

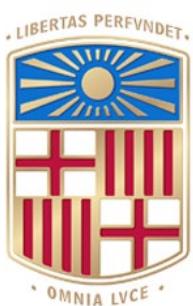
Impact of 5-FU treatment on tumor cell plasticity in colorectal cancer

UOC

A scRNAseq analysis

Universitat
Oberta
de Catalunya

Miquel Gratacós i Aurich



UNIVERSITAT DE
BARCELONA

MU Bioinf. i Bioest.

Àrea de treball final

Tutor/a de TFM

Laia Bassaganyas

**Professor/a responsable de
l'assignatura**

Nom i Cognoms

Data Lliurament



Aquesta obra està subjecta a una llicència de
[Reconeixement-NoComercial-](#)
[SenseObraDerivada 3.0 Espanya de Creative](#)
[Commons](#)

**Llicències alternatives (triar alguna de les següents i substituir la de la
pàgina anterior)**

A) Creative Commons:



Aquesta obra està subjecta a una llicència de
[Reconeixement-NoComercial-
SenseObraDerivada 3.0 Espanya de Creative
Commons](#)



Aquesta obra està subjecta a una llicència de
[Reconeixement-NoComercial-Compartirlqual
3.0 Espanya de Creative Commons](#)



Aquesta obra està subjecta a una llicència de
[Reconeixement-NoComercial 3.0 Espanya de
Creative Commons](#)



Aquesta obra està subjecta a una llicència de
[Reconeixement-SenseObraDerivada 3.0
Espanya de Creative Commons](#)



Aquesta obra està subjecta a una llicència de
[Reconeixement-CompartirIgual 3.0 Espanya de Creative Commons](#)



Aquesta obra està subjecta a una llicència de
[Reconeixement 3.0 Espanya de Creative Commons](#)

B) GNU Free Documentation License (GNU FDL)

Copyright © ANY EL-TEU-NOM.

Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.3 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts.

A copy of the license is included in the section entitled "GNU Free Documentation License".

C) Copyright

© (l'autor/a)

Reservats tots els drets. Està prohibit la reproducció total o parcial d'aquesta obra per qualsevol mitjà o procediment, compresos la impressió, la reprografia, el microfilm, el tractament informàtic o qualsevol altre sistema, així com la distribució d'exemplars mitjançant lloguer i préstec, sense l'autorització escrita de l'autor o dels límits que autoritzi la Llei de Propietat Intel·lectual.

FITXA DEL TREBALL FINAL

Títol del treball:	<i>Impact of 5-FU treatment on tumor cell plasticity in colorectal cancer</i>
Nom de l'autor:	<i>Miquel Gratacós i Aurich</i>
Nom del director/a:	<i>Laia Bassaganyas</i>
Nom del PRA:	<i>Nom i dos cognoms</i>
Data de lliurament (mm/aaaa):	<i>MM/AAAA</i>
Titulació o programa:	<i>Pla d'estudis</i>
Àrea del Treball Final:	<i>Area de TF</i>
Idioma del treball:	<i>anglès</i>
Paraules clau	<i>Màxim 3 paraules clau, validades pel tutor/a del TF</i>
Resum del Treball	
<i>Màxim 250 paraules, amb la finalitat, context d'aplicació, metodologia, resultats i conclusions del treball</i>	
Abstract	
<i>A maximum of 250 words, detailing the purpose, context of application, methodology, results and conclusions of the work</i>	

Index

1.	Introduction	1
1.1.	Context and Justification of the Work.....	1
1.2.	Objectives	3
1.2.1	General objectives	3
1.2.2	Specific objectives	3
1.3.	Impact on sustainability, ethical-social and diversity	4
1.4.	Approach and method followed.....	4
1.5.	Planning	5
1.5.1	Tasks.....	5
1.5.2.	Callendar	6
1.6.	Brief summary of products obtained.....	7
1.7.	Brief description of the other chapters of the report	7
2.	Material and Methods.....	8
3.	Results	9
4.	Conclusion and Discussion	13
5.	Glossari	 Error! Marcador no definido.
6.	Bibliografia	 Error! Marcador no definido.
7.	Annex	16

List of Figures

Figure 1 Gantt Diagram of the followed timeline during the project.....	6
Figure 2. UMAP from scRNAseq analysis by time points. Clustering defined by Seurat v5 Algorithm A) Represents the initial timepoint of the experiment, before the 5-FU treatment, 6 different clusters are identified. B) Time point 2 of the experiment, HCT116 cells after the 5-FU treatment, 6 different clusters represented. C) 7 days of experiment with 6 different clusters. D) Day 14 of the experiment with 5 cell populations.....	9
Figure 3. UMAP from scRNAseq analysis by time points. Clustering defined by Cell Cycle Ranked, G2M, G1 and S phase are labeled. A) Initial timepoint of the experiment, before the 5-FU treatment. B) Time point 2 of the experiment, HCT116 cells after the 5-FU treatment. C) 7 days of experiment. D) Day 14 of the experiment.....	10
Figure 4. UMAP from scRNAseq analysis by time points. Clustering defined by CancerSEA cancer cell marker dataset. A) Initial timepoint of the experiment, before the 5-FU treatment. B) Time point 2 of the experiment, HCT116 cells after the 5-FU treatment. C) 7 days of experiment. D) Day 14 of the experiment.	11
Figure 5. Proportion Plot of the cell populations of each time point based on Cancer Cell Markers dataset CancerSEA.	12

List of Tables

Table 1. Deadlines and duration of the project planification.	6
--------------------------------------------------------------------	---

1. Introduction

1.1. Context and Justification of the Work

Colorectal cancer (CRC) is considered the third most diagnosed cancer, affecting both men and women worldwide. Unfortunately, drug resistance is one of the main reasons for the low survival rates of CRC patients[1].

5-Fluorouracil (5-FU)-based chemotherapy, the mainstay therapy for patients with CRC, is often associated with resistance and consequently tumor recurrence. Despite the advancements in various modulation strategies, nearly half of CRC patients experience recurrence within 5 years of their treatment. In response to anticancer treatments, including 5-FU, tumor cells develop tolerance, promoting tumor recurrence and metastasis development[2], [3].

Cancer cells are heterogeneous in morphology, inheritance, and functions. Throughout cancer cells, the clinical outcomes of 5-FU treatments are often associated with cancer stem cells (CSCs), which are a small population of cancer cells that own the ability to self-renew and generate the tumor. It has been considered that CSCs might regulate the mechanisms of intrinsic or acquired drug resistance[4], enhancing the relapse of tumors treated with chemotherapy[2].

Even though, mechanisms of drug tolerance in cancer are still poorly understood, for this reason, revealing the underlying mechanism, is necessary for improving the outcome of CRC patients. Additionally, transcriptomic, and epigenetic studies describe cellular plasticity as one of the hallmarks of treatment tolerance. Specially, epigenetic modifications promote drug efflux, and the regulation of drug targets and associated signaling pathways. Nevertheless, in the context of tumor drug resistance, the translation of messenger RNA into proteins remains unexplained[5].

In recent years, with the development of single-cell RNA sequencing (scRNA-seq) technology have emerged powerful new tools and techniques that have provided new insights into the definition of molecular signatures of cell types, subtypes, data analysis, visualization, and mining of scRNA-seq datasets, such as Seurat analysis platform [6], [7]. For these reasons, these new technologies

are crucial for the understanding and study of transcriptomic and cell population development.

The tolerance phase is followed by a significant resumption of cell division. Furthermore, single-cell RNA sequencing (scRNAseq) analyses were performed on CRC cells before and after 5-FU treatment. This highlights a yet poorly evaluated role of tumor cell plasticity in the response to chemotherapy in CRC.

1.2. Objectives

1.2.1 General objectives

The study was planned to be divided in the following objectives (5):

1. Demonstrate that translational control is one of the molecular mechanisms responsible for cellular reprogramming in response to 5-FU.
2. Isolate relevant cellular subpopulations for study.
3. Detail the translatome of these subpopulations.
4. Phenotypically characterize the cellular subpopulations
5. Identify the role of these subpopulations in treatment tolerance and tumor recurrence.

1.2.2 Specific objectives

To establish a more detailed scheme and create the workflow needed for this project, we defined specific objectives based on the original general objectives:

1. Investigate the role of translational control in 5-FU induced cellular reprogramming.
 2. Employ and optimize single-cell RNA-seq workflow using Seurat Package from R to obtain relevant cellular subpopulations.
 3. Perform detailed analysis of the translatome of these subpopulations using appropriate methodologies, thanks to Seurat Package workflow and optimization.
- 3.2 Comparison between integrative analysis against the non-integrative analysis.
4. Phenotypically characterize the cellular subpopulations using cell and molecular biology techniques.
 5. Evaluate the contribution of the identified cellular subpopulations to 5-FU treatment tolerance and tumor recurrence.

1.3. Impact on sustainability, ethical-social and diversity

Aquesta secció hauria d'identificar els impactes positius i/o negatius del TF en les tres dimensions de la competència transversal UOC "Compromís ètic i global".

La Guia transversal sobre la Competència Ètica i Global us ajudarà a redactar aquests apartats.

1.4. Approach and method followed

To understand the CRC cell plasticity, an experiment was performed on a CRC cancer cell line (HCT116), which was treated with 5-FU, and the evolution of the cells was observed until day 20. At different time points, days 0, 2, 7, and 14, RNA from the colonies was extracted and sequenced on single-cell. So, the data is a longitudinal dataset of scRNAseq, from before the treatment until the last day of the culturing, to obtain a map of the populations of the different cell lines during the 5-FU treatment on CRC cells.

Based on this data set, we performed a scRNAseq analysis throughout the Seurat v5 package[7] on R-studio[8], where two different studies were performed, one script was designed for analyzing scRNAseq data by time points, and another script was planned for making and integrative analysis of all datasets. For the analysis by time points and the integration study, we identified different cluster types, clustering based on Seurat algorithms, clustering based on the Cell Cycle of each cell (G2M, G1, and S phase), and once cell markers from each cluster are defined, an identification of Cancer Cell Markers based on CancerSEA dataset[9] throughout the use of Scina[10] is performed, an algorithm that is used for automatic Cell Type Detection and Assignment. Moreover, an analysis of ligand-receptors throughout the LIANA package[11].

Finally, once the significant markers are addressed, a biological analysis based on the results of the study is performed to assess the mechanisms behind the cell population changes throughout the 5-FU treatment.

1.5. Planning

1.5.1 Tasks

The tasks developed during this project are described in Figure1, resulting in the following schedule:

Literature Review and Preliminary Analysis:

Duration: 2 weeks

Task: Conduct a comprehensive review of literature related to translational control in CRC and perform preliminary analysis of existing scRNAseq data.

Workflow Optimization:

Duration: 3 weeks

Task: Optimize the scRNAseq workflow using the Seurat package from R, including parameter tuning and integration of multiple datasets.

- By timepoints analysis
- Integrative analysis

Translatome Analysis:

Duration: 4 weeks

Task: Perform detailed analyses of the translatome of identified cellular subpopulations to understand treatment tolerance mechanisms.

Cluster markers analysis:

Duration: 2 weeks

Task:

Phenotypic Characterization:

Duration: 4 weeks

Task: Phenotypically characterize cellular subpopulations using cell and molecular biology techniques to understand their functional properties.

Biological description:

Duration: 2 weeks

Task:

Preparation of the Manuscript:

Duration: 4 weeks.

Task: Writing and developing the conclusions and discussion of the project.

1.5.2. Callendar

The time schedule for this project was defined in order to achieve all the objectives of the study.

Table 1. Deadlines and duration of the project planification.

Task	Start date	Duration (days)	End date
Literature Review and Preliminary Analysis	29/02/2024	14	14/03/2024
PAC01 - Definition and Work Plan	12/03/2024	1	12/03/2024
Workflow Optimization	15/03/2024	21	05/04/2024
Translatome Analysis	22/03/2024	28	19/04/2024
PAC02 - Development of the work Phase 1	09/04/2024	1	09/04/2024
Cluster markers analysis	12/04/2024	14	26/04/2024
Phenotypic Characterization	19/04/2024	14	03/05/2024
Biological description	03/05/2024	14	17/05/2024
PEC03 - Development of the work Phase 2	14/05/2024	1	14/05/2024
Preparation of the Manuscript	17/05/2024	32	18/06/2024
PEC04 - End of the Project	18/06/2024	1	18/06/2024

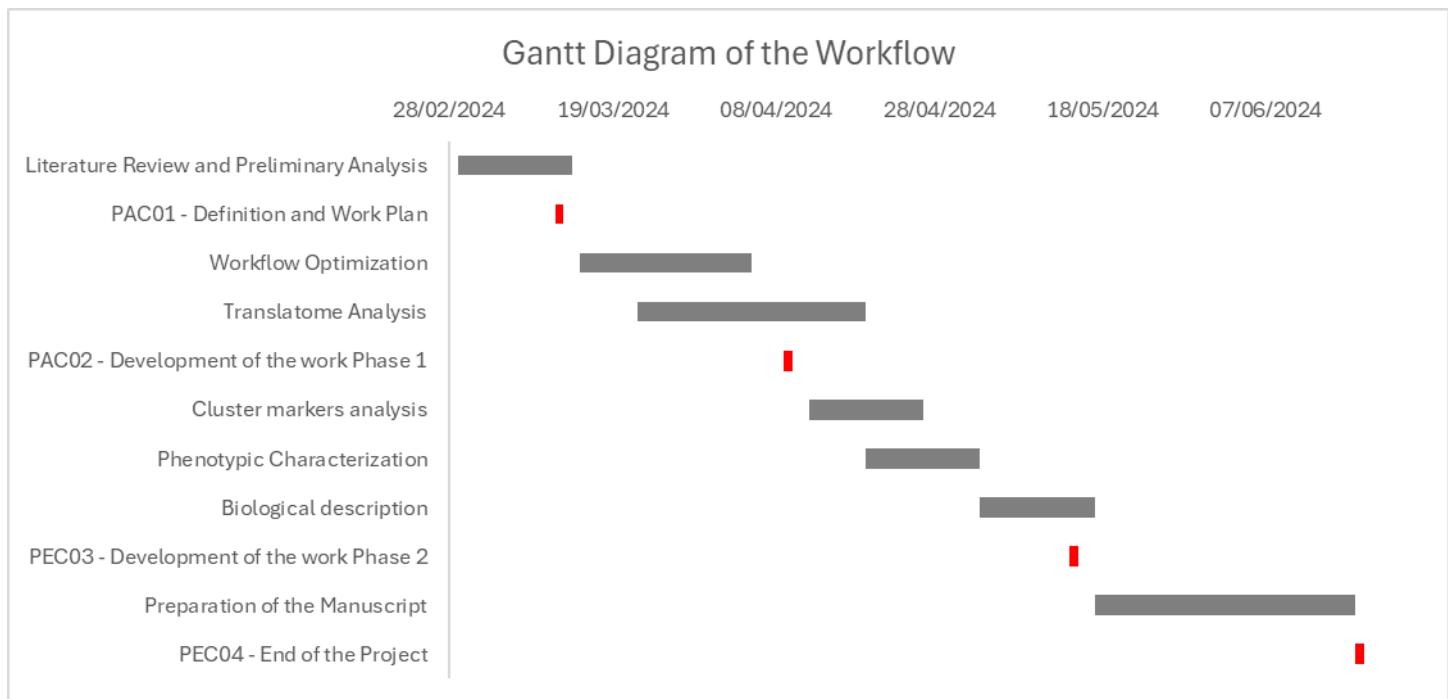


Figure 1 Gantt Diagram of the followed timeline during the project.

1.6. Brief summary of products obtained

- UMAP and table of the most significant markers, based on p-adj and Log2 FC values, from each time point (0, 2, 7 and 14), on 3 different clustering methodologies:
 - o Clustering by Seurat v5 Algorithm.
 - o Grouping by Cell Cycle, G2M, G1 and S phase.
 - o Clustering with labels based on CancerSEA database of Cancer Cell Markers.
 - Proportion plot based on this Cell Markers.
- Integrative analysis of the experiment.
 - o UMAP and table of the most significant markers.
- R-scripts for Integration and time point analysis for Seurat v5 scRNA-seq analysis.
- LIANA RESULTS
- MILOR RESULTS

1.7. Brief description of the other chapters of the report

Breu explicació dels continguts de cada capítol i la seva relació amb el projecte global.

2. Material and Methods

En aquests capítols, cal descriure:

- els aspectes més rellevant del disseny i desenvolupament del treball
- la metodologia triada per a fer aquest desenvolupament, descrivint les alternatives possibles, les decisions preses, i els criteris utilitzats per prendre aquestes decisions.
- descripció dels productes obtinguts.

L'estructuració dels capítols pot variar segons el tipus de treball.

En cas que s'escaigui, s'inclourà un apartat de "Valoració econòmica del treball". Aquest apartat indicarà les despeses associades al desenvolupament i manteniment del treball, així com els beneficis econòmics obtinguts i una anàlisi final sobre la viabilitat del producte.

3. Results

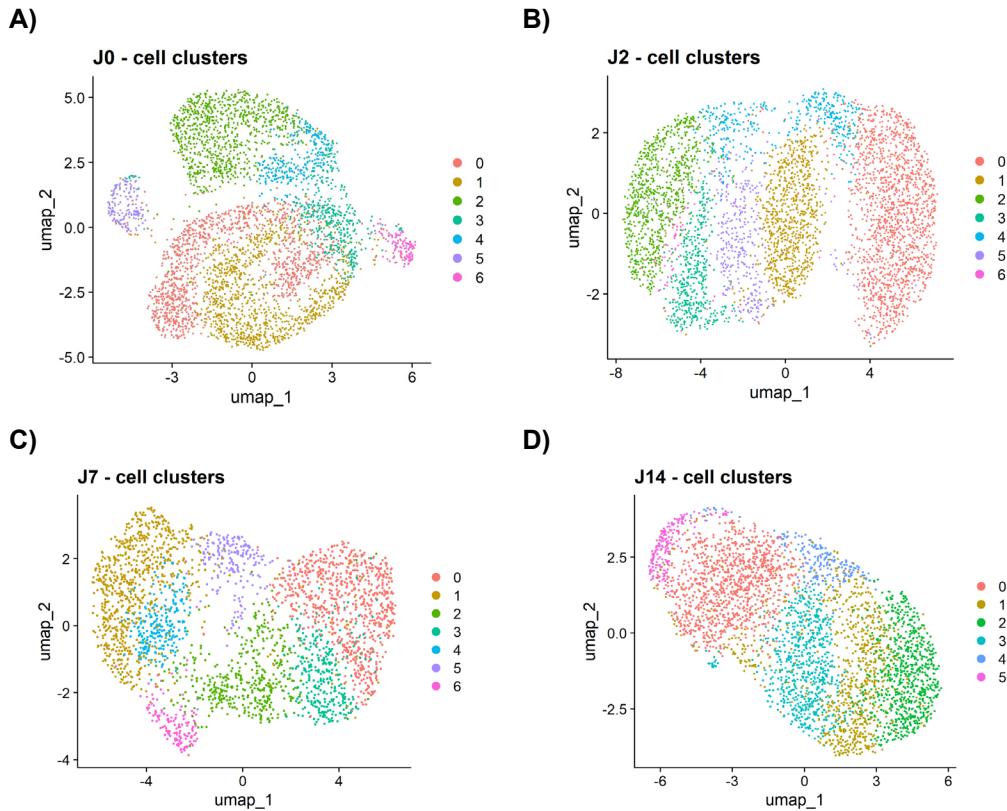


Figure 2. UMAP from scRNAseq analysis by time points. Clustering, defined by Seurat v5 Algorithm **A)** Represents the initial time point of the experiment; before the 5-FU treatment, 6 different clusters are identified. **B)** Time point 2 of the experiment: HCT116 cells after the 5-FU treatment, 6 different clusters represented. **C)** 7 days of experiment with 6 different clusters. **D)** Day 14 of the experiment with 5 cell populations.

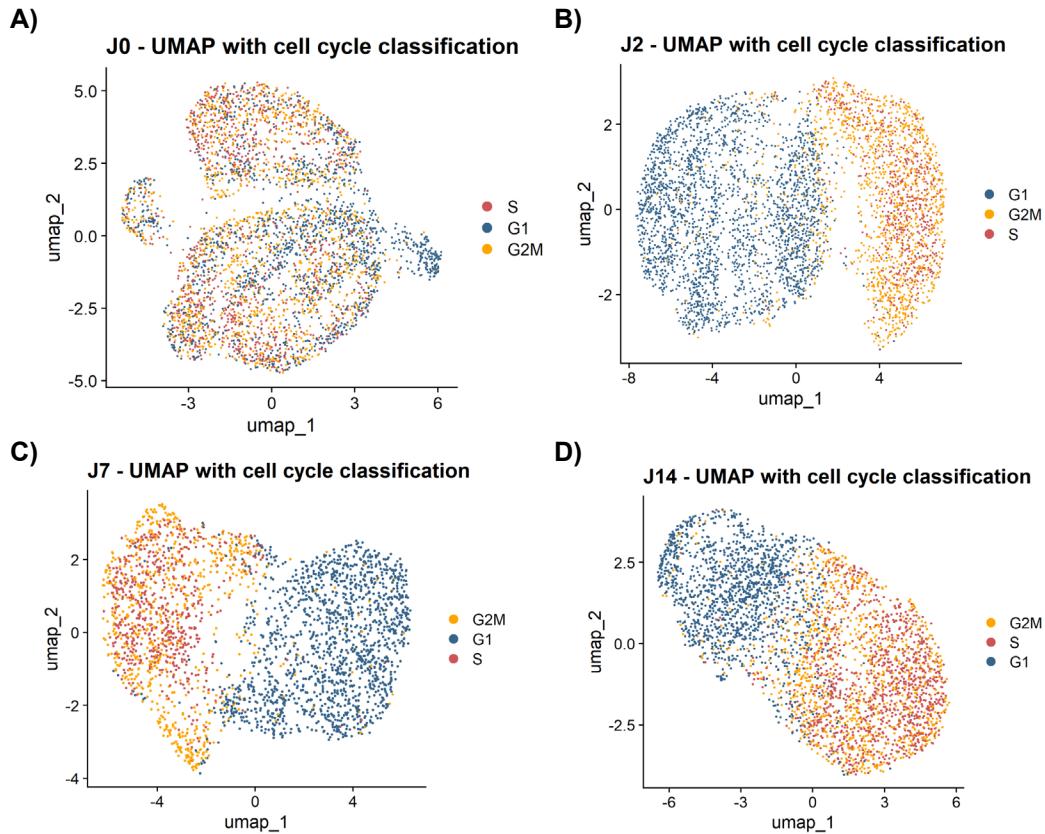


Figure 3. UMAP from scRNAseq analysis by time points. Clustering defined by Cell Cycle Ranked, G2M, G1 and S phase are labeled. **A)** Initial time point of the experiment, before the 5-FU treatment. **B)** Time point 2 of the experiment, HCT116 cells after the 5-FU treatment. **C)** 7 days of experiment. **D)** Day 14 of the experiment.

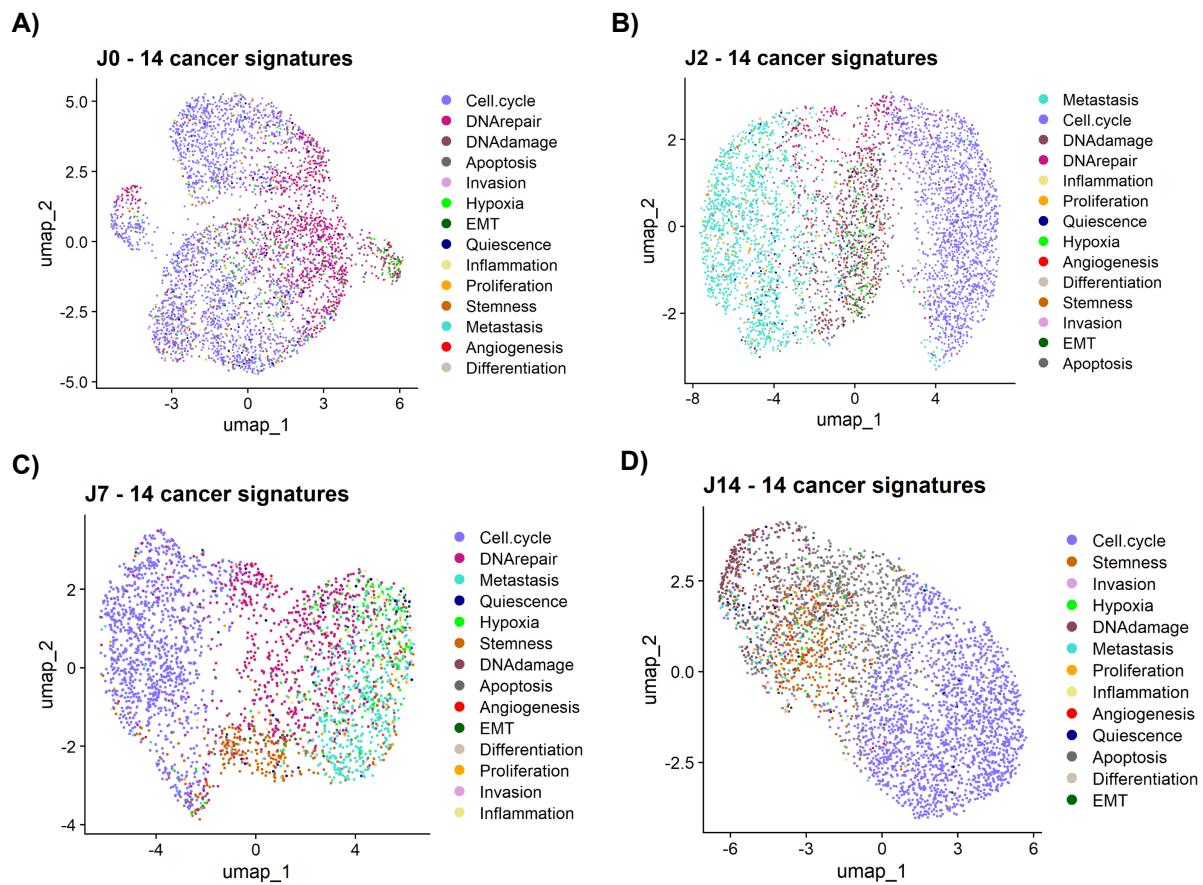


Figure 4. UMAP from scRNAseq analysis by time points. Clustering is defined by the CancerSEA cancer cell marker dataset. **A)** Initial time point of the experiment, before the 5-FU treatment. **B)** Time point 2 of the experiment: HCT116 cells after the 5-FU treatment. **C)** 7 days of experiment. **D)** Day 14 of the experiment.

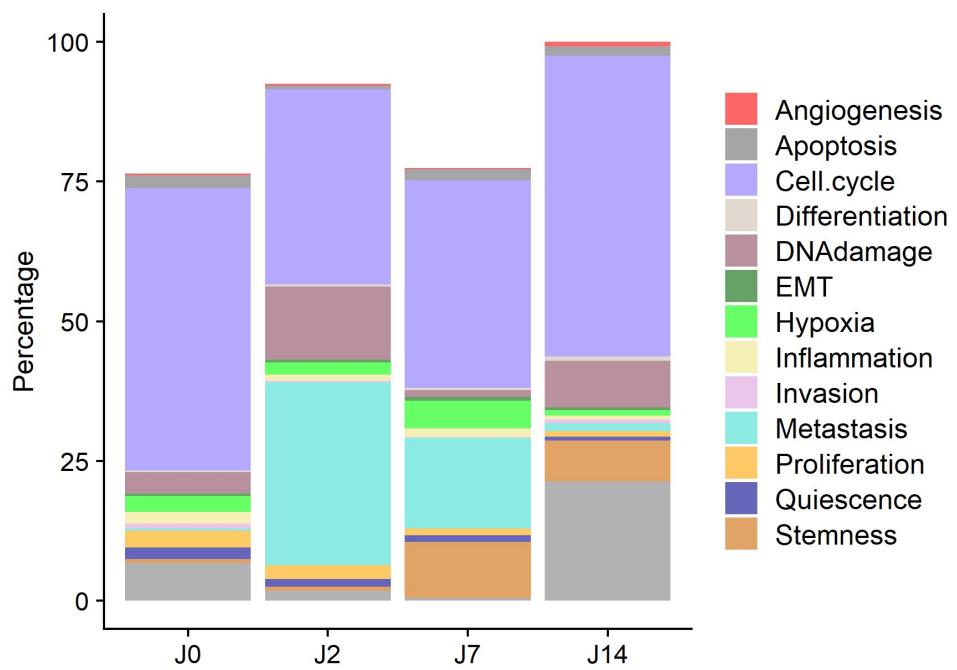


Figure 5. Proportion Plot of the cell populations at each time point based on the Cancer Cell Markers dataset, CancerSEA.

FALTEN PLOTS ANÀLISI INTEGRAT

4. Conclusion and Discussion

Aquest capítol ha d'incloure:

- Una descripció de les conclusions del treball:
 - Un cop s'han obtingut els resultats quines conclusions s'estreu?
 - Aquests resultats són els esperats? O han estat sorprenents? Per què?
- Una reflexió crítica sobre l'assoliment dels objectius plantejats inicialment:
 - Hem assolit tots els objectius? Si la resposta és negativa, per quin motiu?
- Una anàlisi crítica del seguiment de la planificació i metodologia al llarg del producte:
 - S'ha seguit la planificació?
 - La metodologia prevista ha estat prou adequada?
 - Ha calgut introduir canvis per garantir l'èxit del treball? Per què?
- Dels impactes previstos a 1.3 (ètic-socials, de sostenibilitat i de diversitat), avaluar/esmentar si s'han mitigat (si eren negatius) o si s'han aconseguit (si eren positius).
- Si han aparegut impactes no previstos a 1.3, avaluar/esmentar com s'han mitigat (si eren negatius) o què han aportat (si eren positius).
- Les línies de treball futur que no s'han pogut explorar en aquest treball i han quedat pendents.

5. Glossary

Definició dels termes i acrònims més rellevants utilitzats dins la Memòria.

6. Reference

- [1] H. Sung *et al.*, “Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries,” *CA. Cancer J. Clin.*, vol. 71, no. 3, pp. 209–249, 2021, doi: 10.3322/caac.21660.
- [2] Y.-H. Cho *et al.*, “5-FU promotes stemness of colorectal cancer via p53-mediated WNT/β-catenin pathway activation,” *Nat. Commun.*, vol. 11, no. 1, p. 5321, Oct. 2020, doi: 10.1038/s41467-020-19173-2.
- [3] K. V. D. Jeught, H.-C. Xu, Y.-J. Li, X.-B. Lu, and G. Ji, “Drug resistance and new therapies in colorectal cancer,” *World J. Gastroenterol.*, vol. 24, no. 34, pp. 3834–3848, Sep. 2018, doi: 10.3748/wjg.v24.i34.3834.
- [4] M. Elbadawy, T. Usui, H. Yamawaki, and K. Sasaki, “Development of an Experimental Model for Analyzing Drug Resistance in Colorectal Cancer,” *Cancers*, vol. 10, no. 6, Art. no. 6, Jun. 2018, doi: 10.3390/cancers10060164.
- [5] M. Luo, X. Yang, H.-N. Chen, E. C. Nice, and C. Huang, “Drug resistance in colorectal cancer: An epigenetic overview,” *Biochim. Biophys. Acta BBA - Rev. Cancer*, vol. 1876, no. 2, p. 188623, Dec. 2021, doi: 10.1016/j.bbcan.2021.188623.
- [6] S. E. Mohr, S. G. Tattikota, J. Xu, J. Zirin, Y. Hu, and N. Perrimon, “Methods and tools for spatial mapping of single-cell RNAseq clusters in *Drosophila*,” *Genetics*, vol. 217, no. 4, p. iyab019, Apr. 2021, doi: 10.1093/genetics/iyab019.
- [7] Y. Hao *et al.*, “Dictionary learning for integrative, multimodal and scalable single-cell analysis,” *Nat. Biotechnol.*, vol. 42, no. 2, pp. 293–304, Feb. 2024, doi: 10.1038/s41587-023-01767-y.
- [8] “R: The R Project for Statistical Computing.” Accessed: May 13, 2024. [Online]. Available: <https://www.r-project.org/>
- [9] H. Yuan *et al.*, “CancerSEA: a cancer single-cell state atlas,” *Nucleic Acids Res.*, vol. 47, no. D1, pp. D900–D908, Jan. 2019, doi: 10.1093/nar/gky939.
- [10] “SCINA: Semi-supervised Category Identification and Assignment | single cell r package | cell clustering tools | deconvolution software.” Accessed: May 13, 2024. [Online]. Available: <https://dbai.biohpc.swmed.edu/scina/>
- [11] D. Dimitrov *et al.*, “Comparison of methods and resources for cell-cell communication inference from single-cell RNA-Seq data,” *Nat. Commun.*, vol. 13, no. 1, p. 3224, Jun. 2022, doi: 10.1038/s41467-022-30755-0.

7. Annex

Llistat d'apartats que són massa extensos per incloure dins la memòria i tenen un caràcter autocontingut (per exemple, manuals d'usuari, manuals d'instal·lació, etc.)

Dependent del tipus de treball, és possible que no calgui afegir cap annex.