.102972.

124. Wang R, Dang M, Harada K, Han G, Wang F, Pool Pizzi M, Zhao M, Tatlonghari G, Zhang S, Hao D, Lu Y, Zhao S, Badgwell BD, Blum Murphy M, Shanbhag N, Estrella JS, Roy-Chowdhuri S, Abdelhakeem AAF, Wang Y, Peng G, Hanash S, Calin GA, Song X, Chu Y, Zhang J, Li M, Chen K, Lazar AJ, Futreal A, Song S, Ajani JA, Wang L. Single-cell dissection of intratumoral heterogeneity and lineage diversity in

metastatic gastric adenocarcinoma. Nat Med. 2021 Jan;27(1):141-151. doi: 10.1038/s41591-020-1125-8.

125. Borcherding N, Voigt AP, Liu V, Link BK, Zhang W, Jabbari A. Single-Cell Profiling of Cutaneous T-Cell Lymphoma Reveals Underlying Heterogeneity Associated with Disease Progression. Clin Cancer Res. 2019 May 15;25(10):2996-3005. doi: 10.1158/1078-0432.CCR-18-3309.

126. Davis RT, Blake K, Ma D, Gabra MBI, Hernandez GA, Phung AT, Yang Y, Maurer D, Lefebvre AEYT, Alshetaiwi H, Xiao Z, Liu J, Locasale JW, Digman MA, Mjolsness E, Kong M, Werb Z, Lawson DA. Transcriptional diversity and bioenergetic shift in human breast cancer metastasis revealed by single-cell RNA sequencing. Nat Cell Biol. 2020 Mar;22(3):310-320. doi: 10.1038/s41556-020-0477-0.

127. Cao X, Wang Y, Zhang W, Zhong X, Gunes EG, Dang J, Wang J, Epstein AL, Querfeld C, Sun Z, Rosen ST, Feng M. Targeting macrophages for enhancing CD47 blockade-elicited lymphoma clearance and overcoming tumor-induced immunosuppression. Blood. 2022 Jun 2;139(22):3290-3302. doi: 10.1182/blood.2021013901.

128. Lee, HO., Hong, Y., Etlioglu, H.E. et al. Lineage-dependent gene expression programs influence the immune landscape of colorectal cancer. Nat Genet 52, 594–603 (2020). doi: 10.1038/s41588-020-0636-z.

129. Li, H., Courtois, E., Sengupta, D. et al. Reference component analysis of single-cell transcriptomes elucidates cellular heterogeneity in human colorectal tumors. Nat Genet 49, 708–718 (2017). doi: 10.1038/ng.3818.

130. Lin K, Chowdhury S, Zeineddine MA, Zeineddine FA, Hornstein NJ, Villarreal OE, Maru DM, Haymaker CL, Vauthey JN, Chang GJ, Bogatenkova E, Menter D, Kopetz S, Shen JP. Identification of Colorectal Cancer Cell Stemness from Single-Cell

RNA Sequencing. Mol Cancer Res. 2024 Apr 2;22(4):337-346. doi: 10.1158/1541-7786.MCR-23-0468.

131. Di Tommaso, P., Chatzou, M., Floden, E. et al. Nextflow enables reproducible computational workflows. Nat Biotechnol 35, 316–319 (2017). doi: 10.1038/nbt.3820.

132. Montgomery SB, Astakhova T, Bilenky M, Birney E, Fu T, Hassel M, Melsopp C, Rak M, Robertson AG, Sleumer M, Siddiqui AS, Jones SJ. Sockeye: a 3D environment for comparative genomics. Genome Res. 2004 May;14(5):956-62. doi: 10.1101/gr.1890304.

133. Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics. 2018 Sep 15;34(18):3094-3100. doi: 10.1093/bioinformatics/bty191.

134. Düchting H, Seurat G (2000) Seurat. Taschen.

135. Wolf, F., Angerer, P. & Theis, F. SCANPY: large-scale single-cell gene expression data analysis. Genome Biol 19, 15 (2018). doi: 10.1186/s13059-017-1382-0.

136. James T Robinson, Helga Thorvaldsdottir, Douglass Turner, Jill P Mesirov, igv.js: an embeddable JavaScript implementation of the Integrative Genomics Viewer (IGV), Bioinformatics, Volume 39, Issue 1, January 2023, btac830. doi: 10.1093/bioinformatics/btac830.

137. Chawla, S., Rockstroh, A., Lehman, M. et al. Gene expression based inference of cancer drug sensitivity. Nat Commun 13, 5680 (2022). doi: 10.1038/s41467-022-33291-z.

138. La Manno, G., Soldatov, R., Zeisel, A. et al. RNA velocity of single cells. Nature 560, 494–498 (2018). doi: 10.1038/s41586-018-0414-6.

139. Bergen V, Soldatov RA, Kharchenko PV, Theis FJ. RNA velocity-current challenges and future perspectives. Mol Syst Biol. 2021 Aug;17(8):e10282. doi: 10.15252/msb.202110282.

140. Kinker GS, Greenwald AC, Tal R, Orlova Z, Cuoco MS, McFarland JM, Warren A, Rodman C, Roth JA, Bender SA, Kumar B, Rocco JW, Fernandes PACM, Mader CC, Keren-Shaul H, Plotnikov A, Barr H, Tsherniak A, Rozenblatt-Rosen O, Krizhanovsky V, Puram SV, Regev A, Tirosh I. Pan-cancer single-cell RNA-seq identifies recurring programs of cellular heterogeneity. Nat Genet. 2020 Nov;52(11):1208-1218. doi: 10.1038/s41588-020-00726-6.

141. Ooft SN, Weeber F, Dijkstra KK, McLean CM, Kaing S, van Werkhoven E, Schipper L, Hoes L, Vis DJ, van de Haar J, Prevoo W, Snaebjornsson P, van der Velden D, Klein M, Chalabi M, Boot H, van Leerdam M, Bloemendal HJ, Beerepoot LV, Wessels L, Cuppen E, Clevers H, Voest EE. Patient-derived organoids can predict response to chemotherapy in metastatic colorectal cancer patients. Sci Transl Med. 2019 Oct 9;11(513):eaay2574. doi: 10.1126/scitranslmed.aay2574.

142. Yoshimoto S, Taguchi M, Sumi S, Oka K, Okamura K. Establishment of a novel protocol for formalin-fixed paraffin-embedded organoids and spheroids. Biol Open. 2023 May 15;12 doi: 10.1242/bio.059882.

143. Puglisi, C.; Giuffrida, R.; Borzì, G.; Di Mattia, P.; Costa, A.; Colarossi, C.; Deiana, E.; Picardo, M.C.; Colarossi, L.; Mare, M.; et al. Radiosensitivity of Cancer Stem Cells Has Potential Predictive Value for Individual Responses to Radiotherapy in Locally Advanced Rectal Cancer. Cancers 2020, 12, 3672. doi: 10.3390/cancers12123672.

144. Yao Y, Xu X, Yang L, Zhu J, Wan J, Shen L, Xia F, Fu G, Deng Y, Pan M, Guo Q, Gao X, Li Y, Rao X, Zhou Y, Liang L, Wang Y, Zhang J, Zhang H, Li G, Zhang L, Peng J, Cai S, Hu C, Gao J, Clevers H, Zhang Z, Hua G. Patient-Derived Organoids Predict Chemoradiation Responses of Locally Advanced Rectal Cancer. Cell Stem Cell. 2020 Jan 2;26(1):17-26.e6. doi: 10.1016/j.stem.2019.10.010.

145. R Core Team (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/).

146. Anders S, Reyes A, Huber W. Detecting differential usage of exons from RNA-seq data. Genome Res. 2012 Oct;22(10):2008-17. doi: 10.1101/gr.133744.111.

147. Van den Berge K, Soneson C, Robinson MD, Clement L. stageR: a general stage-wise method for controlling the gene-level false discovery rate in differential expression and differential transcript usage. Genome Biol. 2017 Aug 7;18(1):151. doi: 10.1186/s13059-017-1277-0.

148. Gene Ontology Consortium. The Gene Ontology resource: enriching a GOld mine. Nucleic Acids Res. 2021 Jan 8;49(D1):D325-D334. doi: 10.1093/nar/gkaa1113. PMID: 33290552; PMCID: PMC7779012.), KEGG database (Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. KEGG for taxonomy-based analysis of pathways and genomes. Nucleic Acids Res. 2023 Jan 6;51(D1):D587-D592. doi: 10.1093/nar/gkac963.

149. Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. KEGG for taxonomy-based analysis of pathways and genomes. Nucleic Acids Res. 2023 Jan 6;51(D1):D587-D592. doi: 10.1093/nar/gkac963.

150. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005 Oct 25;102(43):15545-50. doi: 10.1073/pnas.0506580102.

151. Hao Y, Hao S, Andersen-Nissen E, Mauck WM 3rd, Zheng S, Butler A, Lee MJ, Wilk AJ, Darby C, Zager M, Hoffman P, Stoeckius M, Papalexi E, Mimitou EP, Jain J, Srivastava A, Stuart T, Fleming LM, Yeung B, Rogers AJ, McElrath JM, Blish CA,

Gottardo R, Smibert P, Satija R. Integrated analysis of multimodal single-cell data. Cell. 2021 Jun 24;184(13):3573-3587.e29. doi: 10.1016/j.cell.2021.04.048.

152. Zhang X, Lan Y, Xu J, Quan F, Zhao E, Deng C, Luo T, Xu L, Liao G, Yan M, Ping Y, Li F, Shi A, Bai J, Zhao T, Li X, Xiao Y. CellMarker: a manually curated resource of cell markers in human and mouse. Nucleic Acids Res. 2019 Jan 8;47(D1):D721-D728. doi: 10.1093/nar/gky900.

153. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, Feng T, Zhou L, Tang W, Zhan L, Fu X, Liu S, Bo X, Yu G. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation (Camb). 2021 Jul 1;2(3):100141. doi: 10.1016/j.xinn.2021.100141.

154. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005 Oct 25;102(43):15545-50. doi: 10.1073/pnas.0506580102.

155. Sachs N, de Ligt J, Kopper O, Gogola E, Bounova G, Weeber F, Balgobind AV, Wind K, Gracanin A, Begthel H, Korving J, van Boxtel R, Duarte AA, Lelieveld D, van Hoeck A, Ernst RF, Blokzijl F, Nijman IJ, Hoogstraat M, van de Ven M, Egan DA, Zinzalla V, Moll J, Boj SF, Voest EE, Wessels L, van Diest PJ, Rottenberg S, Vries RGJ, Cuppen E, Clevers H. A Living Biobank of Breast Cancer Organoids Captures Disease Heterogeneity. Cell. 2018 Jan 11;172(1-2):373-386.e10. doi: 10.1016/j.cell.2017.11.010.

156. van de Wetering M, Francies HE, Francis JM, Bounova G, Iorio F, Pronk A, van Houdt W, van Gorp J, Taylor-Weiner A, Kester L, McLaren-Douglas A, Blokker J, Jaksani S, Bartfeld S, Volckman R, van Sluis P, Li VS, Seepo S, Sekhar Pedamallu C, Cibulskis K, Carter SL, McKenna A, Lawrence MS, Lichtenstein L, Stewart C, Koster

J, Versteeg R, van Oudenaarden A, Saez-Rodriguez J, Vries RG, Getz G, Wessels L, Stratton MR, McDermott U, Meyerson M, Garnett MJ, Clevers H. Prospective derivation of a living organoid biobank of colorectal cancer patients. Cell. 2015 May 7;161(4):933-45. doi: 10.1016/j.cell.2015.03.053.

157. Cristobal A, van den Toorn HWP, van de Wetering M, Clevers H, Heck AJR, Mohammed S. Personalized Proteome Profiles of Healthy and Tumor Human Colon Organoids Reveal Both Individual Diversity and Basic Features of Colorectal Cancer. Cell Rep. 2017 Jan 3;18(1):263-274. doi: 10.1016/j.celrep.2016.12.016.

158. Sachs N, de Ligt J, Kopper O, Gogola E, Bounova G, Weeber F, Balgobind AV, Wind K, Gracanin A, Begthel H, Korving J, van Boxtel R, Duarte AA, Lelieveld D, van Hoeck A, Ernst RF, Blokzijl F, Nijman IJ, Hoogstraat M, van de Ven M, Egan DA, Zinzalla V, Moll J, Boj SF, Voest EE, Wessels L, van Diest PJ, Rottenberg S, Vries RGJ, Cuppen E, Clevers H. A Living Biobank of Breast Cancer Organoids Captures Disease Heterogeneity. Cell. 2018 Jan 11;172(1-2):373-386.e10. doi: 10.1016/j.cell.2017.11.010. Epub 2017 Dec 7.

159. van de Wetering M, Francies HE, Francis JM, Bounova G, Iorio F, Pronk A, van Houdt W, van Gorp J, Taylor-Weiner A, Kester L, McLaren-Douglas A, Blokker J, Jaksani S, Bartfeld S, Volckman R, van Sluis P, Li VS, Seepo S, Sekhar Pedamallu C, Cibulskis K, Carter SL, McKenna A, Lawrence MS, Lichtenstein L, Stewart C, Koster J, Versteeg R, van Oudenaarden A, Saez-Rodriguez J, Vries RG, Getz G, Wessels L, Stratton MR, McDermott U, Meyerson M, Garnett MJ, Clevers H. Prospective derivation of a living organoid biobank of colorectal cancer patients. Cell. 2015 May 7;161(4):933-45. doi: 10.1016/j.cell.2015.03.053.

160. Vlachogiannis G, Hedayat S, Vatsiou A, Jamin Y, Fernández-Mateos J, Khan K, Lampis A, Eason K, Huntingford I, Burke R, Rata M, Koh DM, Tunariu N, Collins D, Hulkki-Wilson S, Ragulan C, Spiteri I, Moorcraft SY, Chau I, Rao S, Watkins D,

Fotiadis N, Bali M, Darvish-Damavandi M, Lote H, Eltahir Z, Smyth EC, Begum R, Clarke PA, Hahne JC, Dowsett M, de Bono J, Workman P, Sadanandam A, Fassan M, Sansom OJ, Eccles S, Starling N, Braconi C, Sottoriva A, Robinson SP, Cunningham D, Valeri N. Patient-derived organoids model treatment response of metastatic gastrointestinal cancers. Science. 2018 Feb 23;359(6378):920-926. doi: 10.1126/science.aao2774.