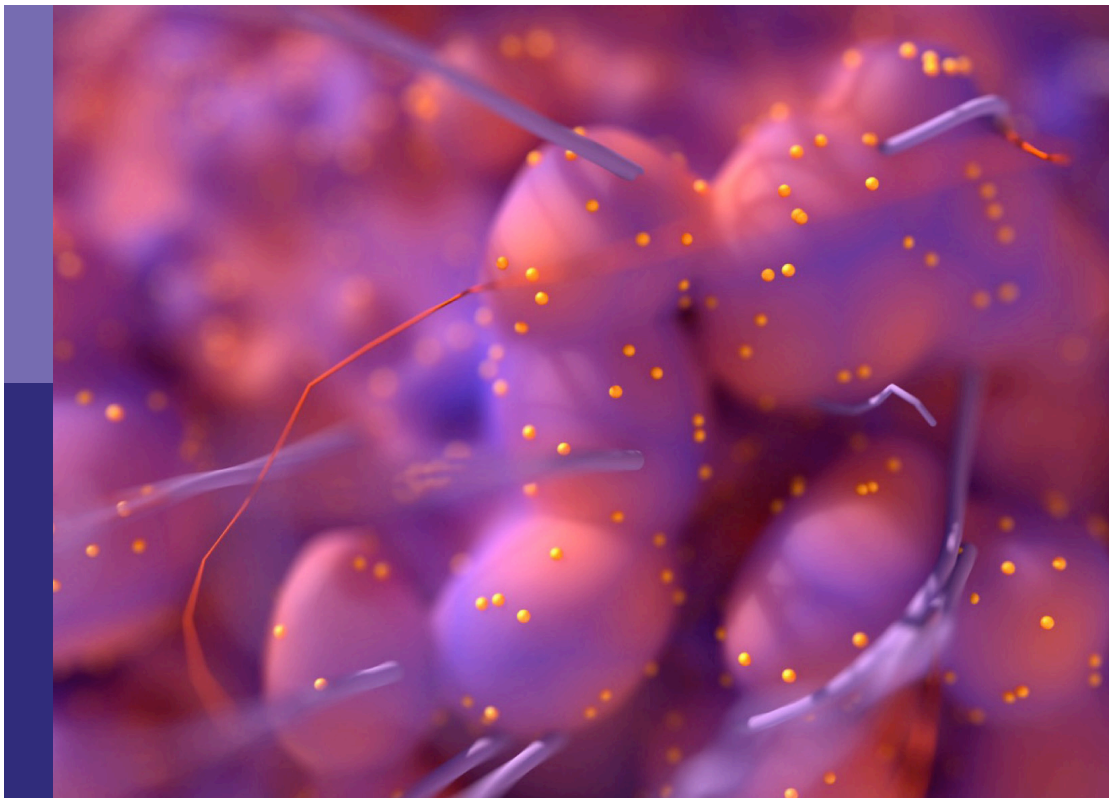
frontiers | Research Topics

Systems biology and single-cell analysis of cancer metabolism and its role in cancer emergent properties

Edited by
Dongya Jia, Yapeng Su, Mingyang Lu and Xuefei Li

Published in
Frontiers in Oncology



ISSN 1664-8714
ISBN 978-2-8325-2696-5
DOI 10.3389/978-2-8325-2696-5

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Systems biology and single-cell analysis of cancer metabolism and its role in cancer emergent properties

Topic editors

Dongya Jia — Center for Cancer Research, National Cancer Institute (NIH), United States

Yapeng Su — Fred Hutchinson Cancer Research Center, United States

Mingyang Lu — Rice University, United States

Xuefei Li — Institute of Synthetic Biology, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences (CAS), China

Citation

Jia, D., Su, Y., Lu, M., Li, X., eds. (2023). *Systems biology and single-cell analysis of cancer metabolism and its role in cancer emergent properties*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-2696-5

Table of contents

frontiers | Frontiers in Oncology

TYPE Editorial
PUBLISHED 31 May 2023
DOI 10.3389/fonc.2023.1217212

OPEN ACCESS

EDITED AND REVIEWED BY
Michael P Lisanti,
University of Salford, United Kingdom

*CORRESPONDENCE
Dongya Jia
dyajia@gmail.com
Xuefei Li
xuefei.li@siat.ac.cn
Yapeng Su
suyapeng.tju@gmail.com

†PRESENT ADDRESS
Dongya Jia, Immunodynamics Group, Laboratory of Integrative Cancer Immunology, Center for Cancer Research, National Cancer Institute, Bethesda, MD, United States

‡These authors have contributed equally to this work

RECEIVED 05 May 2023
ACCEPTED 24 May 2023
PUBLISHED 31 May 2023

CITATION
Jia D, Li X and Su Y (2023) Editorial: Systems biology and single-cell analysis of cancer metabolism and its role in cancer emergent properties.
Front. Oncol. 13:1217212.
doi: 10.3389/fonc.2023.1217212

Editorial: Systems biology and single-cell analysis of cancer metabolism and its role in cancer emergent properties

Dongya Jia[1*†‡], Xuefei Li[2*‡] and Yapeng Su[3*‡]

[1]Center for Theoretical Biological Physics, Rice University, Houston, TX, United States, [2]CAS Key Laboratory for Quantitative Engineering Biology, Shenzhen Institute of Synthetic Biology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China, [3]Program in Immunology & Herbold Computational Biology Program, Fred Hutch Cancer Center, Seattle, WA, United States

KEYWORDS

systems biology, single cell analyses, cancer metabolism, metabolic plasticity, mathematical modeling

Editorial on the Research Topic

Systems biology and single-cell analysis of cancer metabolism and its role in cancer emergent properties

Cancer cells can reprogram their metabolic activities to adapt to heterogeneous tumor microenvironments and to survive various treatments, referred to as metabolic plasticity (1, 2). The past two decades have witnessed our advanced understanding in how cancer cells can acquire multiple metabolic phenotypes, such as glycolysis, oxidative phosphorylation (OXPHOS), the hybrid metabolic phenotype and the metabolically "low-low" phenotype (2, 3). The different cancer metabolic phenotypes are associated with varying cancer metastatic (2) and drug-tolerant potentials (4–6). To decipher the sheer complexity and the multi-faceted nature of cancer metabolism, systems biology approaches that emphasize the interactions between genes, proteins, and metabolites have been developed to create a synergy between theoretical/computational and experimental biology and have led to discoveries at a rapid pace.

Recent advancement in technologies have provided researchers with tools to rigorously quantify cancer metabolism, particularly at the single-cell and the subcellular level. In this Research Topic, Wei et al. reviewed the recently developed technologies for single-cell metabolomics measurement and the integration of multi-omic measurements. To analyze the rapidly increasing cellular metabolism data, Ng et al. provided a summary of the open-source python-based computational toolboxes. Through integrating newly developed technologies with computational toolboxes, we acquired a better understanding of individual metabolic genes and metabolites. Fung et al. used non-invasive Raman-based optical imaging techniques to conduct the 3D spatial and chemometric analyses of triple negative breast cancer (TNBC) cells under the tandem modulation of two key metabolites – insulin and methionine. The authors observed altered *de novo* lipogenesis, 3D lipid droplet morphology, and lipid peroxidation under various methionine and insulin concentrations and verified significant interaction of insulin and methionine metabolism. Altered fatty

acid metabolism is often associated with cancer development (7). Wang et al. provided a review of Acetyl-CoA Carboxylases (ACC), the first rate-limiting enzyme in fatty acid synthesis, covering its structural feature, regulatory mechanism, and roles in cancer development and other diseases. The authors highlighted the regulation of ACC by AMPK and PI3K/Akt/mTOR pathways and its role in post-translational modifications (e.g., acetylation). In addition to ACC, other metabolic genes have been identified for prognosis. Liu et al. showed that the PLA2G2D expression exhibits a positive correlation with immune cell infiltration and favorable immune checkpoint blockade therapy. Chedere et al. showed that nicotinamide adenine dinucleotide (NAD) can classify hepatocellular carcinoma (HCC) patients into three categories depending on their nicotinate phosphoribosyltransferase status, providing a better resolution in understanding the heterogeneity of HCC patients. As we acquire a better understanding of metabolites and metabolic pathways, we may intervene and suppress cancer metabolism. Hou et al. reviewed the current state of arginine deprivation and replenishment therapies for glioma, and the authors emphasized the importance of assessing cancer metabolic state to ensure the effectiveness of deprivation therapy.

In addition to the advancement of experimental technologies and data analysis methods, mathematical modeling approach has been widely used to elucidate the mechanisms underlying cell-fate decision-making in metabolism. Yu and Wang applied the landscape and flux theory to identify the normal and disease states emerging from a core gene regulatory network, during the development of intestinal-gastric cancer. Yu and Wang elucidated the key regulations that are essential for the transition between normal and disease states. To evaluate therapeutic strategies for melanoma treatment, Hodgkinson et al. applied a data-driven multi-dimensional mathematical modeling approach to simulate melanoma response to different therapeutic strategies – combination therapy, continuous therapy, adaptive therapy. Hodgkinson et al. showed that the order in combination therapy matters and that different therapeutic strategies can lead to different tumor heterogeneity. Furthermore, in the computational analysis conducted by Pillai et al., the authors showed that the dedifferentiation of melanoma cells is accompanied by upregulation of mesenchymal genes, but not a concomitant loss of an epithelial program. Interestingly, progression along the mesenchymal axis correlates with the downregulation of OXPHOS, while glycolytic capacity is largely maintained. The dedifferentiation of melanoma cells is closely-linked with its resistance to BRAF inhibitors (8).

Single-cell analysis of cancer metabolism has been a fast-evolving field. With the increasing capacity of acquiring richer single-cell multi-omics data at a better resolution, it is important to leverage the systems biology approaches to make sense of the data and to decode the mechanisms underlying cancer metabolism and its coupling with other cancer hallmarks. With a better understanding of cancer metabolic plasticity, therapies targeting cancer metabolic dependency in principle can be made more effective.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Conflict of interest

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

Publisher's note

References

1. Hanahan D. Hallmarks of cancer: new dimensions. *Cancer Discovery* (2022) 12:31–46. doi: 10.1158/2159-8290.CD-21-1059

2. Jia D, Park JH, Kaur H, Jung KH, Yang S, Tripathi S, et al. Towards decoding the coupled decision-making of metabolism and epithelial-to-mesenchymal transition in cancer. *Br J Cancer* (2021) 124:1902–11. doi: 10.1038/s41416-021-01385-y

3. Jia D, Lu M, Jung KH, Park JH, Yu L, Onuchic JN, et al. Elucidating cancer metabolic plasticity by coupling gene regulation with metabolic pathways. *Proc Natl Acad Sci USA* (2019) 116:3909–18. doi: 10.1073/pnas.1816391116

4. Su Y, Ko ME, Cheng H, Zhu R, Xue M, Wang J, et al. Multi-omic single-cell snapshots reveal multiple independent trajectories to drug tolerance in a melanoma cell line. *Nat Commun* (2020) 11:2345. doi: 10.1038/s41467-020-15956-9

5. Du J, Su Y, Qian C, Yuan D, Miao K, Lee D, et al. Raman-guided subcellular pharmaco-metabolomics for metastatic melanoma cells. *Nat Commun* (2020) 11:4830. doi: 10.1038/s41467-020-18376-x

6. Jia D, Paudel BB, Hayford CE, Hardeman KN, Levine H, Onuchic JN, et al. Drug-tolerant idling melanoma cells exhibit theory-predicted metabolic low-low phenotype. *Front Oncol* (2020) 10:1426. doi: 10.3389/fonc.2020.01426

7. Koundouros N, Poulogiannis G. Reprogramming of fatty acid metabolism in cancer. *Br J Cancer* (2020) 122:4–22. doi: 10.1038/s41416-019-0650-z

8. Su Y, Wei W, Robert L, Xue M, Tsoi J, Garcia-Diaz A, et al. Single-cell analysis resolves the cell state transition and signaling dynamics associated with melanoma drug-induced resistance. *Proc Natl Acad Sci USA* (2017) 114:13679–84. doi: 10.1073/pnas.1712064115

frontiers in Oncology

ORIGINAL RESEARCH
published: 18 October 2021
doi: 10.3389/fonc.2021.755668

Metabolic Molecule PLA2G2D Is a Potential Prognostic Biomarker Correlating With Immune Cell Infiltration and the Expression of Immune Checkpoint Genes in Cervical Squamous Cell Carcinoma

Hong Liu[1,2†], Ruiyi Xu[3†], Chun Gao[1,2], Tong Zhu[1,2], Liting Liu[1,2], Yifan Yang[1,2], Haihong Zeng[4], Yafei Huang[4] and Hui Wang[1,2,3*]*

[1] Cancer Biology Research Center (Key Laboratory of the Ministry of Education), Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, [2] Department of Gynecologic Oncology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, [3] Department of Gynecologic Oncology, Women's Hospital, Zhejiang University School of Medicine, Hangzhou, China, [4] Department of Pathogen Biology, School of Basic Medicine, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

Cervical squamous cell carcinoma (CSCC) is the major pathological type of cervical cancer (CC), the second most prevalent reproductive system malignant tumor threatening the health of women worldwide. The prognosis of CSCC patients is largely affected by the tumor immune microenvironment (TIME); however, the biomarker landscape related to the immune microenvironment of CSCC and patient prognosis is less characterized. Here, we analyzed RNA-seq data of CSCC patients from The Cancer Genome Atlas (TCGA) database by dividing it into high- and low-immune infiltration groups with the MCP-counter and ESTIMATE R packages. After combining weighted gene co-expression network analysis (WGCNA) and differentially expressed gene (DEG) analysis, we found that *PLA2G2D*, a metabolism-associated gene, is the top gene positively associated with immune infiltration and patient survival. This finding was validated using data from The Cancer Genome Characterization Initiative (CGCI) database and further confirmed by quantitative reverse transcription-polymerase chain reaction (qRT-PCR). Finally, multiplex immunohistochemistry (mIHC) was performed to confirm the differential infiltration of immune cells between *PLA2G2D*-high and *PLA2G2D*-low tumors at the protein level. Our results demonstrated that *PLA2G2D* expression was significantly correlated with the infiltration of immune cells, especially T cells and macrophages. More importantly, *PLA2G2D*-high tumors also exhibited higher infiltration of $CD8^+$ T cells inside the tumor region than *PLA2G2D*-low tumors. In addition, PLA2G2D expression was found to be positively correlated with the expression of multiple immune checkpoint genes (ICPs). Moreover, based on other immunotherapy cohort data, *PLA2G2D* high expression is

OPEN ACCESS

Edited by:
Yapeng Su,
Fred Hutchinson Cancer Research Center, United States

Reviewed by:
Haitang Yang,
Shanghai Jiaotong University, China
Jinhui Liu,
Nanjing Medical University, China

***Correspondence:**
Hui Wang
wang71hui@aliyun.com
Yafei Huang
huangy2018@hust.edu.cn

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Cancer Metabolism, a section of the journal Frontiers in Oncology

Received: *09 August 2021*
Accepted: *24 September 2021*
Published: *18 October 2021*

Citation:
Liu H, Xu R, Gao C, Zhu T, Liu L, Yang Y, Zeng H, Huang Y and Wang H (2021) Metabolic Molecule PLA2G2D Is a Potential Prognostic Biomarker Correlating With Immune Cell Infiltration and the Expression of Immune Checkpoint Genes in Cervical Squamous Cell Carcinoma. Front. Oncol. 11:755668. doi: 10.3389/fonc.2021.755668

correlated with increased cytotoxicity and favorable response to immune checkpoint blockade (ICB) therapy. Hence, PLA2G2D could be a novel potential biomarker for immune cell infiltration, patient survival, and the response to ICB therapy in CSCC and may represent a promising target for the treatment of CSCC patients.

Keywords: cervical squamous cell carcinoma, PLA2G2D, tumor immune microenvironment, immune infiltration, metabolism, multiplex immunohistochemistry

INTRODUCTION

Cervical cancer (CC) is the fourth common malignant disease for female with an estimated 604,000 new cases and 342,000 deaths worldwide in 2020 (1). Cervical squamous cell carcinoma (CSCC) and adenocarcinoma are the most common pathological types accounting for approximately 70% and 25% of all CC, respectively (2). Among them, CSCC is mainly related to human papillomavirus (HPV) 16 subtype infection and viral gene integration, whereas adenocarcinoma is often complicated with HPV18 infection (3). Despite the substantial efforts being made in promoting HPV vaccination and early screening, the incidence of CC remains alarming in developing countries (2). In fact, CC is the most frequently diagnosed female cancer in 23 countries according to the latest report (2). Current therapies for CC including surgical treatment, chemotherapy, and radiotherapy have greatly improved the clinical outcome of CC; however, the therapeutic efficacy remains limited for patients with advanced and distant conditions, which estimated a median overall survival of 17 months and 5-year survival of 17% (4, 5). Therefore, there is an urgent need for developing novel therapeutic strategies that can effectively treat these patients (6, 7).

Immunotherapy is one of such strategies that has become a rapidly developing field for cancer treatment including cervical cancer. Through replenishing a sufficient number of expanded autologous T cells that can specifically kill cancer cells, adoptive cell transfer (ACT) has shown great promise in treating CC patients with advance diseases. Two major strategies of ACT, tumor infiltrating lymphocyte (TIL) and T-cell receptor-engineered T cell (TCR-T), were reported to have the objective response rate (ORR) of 44.4% and 50% (8, 9), respectively. Immune checkpoint inhibitors (ICBs), another type of immunotherapy that target programmed death-1(PD-1) receptor or ligand and cytotoxic T-lymphocyte-associated protein 4 (CTLA4), have achieved great success in various kinds of tumors. However, in several clinical trials for advanced cervical cancer patients, the response rate to ICBs was relatively low (10–13). Previous studies have shown that multiple factors are correlated with the efficacy of ICB therapy. Among them, more immune cell infiltration, especially the density and localization of $CD8^+$ T cells in the tumor immune microenvironment (TIME) has been demonstrated to be correlated with a favorable response to ICBs in various cancer types (14, 15). However, the prognostic value and underlying molecular mechanism of immune infiltration in CC with or without immunotherapy remain less characterized.

Here, we utilized the RNA-seq data of CSCC patients from The Cancer Genome Atlas (TCGA) database for the sake of finding biomarkers related to prognosis and immune cell infiltration. To this end, we used two immune scoring algorithms to divide them into two groups with high- and low-immune infiltration. By applying weighted gene co-expression network analysis (WGCNA) and differentially expressed gene (DEG) analysis, phospholipase A_2 Group IID (*PLA2G2D*), an immune- and metabolism-associated molecule, was identified to be the biomarker which is predictive for patient prognosis and immune cell infiltration of CC. Furthermore, five immune-related genes (i.e., *SLAMF6*, *SLAMF1*, *SH2D1A*, *TRAT1*, and *ZNF831*) were found to be co-expressed with *PLA2G2D*. In addition, *PLA2G2D* expression was also found to be positively correlated with the expression of multiple ICP genes. Finally, through a series of bioinformatics analysis and experimental verification approaches at both the transcriptional and protein levels, we proved that PLA2G2D could be a novel biomarker correlating with immune infiltration, especially $CD8^+$ T cells, in CSCC.

MATERIALS AND METHODS

Data Source and Processing

The Cancer Genome Atlas (TCGA) CSCC RNA-seq gene expression matrix based on Illumina platform and phenotype data were downloaded from UCSC Xena (https://xenabrowser.net/datapages/). Two hundred fifty CSCC patients were selected for subsequent analysis according to pathological diagnosis of squamous cell carcinoma. Two types of data including raw counts and fragments per kilo base per million mapped (FPKM) reads were applied. FPKM data were converted to transcripts per kilobase of per million mapped (TPM) data. In addition, the validation cohort dataset also based on Illumina high-throughput platform was downloaded from The Cancer Genome Characterization Initiative (CGCI) database (16). RNA-seq data from two immunotherapy cohorts for melanoma (17) and urothelial cancer (18) were downloaded from a public database. The gene names of all expression matrix were de-annotated through "ClusterProfiler" and "org.Hs.eg.db" package based on R language (V4.1.1) (19).

Estimation of Tumor Immune Infiltration and Cytolytic Activity Score

The Microenvironment Cell Populations-counter (MCP-counter) algorithm was applied to score the level of immune

cell infiltration in tumor, based on which samples were classified into high- and low-immune infiltration groups using the hierarchical clustering method (20). Alternatively, the level of immune and stromal fraction was scored by Estimation of Stromal Immune cells in MAlignant Tumor tissues using Expression data (ESTIMATE) based on log_2-transformed TPM data, and samples were further divided into high- and low-immune infiltration groups equally with the cutoff score set at median (21). In addition, EPIC (22) and quanTIseq (23) algorithm were also utilized to calculate immune cell proportion through TPM expression matrix. CYT score was calculated by the log-transformed geometric mean of GZMA and PRF1 TPM value (24).

Gene Screening by WGCNA

The log_2-transformed TPM value expression matrix was put into WGCNA R package to select immune-associated genes (25). Firstly, 12,417 genes with coefficient of variation (CV) >0.5 were selected for further analysis by WGCNA R package. Of note, five samples were excluded for the abnormal height value in sample dendrogram analysis. The power of β value was set at 3 to construct the topological overlap matrix (TOM). Next, we set the minimum module gene size at 30 and generated 63 gene modules with different colors based on hierarchical clustering method, then merged multiple similar gene modules into one. Finally, the correlation between gene modules and traits was calculated to determine the most relevant module and WGCNA filtered genes were screened by setting the standard of module membership (MM) >0.8 and gene significance (GS) >0.5. Genes network was constructed by Cytoscape (V3.8.0) based on the genes with weight value >0.20 (26).

DIFFERENTIALLY EXPRESSED GENE ANALYSIS AND PATHWAYS ENRICHMENT ANALYSIS

In order to determine the DEGs between the high- and low-immune infiltration groups, the raw read count matrix of 250 CSCC patients from TCGA database was brought into DESeq2 (27). Genes with $|Log_2(FoldChange)|$ >1 and adjusted *P*-value <0.01 were considered as DEGs. ClusterProfiler R package was used to perform Gene Ontology (GO) term enrichment analysis for biological pathway. CGCI data matrix was imported into GSEA software (V4.1.0) for gene set enrichment analysis using the Hallmark gene sets database.

Tissue Collection and Processing

Tumor samples from 18 CSCC patients were collected in Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology. No patient received radiotherapy and chemotherapy before tissue collection. Each tissue was divided into two parts for RNA extraction and formalin-fixed and paraffin-embedded (FFPE), respectively. FFPE tissues were cut into 4-μm-thick sections on slides. This research was approved by the Ethical Committee of Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology (approval number: TJ-IRB2021207).

RNA Extraction and Quantitative Real-Time PCR

Total RNA was extracted and dissolved by RNA Isolation Kit (Vazyme, RC-112-01). cDNA library was obtained using a quantitative real-time reverse transcription-polymerase chain reaction (qRT-PCR) reagent kit (Vazyme, R223-01). The qRT-PCR reaction system contained 1 μg cDNA, 0.4 μl forward primer, 0.4 μl reverse primer, 8.2 μl H_2O, and 10 μl 2× ChamQ Universal SYBR qPCR Master Mix (Vazyme, R223-01). GAPDH was set as internal control for gene quantification. The expression level of each gene was detected at least three times. The primer sequences are listed in **Supplementary Table 1**.

Multiplex Immunohistochemistry

All FFPE slides were deparaffinized by dipping in xylene for 1 h and then rehydrated using the gradient ethanol method. After deparaffinization and rehydration, all slides were put into boiled AR6 retrieval solution for heat-induced epitope retrieval in a microwave for 15 min. Endogenous peroxidase was eliminated with 3% H_2O_2 for 15 min. Slides were cooled to room temperature followed by washing with 1× Tris-buffered saline-Tween 20 (TBST) buffer. Then, Opal 7-color manual IHC kit (Akoya Biosciences, NEL811001KT) was used to stain several markers in a single FFPE slide. To block non-specific protein binding sites, slides were incubated with blocking buffer (Antgene, ANT041) at room temperature for 10 min. Subsequently, FFPE tissue slides were incubated with primary antibody at 37°C for 30 min or 4°C overnight, HRP-labeled secondary antibody for 10 min, and Opal fluorescein for 10 min, successively. To stain four markers in a single slide, four rounds of the above staining procedure were performed with indicated primary antibody and matched Opal fluorescein pairs (**Supplementary Table 2**). The slides were always washed three times between each step with 1× TBST buffer. After four rounds, samples were stained for cellular DNA with 4′,6-diamidino-2-phenylindole (DAPI) (1:10, Servicebio, G1012-10ML) for 10 min followed by mounting with Fluoromount-G (SouthernBiotech, 0100-01) and preserved in the dark at 4°C. All slides were scanned by Verctra V3.0 System, and ×4 and ×20 objective lens were used to acquire low-power and high-power images, respectively. Images were analyzed with Phenochart V1.0 software and inForm V2.4 software for tissue component segmentation of tumor, stroma, and glass regions, respectively. To identify cell phenotypes, a threshold of fluorescein intensity was set for each marker. The ratio of the targeted cell counts and all cell counts in a certain region was used to calculate the percentage of immune cells in each region. The ratio of cell counts and area (cell/mm^2) was utilized to reveal the density of a certain cell population.

Survival Analysis and Statistical Analysis

Survival information of TCGA CSCC patients was downloaded from the UCSC Xena database. Patients who survived less than 30 days were excluded from this analysis. Kaplan–Meier survival

analysis was applied to educe the correlation between 5-year overall survival and the expression level of indicated genes. Based on the relationship with survival time, survival status, and minus PLA2G2D TPM value, ROC curve analysis was applied by survivalROC R package. Statistical analysis was performed with R statistical package. Wilcoxon rank-sum test was employed for comparison between two groups. *P*-values <0.05 were considered significant: $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, and $^{****}P < 0.0001$.

RESULTS

Classification of CSCC by Immune Cell Infiltration

The RNA-seq data and clinical information of a CSCC cohort including 250 patients were downloaded from TCGA database for bioinformatics analysis. Firstly, the MCP-counter, a widely used algorithm, was used to evaluate the level of immune infiltration for each sample, according to which CSCC patients were divided into high- and low-immune infiltration clusters based on the unsupervised stratification method (**Figure 1A**). Subsequently, the 5-year overall survival of the high- and low-immune infiltration groups was compared. The results demonstrated that the high-immune infiltration cluster determined by the MCP-counter had significantly higher survival than the low-immune infiltration cluster in CSCC (**Figure 1B**). Next, ESTIMATE, another algorithm which can similarly stratify CSCC patients into high- and low-immune fraction clusters by a median cutoff (**Figure 1C**), was chosen to verify this finding. Again, the high-immune infiltration cluster classified by ESTIMATE was found to have higher survival than the low-immune infiltration cluster in CSCC (**Figure 1D**). Of note, ESTIMATE can also classify CSCC patients into high- and low-stromal fraction clusters (**Figure 1C**). However, no difference was found between the two stromal fraction-stratified clusters in terms of patient survival (**Figure 1E**, $P = 0.22$). Together, our results demonstrated that immune infiltration, but not stromal fraction, was correlated with patient survival in CSCC. Moreover, the other two algorithms, EPIC and quanTIseq, were used to calculate several immune cell proportions compared with the MCP-counter and ESTIMATE. Moreover, the immune cell infiltration results of EPIC and quanTIseq were consistent with the MCP-counter and ESTIMATE (**Figure S1**).

Construction of Weighted Gene Co-Expression Network and Genes Filtering Analysis

CSCC RNA-seq data from TCGA was imported into WGCNA R package to identify gene modules containing similarly expressed genes, especially immune-related genes. Sample clustering and trait distribution are shown in **Figure 2A**. Power value was set at $\beta = 3$ to build a scale-free network (**Figure 2B**). Gene modules were calculated and the cutoff height was set at 0.5 to merge similar gene modules into one (**Figure 2C**). This analysis resulted in 63 merged dynamic gene modules with each module

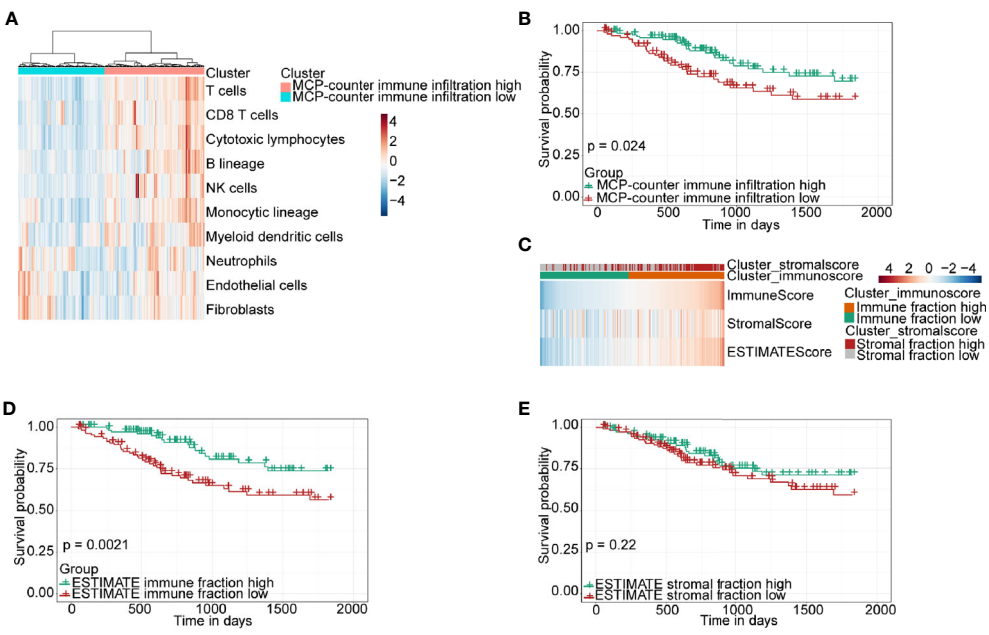


FIGURE 1 | Immune classification and survival analysis of 250 CSCC patients selected from TCGA database. **(A)** Heatmap of TME composition as defined by the MCP-counter algorithm. **(B)** Kaplan–Meier survival curve for immune infiltration-high and infiltration-low groups determined by the MCP-counter. **(C)** Heatmap showing immune and stromal fraction by the ESTIMATE algorithm. **(D, E)** Kaplan–Meier survival curve for patients with high and low fractions of immune cells and stromal cells classified by the ESTIMATE algorithm.

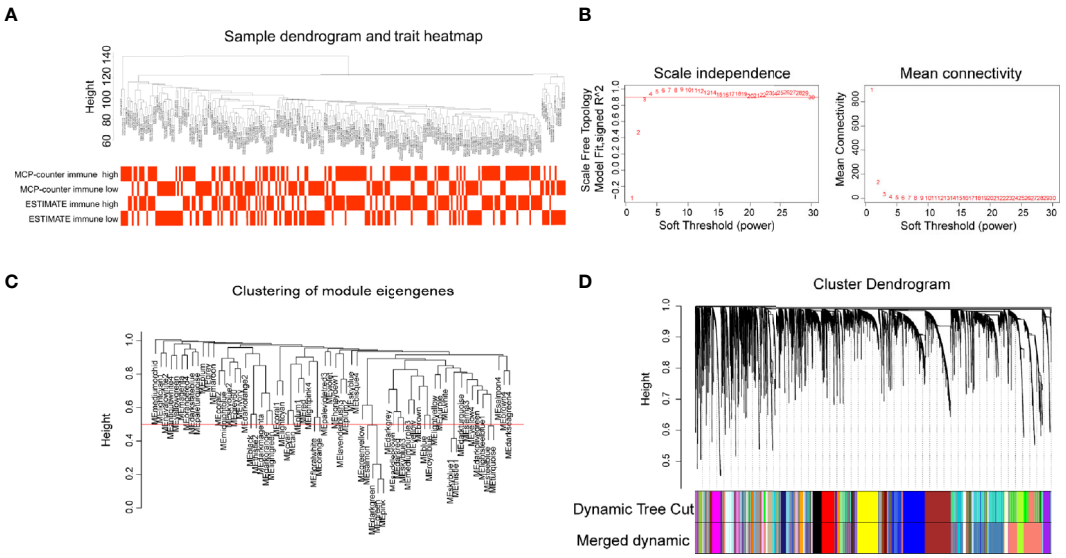


FIGURE 2 | Gene modules construction by weighted gene co-expression network analysis (WGCNA). **(A)** Sample dendrogram and trait heatmap. **(B)** The relationship of soft threshold with scale independence and mean connectivity. **(C)** Clustering of module eigengenes, height set at 0.5 to merge modules. **(D)** Cluster dendrogram showing the hierarchical cluster tree for the identified co-expression gene modules with different colors.

containing dozens to more than a thousand genes (**Figure 2D**, **Supplementary Table 3**).

In order to identify immune-related modules and genes, we constructed the relationship between gene modules and immune infiltration traits (**Figure 3A**). We focused on the brown module which was most relevant to immune infiltration or immune fraction determined by the MCP-counter and ESTIMATE, respectively. To further identify genes in the brown module that are most correlated with the MCP-counter immune infiltration and ESTIMATE immune fraction, the screening criteria of MM >0.8 and GS >0.5 were applied, through which 77 genes (brown module *vs.* MCP-counter immune infiltration) and 73 genes (brown module *vs.* ESTIMATE immune fraction) were filtered (**Figures 3B, C**). Next, Venn plot was used to reveal the overlapped filtered genes of these two groups. As shown in **Figure 3D**, all 73 brown module genes contained in the ESTIMATE high-immune fraction group were found to be presented in the MCP-counter high-immune infiltration group as well. Finally, GO pathways enrichment analysis for the shared 73 WGCNA genes showed that these genes were mostly involved in immune-associated pathways, such as T-cell activation, lymphocyte differentiation, and regulation of T-cell activation (**Figure 3E**).

PLA2G2D Is Positively Correlated With Immune infiltration and Patient Survival

To further explore hub genes from these 73 genes, DEG analysis was performed using DESeq2 R package. DEGs were determined between the high- and low-immune infiltration groups determined by the MCP-counter, through which 968 upregulated genes and 323 downregulated genes were found in the high- and low-immune infiltration groups, respectively (**Figure 4A**). Similarly, 1,060 upregulated genes and 650 downregulated genes were found in the high- and low-immune fraction groups determined by ESTIMATE, respectively (**Figure 4B**). Interestingly, all 73 WGCNA filtered genes were presented in these two upregulated gene sets. Notably, *PLA2G2D*, a metabolism-related gene, was identified as the most differentially expressed gene among the 73 genes, as determined by the $\log_2$ (FoldChange) value of these genes (**Table 1** and **Figures 4A, B**).

Next, we sought to identify the genes that were co-expressed with *PLA2G2D*. The co-expression network between *PLA2G2D* and other genes was constructed by using the criteria of weight >0.20 (**Figure 4C** and **Table 2**). To gain insight into the function of these co-expressed genes, GO analysis of the biological process was conducted and the most co-expressed genes were enriched in immune-related pathways (**Figure 4D**). Of note, several genes with weight >0.3 including *SLAMF6*, *SLAMF1*, *SH2D1A*, *TRAT1*, and *ZNF831* were most co-expressed with *PLA2G2D* (**Figure 4E**). Importantly, Kaplan–Meier survival curve analysis showed that the expression level of *PLA2G2D* was positively correlated with patient survival (**Figure 4F**, $P = 0.00017$). Finally, ROC curve was constructed to demonstrate the prognostic ability of *PLA2G2D* in TCGA-CSCC cohort. The AUCs of PLA2G2D TPM minus value for 2, 3, and 5 years were 0.773, 0.683, and 0.639, respectively (**Figure 4G**).

Bioinformatic Validation Using Data From the CGCI Database and PCR Validation Using Tumor Tissues

To validate the correlation between overexpressed *PLA2G2D* and more immune cell infiltration demonstrated in **Figure 1**, we analyzed RNA-seq data of CSCC from the CGCI database that

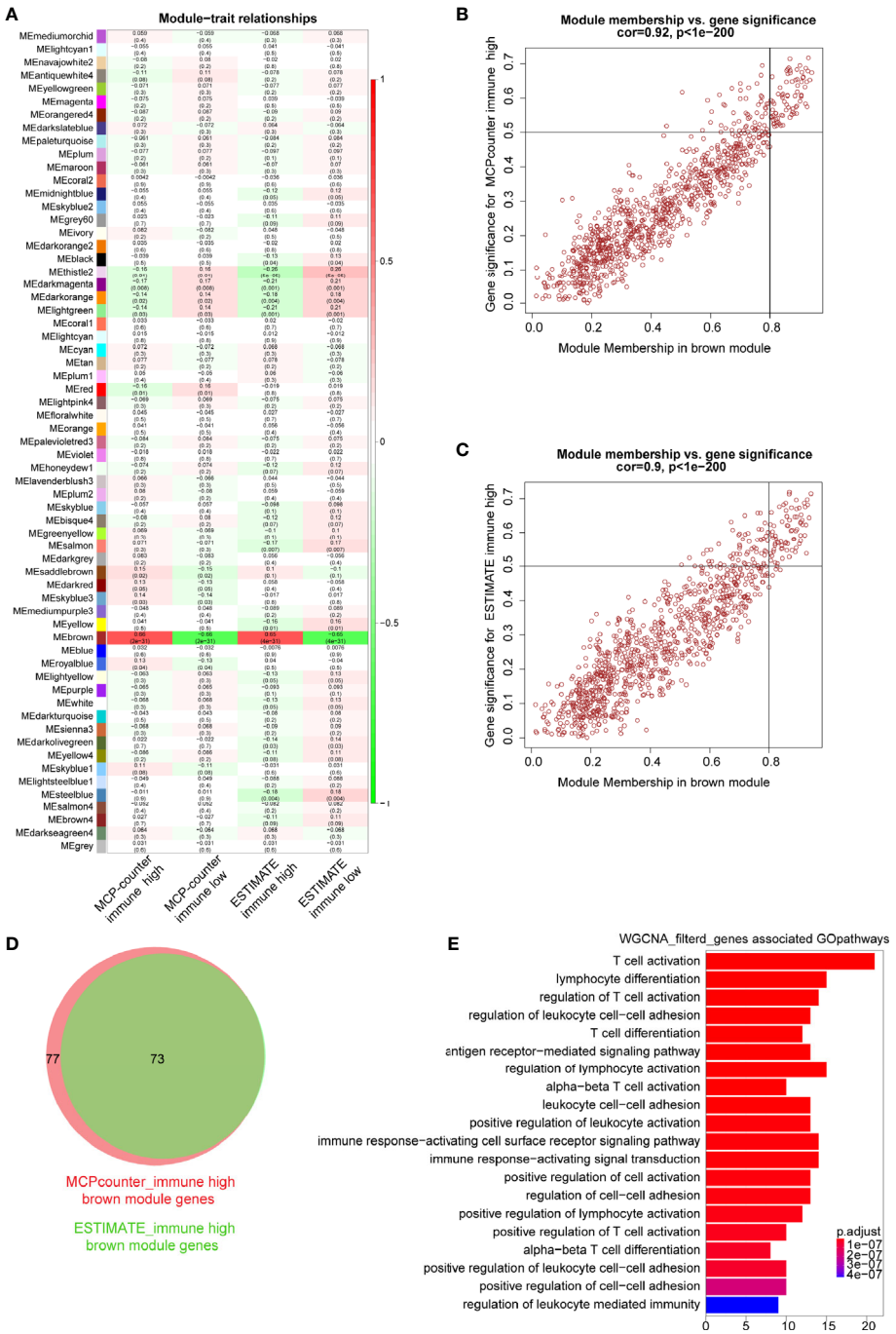


FIGURE 3 | Key modules and genes identified based on module–trait relationship. **(A)** The relationship between co-expression modules and traits; each grid includes the degree of correlation and *P*-value. **(B, C)** Scatter plots showing the relationship of module membership (MM) in brown module, with gene significance (GS) for high-immune infiltration determined by the MCP-counter and ESTIMTE algorithms, for which MM >0.8 and GS >0.5 were set as gene filtered standard, respectively. **(D)** Venn plot showing common filtered genes identified in the high-immune infiltration groups determined by the MCP-counter and ESTIMATE. **(E)** GO analysis for WGCNA filtered genes in the biological process.

was also sequenced on an Illumina high-throughput sequence platform. The immune characteristics of patients in the *PLA2G2D* high- and low-expression clusters are illustrated in **Figure 5A** by the analysis with the four algorithms. Overexpressed *PLA2G2D* was bound up with more immune cell infiltration such as T cells, B cells, NK cells, and myeloid dendritic cells (**Figure 5B**). Similarly, ImmuneScore, StromalScore, and ESTIMATEScore were higher in the

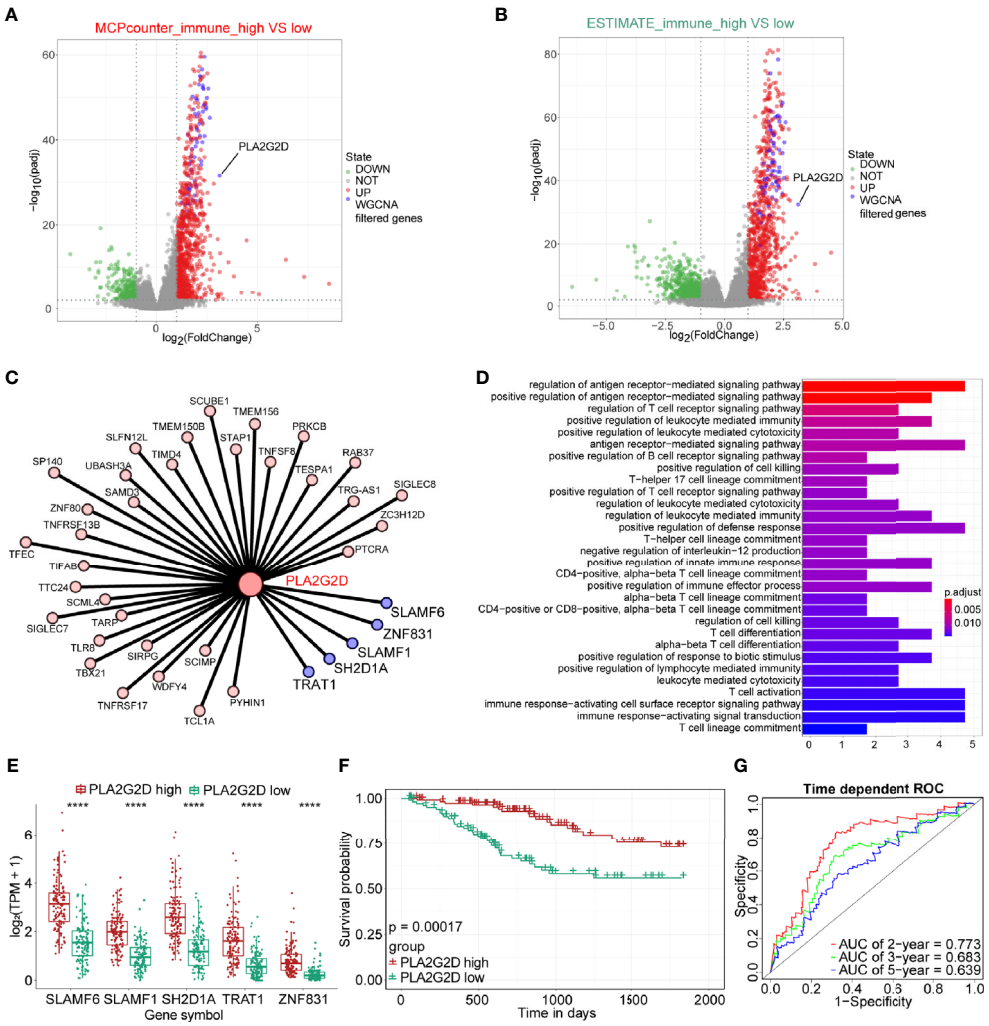


FIGURE 4 | The expression of PLA2G2D is positively correlated with immune infiltration and patient survival. Volcano plot for DEGs identified for high- and low-immune infiltration groups based on the MCP-counter **(A)** and ESTIMATE **(B)**. The green and red dots represented the significantly downregulated and upregulated genes, respectively, and the gray dots represented the undifferentiated genes; the purple dots showed the WGCNA filtered genes, with $|Log_2(FoldChange)|$ >1 and adjusted *P*-value <0.01. **(C)** Cytoscape network plot showed the relationship between *PLA2G2D* and other co-expressed genes with weight value >0.2. The purple nodes showing the co-expressed genes with weight value >0.3. **(D)** GO pathways enrichment analysis for genes co-expressed with *PLA2G2D*. **(E)** Histogram showing the relationship between the expression levels of *PLA2G2D* and other five co-expressed genes, ****$P < 0.0001$. **(F)** Kaplan–Meier survival curve plot showing the relationship between PLA2G2D expression level and patient survival. **(G)** Time-dependent ROC curve analysis for minus *PLA2G2D* TPM value at 2-, 3-, and 5-year cutoffs.

PLA2G2D high-expression cluster than in the *PLA2G2D* low-expression cluster (**Figure 5C**). $CD8^+$ T cells, macrophages, and B cells were also higher in the PLA2G2D-high group based on quanTIseq and EPIC algorithms (**Figure 5D**). GSEA analysis showed that the *PLA2G2D* high-expression cluster was enriched in several immune-associated pathways such as allograft rejection, interferon gamma response, interferon alpha response, complement, and oxidative phosphorylation pathways (**Figure S2A**), whereas *PLA2G2D* low-expression cluster was enriched in hypoxia, apical junction, glycolysis, mitotic spindle, and angiogenesis pathways (**Figure S2B**). We also validated whether *PLA2G2D* was co-expressed with other genes as indicated by the aforementioned WGCNA analysis. Our analysis demonstrated that the expression level of *PLA2G2D* was also positively associated with *SLAMF6*, *SLAMF1*, *SH2D1A*, *TRAT1*, and *ZNF831* when the CGCI database was used (**Figure 5E**). To further validate these results using fresh-isolated clinical samples, 18 cervical squamous carcinoma samples were collected and the corresponding clinical data are shown in **Table 3**. Total RNA per sample was isolated and qRT-PCR was performed to examine the gene expression levels of *PLA2G2D* and other five genes (**Table 4**). Our cohort was classified into *PLA2G2D* high- and low-expression groups according to the median expression level of *PLA2G2D*. Similarly, the expression levels of four immune-related genes were significantly higher in the

TABLE 1 | Top 20 WGCNA filtered genes determined by DEG analysis of high- and low-immune infiltration groups classified by the MCP-counter and ESTIMATE algorithms.

Order	Genes	MCP-counter immune high *vs.* low		Genes	ESTIMATE immune high *vs.* low	
		Log2FoldChange	*P*adj		Log2FoldChange	*P*adj
1	*PLA2G2D*	3.120339747	2.46E-32	*PLA2G2D*	3.129951852	3.07E-33
2	*GZMK*	2.633961849	6.52E-46	*CRTAM*	2.594071877	3.75E-59
3	*TRAT1*	2.565309728	8.77E-53	*TRAT1*	2.543058375	7.56E-53
4	*CRTAM*	2.493534917	1.38E-50	*GZMK*	2.526206854	1.77E-41
5	*SCML4*	2.461018797	1.85E-48	*SH2D1A*	2.46209327	2.63E-65
6	*GPR174*	2.451822301	5.68E-48	*SCML4*	2.456859681	1.78E-49
7	*ZNF831*	2.379191341	4.28E-49	*LINC01857*	2.45160968	1.25E-36
8	*LINC00861*	2.373752175	1.36E-41	*SIRPG*	2.414521914	1.33E-64
9	*SLAMF6*	2.370123928	2.92E-60	*FCRL3*	2.36471631	1.57E-43
10	*GPR171*	2.345873252	6.25E-49	*TBX21*	2.359189375	9.61E-61
11	*SH2D1A*	2.329233049	1.14E-52	*SLAMF6*	2.334677345	2.09E-59
12	*LY9*	2.298409977	2.26E-47	*LY9*	2.327353913	1.71E-50
13	*SIRPG*	2.296040614	5.79E-53	*TLR8*	2.325276706	5.66E-45
14	*TBX21*	2.2919668	1.58E-53	*TIFAB*	2.310890798	6.39E-41
15	*FCRL6*	2.28625033	9.57E-45	*LINC00861*	2.299844149	8.68E-39
16	*TTC24*	2.274489399	3.53E-40	*CD3G*	2.277158384	2.71E-61
17	*CD3G*	2.254820652	2.70E-57	*SP140*	2.272328994	4.91E-79
18	*EOMES*	2.21785005	9.37E-36	*ZNF831*	2.271868811	1.21E-43
19	*FCRL3*	2.206529839	7.39E-35	*GPR174*	2.269780489	3.33E-39
20	*LINC01857*	2.184645997	2.81E-26	*LINC00426*	2.223943929	5.04E-61

PLA2G2D high-expression group, except for *ZNF831*, which could hardly be detected by qRT-PCR in both groups (**Figure 5F** and **Table 4**).

More Immune Cell Infiltration Determined by Multiplex Immunohistochemistry in *PLA2G2D* High-Expression Samples

Next, the multiplex immunohistochemistry (mIHC) method was performed to validate whether immune cell infiltration was correlated with *PLA2G2*D expression in our cohort of 18 CSCC clinical samples. As our aforementioned bioinformatic analysis indicated that T cells and macrophages were the most prevalent cells among the infiltrated immune cells positively correlated with *PLA2G2D* expression, and previous studies have shown that the number and percentage of T cells and macrophages were more than the other kind of immune cells in cervical cancer (28, 29), we chose these two cells along with tumor cells to determine the percentage and density of individual immune cell populations in each slide. For this purpose, the mIHC panel (**Supplementary Table 2**) was designed to simultaneously stain several markers including CD3 (for all T cells), CD8 (for cytotoxic T cells), CD68 (for macrophages), PCK (for tumor cells), and DAPI (for nucleic identification). Of note, $CD3^+CD8^-$ TILs were defined as $CD4^+$ TILs. Our results indicated that $CD3^+$ T cells, including both $CD4^+$ and $CD8^+$ TILs, and $CD68^+$ macrophages were more abundant in the *PLA2G2D*-high group than in the *PLA2G2D*-low group, as shown by representative merged images at low magnification

TABLE 2 | List of genes co-expressed with *PLA2G2D* with weight value >0.20.

***PLA2G2D* co-expressed genes**	**Weight**	***PLA2G2D* co-expressed genes**	**Weight**
SLAMF6	0.316522209	*TLR8*	0.224623497
SH2D1A	0.313611578	*ZC3H12D*	0.218210964
TRAT1	0.308061949	*TIFAB*	0.218179744
ZNF831	0.306000355	*SIGLEC8*	0.217411803
SLAMF1	0.300894868	*SLFN12L*	0.215513035
TESPA1	0.295561708	*TTC24*	0.214768858
PYHIN1	0.29293532	*TNFRSF13B*	0.213697367
WDFY4	0.279759397	*TIMD4*	0.211705633
UBASH3A	0.274771647	*TCL1A*	0.211489274
SAMD3	0.270901098	*PTCRA*	0.208689904
SCML4	0.270390467	*TARP*	0.208642553
PRKCB	0.26625189	*TMEM156*	0.207822287
SIRPG	0.262965895	*ZNF80*	0.206555702
TFEC	0.260430877	*TMEM150B*	0.20555541
TBX21	0.25827151	*STAP1*	0.204952535
SP140	0.256619486	*RAB37*	0.204841733
TRG-AS1	0.250638121	*TNFRSF17*	0.203896246
TNFSF8	0.246175611	*SCUBE1*	0.203302527
SCIMP	0.240444493	*SIGLEC7*	0.200210069

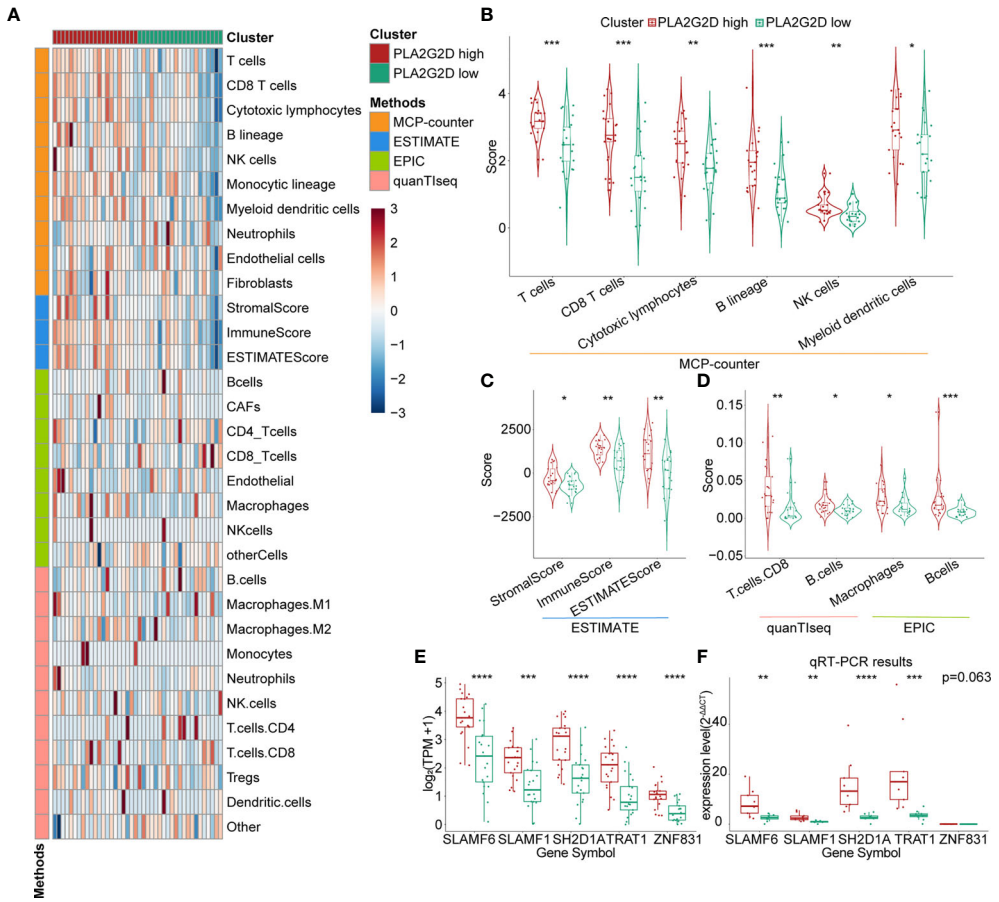


FIGURE 5 | Further validation of the relationship between *PLA2G2D* expression and immune infiltration. **(A)** Heatmap for PLA2G2D high- and low-expression groups based on the MCP-counter, ESTIMATE, EPIC, and quanTIseq algorithms using data from the CGCI database. **(B)** Violin plot showing the differential scores of six immune cell types determined by the MCP-counter algorithm. **(C)** Violin plot showing the differential StromalScore, ImmuneScore, and ESTIMATEScore determined by the ESTIMATE algorithm. **(D)** Violin plot showing the differential immune cell types determined by quanTIseq and EPIC algorithms. **(E, F)** Histograms showing the relationship between the expression levels of PLA2G2D and WGCNA filtered five co-expressed genes using data from the CGCI database **(D)** and freshly isolated clinical samples **(E)**. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, and ****$P < 0.0001$.

TABLE 3 | Clinical characteristics of patients from our CSCC cohort.

Classification	Total	Low *PLA2G2D* expression	High *PLA2G2D* expression
Age			
<40	3	2	1
40–49	8	5	3
≥50	7	2	5
FIGO stage			
I–IIA	10	6	4
IIB–IV	8	3	5
Vascular invasion			
Positive	6	4	2
Negative	12	5	7
Lymphatic metastasis			
Positive	8	3	5
Negative	10	6	4

(**Figures 6A, D**) and high magnification (**Figures 6B, E**) and by single spectrum images for each marker (**Figures 6C, F**). After comparing *PLA2G2D*-high and *PLA2G2D*-low groups for the compositions of individual cell populations in different regions (**Figure 7A**) and the percentages of individual cell populations in each sample (**Figure S3**), we found that the percentage values of $CD3^+$ TILs, $CD8^+$ TILs, and macrophages in the stromal region and all regions of the *PLA2G2D* high-expression group were significantly higher compared with that of the *PLA2G2D* low-expression group (**Figure 7B**). Importantly, in the tumor region, $CD8^+$ TILs were the only cell population that were more frequently observed in the *PLA2G2D* high-expression group (**Figure 7B**).

Next, cell density (cell counts/mm^2 area) was analyzed in the tumor and stromal region, respectively. Similar to the comparison of cell percentage, cell densities of $CD3^+$ TILs, $CD8^+$ TILs, and macrophages were significantly higher in the stromal region of the *PLA2G2D* high-expression group than those of the *PLA2G2D* low-expression group. Moreover, in the tumor region, only $CD8^+$ TIL density in the *PLA2G2D* high-expression group was higher compared with that in the *PLA2G2D* low-expression group (**Figure 7C**).

TABLE 4 | qRT-PCR ΔCT value of *PLA2G2D* and other five co-expressed genes.

Samples	Genes						*PLA2G2D* group
	SLAMF6	*SLAMF1*	*SH2D1A*	*TRAT1*	*ZNF831*	*PLA2G2D*	
SCC11	16.24	16.25	16.25	15.5	20.23	20.06	Low
SCC5	10.59	11.81	10.48	10.78	19.92	15.8	Low
SCC4	10.24	12.03	10.47	9.93	19	14.61	Low
SCC17	11.46	11.97	10.33	10.05	21.07	13.19	Low
SCC6	9.5	11.68	9.89	9.58	18.09	13.15	Low
SCC10	10.55	11.59	10.23	9.28	19.01	12.99	Low
SCC9	9.88	10.95	9.57	9.57	19.72	12.62	Low
SCC1	9.75	11.23	9.34	8.78	18.81	12.39	Low
SCC7	10.15	11.59	10.04	9.81	17.72	12.09	Low
SCC2	10.44	11.47	9.3	8.97	19.58	11.87	High
SCC18	9.48	10.78	7.89	7.82	19.14	11.8	High
SCC15	10.08	10.74	8.79	8.85	17.65	11.69	High
SCC3	8.84	10.11	7.67	7.53	19.53	11.11	High
SCC16	8.77	9.89	8.64	8.32	18.3	11.1	High
SCC14	8.09	10.35	8.11	7.22	16.35	10.72	High
SCC13	8.45	9.15	7.41	7.39	18.16	10.23	High
SCC8	7.91	11.25	7.06	6.22	16.8	9.99	High
SCC12	7.36	9.31	6.31	5.81	17.32	9.54	High

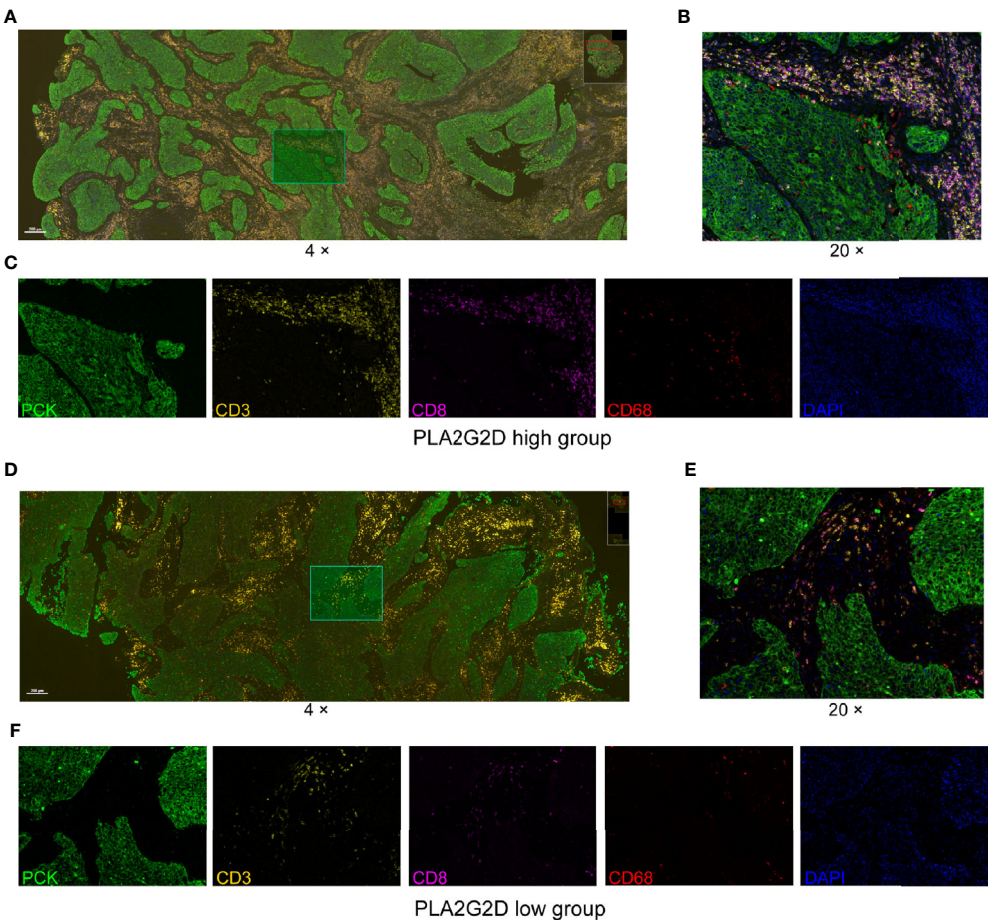


FIGURE 6 | mIHC staining for CD3, CD8, CD68, PCK, and DAPI in cervical squamous cancer specimens with high **(A–C)** and low expression **(D–F)** of the *PLA2G2D* gene. **(A, D)** Merged mIHC images at low magnification (×4). **(B, E)** Merged mIHC images at high magnification (×20) indicating the same filed in **(A, D)**. **(C, F)** Single spectral images indicating the same filed in **(B, E)**.

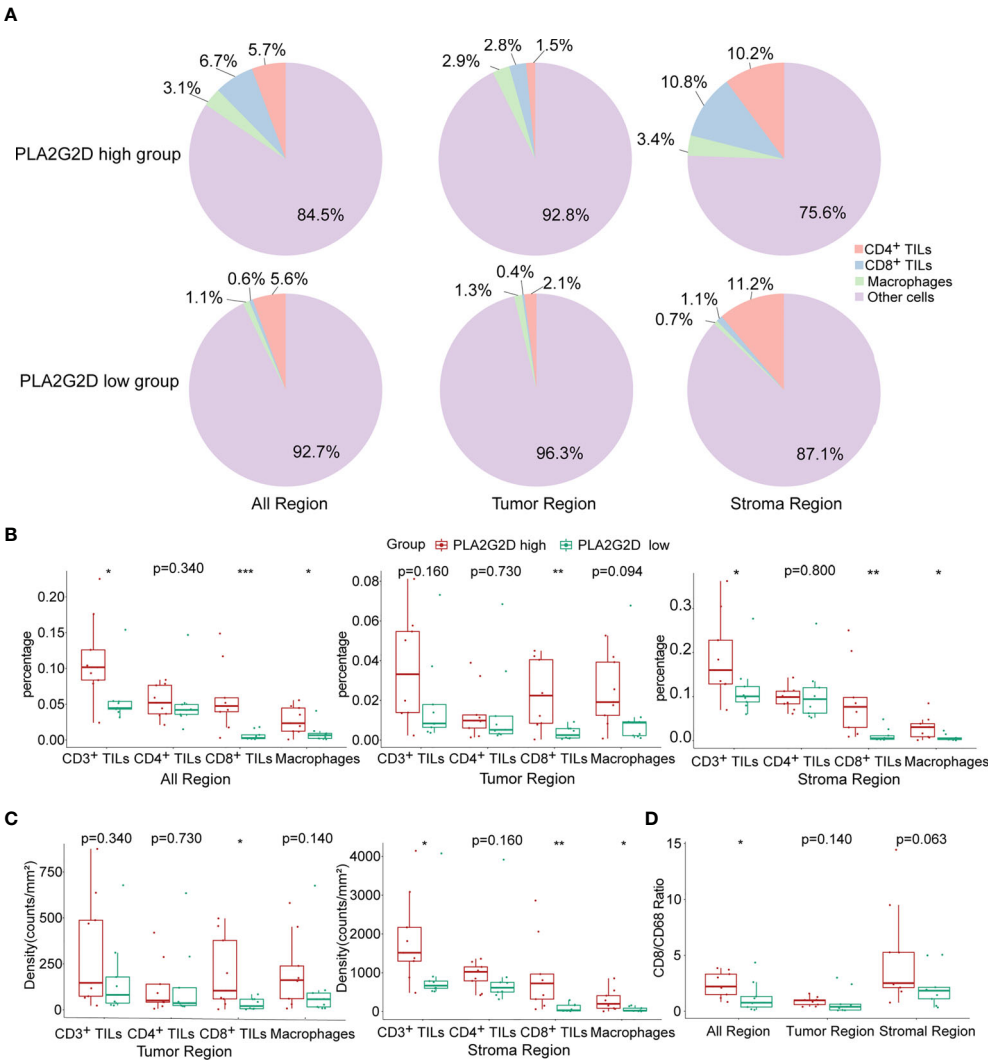


FIGURE 7 | Statistical analysis for mIHC staining. **(A)** Pie plots showing the composition of immune cells within different regions in *PLA2G2D* high- and low-expression groups. **(B)** Histograms showing the comparison of the percentage of individual immune cell populations between *PLA2G2D* high- and low- expression groups in different regions. **(C)** Histograms showing the comparison of the density of individual immune cell populations between *PLA2G2D* high- and low-expression groups in different regions. **(D)** Histogram showing the comparison of the CD8/CD68 ratio between *PLA2G2D* high- and low-expression groups in different regions. $^{*}P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$.

Tumor-resident macrophages were previously considered as a major cell population in suppressing the function of cytotoxic T cells and were reported to be associated with poor prognosis of cancer (30). Therefore, we assessed the CD8/CD68 ratio between these two groups. The ratio was higher in all areas in the *PLA2G2D* high-expression group (**Figure 7D**).

Positive Correlation Between the Expression of *PLA2G2D* and ICPs

The expression of immune checkpoint genes [e.g., *PDCD1* (*PD1*), *CD274* (*PDL1*), *CTLA4*, *LAG3*, and *HAVCR2* (*TIM3*)] has been utilized in predicting the response of patients to ICB therapy in a variety of cancers including CC. To evaluate the possible relationship between the expression of *PLA2G2D* and these ICPs, correlation analysis was performed using transcriptomic data from TCGA database. Our analysis indicated that the expression of ICP genes was significantly higher in the *PLA2G2D* high-expression group compared with that in the *PLA2G2D* low-expression group. Through Pearson correlation analysis, we found that these common ICP genes were positively correlated with *PLA2G2D* expression (**Figures 8A, C**). Similar findings were noted when using the CGCI validation cohort except for *CD274*, the expression of which was comparable in the *PLA2G2D* low-expression and high-expression groups and was not correlated with the expression of *PLA2G2D*. This discrepancy may be attributable to the small sample size in this cohort (**Figures 8B, D**).

Positive Correlation Between the Expression of *PLA2G2D* and the Response to ICB Therapy in Immunotherapy Cohorts

ICB therapy has been performed in patients with advanced cervical cancer, but transcriptional data for these cohorts are not accessible. To validate the relationship between *PLA2G2D* and the response to ICB treatment, we analyzed the transcriptome data from immunotherapy cohorts of melanoma and urothelial carcinoma. In the melanoma anti-PD1 therapy cohort, we selected patients who have not previously received immunotherapy for enrollment analysis and divided them based on *PLA2G2D* expression at pretreatment stage. The percentage of compete response or partial response (CR/PR) was higher in *PLA2G2D*-high patients (**Figure 9A**). PLA2G2D expression level of CR/PR patients also showed a higher tendency compared with stable disease (SD) and progressive disease (PD) patients; however, these differences failed to reach statistical significance (**Figure 9B**). In this cohort, 9 and 10 pairs of pre- and on-treatment patients were assigned to the PLA2G2D-high and PLA2G2D-low cluster, respectively. Before ICB therapy, *PLA2G2D*-high patients show higher ICP expression and CYT score (**Figures 9C–E**). Interestingly, after ICB therapy, *PLA2G2D*-high patients also have higher ICP expression and CYT score (**Figures 9C, F, G**). Importantly, compared with pretreatment patients, on-treatment patients have stronger cytotoxicity and higher ICP expression after anti-PD-1 therapy regardless of *PLA2G2D*-high or *PLA2G2D*-low groups (**Figure 9C**). Furthermore, in the urothelial carcinoma anti-PDL1 therapy cohort, the percentage of CR/PR patients was higher in *PLA2G2D*-high patients, which is similar to that in melanoma immunotherapy cohort (**Figure 9H**). CR/PR patients showed a higher tendency of *PLA2G2D* expression level than SD and PD patients (**Figure 9I**). Unfortunately, only pretreatment transcriptome data are available for this cohort, thus preventing us from performing similar analysis for on-treatment data. Nevertheless, *PLA2G2D*-high expression was also accompanied by high expression of ICPs despite no statistical significance (**Figures 9J, K**), while expression level of cytotoxicity markers (*GZMA* and *PRF1*) and the CYT score were significantly higher in the *PLA2G2D*-high group (**Figures 9J–L**).

DISCUSSION

The tumor microenvironment (TME) is a complex milieu that comprises diverse cell populations including malignant cells, immune cells, stromal cells, and other cell types (31). Each cell population can impact others through releasing soluble molecules and/or by direct cell–cell interaction. As a crucial part of the tumor microenvironment, TIME plays a pivotal role

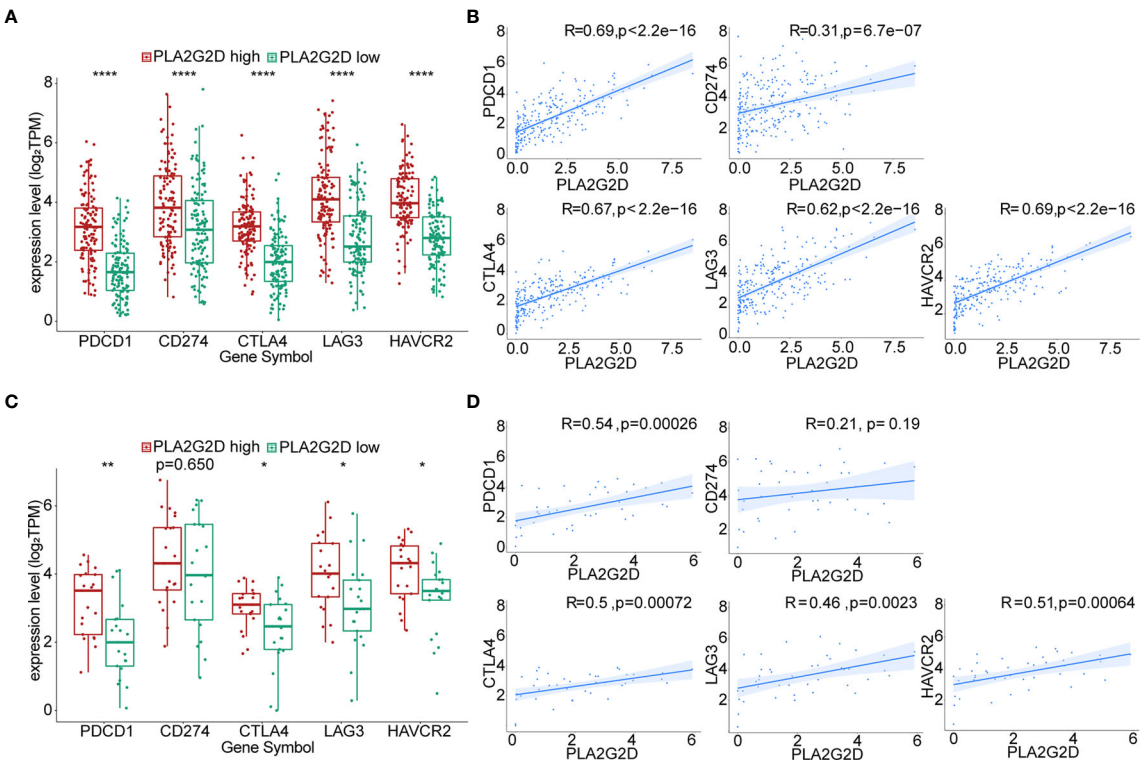


FIGURE 8 | Correlation of the expression level of *PLA2G2D* and ICPs. Histograms showing the comparison of ICP expression level between *PLA2G2D*-high and *PLA2G2D*-low groups in TCGA **(A)** and CGCI database **(C)**. Pearson correlation analysis for the expression level of *PLA2G2D* and ICPs in TCGA **(B)** and CGCI **(D)** database. *$P < 0.05$, **$P < 0.01$, and ****$P < 0.0001$.

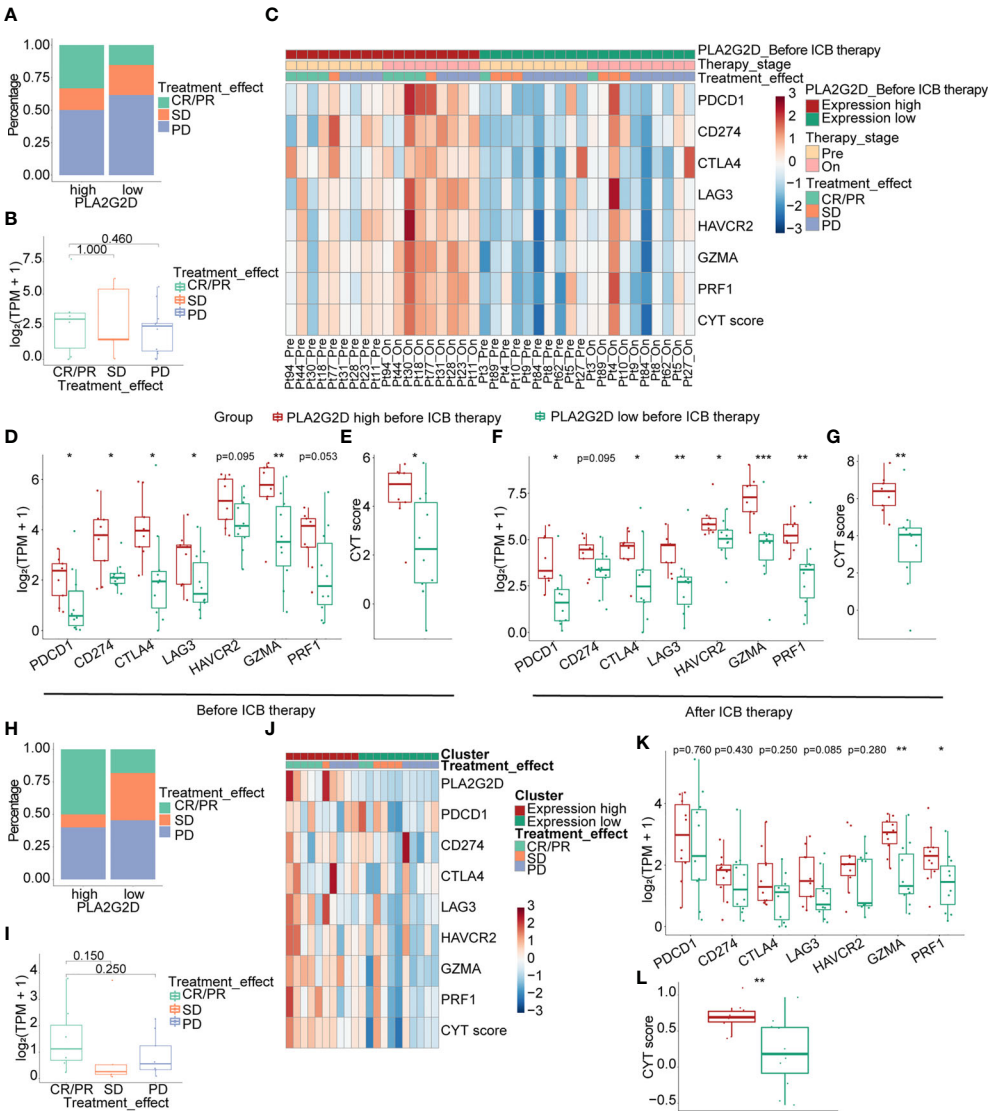


FIGURE 9 | Correlation of the expression level of *PLA2G2D* and response to ICB treatment in melanoma **(A–G)** and urothelial carcinoma immunotherapy cohorts **(H–L)**. **(A, B)** Bar and boxplot for the relationship between *PLA2G2D* expression and response to ICB treatment in melanoma cohort. **(C)** Heatmap of the relationship of ICPs, cytotoxic genes, and CYT score with different *PLA2G2D* expression levels and response to treatment in paired pre- and on-treatment patients. Boxplots showing the differential expression of ICPs, cytotoxic genes, and CYT scores between patients with different *PLA2G2D* expression levels before ICB therapy **(D, E)** and after ICB therapy **(F, G)**. **(H, I)** Bar and boxplot for the relationship between *PLA2G2D* expression and response to ICB treatment in urothelial carcinoma cohort. **(J)** Heatmap of the relationship of ICPs, cytotoxic genes, and CYT score with different *PLA2G2D* expression levels. **(K, L)** Boxplots showing the differential expression of ICPs, cytotoxic genes, and CYT score between patients with different *PLA2G2D* expression levels. $^{*}P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$.

in tumorigenesis and the clinical outcomes of cancer patients (32). Accumulating lines of evidence have suggested that a certain type of TIME characterized by more immune cell infiltration is indicative of a better prognosis and a favorable response to ICBs (15, 33). Furthermore, several factors from tumor cells themselves, immune cells, and stromal cells were reported to be implicated in shaping the landscape of TIME (34–36). However, for CSCC, the connection between TIME (especially immune infiltration) and cancer prognosis and the possible underpinning molecular mechanism are far from elucidated. This limitation greatly prevents the identification of new therapeutic targets that can be utilized to improve patient prognosis through transforming the TIME with poor-immune infiltration to high-immune infiltration. Therefore, this study was designed to address this issue. Through bioinformatics analyses, we identified a key immune- and metabolism-associated molecule PLA2G2D that was significantly related to better prognosis of CSCC patients and more immune cell infiltration.

Metabolism-associated molecules are critical components of TIME (37). Metabolic reprograming of tumor cells may promote tumor growth and metastasis directly or indirectly

through influencing the functions of a variety of immune cells, especially T cells (38–40), and the shape of TIME. Phospholipase A2 (PLA_2) proteins are a group of lipid metabolism-associated molecules that can catalyze the hydrolysis of the sn-2 position of membrane glycerophospholipids to release unsaturated fatty acid and lysophospholipid, thereby playing a pivotal role in multiple biochemical processes including inflammatory response. While both cytosolic and secreted forms of PLA_2 have been identified, more than one-third of the PLA_2 enzymes belong to the secreted PLA_2 ($sPLA_2$) family, which have both pro- and anti-inflammatory functions as a result of the production of diverse lipid mediators (41, 42). As a member of $sPLA_2$, group IID PLA_2 (encoded by *PLA2G2D*), is broadly expressed in human body, such as the spleen, lymph nodes, squamous epithelium, and colorectal cancer tissue (43–45). In the current investigation, PLA2G2D was also found to be expressed in CSCC specimens at the transcriptional level (**Figure 5E**). Previous studies have also shown that PLA2G2D was preferentially expressed in lymphoid tissue-resident dendritic cells and macrophages and implicated in anti-inflammation response in an array of inflammation-related conditions including contact hypersensitivity (43), viral infection-associated inflammation (46), experimental encephalomyelitis and colitis (47), and contact dermatitis and psoriasis (48). The role of *Pla2g2d* in cancer has been investigated as well. Using both *Pla2g2d*-deficient and *Pla2g2d*-transgenic mouse models, Miki et al. found that *Pla2g2d* could facilitate the development of chemical-induced skin cancer accompanied by macrophage polarization toward M2 phenotype and decreased number of cytotoxic T cells (48). The cancer-promoting effect of PLA2G2D has also been suggested in human. Recently, *PLA2G2D* was identified as one of the high-risk genes for colorectal and rectum adenocarcinoma that were negatively correlated with patient survival. However, *PLA2G2D* expression was also found to be positively correlated with immune infiltration and better prognosis in head and neck squamous cell carcinoma and breast cancer in human (49, 50), which is consistent with our current findings in CSCC. The opposite effects of PLA2G2D in different cancer types may result from the distinct microenvironments and different downstream lipid mediators presented in each cancer type, which in turn result in either an elevated or dampened inflammatory responses, thereby leading to the distinct clinical outcomes reported in these investigations (48).

In this study, the relationship between *PLA2G2D* overexpression and more immune cell infiltration in CSCC was validated through two approaches. Firstly, another dataset from the CGCI database was used for validation at the transcriptional level. Secondly, mIHC, a multispectral microscopy technique that can reveal the TIME profile on FFPE slides, was applied to determine the major immune cell populations infiltrated in tumor including $CD4^+$ T cells, $CD8^+$ T cells, and macrophages, which were subsequently used to confirm their relationship with *PLA2G2D* expression. Our results demonstrated that *PLA2G2D* high-expressed samples had more immune cells infiltrated in the stroma region. More importantly, $CD8^+$ TILs were the only cell population that were more frequently found inside the tumor region in *PLA2G2D* high-expressed samples, suggesting that PLA2G2D could participate in the recruitment of $CD8^+$ TILs into the cancer nests, either directly or indirectly. The presence of tumor-associated macrophages is often associated with tumorigenesis, immune suppression, and poor prognosis of patients (35, 51). Indeed, macrophages can directly suppress $CD8^+$ T-cell proliferation *in vitro* (30, 52–54). Therefore, we compared the CD8/CD68 ratio between *PLA2G2D*-high and *PLA2G2D*-low samples. Again, a significantly higher ratio of CD8/CD68 was found in *PLA2G2D*-high samples compared with that in *PLA2G2D*-low samples. Taken together, these results not only confirmed the aforementioned relationship between immune infiltration and *PLA2G2D* expression, but also revealed the spatial characteristics of immune infiltration in *PLA2G2D*-high samples.

The mechanism by which PLA2G2D affects immune infiltration in CSCC is currently unclear, and lipid mediators downstream of PLA2G2D might contribute. For example, the enzyme activity of PLA2G2D could result in the production of leukotriene B4, a potent chemotactic molecule with the capacity to recruit neutrophils, monocytes/macrophages, $CD8^+$ cytotoxic T lymphocyte, and Th17 cells (55). However, molecules other than these lipid mediators may be involved as well. To explore the possible candidates, genes that co-expressed with *PLA2G2D* were determined by WGCNA analysis, which resulted in the identification of the top 5 genes, namely, *SLAMF6*, *SLAMF1*, *SH2D1A*, *TRAT1*, and *ZNF831*. These genes were successfully validated using another dataset from the CGCI database and confirmed by qRT-PCR using fresh-isolated CSCC samples, with the exception of *ZNF831*, the expression of which was extremely low and no statistical significance was found between *PLA2G2D*-high and *PLA2G2D*-low samples. It is noteworthy that these five genes are all associated with immunity. *SLAMF1 (CD150)* and *SLAMF6* belong to the signaling lymphocytic activation molecule (SLAM) family and express on several immune cells including T cells, B cells, and NK cells (56). Previous studies have shown that the expressions of *SLAMF1* in cervical cancer and *SLAMF6* in liver cancer were closely related to tumor infiltration of immune cells and patient prognosis (57, 58). The *SH2D1A* gene encodes SH2 domain-containing protein 1A, which is often known as SLAM-associated protein. SLAM can transduce the tyrosine signaling pathway to promote interferon-γ production in the presence of SH2D1A (56, 59). *TRAT1* encodes the tripartite motif (TRIM) protein which is essential for T-cell activation and positively correlated with the survival of patients with metastatic melanoma (59). *ZNF831* is a transcription factor gene and was also significantly correlated with immune cell infiltration in triple negative breast cancer (60). Together, in light of the role of the aforementioned co-expressed genes in various immune-related pathways, PLA2G2D may interact with these co-expressed genes indirectly or directly, thereby influencing immune cell infiltration.

Immune cell infiltration, along with PDL1 expression and tumor mutation burden (TMB), has been frequently used as independent biomarkers in predicting the response of patients to ICB therapy in multiple cancer types (61). Therefore, the positive correlation between immune cell infiltration and the expression

of *PLA2G2D* in CSCC identified in this study indicates that *PLA2G2D* expression may also represent another predictive biomarker for ICB in CSCC. In support of this notion, we found that the levels of ICP genes, another independent biomarker for ICB therapy, were markedly higher in the *PLA2G2D* high-expression group than those in the *PLA2G2D* low-expression group at the transcriptional level. Furthermore, a positive correlation between the expression of *PLA2G2D* and ICP genes was noted in both TCGA and CGCI cohorts. Finally, we directly examined the predictive value of *PLA2G2D* expression for patient response to ICB therapy in melanoma and urothelial carcinoma immunotherapy cohorts, given the lack of transcriptional data in CSCC immunotherapy cohorts. Before ICB therapy, *PLA2G2D*-high patients have higher ICP expression and CYT score. After ICB therapy, the expression level of ICPs and CYT score were further increased compared with the pretreatment stage, as indicated by paired analysis. Hence, *PLA2G2D*-high patients have higher cytotoxicity and favorable response to ICB treatment. Taken together, *PLA2G2D* expression may represent a novel biomarker with a better predictive power for ICB therapy.

In conclusion, through integrating bioinformatics and experimental verification, we demonstrated that the metabolic molecule PLA2G2D was positively correlated with immune infiltration and patient prognosis in CSCC, suggesting that PLA2G2D could be a novel prognosis biomarker for CSCC patients. Furthermore, PLA2G2D might be a promising biomarker for the evaluation of immune infiltration situation across different patients and represent another independent predictive biomarker for ICB therapy in CSCC. However, the underpinning mechanism regarding how PLA2G2D works in TME is unclear and warrants further investigations.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethical Committee of Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology (approval number: TJ-IRB2021207). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

HW, YH, HL, and RX conceived and designed the study. HL did bioinformatics analysis and wrote the manuscript. HL and RX did the PCR and mIHC experiments and analyzed the data. CG designed the mIHC panel. HL, CG, TZ, LL, YY, and HZ handled the clinical samples and collected clinical data. HW and YH had supervised this project and contributed to writing and revision of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by grants from the National Natural Science Foundation of China (NSFC No. 81772786, No. 81830074) and the Research Funds from Tongji Hospital (No. 20185BJRC004, No. 2019BJRC008).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.755668/full#supplementary-material

Supplementary Figure 1 | Similarity of MCP-counter, ESTIMATE, EPIC and quanTIseq in characterizing immune infiltrations. Heatmap and boxplot showing the similarity between MCP-counter, EPIC and quanTIseq **(A, C)**. Heatmap and boxplot showing the similarity between ESTIMATE, EPIC and quanTIseq **(B, D)**. Pie plots showing the percentage of several kinds of immune cells based on EPIC **(E)** and quanTIseq **(F)** algorithms. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, and $^{****}P < 0.0001$.

Supplementary Figure 2 | GSEA analysis for RNA-seq data from the CGCI database. **(A)** Several pathways enriched in *PLA2G2D* high-expression cluster. **(B)** Several pathways enriched in *PLA2G2D* low-expression cluster.

Supplementary Figure 3 | Frequencies of the major cell populations for each sample calculated by mIHC method. Composition of different kind of cells for each sample in all region **(A)**, stromal region **(B)** and tumor region **(C)**.

REFERENCES

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* (2021) 71:209–49. doi: 10.3322/caac.21660
2. Cohen PA, Jhingran A, Oaknin A, Denny L. Cervical Cancer. *Lancet* (2019) 393:169–82. doi: 10.1016/S0140-6736(18)32470-X
3. Hu Z, Zhu D, Wang W, Li W, Jia W, Zeng X, et al. Genome-Wide Profiling of HPV Integration in Cervical Cancer Identifies Clustered Genomic Hot Spots and a Potential Microhomology-Mediated Integration Mechanism. *Nat Genet* (2015) 47:158–63. doi: 10.1038/ng.3178
4. Van Meir H, Kenter GG, Burggraaf J, Kroep Jr, Welters MJ, Melief CJ, et al. The Need for Improvement of the Treatment of Advanced and Metastatic Cervical Cancer, the Rationale for Combined Chemo-Immunotherapy. Anti-Cancer Agents in Medicinal Chemistry. *Anticancer Agents Med Chem* (2014) 14:190–203. doi: 10.2174/1871520611313660372
5. National Cancer Institute. Cancer Stat Facts: Cervical Cancer. *Cancer Statistics*. (2021). Available at https://seer.cancer.gov/statfacts/html/cervix.html [Accessed October 9, 2021]
6. Tewari KS, Sill MW, Long HJ 3rd, Penson RT, Huang H, Ramondetta LM, et al. Improved Survival With Bevacizumab in Advanced Cervical Cancer. *N Engl J Med* (2014) 370:734–43. doi: 10.1056/NEJMoa1309748

7. Tewari KS, Monk BJ. New Strategies in Advanced Cervical Cancer: From Angiogenesis Blockade to Immunotherapy. *Clin Cancer Res* (2014) 20:5349–58. doi: 10.1158/1078-0432.CCR-14-1099
8. Nagarsheth NB, Norberg SM, Sinkoe AL, Adhikary S, Meyer TJ, Lack JB, et al. TCR-Engineered T Cells Targeting E7 for Patients With Metastatic HPV-Associated Epithelial Cancers. *Nat Med* (2021) 27:419–25. doi: 10.1038/s41591-020-01225-1
9. Jazaeri AA, Zsiros E, Amaria RN, Artz AS, Edwards RP, Wenham RM, et al. Safety and Efficacy of Adoptive Cell Transfer Using Autologous Tumor Infiltrating Lymphocytes (LN-145) for Treatment of Recurrent, Metastatic, or Persistent Cervical Carcinoma. *J Clin Oncol* (2019) 37:2538–8. doi: 10.1200/JCO.2019.37.15_suppl.2538
10. Lheureux S, Butler MO, Clarke B, Cristea MC, Martin LP, Tonkin K, et al. Association of Ipilimumab With Safety and Antitumor Activity in Women With Metastatic or Recurrent Human Papillomavirus-Related Cervical Carcinoma. *JAMA Oncol* (2018) 4:e173776. doi: 10.1001/jamaoncol.2017.3776
11. Petitprez F, Meylan M, De Reynies A, Sautes-Fridman C, Fridman WH. The Tumor Microenvironment in the Response to Immune Checkpoint Blockade Therapies. *Front Immunol* (2020) 11:784. doi: 10.3389/fimmu.2020.00784
12. Chung HC, Ros W, Delord JP, Perets R, Italiano A, Shapira-Frommer, et al. Efficacy and Safety of Pembrolizumab in Previously Treated Advanced Cervical Cancer: Results From the Phase II KEYNOTE-158 Study. *J Clin Oncol* (2019) 37:1470–8. doi: 10.1200/JCO.18.01265
13. Frenel JS, Le Tourneau C, O'neil B, Ott PA, Piha-Paul SA, Gomez-Roca C, et al. Safety and Efficacy of Pembrolizumab in Advanced, Programmed Death Ligand 1-Positive Cervical Cancer: Results From the Phase Ib KEYNOTE-028 Trial. *J Clin Oncol* (2017) 35:4035–41. doi: 10.1200/JCO.2017.74.5471
14. Tumeh PC, Harview CL, Yearley JH, Shintaku IP, Taylor EJ, Robert L, et al. PD-1 Blockade Induces Responses by Inhibiting Adaptive Immune Resistance. *Nature* (2014) 515:568–71. doi: 10.1038/nature13954
15. Herbst RS, Soria JC, Kowanetz M, Fine GD, Hamid O, Gordon MS, et al. Predictive Correlates of Response to the Anti-PD-L1 Antibody MPDL3280A in Cancer Patients. *Nature* (2014) 515:563–7. doi: 10.1038/nature14011
16. Gagliardi A, Porter VL, Zong Z, Bowlby R, Titmuss E, Namirembe C, et al. Analysis of Ugandan Cervical Carcinomas Identifies Human Papillomavirus Clade-Specific Epigenome and Transcriptome Landscapes. *Nat Genet* (2020) 52:800–10. doi: 10.1038/s41588-020-0673-7
17. Riaz N, Havel JJ, Makarov V, Desrichard A, Urba WJ, Sims JS, et al. Tumor and Microenvironment Evolution During Immunotherapy With Nivolumab. *Cell* (2017) 171:934–49.e916. doi: 10.1016/j.cell.2017.09.028
18. Snyder A, Nathanson T, Funt SA, Ahuja A, Buros Novik J, Hellmann MD, et al. Contribution of Systemic and Somatic Factors to Clinical Response and Resistance to PD-L1 Blockade in Urothelial Cancer: An Exploratory Multi-Omic Analysis. *PLoS Med* (2017) 14:e1002309. doi: 10.1371/journal.pmed.1002309
19. Yu G, Wang LG, Han Y, He QY. Clusterprofiler: An R Package for Comparing Biological Themes Among Gene Clusters. *OMICS* (2012) 16:284–7. doi: 10.1089/omi.2011.0118
20. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, et al. Estimating the Population Abundance of Tissue-Infiltrating Immune and Stromal Cell Populations Using Gene Expression. *Genome Biol* (2016) 17:218. doi: 10.1186/s13059-016-1070-5
21. Yoshihara K, Shahmoradgoli M, Martinez E, Vegesna R, Kim H, Torres-Garcia W, et al. Inferring Tumour Purity and Stromal and Immune Cell Admixture From Expression Data. *Nat Commun* (2013) 4:2612. doi: 10.1038/ncomms3612
22. Racle J, De Jonge K, Baumgaertner P, Speiser DE, Gfeller D. Simultaneous Enumeration of Cancer and Immune Cell Types From Bulk Tumor Gene Expression Data. *Elife* (2017) 6:e26476. doi: 10.7554/eLife.26476
23. Finotello F, Mayer C, Plattner C, Laschober G, Rieder D, Hackl H, et al. Molecular and Pharmacological Modulators of the Tumor Immune Contexture Revealed by Deconvolution of RNA-Seq Data. *Genome Med* (2019) 11:34. doi: 10.1186/s13073-019-0638-6
24. Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and Genetic Properties of Tumors Associated With Local Immune Cytolytic Activity. *Cell* (2015) 160:48–61. doi: 10.1016/j.cell.2014.12.033
25. Langfelder P, Horvath S. WGCNA: An R Package for Weighted Correlation Network Analysis. *BMC Bioinf* (2008) 9:559. doi: 10.1186/1471-2105-9-559
26. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res* (2003) 13:2498–504. doi: 10.1101/gr.1239303
27. Love MI, Huber W, Anders S. Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data With Deseq2. *Genome Biol* (2014) 15:550. doi: 10.1186/s13059-014-0550-8
28. Zou R, Gu R, Yu X, Hu Y, Yu J, Xue X, et al. Characteristics of Infiltrating Immune Cells and a Predictive Immune Model for Cervical Cancer. *J Cancer* (2021) 12:3501–14. doi: 10.7150/jca.55970
29. Wang J, Li Z, Gao A, Wen Q, Sun Y. The Prognostic Landscape of Tumor-Infiltrating Immune Cells in Cervical Cancer. *BioMed Pharmacother* (2019) 120:109444. doi: 10.1016/j.biopha.2019.109444
30. Ruffell B, Coussens LM. Macrophages and Therapeutic Resistance in Cancer. *Cancer Cell* (2015) 27:462–72. doi: 10.1016/j.ccell.2015.02.015
31. Lei X, Lei Y, Li JK, Du WX, Li RG, Yang J, et al. Immune Cells Within the Tumor Microenvironment: Biological Functions and Roles in Cancer Immunotherapy. *Cancer Lett* (2020) 470:126–33. doi: 10.1016/j.canlet.2019.11.009
32. Su D, Wu G, Xiong R, Sun X, Xu M, Mei Y, et al. Tumor Immune Microenvironment Characteristics and Their Prognostic Value in Non-Small-Cell Lung Cancer. *Front Oncol* (2021) 11:634059. doi: 10.3389/fonc.2021.634059
33. Binnewies M, Roberts EW, Kersten K, Chan V, Fearon DF, Merad M, et al. Understanding the Tumor Immune Microenvironment (TIME) for Effective Therapy. *Nat Med* (2018) 24:541–50. doi: 10.1038/s41591-018-0014-x
34. Li J, Byrne KT, Yan F, Yamazoe T, Chen Z, Baslan T, et al. Tumor Cell-Intrinsic Factors Underlie Heterogeneity of Immune Cell Infiltration and Response to Immunotherapy. *Immunity* (2018) 49:178–93.e177. doi: 10.1016/j.immuni.2018.06.006
35. Denardo DG, Ruffell B. Macrophages as Regulators of Tumour Immunity and Immunotherapy. *Nat Rev Immunol* (2019) 19:369–82. doi: 10.1038/s41577-019-0127-6
36. Khalili JS, Liu S, Rodriguez-Cruz TG, Whittington M, Wardell S, Liu C, et al. Oncogenic BRAF(V600E) Promotes Stromal Cell-Mediated Immunosuppression via Induction of Interleukin-1 in Melanoma. *Clin Cancer Res* (2012) 18:5329–40. doi: 10.1158/1078-0432.CCR-12-1632
37. Cassim S, Pouyssegur J. Tumor Microenvironment: A Metabolic Player That Shapes the Immune Response. *Int J Mol Sci* (2019) 21:157. doi: 10.3390/ijms21010157
38. Kolb D, Kolishetti N, Surnar B, Sarkar S, Guin S, Shah AS, et al. Metabolic Modulation of the Tumor Microenvironment Leads to Multiple Checkpoint Inhibition and Immune Cell Infiltration. *ACS Nano* (2020) 14:11055–66. doi: 10.1021/acsnano.9b10037
39. Leone RD, Zhao L, Englert JM, Sun IM, Oh MH, Sun IH, et al. Glutamine Blockade Induces Divergent Metabolic Programs to Overcome Tumor Immune Evasion. *Science* (2019) 366:1013–21. doi: 10.1126/science.aav2588
40. Kishton RJ, Sukumar M, Restifo NP. Metabolic Regulation of T Cell Longevity and Function in Tumor Immunotherapy. *Cell Metab* (2017) 26:94–109. doi: 10.1016/j.cmet.2017.06.016
41. Lindbom J, Ljungman AG, Lindahl M, Tagesson C. Increased Gene Expression of Novel Cytosolic and Secretory Phospholipase A(2) Types in Human Airway Epithelial Cells Induced by Tumor Necrosis Factor-Alpha and IFN-Gamma. *J Interferon Cytokine Res* (2002) 22:947–55. doi: 10.1089/10799900260286650
42. Lambeau G, Lazdunski M. Receptors for a Growing Family of Secreted Phospholipases A2. *Trends Pharmacol Sci* (1999) 20:162–70. doi: 10.1016/S0165-6147(99)01300-0
43. Miki Y, Yamamoto K, Taketomi Y, Sato H, Shimo K, Kobayashi T, et al. Lymphoid Tissue Phospholipase A2 Group IID Resolves Contact Hypersensitivity by Driving Antiinflammatory Lipid Mediators. *J Exp Med* (2013) 210:1217–34. doi: 10.1084/jem.20121887
44. Mounier CM, Wendum D, Greenspan E, Flejou JF, Rosenberg DW, Lambeau G. Distinct Expression Pattern of the Full Set of Secreted Phospholipases A2 in Human Colorectal Adenocarcinomas: Spla2-III as a Biomarker Candidate. *Br J Cancer* (2008) 98:587–95. doi: 10.1038/sj.bjc.6604184
45. Haas U, Podda M, Behne M, Gurrieri S, Alonso A, Fürstenberger G, et al. Characterization and Differentiation-Dependent Regulation of Secreted

Phospholipases A2 in Human Keratinocytes and in Healthy and Psoriatic Human Skin. *J Invest Dermatol* (2005) 124:204–11. doi: 10.1111/j.0022-202X.2004.23513.x
46. Vijay R, Hua X, Meyerholz DK, Miki Y, Yamamoto K, Gelb M, et al. Critical Role of Phospholipase A2 Group IID in Age-Related Susceptibility to Severe Acute Respiratory Syndrome-CoV Infection. *J Exp Med* (2015) 212:1851–68. doi: 10.1084/jem.20150632
47. Von Allmen CE, Schmitz N, Bauer M, Hinton HJ, Kurrer MO, Buser RB, et al. Secretory Phospholipase A2-IID Is an Effector Molecule of CD4+CD25+ Regulatory T Cells. *Proc Natl Acad Sci USA* (2009) 106:11673–8. doi: 10.1073/pnas.0812569106
48. Miki Y, Kidoguchi Y, Sato M, Taketomi Y, Taya C, Muramatsu K, et al. Dual Roles of Group IID Phospholipase A2 in Inflammation and Cancer. *J Biol Chem* (2016) 291:15588–601. doi: 10.1074/jbc.M116.734624
49. Ye Z, Zou S, Niu Z, Xu Z, Hu Y. A Novel Risk Model Based on Lipid Metabolism-Associated Genes Predicts Prognosis and Indicates Immune Microenvironment in Breast Cancer. *Front Cell Dev Biol* (2021) 9:691676. doi: 10.3389/fcell.2021.691676
50. Xiong Y, Si Y, Feng Y, Zhuo S, Cui B, Zhang Z. Prognostic Value of Lipid Metabolism-Related Genes in Head and Neck Squamous Cell Carcinoma. *Immun Inflamm Dis* (2021) 9:196–209. doi: 10.1002/iid3.379
51. Mantovani A, Marchesi F, Malesci A, Laghi L, Allavena P. Tumour-Associated Macrophages as Treatment Targets in Oncology. *Nat Rev Clin Oncol* (2017) 14:399–416. doi: 10.1038/nrclinonc.2016.217
52. Doedens AL, Stockmann C, Rubinstein MP, Liao D, Zhang N, Denardo DG, et al. Macrophage Expression of Hypoxia-Inducible Factor-1 Alpha Suppresses T-Cell Function and Promotes Tumor Progression. *Cancer Res* (2010) 70:7465–75. doi: 10.1158/0008-5472.CAN-10-1439
53. Denardo DG, Brennan DJ, Rexhepaj E, Ruffell B, Shiao SL, Madden SF, et al. Leukocyte Complexity Predicts Breast Cancer Survival and Functionally Regulates Response to Chemotherapy. *Cancer Discov* (2011) 1:54–67. doi: 10.1158/2159-8274.CD-10-0028
54. Ruffell B, Chang-Strachan D, Chan V, Rosenbusch A, Ho CM, Pryer N, et al. Macrophage IL-10 Blocks CD8+ T Cell-Dependent Responses to Chemotherapy by Suppressing IL-12 Expression in Intratumoral Dendritic Cells. *Cancer Cell* (2014) 26:623–37. doi: 10.1016/j.ccell.2014.09.006
55. Jiang X, Nicolls MR, Tian W, Rockson SG. Lymphatic Dysfunction, Leukotrienes, and Lymphedema. *Annu Rev Physiol* (2018) 80:49–70. doi: 10.1146/annurev-physiol-022516-034008
56. Veillette A, Latour S. The SLAM Family of Immune-Cell Receptors. *Curr Opin Immunol* (2003) 15:277–85. doi: 10.1016/S0952-7915(03)00041-4
57. Chen Q, Qiu B, Zeng X, Hu L, Huang D, Chen K, et al. Identification of a Tumor Microenvironment-Related Gene Signature to Improve the Prediction of Cervical Cancer Prognosis. *Cancer Cell Int* (2021) 21:182. doi: 10.1186/s12935-021-01867-2
58. Liu X, Niu X, Qiu Z. A Five-Gene Signature Based on Stromal/Immune Scores in the Tumor Microenvironment and Its Clinical Implications for Liver Cancer. *DNA Cell Biol* (2020) 39:1621–38. doi: 10.1089/dna.2020.5512
59. Bogunovic D, O'neill DW, Belitskaya-Levy I, Vacic V, Yu YL, Adams S, et al. Immune Profile and Mitotic Index of Metastatic Melanoma Lesions Enhance Clinical Staging in Predicting Patient Survival. *Proc Natl Acad Sci U S A* (2009) 106:20429–34. doi: 10.1073/pnas.0905139106
60. He Y, Jiang Z, Chen C, Wang X. Classification of Triple-Negative Breast Cancers Based on Immunogenomic Profiling. *J Exp Clin Cancer Res* (2018) 37:327. doi: 10.1186/s13046-018-1002-1
61. Gibney GT, Weiner LM, Atkins MB. Predictive Biomarkers for Checkpoint Inhibitor-Based Immunotherapy. *Lancet Oncol* (2016) 17:e542–51. doi: 10.1016/S1470-2045(16)30406-5

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

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Oncology

REVIEW
published: 10 January 2022
doi: 10.3389/fonc.2021.814085

The Development of Single-Cell Metabolism and Its Role in Studying Cancer Emergent Properties

Dingju Wei[1†], *Meng Xu*[1†], *Zhihua Wang*[2,3] *and Jingjing Tong*[1*]

[1] School of Life Science, Central China Normal University, Wuhan, China, [2] Shenzhen Key Laboratory of Cardiovascular Disease, Fuwai Hospital Chinese Academy of Medical Sciences, Shenzhen, China, [3] State Key Laboratory of Cardiovascular Disease, Fuwai Hospital, National Center for Cardiovascular Disease, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

Metabolic reprogramming is one of the hallmarks of malignant tumors, which provides energy and material basis for tumor rapid proliferation, immune escape, as well as extensive invasion and metastasis. Blocking the energy and material supply of tumor cells is one of the strategies to treat tumor, however tumor cell metabolic heterogeneity prevents metabolic-based anti-cancer treatment. Therefore, searching for the key metabolic factors that regulate cell cancerous change and tumor recurrence has become a major challenge. Emerging technology—single-cell metabolomics is different from the traditional metabolomics that obtains average information of a group of cells. Single-cell metabolomics identifies the metabolites of single cells in different states by mass spectrometry, and captures the molecular biological information of the energy and substances synthesized in single cells, which provides more detailed information for tumor treatment metabolic target screening. This review will combine the current research status of tumor cell metabolism with the advantages of single-cell metabolomics technology, and explore the role of single-cell sequencing technology in searching key factors regulating tumor metabolism. The addition of single-cell technology will accelerate the development of metabolism-based anti-cancer strategies, which may greatly improve the prognostic survival rate of cancer patients.

Keywords: cancer metabolism, metabolic reprogramming, metabolic heterogeneity, single-cell metabolomics, tumor drug resistance

OPEN ACCESS

Edited by:
Yapeng Su,
Fred Hutchinson Cancer Research Center, United States

Reviewed by:
Zhonghan Li,
University of California, Riverside, United States
Chunxiu Hu,
Dalian Institute of Chemical Physics (CAS), China

****Correspondence:***
Jingjing Tong
tongjj@mail.ccnu.edu.cn

[†]These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Cancer Metabolism, a section of the journal Frontiers in Oncology

Received: *12 November 2021*
Accepted: *16 December 2021*
Published: *10 January 2022*

Citation:
Wei D, Xu M, Wang Z and Tong J (2022) The Development of Single-Cell Metabolism and Its Role in Studying Cancer Emergent Properties. Front. Oncol. 11:814085. doi: 10.3389/fonc.2021.814085

BACKGROUND

Cancer is one of the top lethal factors, and cancer patients bear a heavy burden of life expectancy globally. According to the latest data reported by World Health Organization's International Agency for Research on Cancer (IARC), up to 19.3 million new cancer cases and approximately 10.0 million cancer deaths occurred in 2020 worldwide (1). Therefore, the development of early diagnosis and effective treatment of cancer is urgent.

One of the most significant features of cancer is metabolic reprogramming, and increasing evidence suggests that dysregulated cell metabolite facilitates tumor initiation, progression, metastasis, and drug resistance. In 1924, Otto Warburg firstly identified that cancer utilizes

glycolysis instead of oxidative tricarboxylic acid cycle (TCA) to provide energy, nucleotide, lipid, and amino acid for the growth even under aerobic conditions (2). Later in 1975, Ambanelli reported that malignant tumors exhibited congenital errors of the degradation of tryptophan (3). Several years later, it was found that glutamine is the most rapidly consumed amino acid in proliferating Ehrlich ascites carcinomas and also a number of hepatomas and carcinosarcomas (4). Therefore, understanding metabolic alterations in cancer cells may give us a hint to discover new therapeutic targets and facilitate oncology drug research and development of cancer therapy. The idea of cancer metabolism-based therapy has been raised for decades, but significant side effects of antimetabolite drugs, which are caused by destroying normal rapidly proliferating cells, made them limited in preclinical studies (5). For instance, although numerous preclinical studies have manifested the anti-proliferative effects of 2-deoxyglucose (2-DG) (6), the usage of 2-DG was limited by its toxicity related to hypoglycemia symptoms (7). Recent clinical trials have illustrated that lower doses of 2-DG are insufficient to inhibit disease progression (8, 9). Moreover, although lactate dehydrogenase-A (LDH-A) was frequently identified as an overexpression gene in human cancers (10), none of the LDH-A inhibitors have reached clinical trials as effective antimetabolite chemotherapy drugs, suggesting either insufficient drug exposure, unacceptable toxicity (11), or a lack of LDH-A dependence in human tumors. Similarly, two diabetes therapeutic biguanide compounds metformin and phenformin have been found to reduce tumor growth (12), however, in cells lacking a functional LKB1 pathway, the biguanide drugs have been demonstrated to result in rapid apoptosis (13).

In this scenario, would interfering cancer cell metabolism be wrong? Looking back at previous studies, we found that on the one hand, since multiple metabolic pathways altered in tumor cell, by only interfering one metabolite is not sufficient for cancer treatment. On the other hand, some metabolic targets of cancer are also essential for the normal cells growth, so it is very important to control the dosage of tumor metabolism therapeutic drugs. In addition, considering the traditional metabolomics analysis method, the proposed therapeutic targets have been extracted from the average information of a cell mixture. In reality, recent research has confirmed that even from the same tissue, individuals of the same group of cells are different because of cell heterogeneity. Cell heterogeneity is particularly common in tumor tissues, as the fact that some specific cells survived by altering their metabolism after chemotherapy, therefore cancer may re-emerge years later (14, 15). And single-cell analysis can identify the intracellular biochemical components and features, and at the same time can study the relationship between metabolism and cell function, cell development, and differentiation. Using this method, the pathological mechanism of cells in disease states can be studied, and it is also helpful for clinical diagnosis and prognosis (16, 17) (**Figure 1**). In addition, this technology can also understand detailed cell information and reduce the problems caused by statistics and data processing in previous research methods.

In this paper, we summarized the research progression of single-cell metabolomics analysis technology, mainly from two aspects: (1) Research status of single cell metabolomics technology; and (2) Application of single cell metabolomics in tumor research.

THE DEVELOPMENT OF SINGLE-CELL METABOLOMICS TECHNIQUES

The first single-cell metabolomics was established by Kennedy and Jorgenson. They used open tubular capillary chromatography to analyze the amino acid composition of a snail single giant neuron (**Figure 2A**) (18, 19). At the same time, Wallingford and Ewing reported sampling the internal contents of a single giant neuron by using a capillary (20). However, considering the different size and volume between the giant neurons and normal cells, the method for extraction of single cell metabolites needs to be suitable for different cell sizes (21). Moreover, the metabolic pathways in cells can be easily affected by both internal and external factors and are highly dynamic, which is very different from the genome, transcriptome, and proteome. Compared with single-cell genomics, transcriptomics, and proteomics, single-cell metabolomics can provide the most sensitive dynamic picture for understanding cell functions, but the measurement of single-cell metabolomics targets is undoubtedly the most difficult. One of the main problems in the preparation for single-cell samples is how to avoid or reduce the impact on cell metabolism during the sample preparation process. In addition, the content of substances in single cells is low, which puts forward higher requirements for the sensitivity of the detection method. Finally, there are many kinds of metabolites in a single cell, and the concentration difference of metabolites can be as high as 10^6~10^9 times. This requires the detection method not only to respond to multiple substances at the same time, but also to have a wide response range. In short, single-cell analysis technology requires high sensitivity, small sample size, good selectivity, fast response speed, and no impact on cell status, while the data analysis requires complex techniques and models.

The Single Cell Metabolite Extraction

Cell sampling is the first and arguably the most critical step in single cell analysis. To avoid or reduce the impact of environmental changes on cell metabolism during the sample preparation process, one of the methods is to maintain cells in a natural environment as much as possible. For example, use microfluidic chips for cell culturing after cell separation, inject quantitative chemical substances into the culture medium, and selectively release the cells for analysis (22, 23). Another method is to perform rapid freezing of the cells before the metabolite determination to prevent the cells from undergoing dramatic changes in the metabolites (24). The extraction tools for single cell metabolites have been optimized over the past decades. Hajime Mizuno successfully established a direct and rapid analysis of the locations and the metabolic pathways of tryptophan and histidine metabolites in a live rat basophil leukemia cell by live single-cell video-mass spectrometry. The contents of the cell were sucked into a nano-electrospray ionization (nano-ESI) tip, dissolved in an ionization solvent, and directly introduced into a quadrupole-time of flight mass

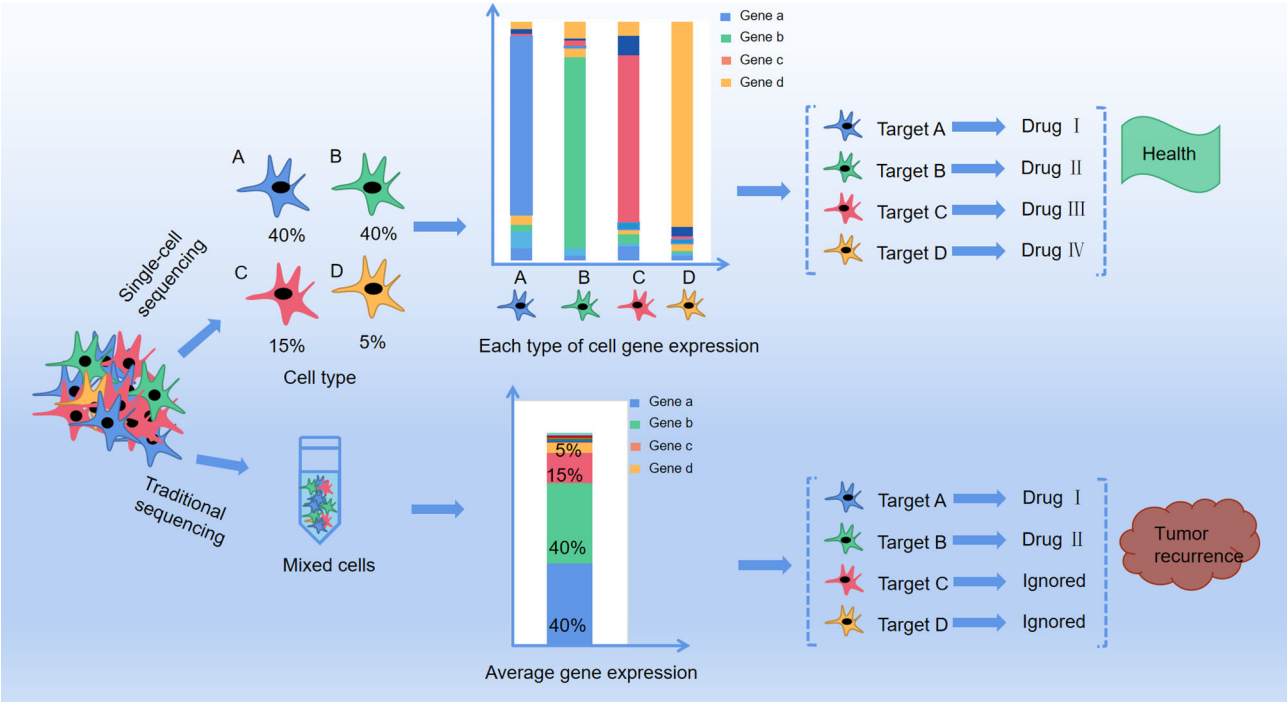


FIGURE 1 | The difference between traditional method and single cell method in profiling anti-cancer targets. A、B、C and D represent four different cell types in cancer tissue with different distribution (A: 40%, B:40%, C:15%, D:5%), and each type of the cells present a target gene as gene a, b, c and d respectively by single cell sequencing. Drug I, II, III and IV are targeted drugs for gene a, b, c and d respectively, and with co-administration of drug I, II, III and IV the cancer is completely curable. However, with traditional RNA-sequencing, mixture of cell A, B, C, and D only presented target gene a and b With co-administration of drug I and II, only group of cell A and B are killed but cell C and D are still survival, leaving a chance for tumor recurrence.

spectrometry (Q-TOF-MS) by nano-spray (**Figure 2B**) (25). Later, based on Mizuno's research, Ning Pan developed a new miniaturized multifunctional sampling and ionization device, the Single-probe, with a sampling tip size smaller than 10 μm, which can be inserted into single cells to extract intracellular compounds. Multiple endogenous and exogenous cellular metabolites in a single living eukaryotic cell could be analyzed in real-time by directly coupling the new probe with the mass spectrometer (**Figure 2D**) (26, 27). Single cell mass spectrometry (SCMS) enables to obtain higher sensitivity and accuracy of chemical information at the single-cell level, which could improve our understanding of biological and pharmaceutical bioanalytical research compared with previous methods.

Raphael proposed a different method for metabolite extraction, he combined glucose oxidase (GOx) covalently with the nanopipettes tip (**Figure 2C**), and the tip was functionalized as glucose nanosensors to quantify single cell intracellular glucose levels. During and after the nanopipette measurement, the cells remain viable. Therefore, nanopipette-based glucose sensors provide a way to compare changes in glucose levels with changes in cell proliferation or metastasis. The nanopipette-based glucose sensors has broad prospects as a diagnostic tool for distinguishing cancer cells from nonmalignant cells in heterogeneous tissue biopsies and a tool for monitoring cancer progression *in situ* (28). Later, Yanlin Zhu has developed a new technology that uses a silica capillary fused micropipette needle, in which can induce Paternò-Büchi (PB) reactions at the C═C bond, and locations of C═C bonds in unsaturated lipids can be determined in cell lysate at the single-cell level (**Figure 2E**). The capillary needle exhibits multiple functions including single cell metabolite extraction probe, cell lysis container, micro-reactor, and nano-ESI emitter during the measurement of metabolites in a single human colon cancer cell HCT-116. This technique is potentially able to apply in other reactive SCMS studies to enhance molecular analysis for broad ranges of single cell metabolites (29).

In addition, mitochondria are important organelles where glucose metabolism happens. Studies on extraction of metabolites in mitochondria are also proposed. Tsuyoshi Esaki combined fluorescence probing with live single-cell mass spectrometry, directly analyzed of mitochondria metabolism in a live HepG2 cell. They stained mitochondria of target cells by fluorescence probe and directly sucked the mitochondria into the nanospray tip under a micromanipulator operation. The sample was then sent to a high-resolution mass spectrometer LTQ-Orbitrap Velos Pro equipped with a nano-electrospray ionization source, and the final result was analyzed by comparing the data gained from stained mitochondria with unstained cytosol blank samples. This fluorescence imaging technique opens the door to analysis of site- and state-specific molecular detection to clarify

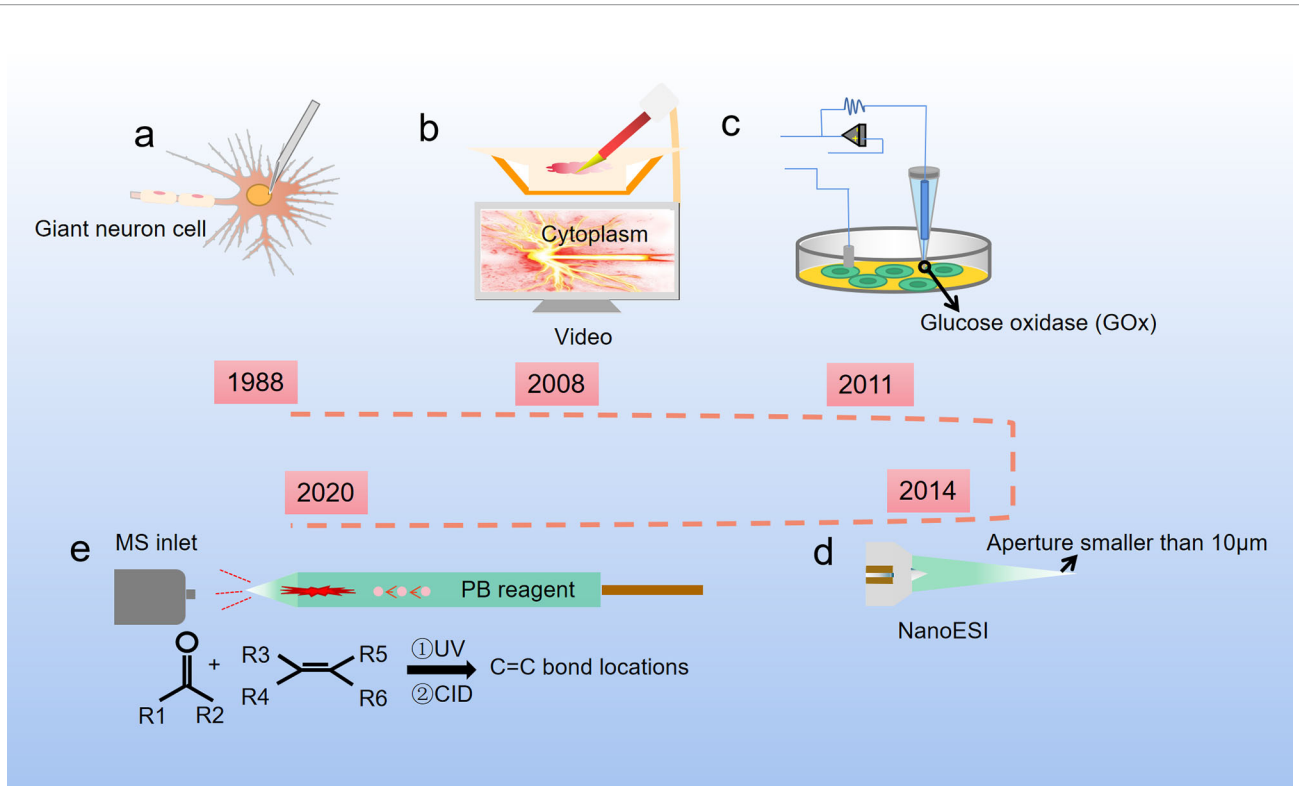


FIGURE 2 | Development of single cell metabolite extraction. **(A)** Microsyringes composed of glass straws were firstly used to extract metabolites from giant neuron cells of snails (1988); **(B)** Nanospray ionization captures the sample from a single living cell, and this process is controlled by a cell manipulation system (2008); **(C)** A nanotube that can measure glucose levels in single-cell level (2011); **(D)** The Single-probe coupled with a sampling tip size < 10 μm, is a miniaturized multifunctional sampling and ionization tool which can achieve *in situ* metabolomic analysis of individual living cells with real-time performance (2014); **(E)** The micropipette needle accommodated with PB reactions inside to determine C=C bond locations in unsaturated lipids at single-cell level (2020).

the precise molecular principles at the level of single-cell and organelle (30).

The Single-Cell Metabolite Flux Detection

Researchers usually use mass spectrometry (MS) or nuclear magnetic resonance (NMR) to carry out metabolomics research, but because NMR technology is not very sensitive, MS has become the main method. However, single-cell analysis is still a tough challenge even with recent technologies, because unlike genes, metabolites cannot be amplified. To analyze such a tiny amount of metabolites in a single cell, many efforts have been tried to improve both detection sensitivity and ionization techniques in MS. On the one hand, considering the small amount and wide dynamic concentration ranges of metabolites in an individual cell, researchers have considered amplifying the signal of the small amount of analyte. Richard B. Keithley administrated three fluorescently labeled glycosphingolipid substrates, GM3-BODIPY-FL, GM1-BODIPY-TMR, and lactosylceramide-BODIPY-650/665 to simultaneous probe metabolism at three different points in the cascade of glycolipid metabolism in HCT 116 spheroids. Finally, they found cells from different regions of HCT 116 spheroids exhibited differences in metabolism, and this three-color fluorescence labeling dramatically amplified the signal of metabolites (31). On the other hand, the direct injection of single cells separated by microfluidic devices or micropipettes into MS suggested the possibility of highly sensitive metabolite analysis in single cells. Hsiao-Wei Liao applied field amplified sample injection (FASI) to capillary electrophoresis electro spray ionization mass spectrometry (CE-ESI-MS) to detect intracellular metabolites from a single neuron, and achieved 100- to 300-fold enhancement of detection limit compared to normal injections. The analytes identification and quantification accuracy were further enhanced through the introduction of internal standards (32). Expect for amplifying the signal of the detected substance, to increase sensitivity of the detection equipment can also achieve the same goal. Takayuki Kawai firstly developed "nano-CESI" emitter, which has a thin conductive wall (10μm) and tapered (5-10μm) end. Compared with a conventional sheathless emitter, the nano-CESI emitter provided up to 3.5-fold increase in sensitivity, and by coupling a sample enrichment method, large-volume dual preconcentration by isotachophoresis and stacking (LDIS), up to 800-fold increase of sensitivity has been achieved compared with normal sheathless CE-MS in total (33).

Mass spectrometry imaging (MSI) is a powerful tool that advances our understanding of complex biological processes by revealing unprecedented details of metabolic biology. MSI does

not require labeling, so it can analyze any compound present on the tissue, which is in stark contrast to most label-based imaging methods, which require prior knowledge of clearly defined targets. In addition, hundreds of compounds can be imaged simultaneously in MSI, which is different from traditional optical imaging in which by using limited number of different fluorescent label colors, only a few targets can be imaged at a time. MSI can be performed by combining any desorption ionization technique with MS in microprobe mode, with the application of matrix-assisted laser desorption ionization (MALDI), the laser can be focused to a micrometer size, thereby reducing the sampling area to a subcellular size. By using the high spatial resolution MALDI-MSI, high-precision metabolite positioning can be obtained at cellular and subcellular levels *in situ* (34).

Single-Cell Metabolomics Data Analysis

In addition to optimize the methods of sample preparation and detection, high-throughput information acquisition is also essential in single cell metabolomics analysis to understand the metabolic process in a single cell. Although the application both of traditional bioinformatics methods and unspecialized software MassLynx (35, 36) to interpret the experimental results of single-cell metabolomics is not a problem, the main challenge is to develop a method that can analyze multiple single cells at the same time and detect the metabolites in each cell while making the results statistically significant. Renmeng Liu analyzed combined SCMS experiments with a generalized integrated data analysis workflow, including data preprocessing, visualization, statistical analysis, machine learning, and pathway enrichment analysis, to conduct single cell metabolomics studies of live cancer cells to discover phenotypic biomarkers and unveil related biological pathways changes during liver cancer chemotherapy (37). Luca Rappez applied SpaceM together with a fluorescence-based readout to detect >100 metabolites from >1,000 individual cells per hour with retention of morpho-spatial features (38). To minimize labor-intensity and enhance the analytical sensitivity, Anqi Chen combined mass spectrometry analyses with a visual serving robotic micromanipulation platform, which sequentially extracted, aspirated, and ionized single cells. This system is the first automated single cell mass spectrometry (SCMS) system. Compared with traditional methods, the automated SCMS system functions without manual operation and facilitates a high-performance single cell metabolic analysis (39).

Chapter Summary

The typical analysis of living cells is a delicate process, mainly start with using nano ESI capillary to extract the contents from the cells, and then use the same capillary to inject the contents directly into the MS. The conceptual schematic diagram of the workflow of MS analysis of single cell metabolites is shown in **Figure 3**. However, the current single-cell metabolomics technology is still under the way to get more refined. Extraction, detection and analysis are the three major difficult problems which must be resolved. With continuous optimization of current single-cell metabolomics techniques, a comprehensive portrait of metabolic features of each unique cell can be expected in the near future.

THE APPLICATION OF SINGLE-CELL METABOLOMICS IN CANCER RESEARCH

The detection and understanding of cancer cells is one of the most important the potential applications of single-cell metabolomics. One application is the discovery of cancer cells with abnormally high metabolic rates within cells undergoing normal metabolism, including circulating tumor cells that cause cancer metastasis (40, 41). In current cancer treatment, single-cell metabolomics technology can be applied to dig out why some cancer cells are still able to survive after being stimulated by the environment or drugs by altering their metabolic pathways (42, 43). Other potential applications single-cell metabolomics technology include understanding the mechanism of tumor metastasis (44, 45) and obtaining the input and output data required to establish a mathematical model of cell metabolism and to learn more about the fate of cancer cells (46).

Single-Cell Metabolomics Reveals Mechanisms of Tumor Drug Resistance

During tumor chemotherapy, part of the patients develop drug resistance, resulting in treatment failure and tumor recurrence, which causes more than 90% of cancer-related deaths. Solid cancers have shown the intratumor heterogeneity spatially and temporally, which has become the biggest obstacle in tumor therapy (47, 48). Single-cell technology is a powerful tool to analyze tumor heterogeneity, and it pulls the dimension of our observation of tumors to the dimension of a single cell (49). By analyzing the genetic and metabolite information of individual cells, we can distinguish genes and regulatory pathways driving drug resistance development (50). For example, Renmeng Liu exposed HCT-116 cells to taxol and vinblastine, which are two mitotic inhibitors, under a series of treatment conditions, then they used Single-probe SCMS system to measure metabolomics change in cells. Phenotypic biomarkers related to the emerging phenotypes resulted from drug treatment were discovered and compared through a series of rigorous statistical analysis with the single cell study and traditional liquid chromatography-MS (LC-MS) study from bulk cell samples. Through pathway enrichment analysis, four biological pathways that may be involved in the drug treatment of colorectal carcinoma have been identified, and this technique can be potentially applied to future pharmaceutical and chemotherapeutic research (51). Clinical studies demonstrated that high concentration (10 μM) erlotinib inhibited cancer proliferation, but beyond the normal tolerance level, while low concentration (1 μM) erlotinib exhibited no treatment effect. Based on Xue Min's research, the low-dose (1 μM) of erlotinib actually increased the energy potential of cells, even if glucose uptake and phosphoprotein signaling were inhibited, which may help explain the resistance of some cancer patients to EGFR inhibitors (52). Mei Sun used the Single-probe mass spectrometry (MS) technique to inspect the metabolic features

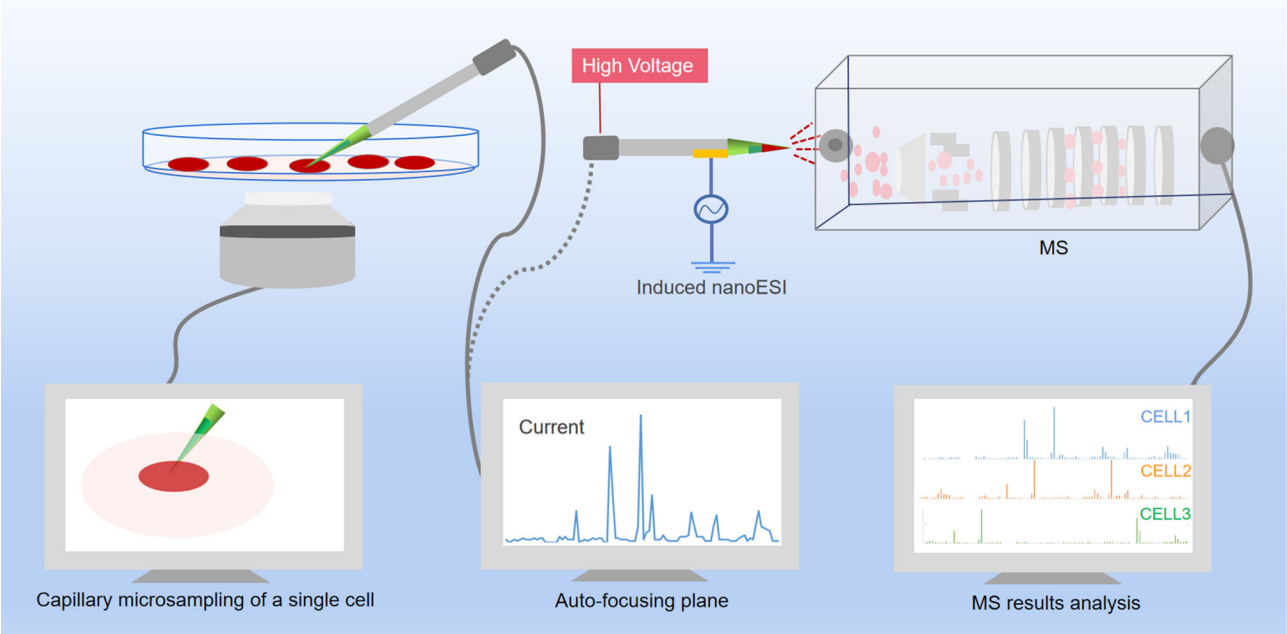


FIGURE 3 | Scheme of single-cell metabolomics analysis of live cell. The content of a live single cell is extracted by micropipette needle under the guidance of an inverted micro-scope, and a cell manipulation system was used to control the micropipette needle to aim the targeted cell. After assistant solvent was added to the sampling tip, the biomolecules were dissolved into the assistant solvent and immediately ejected from the tip into the MS under the impetus of the electric field. The raw data then analyzed by software.

of individual live colorectal cancer stem cells (CSCs). Comparison with non-stem cancer cells (NSCCs), she/he found CSCs contained relatively higher amount of tricarboxylic acid (TCA) cycle metabolites and unsaturated lipids. Application of inhibitors of stearoyl-CoA desaturase-1 (SCD1), nuclear factor κB (NF-κB), and aldehyde dehydrogenases (ALDH1A1) in CSCs significantly contracted the abundance of unsaturated lipids and hindered the formation of tumor spheroids, leading to reduced stemness of CSCs. This indicates that single-cell metabolomics can potentially be used for metabolomics research on rare types of cells, and provides a new method to discover functional biomarkers as therapeutic targets (53). Yapeng Su integrated single-cell flow cytometry with theoretical investigation to study the cell-state transition dynamics associated with BRAF inhibitor drug resistance in BRAF-mutant melanoma cell. They concluded that in certain plastic cancers, the population heterogeneity and evolution of cell phenotypes may be comprehended by explaining the competitive interaction between the epigenetic potential landscape and state-dependent cell proliferation. Their research suggested that experimentally verifiable predictions can potentially determine the trajectories that single BRAFV600E mutant melanoma cancer cells take between drug-naive and drug-tolerant states and guide the design of effective treatment strategies (54). Therefore, the resolved heterogeneous drug-response trajectories by single-cell technique update our current understanding of how drug resistance developed and can provide a powerful methodology for identifying effective combined treatment (55).

Single-Cell Metabolomics Reveals Tumor Metastasis

Much effort had also been made to explore the predictive genomic changes in disease prognosis (56, 57). Recent studies showed that even individual cells from the same clonal can display a broad landscape of different properties, such as different patterns of gene expression (58) and invasive behaviors (59), further increased the challenge of deciphering the mechanism of metastasis in cancer. In this scenario, single-cell omics are applied in tumor metastasis studies. For instance, Ryan T. Davis used single-cell RNA sequencing, flow cytometric and metabolomics to analyze patient-derived-xenograft models of breast cancer, and they found breast cancer micrometastases display a distinct metabolic profile and many of them implicated with metastasis—such as glutamine, fatty-acid and proline metabolism. Most importantly, they found the breast cancer micrometastases converged on or produced critical metabolites to drive oxidative phosphorylation (OXPHOS), and pharmacological inhibition (oligomycin) of OXPHOS substantially attenuates lung metastasis (60). Circulating tumor cells (CTCs) which are released from primary tumor lesion sites into the blood circulation are an important source of tumor metastasis to distant body organs (61). Yasmine Abouleila analyzed untargeted molecular profile of single CTCs collected from gastric cancer (GC) and colorectal cancer (CRC) patients by using live single cell mass spectrometry integrated with microfluidics-based cell enrichment techniques. The authors revealed significant differences in metabolites between the

CTCs and lymphocytes from the same patient, and Principal component analysis-discriminant analysis (PCA-DA) showed obviously different clustering behavior between CTCs and lymphocytes in each cancer. And due to the metabolic differences between GC and CRC, CTCs were clustered into two different groups corresponding to different cancer types, suggesting that the characteristics of CTCs metabolome may become a tool for cancer diagnosis in the future (62). According to the above studies, we learn that at single cell level, it is possible to find new potential biomarkers which promoting tumor metastasis and accelerating tumor deterioration that cannot be detected by traditional methods. In addition, different types of cancer can be classified based on their metabolic fingerprints, which may play a role in identifying clinical targets that may slow down or prevent tumor metastasis.

Single-Cell Metabolomics Reveals Cell Fate

Single-cell technology also allows you to track the state change of the cells under different environment. Qi Zhang published a comprehensive study on the heterogeneity of Hepatocellular carcinoma (HCC) from genome to phenotype and from single-cell level to body level. By single cell sequencing, they classified HCC as the immunocompetent subtype, immunodeficient subtype, and immunosuppressive subtype, respectively. Among three subtypes of cells, the immunosuppressive subtype showed inhibited glycolysis and enhanced mitochondrial respiration, the immunodeficient subtype showed increased nucleotide biosynthesis, while immunocompetent subtype was featured by upregulated urea cycle (**Figure 4**). They also found that although the heterogeneity of tumor cells is significant in all dimensions, the local immune status of HCC is less heterogeneous, therefore they believe that targeting local immunity might be suitable for HCC treatment (63). To track the state of cells, Felix J. Hartmann employed single-cell metabolic regulome profiling (scMEP), an approach applies antibody-based assays to analyze cellular identity and metabolic regulation in the single-cell level. They used a mass spectrometry flow cytometry, time-of-flight flow cytometry (CyTOF), to compare scMEP with a large number of metabolic assays by reconstructing the metabolic remodeling of naive and memory CD8+ T cells activated *in vitro*. Based on the changes in the expression of metabolic characteristics over pseudo-time, three inflection points in the metabolic remodeling process of naive human CD8+ T cells were defined. The first inflection point was marked by the concerted and accelerated up-regulation of metabolic proteins, such as GLUT1, ASCT2, OGDH and VDAC1, leading to the second inflection point, which was characterized by the onset of RNA synthesis and activated cellular stress responses. The third metabolic inflection point was defined by the expression levels of stable or reduced metabolic proteins, such as GLUT1 and ASCT2, and peak translation activity. Therefore, the application of scMEP allows us get a better understanding of tumor–immune boundary and helps identify disease-related metabolic changes, which can be used as potential biomarkers and therapeutic targets for various human diseases (64).

If the cell identity can indeed exist as a continuum, this provides an opportunity to stabilize the transient phenotype and create new cell identities, giving new functions to known cell types. These questions are also related to the field of cell fate reprogramming. At least, we will obtain a high-resolution template to summarize the identities of the main functional cell types. Through the above analysis, we fully understand that single-cell metabolomics techniques serve as a fast, high-throughput predictive tool in cancer research, which can predict metabolite targets in cancer therapy and pharmacology. At the same time, single cell metabolomics is expected to find reliable and effective clinical biomarkers for cancer prognosis and diagnosis.

CONCLUSIONS

In the past decade, single-cell technology has pushed biological research into a new era of exploring cellular and molecular phenotypes at an unprecedented level of resolution. This progress is mainly due to the innovative progress of high-throughput technology and the development of new computing tools, which enable us to capture the genome, transcriptome, proteome, metabolic status, and other multi-dimensional information of thousands of single cells at the same time without relevant information. With these emerging technologies, we can study cell types, cell states, and individual cell responses to external stimuli or internal biological processes. Although metabolomics started late, with the continuous advancement and innovation of detection technology, more breakthroughs will surely be made in the future. Nevertheless, before single-cell metabolomics is truly applied to systems biology and medical diagnosis of cancer research, some challenges of experimental technology or bioinformatics still need to be overcome. There are still many problems need to be solved, including expanding the coverage of metabolites in living cells, faster identification of single-cell metabolites while achieving high-throughput detection, and reduce cost.

In the next few years, greater progress is likely to be made toward new combinations of different single-cell omic techniques to capture of all molecules in an individual cell. We also expect that by further combining existing indexing or microfluidic technologies, technologies that can analyze four or five-layer omics data in parallel can be developed, resulting in single-cell multiomics methods capable of high-throughput processing of thousands of cells. Currently, each omics data set need to be analyzed separately and compare the final results to get the conclusion, therefore novel computational methods and specialist multiomics algorithms that allow the integrated study of two or more omic layers per cell of large heterogeneous populations and data analysis of the different layers together will also need to be developed. Future multi-omics technologies will eventually need to cover all single-cell omics technologies in order to characterize these omnisciently for different layers, three-dimensional coordinates, phenotypes, and cell lineage history.

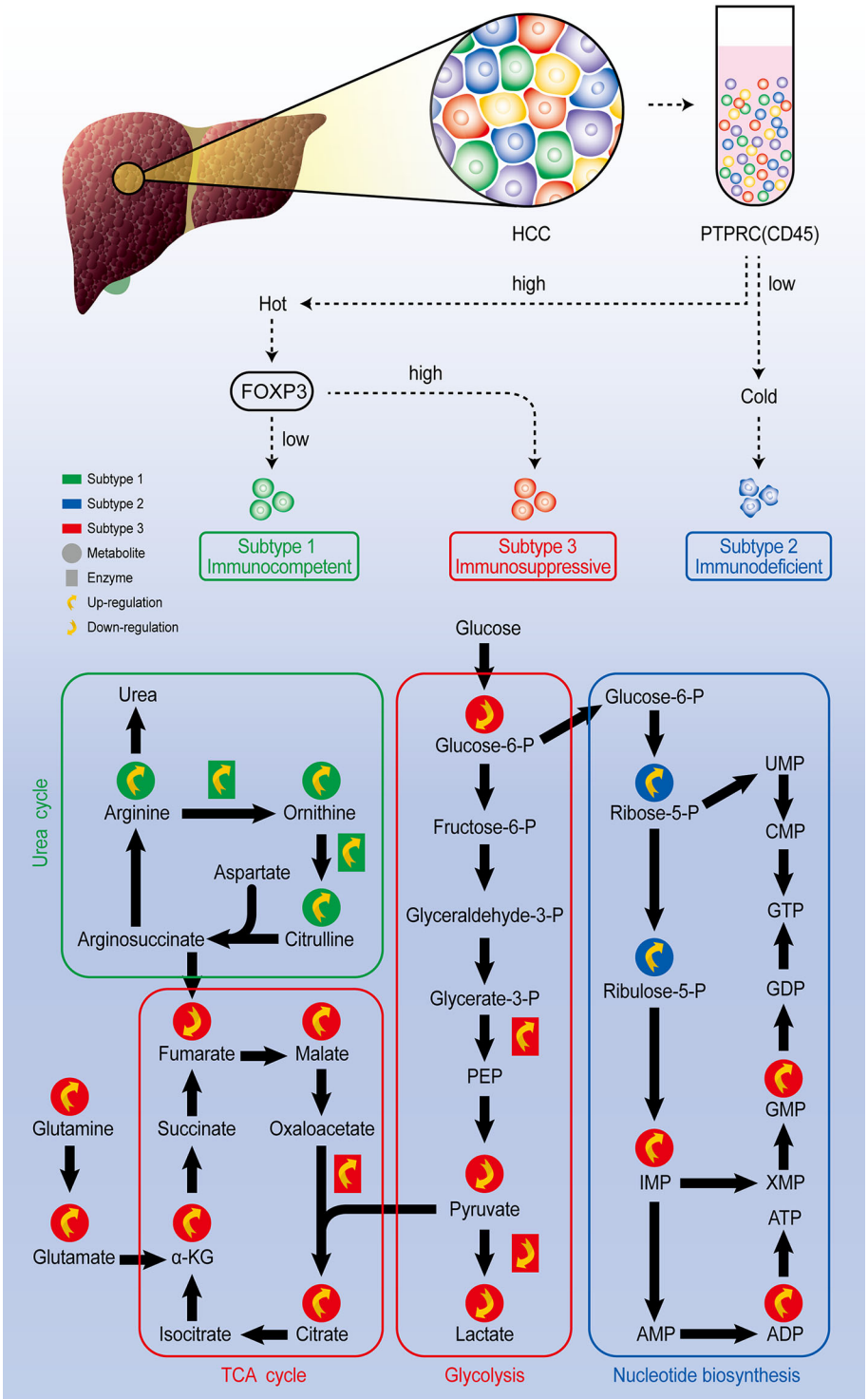


FIGURE 4 | Comprehensive exploration of the heterogeneity of HCC analysis scheme. Through single-cell sequencing, HCC could be classified as three subtypes according to the expression level of marker genes PTPRC and FOXP3. Further single-cell metabolomic analysis revealed that the immunocompetent subtype is with upregulated urea cycle, the immunosuppressive subtype is with upregulated TCA cycle and inhibited glycolysis pathway, and the immunodeficient subtype presents upregulated nucleotide biosynthesis pathway.

AUTHOR CONTRIBUTIONS

JT, DW, and MX conceived and designed the review. DW did paper search and DW and MX wrote the manuscript. JT had supervised this project and contributed to writing and revision of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by grants from the National Natural Science Foundation of China (NSFC No. 32170763), the National Natural Science Foundation of Hubei (ZRMS2020002147) and the Research Funds from Central China Normal University (CCNU20TS017).

REFERENCES

1. Sun L, Zhang H, Gao P. Metabolic Reprogramming and Epigenetic Modifications on the Path to Cancer. *Protein Cell* (2021) 36). doi: 10.1007/s13238-021-00846-7
2. Warburg OA. On the Origin of Cancer Cell. *Science* (1956) 123(3191):309–14. doi: 10.1126/science.123.3191.309
3. Ambanelli U, Manganelli P. Clinical Significance of Changes in Tryptophan Metabolism. *Ateneo Parmense Acta Biomed* (1975) 46(1-2):5–19.
4. Dauchy L.A.S.J.W.S.R.T. Amino Acid, Glucose, and Lactic Acid Utilization in Vivo by Rat Tumors. *Cancer Res* (1982) 42(10):4090–7. doi: 10.1016/0304-3835(82)90136-7
5. Luengo A, Gui DY, Vander Heiden MG. Targeting Metabolism for Cancer Therapy. *Cell Chem Biol* (2017) 24(9):1161–80. doi: 10.1016/j.chembiol.2017.08.028
6. Zhang D, Li J, Wang F, Hu J, Sun Y. 2-Deoxy-D-Glucose Targeting of Glucose Metabolism in Cancer Cells as a Potential Therapy. *Cancer Lett* (2014) 355 (2):176–83. doi: 10.1016/j.canlet.2014.09.003
7. Landau BR, Laszlo J, Stengle J, Burk D. Certain Metabolic and Pharmacologic Effects in Cancer Patients Given Infusions of 2-Deoxy-D-Glucose. *J Natl Cancer Inst* (1958) 21(3):485–94. doi: 10.1016/0014-4827(58)90178-2
8. Mark S, Lin H, Jeyamohan C, Dvorzhinski D, Gounder M. Targeting Tumor Metabolism With 2-Deoxyglucose in Patients With Castrate-Resistant Prostate Cancer and Advanced Malignancies. *Prostate* (2010) 70(13):1388–94. doi: 10.1002/pros.21172
9. Raez LE, Papadopoulos K. A Phase I Dose-Escalation Trial of 2-Deoxy-D-Glucose Alone or Combined With Docetaxel in Patients With Advanced Solid Tumors. *Cancer Chemother Pharmacol* (2013) 71(2):523–30. doi: 10.1007/s00280-012-2045-1
10. Shim H, Dolde C, Lewis BC, Wu CS, Dang G, Jungmann RA, et al. C-Myc Transactivation of LDH-A: Implications for Tumor Metabolismandgrowth. *Proc Natl Acad Sci U S A* (1997) 94(13):6658–63. doi: 10.1073/pnas.94.13.6658
11. Boudreau A, Purkey HE, Hitz A. Metabolic Plasticity Underpins Innate and Acquired Resistance to LDHA Inhibition. *Nat Chem Biol* (2016) 12(10):779–86. doi: 10.1038/nchembio.2143
12. Shackelford DB, Abt E, Gerken L, Vasquez DS, Seki A, Leblanc M, et al. LKB1 Inactivation Dictates Therapeutic Response of Non-Small Cell Lung Cancer to the Metabolism Drug Phenformin. *Cancer Cell* (2013) 23(2):143–58. doi: 10.1016/j.ccr.2012.12.008
13. Shaw RJ, Kosmatka M, Bardeesy N, Hurley RL, Witters LA, DePinho RA, et al. The Tumor Suppressor LKB1 Kinase Directly Activates AMP-Activated Kinase and Regulates Apoptosis in Response to Energy Stress. *Proc Natl Acad Sci U S A* (2004) 101(10):3329–35. doi: 10.1073/pnas.0308061100
14. Altschuler SJ, Wu LF. Cellular Heterogeneity: Do Differences Make a Difference? *Cell* (2010) 141(4):559–63. doi: 10.1016/j.cell.2010.04.033
15. Lyssiotis C, Kimmelman A. Metabolic Interactions in the Tumor Microenvironment. *Trends Cell Biol* (2017) 27(11):863. doi: 10.1016/j.tcb.2017.06.003
16. Fessenden M. Metabolomics: Small Molecules, Single Cells. *Nature* (2016) 540 (7631):153–5. doi: 10.1038/540153a
17. Rubakhin SS, Lanni EJ, Sweedler JV. Progress Toward Single Cell Metabolomics. *Curr Opin Biotechnol* (2013) 24(1):95–104. doi: 10.1016/j.copbio.2012.10.021
18. Kennedy R, Oates M, Cooper B, Nickerson B, Jorgenson J. Microcolumn Separations and the Analysis of Single Cells. *Science* (1988) 246(4926):57–63. doi: 10.1126/science.2675314
19. Kennedy RT, Claire R, White JG, Jorgenson JW. Chemical Analysis of Single Neurons by Open Tubular Liquid Chromatography. *Microchimi Acta* (1987) 92(1-3):37–45. doi: 10.1007/BF01201715
20. Wallingford RA, Ewing AG. Capillary Zone Electrophoresis With Electrochemical Detection in 12.7 μm Diameter Columns [1]. *Analyt Chem* (1988) 60(18):1972–5. doi: 10.1021/ac00169a027
21. Hiyama E, Ali A, Amer S, Harada T, Shimamoto K, Furushima R, et al. Direct Lipido-Metabolomics of Single Floating Cells for Analysis of Circulating Tumor Cells by Live Single-Cell Mass Spectrometry. *Analyt Sci* (2015) 31 (12):1215–7. doi: 10.2116/analsci.31.1215
22. Jen CP, Hsiao JH, Maslov NA. Single-Cell Chemical Lysis on Microfluidic Chips With Arrays of Microwells. *Sensors* (2012) 12(1):347–58. doi: 10.3390/s120100347
23. Mi S, Yang S, Liu T, Du Z, Xu Y, Li B, et al. A Novel Controllable Cell Array Printing Technique on Microfluidic Chips. *IEEE Trans Biomed Eng* (2019) 66 (9)2512–20. doi: 10.1109/TBME.2019.2891016
24. Ibanez AJ, Fagerer SR, Schmidt AM, Urban PL, Jefimovs K, Geiger P, et al. Mass Spectrometry-Based Metabolomics of Single Yeast Cells. *Pnas* (2013) 110(22):8790–4. doi: 10.1073/pnas.1209302110
25. Hajime M, Tsuyama N, Date S, Harada T, Masujima T. Live Single-Cell Metabolomics of Tryptophan and Histidine Metabolites in a Rat Basophil Leukemia Cell. *Analyt Sci Int J Japan Soc Analyt Chem* (2008) 24(12):1525–7. doi: 10.2116/analsci.24.1525
26. Pan N, Rao W, Kothapalli NR, Liu R, Burgett A, Yang Z. The Single-Probe: A Miniaturized Multifunctional Device for Single Cell Mass Spectrometry Analysis. *Analyt Chem* (2014) 86(19):9376–80. doi: 10.1021/ac5029038
27. Pan N, Rao W, Yang Z. Single-Probe Mass Spectrometry Analysis of Metabolites in Single Cells. *Single Cell Metab* (2020) (2064)61–71. doi: 10.1007/978-1-4939-9831-9_5
28. Nascimento R, Ozel RE, Mak WH, Mulato M, Singaram B, Pourmand N. Single Cell "Glucose Nanosensor" Verifies Elevated Glucose Levels in Individual Cancer Cells. *Nano Lett* (2011) 16(2):1194–200. doi: 10.1021/acs.nanolett.5b04495
29. Zhu Y, Wang W, Yang Z. Combining Mass Spectrometry With Paternò–Büchi Reaction to Determine Double-Bond Positions in Lipids at the Single-Cell Level. *Analyt Chem* (2020) 92(16):11380–7. doi: 10.1021/acs.analchem.0c02245
30. Esaki T, Masujima T. Fluorescence Probing Live Single-Cell Mass Spectrometry for Direct Analysis of Organelle Metabolism. *Analyt Sci Int J Japan Soc Analyt Chem* (2015) 31(12):1211. doi: 10.2116/analsci.31.1211
31. Keithley RB, Weaver EM, Metzinger MP, Rosado AM, Dovichi NJ. Single Cell Metabolic Profiling of Tumor Mimics. *Analyt Chem* (2013) 85(19):8910–8. doi: 10.1021/ac402262e
32. Liao HW, Rubakhin SS, Philip MC, Sweedler JV. Enhanced Single-Cell Metabolomics by Capillary Electrophoresis Electrospray Ionization-Mass Spectrometry With Field Amplified Sample Injection. *Analyt Chim Acta* (2020) 1118:36–43. doi: 10.1016/j.aca.2020.04.028
33. Kawai T, Ota N, Okada K, Imasato A, Tanaka Y. Ultrasensitive Single Cell Metabolomics by Capillary Electrophoresis–Mass Spectrometry With a Thin-Walled Tapered Emitter and Large-Volume Dual Sample Preconcentration. *Analyt Chem* (2019) 91(16):10564–72. doi: 10.1021/acs.analchem.9b01578
34. Hansen RL, Lee YJ. High-Spatial Resolution Mass Spectrometry Imaging: Toward Single Cell Metabolomics in Plant Tissues. *Chem Rec* (2017) 18 (1):65–77. doi: 10.1002/tcr.201700027
35. Stolee JA, Shrestha B, Mengistu G, Vertes A. Observation of Subcellular Metabolite Gradients in Single Cells by Laser Ablation Electrospray Ionization Mass Spectrometry. *Angewandte Chemie* (2012) 51(41):10386–9. doi: 10.1002/anie.201205436
36. Zhang L, Sevinsky CJ, Davis BM, Vertes A. Single Cell Mass Spectrometry of Subpopulations Selected by Fluorescence Microscopy. *Analyt Chem* (2018) 90 (7):4626–34. doi: 10.1021/acs.analchem.7b05126

37. Liu R, Zhang G, Sun M, Pan X, Yang Z. Integrating a Generalized Data Analysis Workflow With the Single-Probe Mass Spectrometry Experiment for Single Cell Metabolomics. *Analyt Chim Acta* (2019) 1064:71–9. doi: 10.1016/j.aca.2019.03.006
38. Rappez L, Stadler M, Triana S, Gathungu RM, Alexandrov T. SpaceM Reveals Metabolic States of Single Cells. *Nat Methods* (2021) 18(7):799–805. doi: 10.1038/s41592-021-01198-0
39. Chen A, Yan M, Feng J, Bi L, Wen L. Single Cell Mass Spectrometry With a Robotic Micromanipulation System for Cell Metabolite Analysis. *IEEE Trans Biomed Eng PP* (2021) 99):1–1. doi: 10.1109/TBME.2021.3093097
40. Huang NT, Zhang HL, Chung MT, Seo JH, Kurabayashi K. Recent Advancements in Optofluidics-Based Single-Cell Analysis: Optical on-Chip Cellular Manipulation, Treatment, and Property Detection. *Lab Chip* (2014) 14(7):1230–45. doi: 10.1039/c3lc51211h
41. Shen FM, Zhu L, Ye H, Yang YJ, Pang DW, Zhang ZL. A High Throughput Micro-Chamber Array Device for Single Cell Clonal Cultivation and Tumor Heterogeneity Analysis. *Sci Rep* (2015) 5(1):11937. doi: 10.1038/srep11937
42. Pang L, Ding J, Ge Y, Fan J, Fan SK. Single-Cell-Derived Tumor-Sphere Formation and Drug-Resistance Assay Using an Integrated Microfluidics. *Analyt Chem* (2019) 91(13):8318–25. doi: 10.1021/acs.analchem.9b01084
43. Zhang S, Wang Q, Guldner IM, Golomb SM, Lu X. Single-Cell Profiling Guided Combinatorial Immunotherapy for Fast-Evolving CDK4/6 Inhibitor Resistant HER2-Positive Breast Cancer. *Nat Commun* (2019) 23(1):3817. doi: 10.1038/s41467-019-11729-1
44. Lawson DA, Kessenbrock K, Davis RT, Pervolarakis N, Werb Z. Tumour Heterogeneity and Metastasis at Single-Cell Resolution. *Nat Cell Biol* (2018) 20(12):1349–60. doi: 10.1038/s41556-018-0236-7
45. Wu PH, Gilkes DM, Phillip JM, Narkar A, Wirtz D. Single-Cell Morphology Encodes Metastatic Potential. *Sci Adv* (2020) 6(4):eaaw6938. doi: 10.1126/sciadv.aaw6938
46. Portero EP, Nemes P. Dual Cationic–Anionic Profiling of Metabolites in a Single Identified Cell in a Live Xenopus Laevis Embryo by Microprobe CE-ESI-MS. *Analyst* (2019) 144(3):892–900. doi: 10.1039/c8an01999a
47. Gerlinger M, Rowan AJ, Horswell S, Math M, Larkin J, Endesfelder D, et al. Intratumor Heterogeneity and Branched Evolution Revealed by Multiregion Sequencing. *N Engl J Med* (2012) 366(10):883–92. doi: 10.1056/NEJMoa1113205
48. Network TCGA. Integrated Genomic Analyses of Ovarian Carcinoma. *Nature* (2011) 474(7353):609–15. doi: 10.1038/nature10166
49. Shiokawa D, Sato A, Ohata H, Mutoh M, Sekine S, Kato M, et al. The Induction of Selected Wnt Target Genes by Tcf1 Mediates Generation of Tumorigenic Colon Stem Cells. *Cell Rep* (2017) 19(5):981–94. doi: 10.1016/j.celrep.2017.04.017
50. Saunders NA, Simpson F, Thompson EW, Hill MM, Endo-Munoz L, Leggatt G, et al. Role of Intratumoural Heterogeneity in Cancer Drug Resistance: Molecular and Clinical Perspectives. *EMBO Mol Med* (2012) 4(8):675–84. doi: 10.1002/emmm.201101131
51. Liu R, Sun M, Zhang G, Lan Y, Yang Z. Towards Early Monitoring of Chemotherapy-Induced Drug Resistance Based on Single Cell Metabolomics: Combining Single-Probe Mass Spectrometry With Machine Learning. *Analyt Chim Acta* (2019) 1092:42–8. doi: 10.1016/j.aca.2019.09.065
52. Min X, Wei W, Su Y, Johnson D, Heath JR. Supramolecular Probes for Assessing Glutamine Uptake Enable Semi-Quantitative Metabolic Models in Single Cells. *J Am Chem Soc* (2016) 138(9):3085. doi: 10.1021/jacs.5b12187
53. Sun M, Yang Z. Metabolomic Studies of Live Single Cancer Stem Cells Using Mass Spectrometry. *Analyt Chem* (2018) 91(3):2384–91. doi: 10.1021/acs.analchem.8b05166
54. Su Y, Bintz M, Yang Y, Robert L, Wei W. Phenotypic Heterogeneity and Evolution of Melanoma Cells Associated With Targeted Therapy Resistance. *PLoS Comput Biol* (2019) 15(6):e1007034. doi: 10.1371/journal.pcbi.1007034
55. Su Y, Ko ME, Cheng H, Zhu R, Xue M, Wang J, et al. Multi-Omic Single-Cell Snapshots Reveal Multiple Independent Trajectories to Drug Tolerance in a Melanoma Cell Line. *Nat Commun* (2020) 11(1):2345. doi: 10.1038/s41467-020-15956-9
56. Stratton M, Campbell P, Futreal P. The Cancer Genome. *Nature* (2009) 458 (7239):719–24. doi: 10.1038/nature07943
57. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Kinzler KW. Cancer Genome Landscapes. *Science* (2013) 339(6127):546–58. doi: 10.1126/science.1235122
58. Nguyen A, Yoshida M, Goodarzi H, Tavazoie SF. Highly Variable Cancer Subpopulations That Exhibit Enhanced Transcriptome Variability and Metastatic Fitness. *Nat Commun* (2016) 7:11246. doi: 10.1038/ncomms11246
59. Minn AJ, Kang Y, Serganova I, Gupta GP, Giri DD, Doubrovin M, et al. Distinct Organ-Specific Metastatic Potential of Individual Breast Cancer Cells and Primary Tumors. *J Clin Invest* (2005) 115(1):44–55. doi: 10.1172/JCI22320
60. Davis RT, Blake K, Ma D, Gabra M, Lawson DA. Transcriptional Diversity and Bioenergetic Shift in Human Breast Cancer Metastasis Revealed by Single-Cell RNA Sequencing. *Nat Cell Biol* (2020) 22(3):310–20. doi: 10.1038/s41556-020-0477-0
61. Jie XX, Zhang XY, Xu CJ. Epithelial-To-Mesenchymal Transition, Circulating Tumor Cells and Cancer Metastasis: Mechanisms and Clinical Applications. *Oncotarget* (2017) 8(46):81558–71. doi: 10.18632/oncotarget.18277
62. Abouleila Y, Onidani K, Ali A, Shoji H, Kawai T, Lim CT, et al. Live Single Cell Mass Spectrometry Reveals Cancer-Specific Metabolic Profiles of Circulating Tumor Cells. *Cancer Sci* (2019) 110(2):687–706. doi: 10.1111/cas.13915
63. Zhang Q, Lou Y, Yang J, Wang J, Bai X. Integrated Multiomic Analysis Reveals Comprehensive Tumour Heterogeneity and Novel Immunophenotypic Classification in Hepatocellular Carcinomas. *Gut* (2019) 68(11):2019–31. doi: 10.1136/gutjnl-2019-318912
64. Hartmann FJ, Mrdjen D, Mccaffrey E, Glass DR, Bendall SC. Single-Cell Metabolic Profiling of Human Cytotoxic T Cells. *Nat Biotechnol* (2020) 39 (2):186–97. doi: 10.1038/s41587-020-0651-8

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

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Oncology

ORIGINAL RESEARCH
published: 10 March 2022
doi: 10.3389/fonc.2022.858017

Imaging Sub-Cellular Methionine and Insulin Interplay in Triple Negative Breast Cancer Lipid Droplet Metabolism

*Anthony A. Fung, Khang Hoang, Honghao Zha, Derek Chen, Wenxu Zhang and Lingyan Shi**

Department of Bioengineering, University of California San Diego, La Jolla, CA, United States

Triple negative breast cancer (TNBC) is a particularly aggressive cancer subtype that is difficult to diagnose due to its discriminating epidemiology and obscure metabolome. For the first time, 3D spatial and chemometric analyses uncover the unique lipid metabolome of TNBC under the tandem modulation of two key metabolites – insulin and methionine - using non-invasive optical techniques. By conjugating heavy water (D_2O) probed Raman scattering with label-free two-photon fluorescence (TPF) microscopy, we observed altered *de novo* lipogenesis, 3D lipid droplet morphology, and lipid peroxidation under various methionine and insulin concentrations. Quantitative interrogation of both spatial and chemometric lipid metabolism under tandem metabolite modulation confirms significant interaction of insulin and methionine, which may prove to be critical therapeutic targets, and proposes a powerful optical imaging platform with subcellular resolution for metabolic and cancer research.

Keywords: stimulated Raman scattering, heavy water, TPF, lipid metabolism, methionine, insulin, breast cancer, DO-SRS

OPEN ACCESS

Edited by:
Dongya Jia,
National Cancer Institute (NIH), United States

Reviewed by:
Chenxi Qian,
California Institute of Technology, United States
Lixue Shi,
Columbia University, United States

***Correspondence:**
Lingyan Shi
Lingyanshi@ucsd.edu

Specialty section:
This article was submitted to Cancer Metabolism, a section of the journal Frontiers in Oncology

Received: *19 January 2022*
Accepted: *14 February 2022*
Published: *10 March 2022*

Citation:
Fung AA, Hoang K, Zha H, Chen D, Zhang W and Shi L (2022) Imaging Sub-Cellular Methionine and Insulin Interplay in Triple Negative Breast Cancer Lipid Droplet Metabolism.. Front. Oncol. 12:858017. doi: 10.3389/fonc.2022.858017

INTRODUCTION

Breast cancer is the most reported form of cancer in biological women, but the pathophysiology is rife with subtypes that have material consequences on patient outcomes. Triple negative breast cancer (TNBC) is a particularly aggressive cancer subtype that accounts for approximately 15% of all breast cancer cases and its epidemiology reveals a discriminating predilection for non-Hispanic African women (1, 2) (**Figure S1**). Although the genomes and proteomes of these breast cancer subtypes are distinguishable, little is known about their metabolic phenotypes and the consequential prognoses they manifest.

Recently, lipid metabolism has emerged as a major indicator of cellular stress, phenotypic state, and disease status in biological research and medicine. Dysregulation of lipid metabolism and heightened lipid synthesis are hallmarks of cancer, as varying demands of lipids for energy maintenance, metastasis, and angiogenesis warrant transcriptional changes that contribute to the metabolic phenotype (3–5).

The quantity and diversity of lipids and their functions have been instrumental in profiling cancers as well. For example, membrane lipid compositions of cholesterol, phosphatidylcholine (PC), and phosphatidylethanolamine (PE) are essential to cell membrane fluidity, which has become a target for cancer treatments (6–9). Additionally, the degree of saturation of lipid content in a cell may provide further insight into its state of stress, as breast cancer cells may produce more saturated and monounsaturated membrane lipids to guard against oxidative stress (10–12). To interrogate lipid metabolism, lipid droplets (LD) were the primary organelle of interest since their ubiquitous structures not only serve as energy stores, but are also involved in protein folding and trafficking, signaling pathways, and have diverse spatial and chemical information that may reflect oxidative stress, metabolic flux, and disease status (10–18). However, direct visualization of LD metabolism manipulated by tandem nutritional interventions at a subcellular level has not yet been reported in TNBC cells, which is partially due to a lack of spatial resolution in conventional lipidomic modalities. Optical techniques such as spontaneous Raman spectroscopy and SRS imaging microscopy are well suited to both the chemometric and spatial dimensions for imaging LD metabolism; they can analyze not only the size, number, and distribution of LDs, but also their protein and lipid diversity and metabolism at subcellular resolution.

Despite the many mysteries of TNBC, a documented hallmark is its hyperactivity of mammalian target of Rapamycin (mTOR) pathways, which play important roles in glucose, protein, and lipid metabolism (19–23). Insulin and L-methionine (an essential amino acid involved in protein translation, genetic/epigenetic control, nutrient sensing, and redox homeostasis) (24) are both involved in mTOR pathways but have not been directly studied in tandem to date (25–30). This is due, in part, to previous studies that observed MDA-MB-231 cells to be insulin insensitive to mitogenic effects, despite having many receptors that bind insulin (31). Other studies observe insulin effects in the same cell line, and there is currently no consensus on the independent effects of insulin. With respect to TNBC, insulin and methionine both independently drive cancer proliferation (32–35) and affect lipid metabolism (25, 34, 36–40), and separate studies indicate insulin metabolism directly affects the uptake of amino acids in yeast and dogs (41, 42). Given the well-documented relationships between insulin, methionine, and mTOR, it is possible that TNBC's mTOR hyperactivity exhibits a unique lipidomic response to insulin and methionine manipulation. The

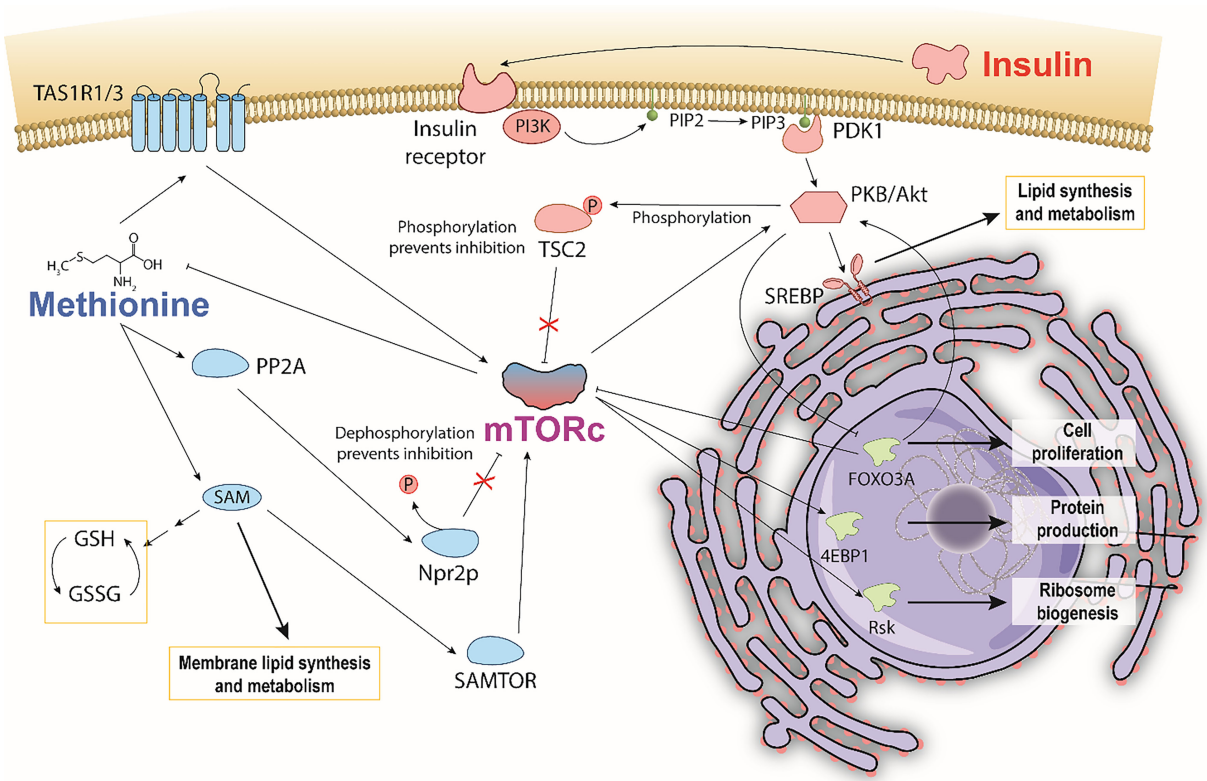


FIGURE 1 | Hypothesized pathway illustrating a potential methionine and insulin interaction mediated through mTOR. Bi-directional control of methionine and mTORc1 depicts general mechanisms by which methionine is sensed by and activates mTORc1. Insulin also activates mTORc1 by phosphorylating TSC2, and consequently affecting mTORc1 regulation of methionine. Insulin stimulates SREBP mediated lipid synthesis and metabolism. Methionine stimulates SAM PC and PE membrane lipid synthesis. Increased production of reduced glutathione *via* SAM is thought to reduce the extent of lipid peroxidation.

conceptual pathway detailing macroscopic mTOR-mediated lipid response to insulin and methionine (**Figure 1**) highlights the points discussed in this paper. Lipid peroxidation, *de novo* synthesis, and chemical diversity can all be investigated using optical techniques that provide subcellular spatial and chemical information. Given that TNBC has been an archetype for methionine dependence (35), and that PI3K/AKT/mTOR is a key driver of the aggressive biology of TNBC (23), the interplay between methionine and insulin, coupled with the perspective of lipid biology, may illuminate promising directions for future therapeutic research.

Non-linear optical techniques such as coherent Raman scattering microscopy and two-photon fluorescence (TPF) microscopy have been used to profile breast cancer metabolism by revealing correlations between cancer metastasis and cellular redox state, and lipid metabolism (43). Recent studies have identified several metabolites implicated in tumorigenesis and lipid metabolism in cancer, such as glutamine (44–46) and serine (47) dependence. Raman spectroscopy/microscopy coupled with D_2O probing allows for direct visualization of metabolic dynamics of a variety of biomolecules including lipids, protein, and DNA in cells, C. elegans, zebrafish, and rodents by highlighting the newly synthesized macromolecules (48). In this study we first employed spontaneous Raman spectroscopy to differentiate molecular signatures within LDs between TNBC and normal cells. Using D_2O probing and SRS (DO-SRS) imaging we then examined the impacts of methionine and insulin on lipid metabolism in cancer cells. The effects of methionine and insulin on cellular respiration and lipid peroxidation were also examined by using TPF microscopy. To analyze the rich chemometric dataset and inspire targeted image analyses, we applied a relative entropy approach to Raman spectra for the first time. This method can quickly highlight distinct or tandem effects of independent variables in any Raman spectroscopy study.

RESULTS

Lipid Droplet Metabolism

We first examined the effects of methionine on LD metabolism in TNBC cells (MDA-MB-231), luminal A breast cancer cells (MCF-7), and normal breast epithelial cells (MCF10A, as a control) by adding excess (20x) methionine to the growth media supplemented with 50% D_2O. Cells were scanned by using a spontaneous Raman spectroscopy, and revealed that TNBC cells most starkly contrasted MCF10A cells with respect to overall lipid content (CH_2 stretching at 2850cm^{-1}). This attenuated lipid:protein ratio difference between excess and physiological methionine concentrations is shown in **Figure 2A**. This absence of marked differences is also personified by poorer ReLu neural network classification between TNBC cells with and without excess methionine (**Figure S3**). Despite the absence of insulin in TNBC cell culture growth media recipes (49), we then added various concentrations of insulin (1mg/L, 10mg/L, and 20mg/L, correspondingly, 0.1x, 1x and 2x) to the media and evaluated its interaction with methionine in both cell lines. In this second part, insulin concentration in growth media was modulated in tandem with methionine, and augmented effects in several Raman spectral regions were observed, including the C-H stretching region, which illustrates the relative contents of CH_2 (lipid) and CH_3 (protein) (**Figure 2B**). TNBC contrasted MCF10A cells which exhibited decreased lipid:protein ratios in the presence of excess methionine at all insulin concentrations. Importantly, it was found that the difference in lipid:protein ratio between excess and physiological methionine increased with the addition of insulin in TNBC. **Figure 2C** highlights this effect, marked by orange arrows in **Figure 2B**, and supports potential interactions between insulin and methionine. A significant interaction term was confirmed by 2-way ANOVA (**Table S1**) in TNBC.

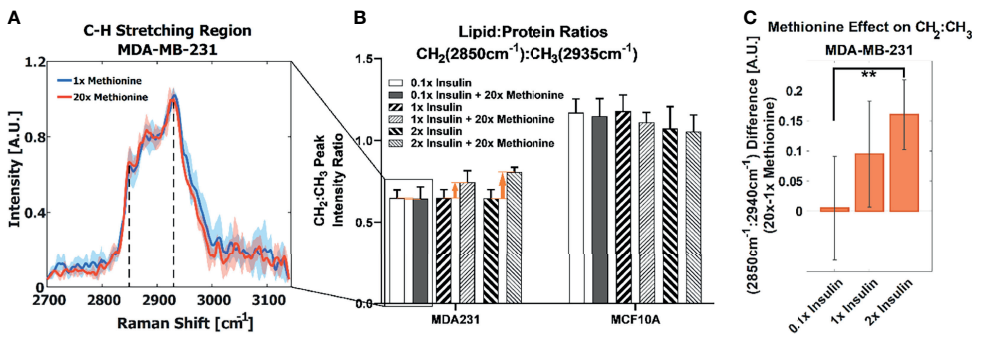


FIGURE 2 | **(A)** Average CH stretching region spectra for TNBC with highlighted CH_2 and CH_3 (2850cm^{-1} and 2935cm^{-1}) levels ascribed to total lipid and protein content, respectively. Regular and excess methionine groups refer to 0.03g/L and 0.6g/L respectively. Of note, TNBC did not exhibit significant relative lipid and protein changes in the presence of excess methionine. One standard deviation is indicated by shaded areas surrounding the lines. **(B)** With tandem insulin control CH_2:CH_3 peak ratios at each insulin and methionine concentration group are shown. 2-way balanced ANOVA results for TNBC cells highlights significance of methionine and insulin-methionine interaction term in lipid:protein ratios. Error bars indicate one standard deviation. **(C)** The difference in CH_2:CH_3 ratios for the 15x methionine and 1x methionine groups of the MDA-MB-231 subtype is negligible at the lower insulin concentration but is increased ten-fold in the 2x insulin case. Error bars are propagated in quadrature from **(B)**. **P < 0.01.

In TNBC cells, the ratio of total lipid-to-protein did not change with the increase of methionine concentration alone, but slightly decreased in normal cells (**Figure 2B**). With the addition of insulin, this difference was augmented with higher concentrations of insulin (**Figure 2C**). At this point, it is still unclear whether *de novo* lipid synthesis increased alone, or if protein synthesis decreased, or some combination of both. Perhaps *de novo* lipid synthesis decreased, but not as much as protein synthesis. This clarity entails DO-SRS, which provides insight into *de novo* synthesis. As cells incorporate deuterium from heavy water into macromolecules such as lipids and proteins, the C-D bonds in the newly synthesized molecules become visible in the cell silent region around 2150 cm^{-1}. Even though lipids and proteins are the main biomolecular constituents of cells, the CH_2 and CH_3 peaks may only paint part of the picture. **Figures 3A, B** shows average Raman spectra of both cell lines treated with D_2O and different concentrations of methionine and insulin. These spectra are consistent with previous LD studies using Raman spectroscopy, which display minute protein peaks in the fingerprint region such as the phenylalanine peak at 1000cm^{-1} and amide I-III peaks at 1660cm^{-1}, 1450cm^{-1}, and 1200-1300cm^{-1}, respectively, as well as elevated CH_2 stretch at 2850cm^{-1}, saturated CH_2 stretch at 2880cm^{-1} (typical of cholesterol and other saturated lipids) (50), and H-C= stretch at 3010cm^{-1} (typical of unsaturated lipids) (51). Common lipid components of LD are shown in **Figure 3C** in descending order of prevalence. The structure of LDs is such that a phospholipid monolayer surrounds a core of neutral lipids such as cholesterols and TAGs. Less prevalent lipid species such as ceramides, sphingolipids, and their metabolites only account for a small percentage of LD composition, but have gained increasing significance in LD physiology and diseases (52). Furthermore, there are hundreds of apo-lipoproteins on or near the surface of LDs, which may contribute to the observed Raman spectra of LDs. The presence of the C-D peak in the spectra confirm *de novo* synthesis. Some Raman shifts of interest are shown, but minute differences may be difficult to discern by raw visual inspection alone.

Although the delineated Raman shifts in **Figures 3A, B** highlight several aspects of lipid and protein metabolism, there are others ascribed to lipids and other important molecules as well. Principal component analysis (PCA) shows that 12 principal components (PCs) account for nearly all the variance in the 6 groups of MDA-MB-231 LD spectra. To visualize this while avoiding over-fitting, a t-SNE diagram of the top 10 PCs is shown in **Figures 4A, B**. There exists at least one dimension that discriminates insulin effects and methionine effects on Raman spectra of TNBC LD. Importantly, this demonstrates that LD, alone, contain sufficient chemometric data to discriminate cell phenotypes. This confirms the ability of LDs to reflect cellular state. To date, label-free chemometric demonstrations of this ability are sparse. PCA initialization can be a robust step to reduce dimensions, denoise data, and preserve global structure in t-SNE visualizations, but even though PCA can vectorize these values, the PCs themselves do not take the form of Raman peaks suitable for direct assignment of methionine and insulin effects individually.

Statistical quantification of independent variable effects at every Raman shift entails a new measure in which the separation of insulin and methionine effects, as well as relative significance in class attribution is shown. To quickly rank and visualize all the wavenumber variables that may have been influenced by a particular treatment, the Kullback-Leibler divergence ($D_{KL)}$, a metric for the distance between two distributions for classification problems (53), at each Raman shift is plotted for each metabolite manipulation (**Figure 4C**). This method is also known as relative entropy.

KL divergences of Raman spectra were plotted on the same axes for MCF10A and MDA-MB-231 with selected wavenumbers labeled for clarity (**Figure 4C**). From **Figure 4C**, it is apparent the lipid peak of MCF10A cells at 2850cm^{-1} was heavily influenced by both insulin and methionine concentrations, while the protein peak at 2940cm^{-1} seems to be more heavily influenced by insulin concentration. This contrasts with the MDA-MB-231 TNBC cells in which divergences at most wavenumbers were dominated by the delineation of methionine concentration. Although these representations are not perfect (see **Supplementary Material**), this is especially useful when simultaneous treatment groups have both compound and independent effects. For example, MCF10A spectra (**Figure 4A**) exhibit changes in the unsaturated lipid peak (3010cm^{-1}) under either insulin or methionine manipulation, while the TNBC spectra (**Figure 4B**) exhibit changes here (3010cm^{-1}) mainly under methionine manipulation. This can be easily seen though the relative entropy at that Raman shift in **Figure 4C**, in which MCF10A has high relative entropy at 3010cm^{-1} when examined along either the insulin or methionine dimension, while TNBC shows a higher relative entropy when examined along the methionine dimension.

While excess methionine appears to decrease the lipid-to-protein ratio in MCF10A cells and increase the ratio in MDA-MB-231 cells, the results do not necessarily indicate discrepant rates of *de novo* lipogenesis since these values are affected by both synthesis and degradation of lipid and protein. For instance, the decreased lipid-to-protein ratio might be due to enhanced lipid utilization. To explore how much lipid and protein were synthesized, we quantitatively examined the carbon-deuterium (CD) peaks at 2135cm^{-1} (*de novo* synthesized lipids, CD_L) relative to 2180 cm^{-1} (*de novo* synthesized proteins, CD_P), and 2850cm^{-1} (total lipids, CH_2) for each treatment group (48) (**Figures 5A, B**). **Figure 5C** shows that excess methionine stimulates *de novo* lipogenesis in TNBC. Together, **Figure 5** illustrates both direct and relative *de novo* lipid and protein synthesis and metabolism, and informs the potential reasons for the discrepant lipid:protein effects of excess methionine on TNBC and normal-like breast cells.

Two-way ANOVA (**Table S2**) confirms a significant interaction term for methionine and insulin concentrations in TNBC for the *de novo* synthesized lipids relative to the total lipids (**Figures 5A, B** Right). Contrarily, only the insulin independent variable was significant for the MCF10A in the *de novo* synthesized lipids relative to the total lipids (**Figures 5A, B** Right), but no interaction term, or even a significant methionine

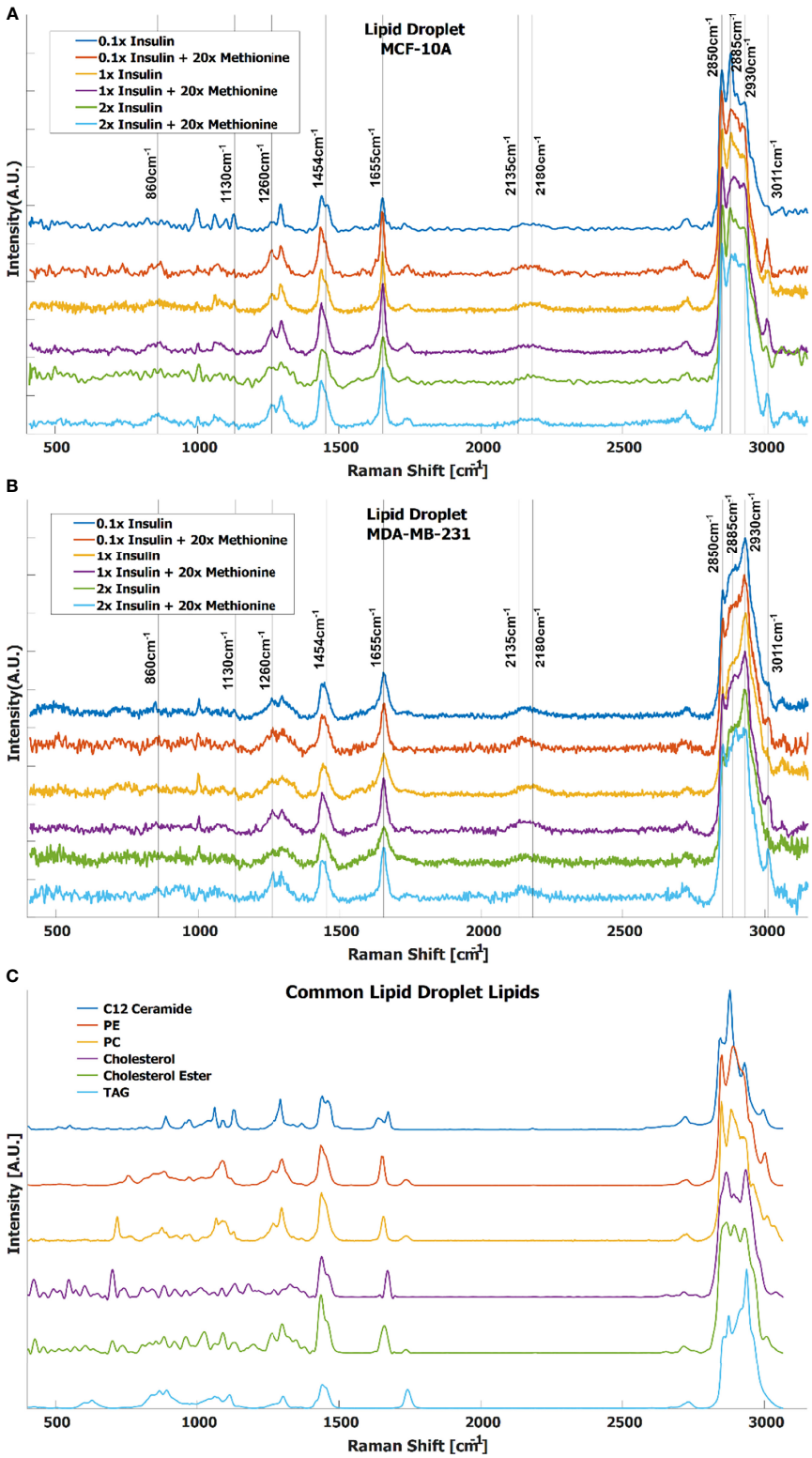


FIGURE 3 | Average spectra of LD from **(A)** MCF-10A and **(B)** MDA-MB-231 cells under various methionine and insulin concentrations. Manual identification of potential Raman peak targets is highlighted with vertical lines and labels. For example, the CD_{Lipid} peak in the cell silent region at 2135cm^{-1} has a noticeable increase relative to the $CD_{Protein}$ peak at 2180cm^{-1} in the excess methionine groups. This could indicate preference for *de novo* lipogenesis in excess methionine environments. **(C)** Raman spectra of common lipid species in LD, in descending order of prevalence.

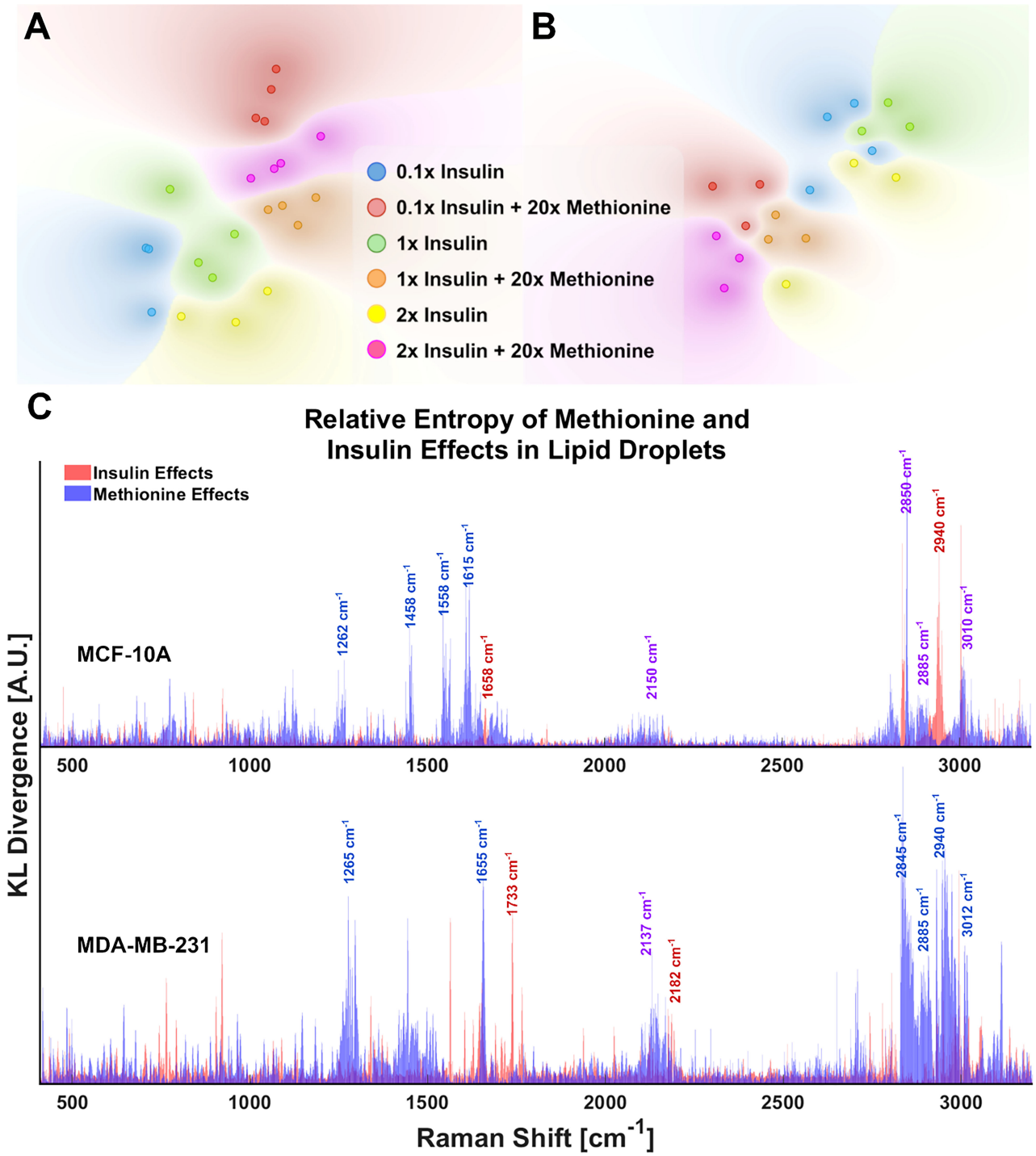


FIGURE 4 | LD chemometric data is sufficient to discriminate cell phenotype **(A)** tSNE separates experimental groups for MCF-10A and **(B)** MDA-MB-231 tSNE plots of the top 10 PCs from PCA of LD spectra. Global structure is preserved, and no exaggeration was applied. Each point represents the average of 5 LD spectra taken from a single cell of the corresponding sample. **(C)** Relative entropy provides a metric for ranking features (Raman peaks) by their ability to classify the spectra as belonging to 20x methionine or 1x methionine groups, as well as 0.1x, 1x, 2x insulin groups. Raman peaks that appear to be influential in both classification schema are denoted in purple text labels for clarity. These cell subtype plots are aligned manually for clarity. Subplots are not generated in such a fashion automatically.

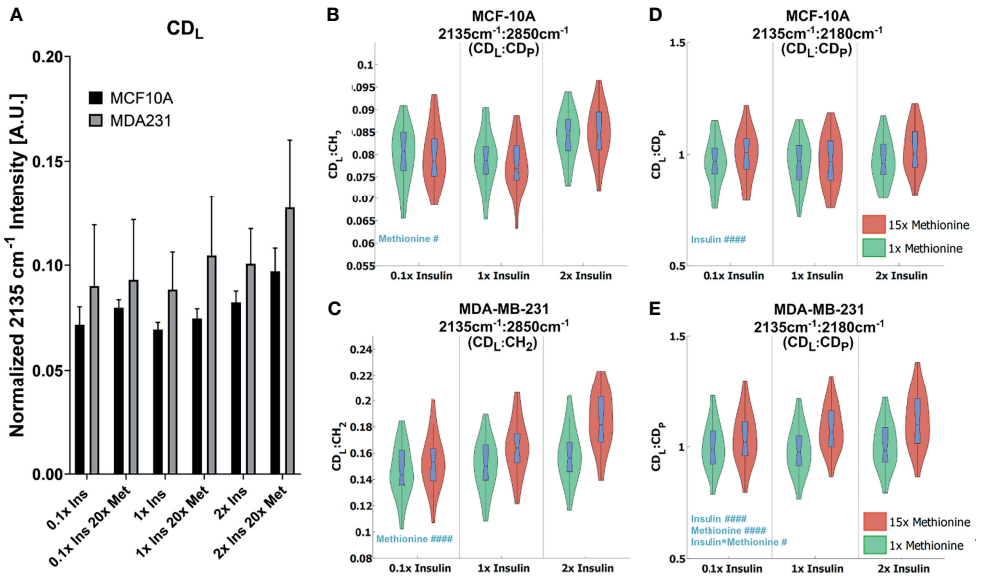


FIGURE 5 | Quantitative *de novo* lipid synthesis **(A)** Normalized CD_L intensities show excess methionine stimulates *de novo* lipogenesis. **(B, C)** CD_L Ratios show violin box-plots of *de novo* lipid synthesis CD_L ratios for MCF10A and MDA-MB-231, respectively. CDL : CH2 illustrates the relative *de novo* lipid synthesized to total ascribable lipid content. Balanced 2-way ANOVA with constrained sum of squares results of CD_L ratios shows methionine concentration significantly influenced the CD_L: CD_P ratio in both MCF10A and MDA-MB-231 lipid droplet spectra with rejection levels of $^{\#}P < 0.05$ and $^{\#\#\#\#}P < 0.0001$, respectively. **(D, E)** CD_L: CD_P illustrates the relative *de novo* lipid and protein synthesized biomolecules for MCF10A and MDA-MB-231, respectively. Values were taken from spectra of lipid droplets only. while. There was no significant evidence of interactions between these two independent variables for these ratios. Balanced 2-way ANOVA with constrained sum of squares results of CD_L ratios indicate insulin significantly influenced the CD_L: CH_2 ratio in MCF-10A lipid droplet spectra with a rejection level of $^{\#\#\#\#}P < 0.0001$, but no significant evidence of interactions between these two independent variables. However, in TNBC insulin, methionine, and the interaction term significantly influenced the CD_L : CH_2 ratio in MDA-MB-231 lipid droplet spectra with a rejection level of $^{\#\#\#\#}P < 0.0001$ and $^{\#}P < 0.05$ for the individual and interaction terms, respectively.

term. Only in TNBC did the methionine term have a significant impact on this ratio, which lead us to believe the discrepant effects on CH_2:CH_3 ratios we observed (**Figure 2B**) might arise from differential *de novo* lipogenesis, rather than protein synthesis and metabolism. Since the excess methionine stimulated *de novo* lipogenesis (**Figure 5C**) and was a significant term in the ratio of *de novo* synthesized lipids to proteins for both cell lines (**Figures 5A, B**), methionine is likely to preferentially stimulate lipid production more than protein production. Despite these findings, the relative proportion of lipids to proteins in MCF10A still decreases under excess methionine (**Figure 2B**). Therefore, either the pool of proteins must be getting larger, or the lipid utilization must increase. Excess methionine did not stimulate lipid utilization because CD_L : CH_2 did not significantly increase (**Figure 5A** Right). This leads us to believe that the protein signal must increase excess methionine. However, excess methionine did not stimulate protein production faster to a greater extent than lipid production since the CD_L : CD_P slightly increased under excess methionine (**Figure 5A** Left). If there was no relative increase in protein nor decrease in lipids, then MCF10A may not breakdown proteins as much in the presence of excess methionine, or uptake and retain the excess methionine itself more efficiently than TNBC. The hydrophobic amino acid can interact with the acyl chains of the fatty acids in lipid droplets, and since the excess methionine supplied was not deuterated, this protein would not appear in the cell silent region. This way, the excess methionine can affect the CH_3 peak without affecting the CD_P peak, and explain the behavior observed in **Figures 2**, **5**. Excess methionine can also incite endoplasmic reticulum stress due to complex interactions with cysteine pathways since both are sulfur containing and are critical in protein folding due to disulfide bonds. These misfolded proteins may be sequestered by LDs differently across subtypes.

Morphological Changes in Lipid Droplet

3D SRS images were taken for each individual cell at 2850cm^{-1} (**Figures 6A–D**) to assess the size and number of LDs more accurately. LDs were computationally segmented using MATLAB (**Figures 6E–H**) to acquire individual LD volume and number of LDs per cell. The addition of excess methionine produced the most noticeable changes in lipid droplet morphology – a decrease in lipid droplet number but increase in volume. This effect was observed in both MCF10A and TNBC cells (**Figures 6I, J**). Of note is the insulin restricted case in TNBC cells, which had no discernible change to lipid droplet number or size. Qualitatively, the lipid droplets also appeared more clustered in excess methionine cases. Lipid droplet volume was also observed to slightly increase from restricted insulin

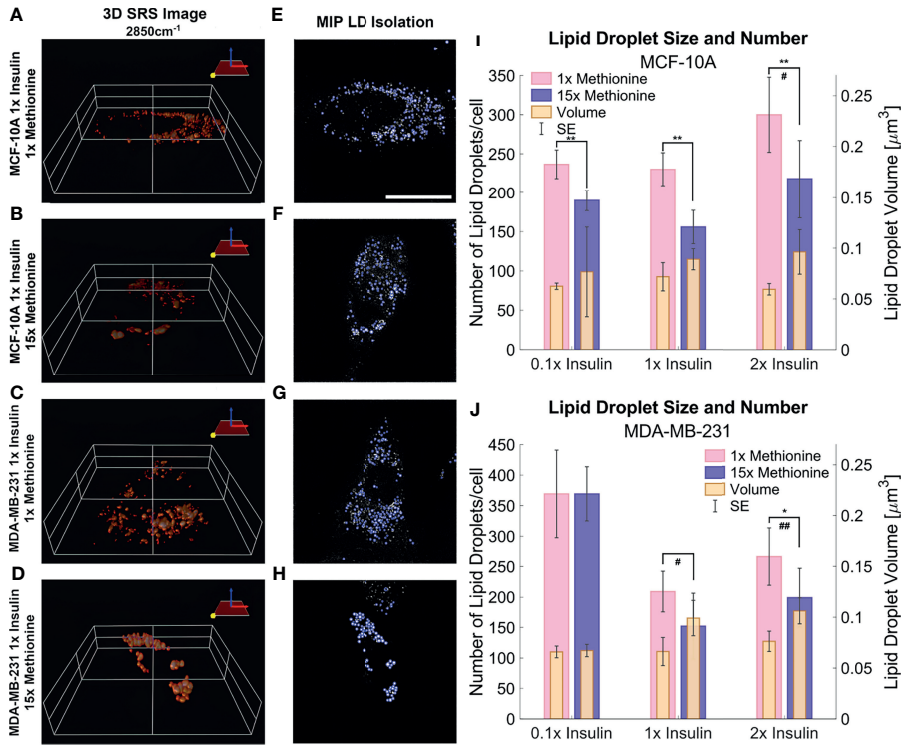


FIGURE 6 | 3D SRS image lipid droplet analysis **(A–D)** 3D isosurface reconstructions of the single cell SRS images taken at 2850cm^{-1}. **(E–H)** LD segmentation shows representative maximum intensity projections of SRS image stacks shown in **(A–D)** with lipid droplets highlighted in blue outlines. **(I, J)** Quantitative LD structure summary shows average lipid droplet number and volume for each experimental group. Excess methionine groups display decreased lipid droplet number and increased size. Lipid droplets also appear qualitatively more clustered in excess methionine as well. Scale bar is 20 μm. Two-tailed t-tests were performed between each pair of bars to highlight excess methionine effects. Asterisks '*' correspond to the following p-values for LD number: *P < 0.05, **P < 0.01. Octothorps '#' correspond to the following p-values for LD volume: $^{\#}$P < 0.05, $^{\#\#}$P < 0.01. Scale bar is 20 μm.

(0.1x) to physiological and excess insulin (1x and 2x) in TNBC under excess methionine conditions. This corroborates the potential interplay between insulin and methionine in TNBC.

Lipid droplet distribution can be a major indicator of cell cycle status, nutrient availability, and ER stress (54). LD size may influence the degree to which beta oxidation occurs in cells and be affected by mitochondrial recruitment during LD expansion in nutrient rich environments. Regardless, the physical contact between these organelles is thought to mediate their proper function (55, 56). A label-free approach to identifying mitochondrial presence near lipid droplets may be the spectral presence of cytochrome C (cytC), which is found in the intermembrane space of mitochondria. Some peaks canonically representative of cytC are the heme backbone at 1558cm^{-1} and the side chains of tryptophan, tyrosine, and phenylalanine in alpha structures at 1610cm^{-1} (57) which were weakly present near the fingerprint region of the spectra in MCF10A and TNBC cells. It was found in **Figures 7A**, **8A** that excess methionine cases displayed a diminished spectral presence of unmixed cytC peaks. This suggests LDs in excess methionine may cluster near other organelles such as lysosomes, or even with other LDs for fusion events, instead of co-localizing with mitochondria for energy. Examples of Gaussian-Lorentzian peak unmixing for MCF10A and TNBC (**Figures 7B**, **8B**), respectively, with quantitative summaries in the form of bar graphs (**Figures 7D**, **8C, D**). The number of unmixed peaks was optimized such that the overall fit is accurate, while the unmixed peaks are easily ascribed to canonical protein and lipid deformations. The Amide II' region contains various CH_2 and CH_3 deformations such as wagging, stretching, scissoring, and twisting (58, 59). The Amide I region contains secondary structure information and has been used to study proteins such as collagen (60). Between these peaks lies the C-C bond of the heme backbone. MCF10A and TNBC exhibited distinct peak shapes in all areas of this region. In MCF10A, the Amide II' peak had a narrower shoulder at 1458cm^{-1} under excess methionine (**Figure 7C**), while TNBC had a narrower Amide I peak under excess methionine (**Figures 8E–G**). The Amide I and II' regions also contain protein and lipid information and have various assignments in the literature. **Figure 8G** quantifies the width and prominence of the Amide I peak in TNBC with and without excess methionine. Results indicate altered protein folding, in which methionine plays crucial roles. Methionine is not only a protein translational initiator, but its metabolism is also involved in purine synthesis, epigenetic control, and secondary disulfide bond formation (24). Misfolded proteins have tangible effects on ER stress and lipid

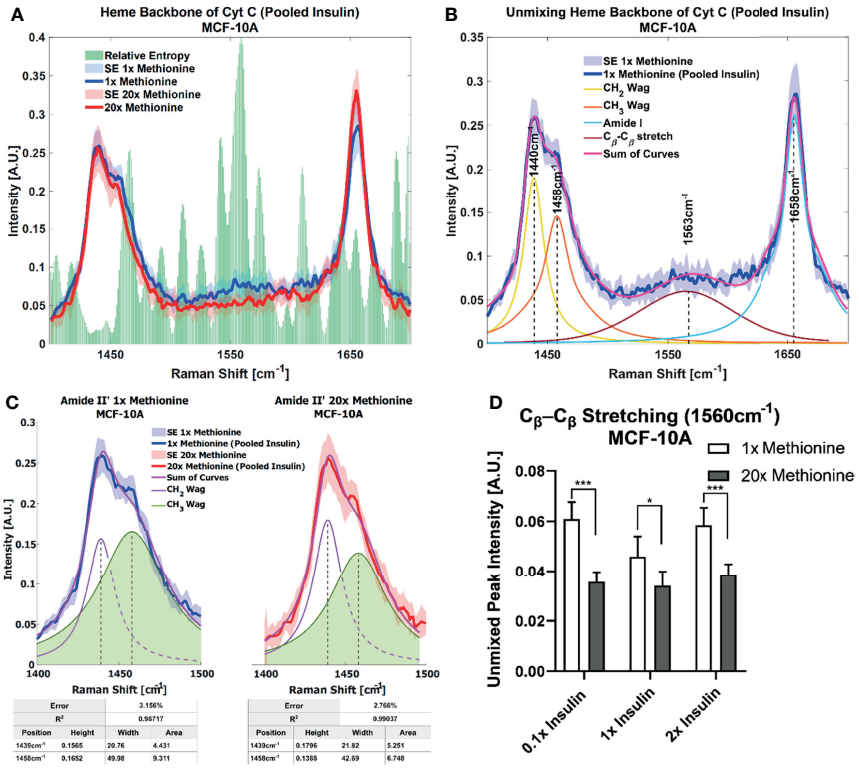
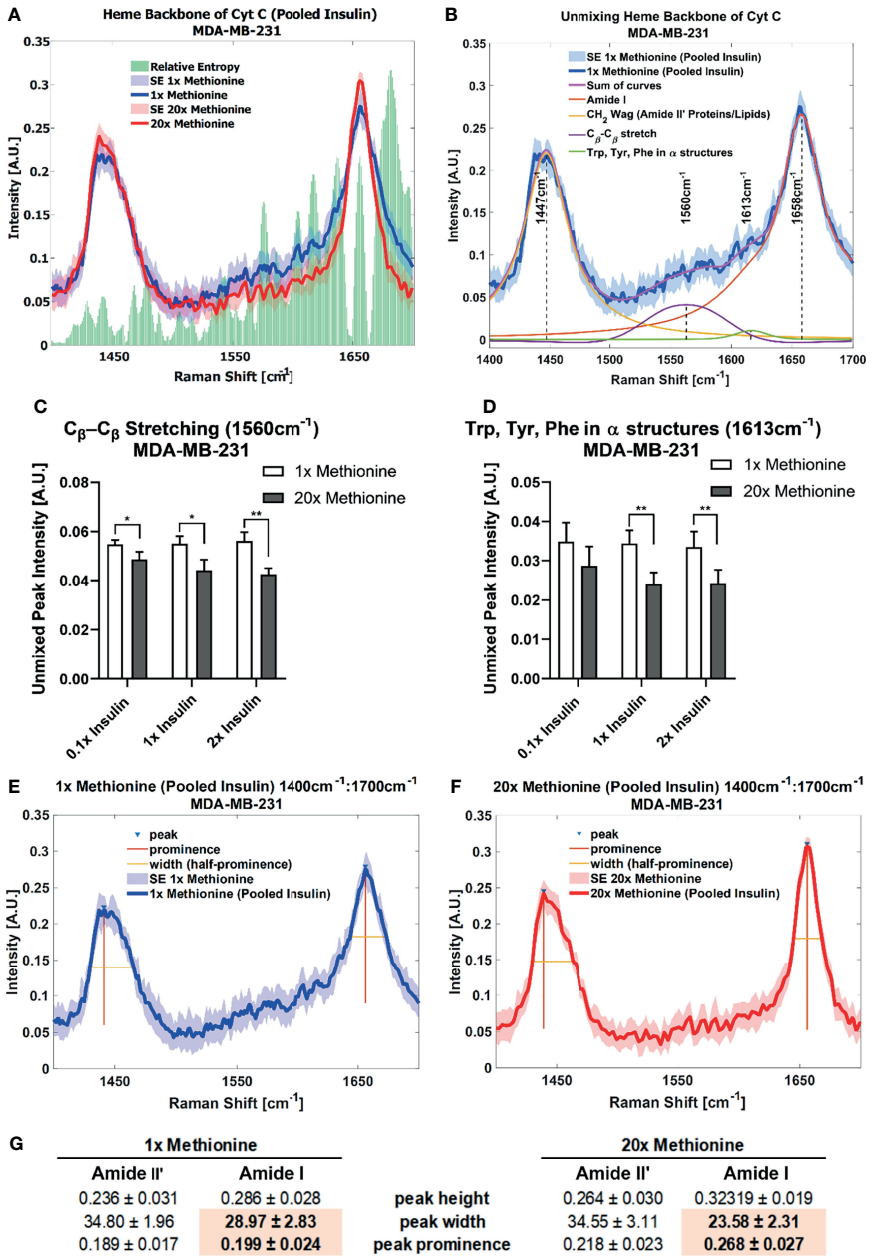


FIGURE 7 | Spontaneous Raman Spectroscopy detects CytC presence and protein folding **(A)** normal vs excess methionine Expanded view of lipid droplet spectra grouped by methionine concentration shows a high relative entropy in the 1550cm^{-1} region, ascribable to the heme backbone of cytochrome C. **(B)** Unmixing Peaks with four peaks using a Gaussian-Lorentzian blend yields an error of 2.367% and an R^2 of 0.98854. **(C)** Amide II' Peak Shoulder shows an **e**xpanded view of normal and excess methionine groups' Amide II' regions highlight a relatively narrowed shoulder at 1458cm^{-1}. Unmixed peaks follow the overall shape of the average Amide II' peaks, with the error and correlation coefficient reported in the table below. Width and area information is also summarized in the table to clearly communicate the disparate shoulder widths. **(D)** Quantitative summary of the heme backbone unmixed peak intensities for each experimental condition of MCF-10A cultures shows decreased spectral presence. Two-tailed t-tests were performed between each pair of bars to highlight excess methionine effects. Asterisks '*' correspond to the following p-values: *P < 0.05, ***P < 0.001.

droplet distribution and chemistry (13), as these proteins have been shown to accumulate in LDs destined for proteasomal breakdown (54). In this manner, LDs may serve as reservoirs and chaperones to mitigate lipid and protein toxicity. Although further investigations are required to confidently assign the phenomenon observed herein, the fact that consistent alterations in these areas were observed using label-free vibrational imaging techniques sets the stage for more in-depth studies of dietary methionine-controlled protein folding in breast cancer cells. **Figure 7C** quantifies the changes in the Amide II' peak of MCF10A and may indicate altered lipid and protein structure as well. Various bond deformations occur at slightly different wavenumbers, with CH_2 scissoring being red-shifted with respect to CH_2 stretching. Acyl chains of different length and saturation may influence the degrees to which each of these deformations take place. Further investigation into purified LD content with other techniques such as gas chromatography and mass spectrometry are warranted. Spectroscopic data are usually sensitive to baseline correction, background subtraction, and normalization methods, and are therefore better suited to relative observations, while chromatography and spectrometry offer absolute quantification and detailed chemical structure. Conjugating these techniques is beyond the scope of this label-free optical platform, but is promising and critical step in progressing this technology.

Lipid Peroxidation Status

Another global lipid response to excess methionine takes form in the lipid peroxidation status. Under oxidative stress, long chain unsaturated fatty acids can undergo a vicious cycle of lipid peroxidation (51). Several Raman shifts have been used to describe the degree of unsaturation of fatty acids, including the one near 3010 cm^{-1} that corresponds to the H-C= stretching region (51). Interrogating the relative entropy plot in **Figure 4C**, we find that the saturated lipid peak at 3010 cm^{-1} and the lipid peak at 2850 cm^{-1} both rank highly for both cell types, but TNBC is more heavily influenced by methionine concentration. That is, we can see from the spectroscopic data that MCF10A, whether L-methionine was normal or in excess, expressed relatively different levels of unsaturated lipids



G

1x Methionine			20x Methionine	
Amide II'	Amide I		Amide II'	Amide I
0.236 ± 0.031	0.286 ± 0.028	peak height	0.264 ± 0.030	0.32319 ± 0.019
34.80 ± 1.96	**28.97 ± 2.83**	peak width	34.55 ± 3.11	**23.58 ± 2.31**
0.189 ± 0.017	**0.199 ± 0.024**	peak prominence	0.218 ± 0.023	**0.268 ± 0.027**

FIGURE 8 | Spontaneous Raman Spectroscopy detects CytC presence and protein folding differences **(A)** normal vs excess methionine shows expanded view of lipid droplet spectra grouped by methionine concentration shows a high relative entropy in the 1550cm^{-1} and 1650cm^{-1} regions, ascribable to the heme backbone of cytochrome C and side chains of tryptophan, tyrosine, and phenylalanine, respectively. **(B)** Unmixing peaks with four peaks using a Gaussian-Lorentzian blend yields an error of 2.620% and an R^2 of 0.98544. **(C, D)** Quantitative summary of the unmixed peak intensities for each experimental condition of MDA-MB-231 cultures shows decreased spectral presence of cytochrome C. **(E, F)** Amide I peak width shows an expanded view of lipid droplet from 1x Methionine (LEFT), and 20x Methionine (RIGHT) experimental conditions. Two-tailed t-tests were performed between each pair of bars to highlight excess methionine effects. Asterisks '*' correspond to the following p-values: *P < 0.05, **P < 0.01. **(G)** Peak analysis shows that the peak prominence and peak width at half prominence is significantly narrower at the Amide I region in excess methionine lipid droplet spectra.

depending on the level of insulin. This suggests that *de novo* synthesis of branched chain fatty acids, or perhaps their accumulation in LDs was upregulated in excess insulin conditions. So, while insulin was critical in influencing *de novo* synthesis of lipids in TNBC, it may not influence lipid peroxidation as much as methionine does. **Figure 9** shows the effects of excess methionine in TNBC using multi-modal optical techniques.

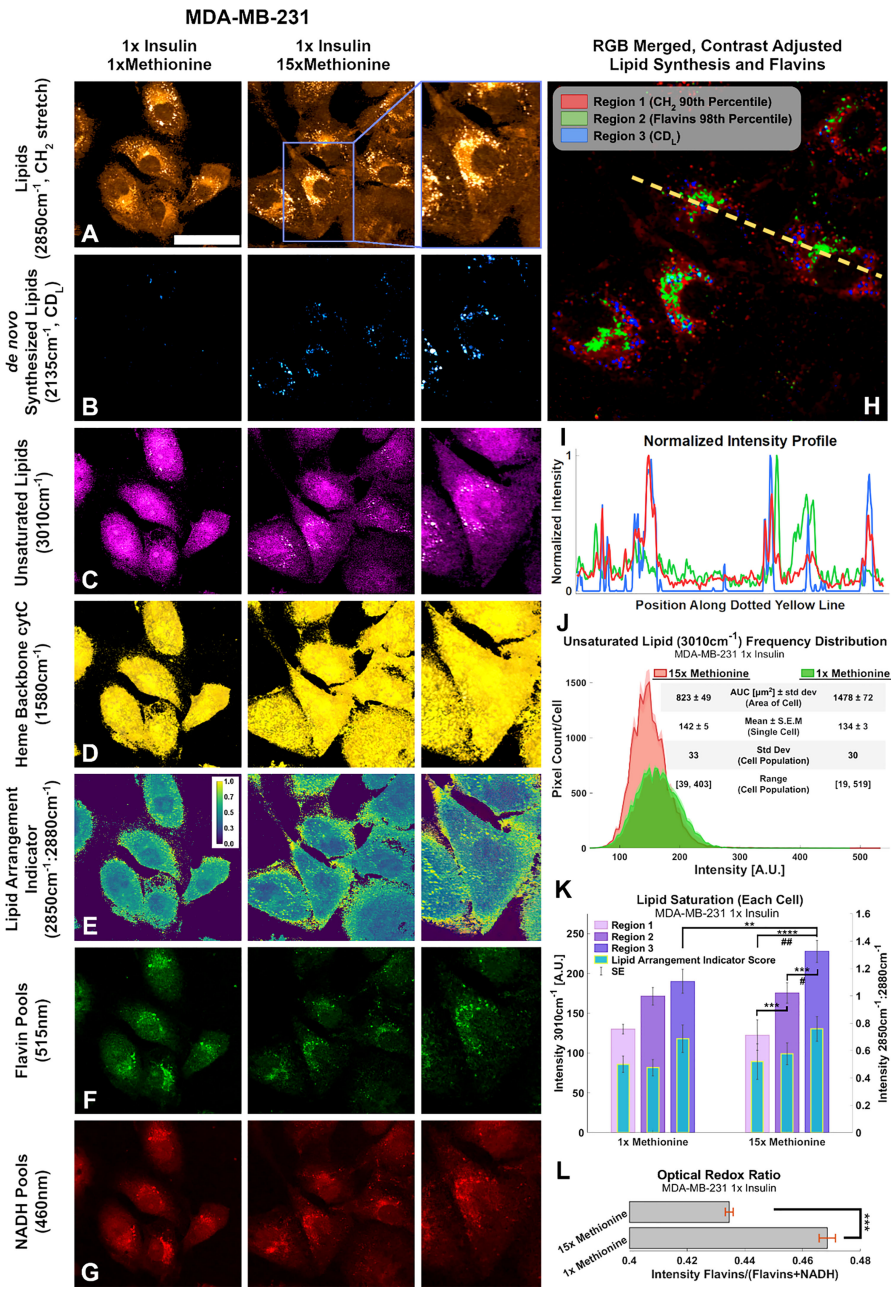


FIGURE 9 | Multi-modal optical analysis depicts MDA-MB-231 1x insulin sample images demonstrate the conjugated SRS and TPF system. **(A-G)** Multichannel Images illustrate SRS and TPF image channels of interest for lipidomic responses to excess methionine. **(H)** Overlaid Composite Regions image of the 15x Methionine lipid (CH_2), flavins, and *de novo* synthesized lipids (CD_L). Channels were masked according to the indicated thresholds using ImageJ and contrast was adjusted for optimal clarity. **(I)** Intensity profile plot depicts the intensities of pixels along the dotted yellow line shown in **(H)** of each of the three composite channels. **(J)** Composite intensity histograms of the unsaturated lipid channel ($3010cm^{-1}$). Bolded distribution outlines represent the average frequency of pixel intensities among the cells in each group. Shaded areas around the bolded distribution outline represent the standard error of the mean of each bin of pixel intensities. Each distribution curve represents the pixel intensities of a cell sampled from the experimental condition. **(K)** Quantitative Lipid saturation summary depicts the $3010cm^{-1}$ pixel intensities (Left axis) in each of the three regions shown in **(H)** of a typical cell from the indicated experimental condition. Additionally, the lipid arrangement indicator ratio ($2850cm^{-1}$:$2880cm^{-1}$) for each of the regions in **(H)** is also depicted for the typical cell from each experimental condition. Two-tailed t-tests were performed between each pair of bars to highlight excess methionine effects. Asterisks '*' correspond to the following p-values for the unsaturated lipid peak ($3010cm^{-1}$) intensities: $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$. Octothorps '#' correspond to the following p-values for the Lipid Arrangement Indicator ($2850cm^{-1}$:$2880cm^{-1}$) intensities: $^{\#}P < 0.05$, $^{\#\#}P<0.01$. **(L)** Optical redox ratio (Flavin/(Flavin+NADH) autofluorescence intensity) for the typical cell from each experimental condition. Results corroborate and extend spectral data findings, as well as previous third-party studies. Scale bar is 50um. Two-tailed t-tests were performed between each pair of bars to highlight excess methionine effects. Asterisks '*' correspond to the following p-values for the unsaturated lipid peak ($3010cm^{-1}$) intensities: $^{***}P < 0.001$.

Conjugated SRS and TPF microscopy display spatial distributions of points of interest regarding excess methionine effects in TNBC cells (**Figures 9A–G**). These results also corroborate with the spectroscopic data. **Figure 9B** reveals that the cells undergo enhanced *de novo* lipogenesis under excess methionine with respect to control groups. Contrary to expectations, the unsaturated lipid signal in the excess methionine group was weaker than the control overall but was stronger near the large lipid droplets (**Figure 9C**). This information is lost in spectral acquisitions alone because spectra were obtained from lipid droplets only. Excess methionine treated cells exhibited larger cross-sectional area (**Figure 9J**) and may be due to the cells being more spread out and flatter. Due to the point spread function of the confocal laser scanning microscope, this spreading out of the cells may contribute to an apparent decrease in concentration of unsaturated fatty acids because the scattering cross section along the beam path is smaller. Consistent contrast makes it difficult to discern the abundance of smaller lipid droplets in the control images without oversaturating the excess methionine images. There were no discernible differences in spatial distribution of heme groups at the 1580 cm^{-1} (**Figure 9D**), but co-localization algorithms may help in future studies. The indicator of crystalline arrangement in lipids corresponding to the symmetric:antisymmetric CH_2 stretching ratio indicates that the excess methionine group may have less lipid saturation near the plasma membrane (**Figure 9E**). Higher ratios would indicate a lower concentration of 2880 cm^{-1} species, which has been ascribed to the Fermi resonance of CH methylene groups (50). This ratio has been found to inversely correlate with thermodynamic stability, and when the ratio is larger, there may be less lattice order in the structure (61). In the context of cell membranes, fluidity and saturation are critical functional properties, and the decreased lipid saturation score near the plasma membrane may also contribute to the observed "flattness" of the cells with excess methionine, as the cells may be able to spread out more easily.

Different areas of the cells provide niche microenvironments, in which lipid peroxidation may vary. Three subcellular regions of interest include where all lipids exist (**Figure 9H**, region 1), where flavins are more present (region 2), and where newly synthesized lipids are present (region 3). As shown in **Figures 9H, I**, these regions do not necessarily overlap. Flavins have been shown to report on oxidative stress, and certain flavin enzymes have been associated with lipid peroxidation as well. The quantitative image analyses of unsaturated lipids (3010cm^{-1}) and the lipid arrangement indicator ratio (2850cm^{-1}:2880cm^{-1}) are summarized in **Figure 9K**, in which distinct regions are separately quantified. **Figure 9J** highlights a larger cross-sectional area of the imaged cells, which may be afforded by a more fluid cell membrane. In certain cells, oxidative stress has been found to increase lipid saturation for protection. Furthermore, the presence of higher ratios near LDs suggests there is less synthesis of saturated lipid species as well. Finally, flavin autofluorescence decreased in the presence of excess methionine (**Figure 9F**), while NADH autofluorescence remained more consistent (**Figure 9G**). The flavin/(NADH + flavin) ratio has been shown to be an indicator of oxidative stress and estimator for NAD^+:NADH (62). Results corroborate with previous studies in which this ratio was used to differentiate breast cancer cell lines (63), with the TNBC having relatively weaker flavin autofluorescence than the normal-like cell type. Under oxidative stress, this ratio has been shown to increase. A decrease here (**Figure 9L**) may demonstrate the antioxidant properties of methionine. Flavin autofluorescence data is summarized in **Figure S2**.

DISCUSSION

For the first time, the unique lipid metabolism of triple negative breast cancer was studied under tandem excess methionine and insulin conditions, and revealed key insights that span the metabolic, spatial, and biochemical dimensions. Not only did this study confirm lipid droplets are reflective of cellular phenotypes and demonstrate their efficacy in classifying breast cancer subtypes, and even phenotypes, it improves morphological analysis using 3D imaging, as opposed to 2D, and efficiently displays relevant chemical disparities using the first demonstration of relative entropy for Raman data. Considering the critical impact lipid metabolism has on the progression of diseases such as cancer, the analyses on lipid saturation and peroxidation, optical redox status, and LD size and distribution solidify the effects of methionine and insulin, which may prove to be therapeutic targets for breast cancer in the future.

These experiments demonstrate the power of nearly label-free optical techniques to probe LD phenotypes for the study of TNBC's unique metabolism. Methionine dependence, also known as the Hoffman effect, has been explored in TNBC and other cancers, but fewer studies explored the effects of excess methionine, and fewer still, the tandem manipulation of methionine and insulin. Upon the addition of insulin in TNBC growth media, macromolecular changes appeared in the CH stretching region of excess methionine treated cells, as the CH_2:CH_3 ratio increased in TNBC, but decreased in MCF10A control cells. A potential pathway that involves both insulin and methionine in LD metabolism may be mediated by TNBC's elevated mTOR activity, and was explored through the chemometric, spatial, and molecular imaging dimensions with subcellular resolution. Currently the stoichiometric mass action of this pathway remains to be investigated in these breast cancer subtypes, but several studies have linked methionine, mTOR, and insulin signaling pathways (25–30), albeit transitively. Paramount in this investigation is the implication of these metabolites in the pursuit of TNBC diagnosis and treatment. Unmixing the interplay between insulin and methionine may afford targeted therapies that address the rampant lipid metabolism that facilitates breast cancer progression.

LD chemical composition also demonstrated excellent classification ability, as lipid and protein Raman cross sections are not only larger, but also very diverse and highly implicated in

metabolic cellular states. Classification of breast cancer subtypes, and even more so their phenotypic states, can be critical in improving patient outcomes due to the necessity of early diagnosis. MCF10A cells may exhibit differential protein metabolism by retaining scavenged methionine or not metabolizing proteins as much as TNBC, which is due, in part, to very different doubling times. Several other considerations including LD size may also contribute to these disparities, as larger LDs have a greater volume:surface area ratio, and thus a greater lipid:protein ratio since the apolipoproteins stud the phospholipid surface of the lipid core. LD fusion can affect this as well, since newly synthesized LDs may have a diluted CD signal if fused with older LDs. Further investigation is warranted to uncover the details of differential lipid metabolism in breast cancer subtypes using LDs, but this platform has set the stage for macroscopic observations using efficient optical techniques.

Both TNBC and MCF10A cells exhibited similar spatial information of LDs under these treatments as excess methionine conditions generally decreased the number of lipid droplets while increasing their volume in both cell types, while insulin generally increased both size and number of LDs. Insulin restriction appeared to increase LD number as well, and in TNBC, appeared to attenuate the effects of excess methionine on LD number. This interesting pattern not only suggests that TNBC has unique insulin-mediated lipid metabolism, but that insulin and methionine may have more complex concentration-dependent interactions in general as well. LD distribution also seemed to be more clustered in the excess methionine cases, and we intuit from the diminished spectral presence of cytC that these lipid droplets were less co-localized with mitochondria. Finally, the antioxidant properties of methionine expectedly diminished flavin autofluorescence and resulting lipid droplet spectra showed higher degrees of lipid unsaturation. In **Figures 9K**, L the optical redox ratios and the lipid arrangement indicator ratios indicate that methionine plays a large role in lipid peroxidation and saturation. The degree of saturation of lipids is a critical consideration for cell membrane fluidity, especially in aggressive cancers that can alter their extra cellular matrix (ECM), or those that metastasize and migrate rapidly. For the first time, the dynamics of lipid saturation and peroxidation under nutritional control has been imaged with label-free subcellular resolution.

To broaden the scope of the investigation and capitalize on the rich chemical data of the Raman spectrum, relative entropy was used to rank the features that exhibited the greatest variance between different groups. As expected, there are several areas other than the CH stretching region that offer strong classification ability despite lower Raman intensities. This may be attributed to the higher intensity deviations at higher intensities typical of multiplicative scattering effects. Additionally, the effects of individual nutritional manipulations become clearer with all Raman peaks being visible simultaneously. From this insight, the interrogation of pathways with Raman spectra can be done more efficiently, as the relative entropy scores for each Raman shift can be seen at once, reducing the number of spectra and subplots that need to be displayed. With this demonstration of efficacy, more critical quantitative analyses, as well as algorithmic improvements will be conducted. For example, incorporation of directional shifts in intensities can be made visible on the relative entropy plot, as opposed to absolute distance metrics alone. This will not only identify discriminating variables but will also circumvent the need to manually determine significant ratios, ratio differences, and other trends as well. Further, this relative entropy plot may be useful in feature reduction, so that fewer hyperspectral images may be required for discriminating LD microenvironments and subpopulations. Other methods more directly identify the wavenumbers that contribute the most to a spectrum's classification, such as the hybrid variable combination population analysis (VCPA) and iterative retaining important variables (IRIV) approach (64). However, due to the large number of variables, IRIV can be time and resource intensive.

The diverse pathophysiology of breast cancer may have important mechanisms involving methionine and insulin that can be studied with optical techniques such as spontaneous Raman spectroscopy and SRS/TPF microscopy. This study also emphasizes that LDs are organelles diverse in structure and function and can yield rich metabolic information when interrogated by Raman techniques. Future studies that involve automated high-throughput acquisitions of spectra and images at more finely tuned concentrations of insulin and methionine may increase the power of the results discussed here. Different distribution fits for the relative entropy algorithm, displaying directionality of peak intensity changes, as well as the multiplexing of dietary manipulations such as glucose, pyruvate, and glutamine may paint a clearer picture of the metabolic dynamics in breast cancer (65, 66). This will also help make hyperspectral imaging more efficient in terms of disk space and clustering ability. Utilizing morphological characteristics and intensity changes to augment classification has not been performed in this study but will be a prudent next step in developing these optical techniques for classification purposes. Additionally, spatial distribution of LD by size and chemometric composition, as well as quantitative descriptions of LD distribution and co-localization would further enrich this investigation. This kind of quantitative hyperspectral image data will bolster the utility of LD analysis in the study of breast cancer, and ultimately improve not only our understanding of the complex disease, but patient outcomes in eventual translation as well.

MATERIAL AND METHODS

Experimental Design

An experimental outline is shown below (**Figure 10).** First, three cell subtypes were grown in media with either 1x methionine (0.03g/L) or 20x methionine (0.6g/L). Then the experiment was repeated with the addition of 3 insulin concentrations for each of the groups to investigate their relationship.

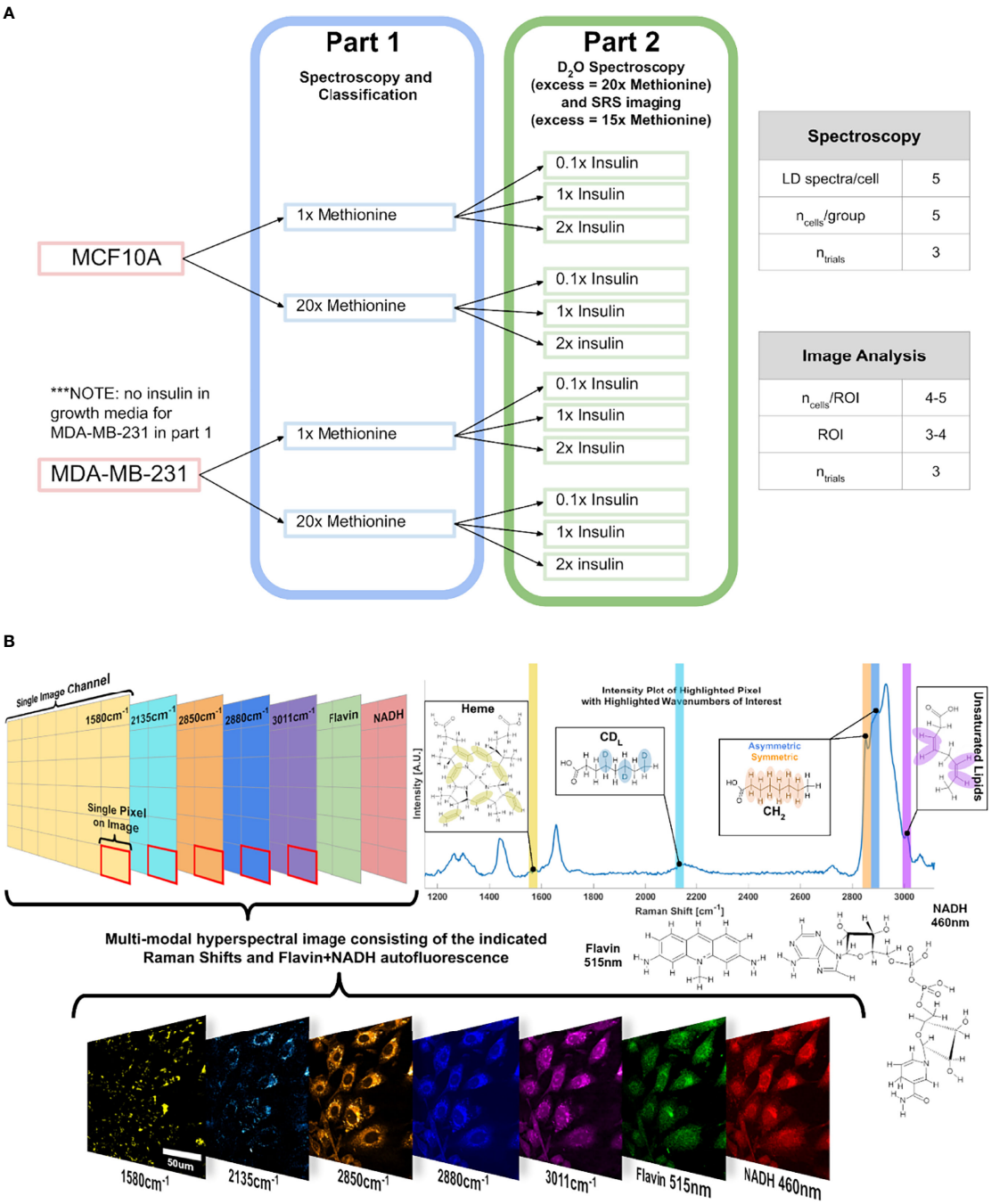


FIGURE 10 | Experimental Design and Points of Interest **A** Experimental design illustration of the groups in this study. Only methionine concentration is modulated at first, and then insulin and methionine were modulated in tandem. DMEM used in this experiment already contains 0.03g/L of L-methionine, which corresponds to the 1x methionine group. NOTE: in part 2, the excess methionine concentration is 20x for the Raman spectroscopy, and 15x for SRS imaging. **(B)** Points of interest where SRS images relating to lipid metabolism are acquired, related to the lipidomic investigation. Hyperspectral Image (HSI) format is shown conceptually to convey the multi-modal approach to quantitative optical analysis. Vibrational modes are color coded, with an example image of a HSI of MCF10A control cells.

Cell Culture

Human triple negative breast cancer cell line (MDA-MB-231) and normal-like breast epithelial cell line (MCF10A) were cultured in Dulbecco's modified Eagles' medium (DMEM), supplemented with 5% fetal bovine serum (FBS) and 1% penicillin/streptomycin (Fisher Scientific, Waltham, MA), and incubated with 5% CO_2 at 37°C. Cell cycles were synchronized using double thymidine block (67). After passaging at 80%

confluence, cells were seeded at a concentration of 2×10^5/mL atop 70% ethanol-soaked cover glass in 24-well plates and incubated for 8 hours. Then the growth media was changed to 50% heavy water (D_2O) and treatment media as follows.

For MDA-MB-231 and MCF10A cell culture media, 57 mg/L and 42 mg/L L-methionine (M8439, Sigma Aldrich) was added to DMEM for the excess methionine group for spontaneous Raman spectroscopy and SRS imaging (20x and 15x concentration), respectively. The DMEM powder used in this study already contains 0.03g/L (1x concentration) L-methionine and corresponds to the physiological concentration group. The reason for lowering the excess methionine concentration for SRS imaging analysis is because the cell morphological changes were more varied and poorer with 20x methionine, making it more difficult to acquire quantitative metabolic activity from images on a per-cell basis. Insulin (Sigma Aldrich, St. Louis, MO) was added at 1μg/mL, 10 μg/mL, and 20 μg/mL for the 0.1x, 1x, and 2x insulin groups, respectively.

Cells were incubated for 48 hours, which corresponds to a deuterium-retarded cell cycle. Cyclin dependent kinase 1 (CDK1) inhibitor (RO-3066, Sigma) was added with 8 hours remaining to arrest growth before mitosis. Cells were gently rinsed with 1x PBS with Calcium and Magnesium ions at 4°C (Fisher Scientific, 14040216), and finally fixed in 4% methanol-free PFA solution (VWR, 15713-S) for 15 minutes. The cover glass was mounted on 1mm thick glass microscope slides with 120 μm spacers filled with 1x PBS without calcium and magnesium ions. These samples are stored at 4°C submerged in PBS when not in use.

Spontaneous Raman Spectroscopy

Spontaneous Raman scattering spectra were obtained by a confocal Raman microscope (XploRA PLUS, Horiba) equipped with a 532 nm diode laser source and 1800 lines/mm grating. The acquisition time is 30 s with an accumulation of 4. The excitation power is ~40 mW after passing through a 100x objective (MPLN100X, Olympus). The background spectra were taken for each LD at the same focus plane as the LD and were subtracted from each LD spectrum immediately. Spectra were preprocessed using vector normalization and simplex normalized. Peaks were normalized to the protein peak at 2940 cm^{-1}. Previous studies suggest Raman microspectroscopy can quantify lipids non-invasively (68).

Stimulated Raman Scattering Imaging Microscopy

An upright laser-scanning microscope (DIY multiphoton, Olympus) with a 25x water objective (XLPLN, WMP2, 1.05 NA, Olympus) was applied for near-IR throughput. Synchronized pulsed pump beam (tunable 720–990 nm wavelength, 5–6 ps pulse width, and 80 MHz repetition rate) and Stokes (wavelength at 1032nm, 6 ps pulse width, and 80MHz repetition rate) were supplied by a picoEmerald system (Applied Physics & Electronics) and coupled into the microscope. The pump and Stokes beams were collected in transmission by a high NA oil condenser (1.4 NA). A high O.D. shortpass filter (950nm, Thorlabs) was used that would completely block the Stokes beam and transmit the pump beam only onto a Si photodiode for detecting the stimulated Raman loss signal. The output current from the photodiode was terminated, filtered, and demodulated in X with a zero phaseshift by a lock-in amplifier (HF2LI, Zurich Instruments) at 20MHz. The demodulated signal was fed into the FV3000 software module FV-OSR (Olympus) to form the image during laser scanning. All 3D lipid droplet images were obtained with a pixel dwell time 40 μs with 3-frame averaging for a total imaging speed of ~10-15 min per image stack. Laser power incident on the sample is approximately 40mW.

Two Photon Fluorescence Microscopy

Autofluorescence of flavins was excited at 820 nm and autofluorescence of NADH was excited at 780nm using the same tunable picosecond laser described in section 2.3. Epi-detected emission of flavin autofluorescence was collected using a 460 nm filter cube (OCT-ET460/50M32, Olympus), and NADH was collected using a 515nm filter. These images were also 512x512 pixels and were acquired with a 12.5 μs/pixel dwell time using a 300mW power at the laser shutter. Autofluorescence images were background subtracted using a rolling ball algorithm with a radius of 50px, which is intended to approximate cell size in these images.

Data Analysis

Spectral Clustering

Previous studies have shown these breast cancer subtypes have unique Raman features (16, 69). Machine learning was conducted to determine the extent to which these features can be used to segregate these subtypes and be augmented using the metabolic dimension of excess methionine. Neural network classification was done using a simple multi-layer perceptron (MLP) model with 100 neurons in the hidden layer and a rectified linear unit (ReLU) activation function for each neuron. An L2 regularization term with hyperparameter α=0.0002 penalizes the model for incorrect classification during learning with cross-entropy loss minimization. The classification is stochastically optimized using an adaptive moment estimation algorithm called Adam. Advantages of this choice of activation function and solver in an MLP include invariance to rescaling gradients, the ability to learn non-linear models, and a natural simulated annealing to optimize the gradient (70). MLP are, however, sensitive to parameter tuning, and all spectra were normalized to have the same range.

The input for the MLP model consists of a matrix of Raman spectra, and a vector of target classes. In this study, target classes are of categorical type and correspond to the cell subtype and metabolite concentration groups. Each spectrum X_i is represented as a vector containing m wavenumbers which are each input into a first layer of neurons. Each neuron in the hidden layer accepts the weighted linear combination of input features and applies the ReLU activation function, outputting the data to the output layer. Softmax is used to probabilistically

determine the target class of the multiclass model. The model is trained *via* backpropagation to minimize cross-entropy loss with a maximum of 200 iterations in which subsequent weight vectors reflect a subtracted loss gradient according to equation 1 below.

$$W^{i+1} = W^{i} - \in \nabla Loss_{W}^{i} \quad (1)$$

where, ϵ corresponds to the learning rate. A python implementation of model is readily available from scikit-learn v0.24.1 (71). The width of the hidden layer, k=100, as well as its depth of 1 single layer, are tunable depending on the dataset. Larger datasets may require more neurons and deeper networks to perform better. The geometric mean of input variables and classes roughly totaled 100, and an underlying assumption of a simple binary effect of excess methionine and very distinct Raman spectra was comfortable with only a single layer. However, multiplexing of variables such as cell subtype, methionine concentration, and other manipulations may intuitively justify additional hidden layers in future experiments. Classification in this study mainly attempts to highlight the higher dimensionality of methionine's non-linear effects on breast cancer subtypes, and discuss potential pathway interactions for further investigation. That is, if all breast cancer subtypes responded similarly and to a similar degree, more complex neural nets may not be necessary to achieve good performance.

Principal component analysis (PCA) is performed using Orange 3.26 on the pre-processed data. The first 10 PCs are used as the vectors for tSNE visualization without any exaggeration and a perplexity of 30. A graphical method outline can be found in **Figure S2**.

Selected Raman Feature Analysis

Spectroscopic data is extracted using MATLAB and is plotted using either MATLAB or Prism 7. To visualize the influence of all Raman peaks on classification simultaneously, relative entropy is employed. For binary classification systems, the amount of data lost in classifying data B as data A is described by the one-dimensional cross entropy equation 2 below.

$$H(A,B) = -\Sigma_{i}^{n}(p_{A} \quad [\![(v]\!]_{i}) \log p_{B}(v_{\mathrm{i}})) \quad (2)$$

H(A,B) is the cross-entropy, $p_{\mathrm{A}}(v_{\mathrm{i}})$ and $p_{\mathrm{A}}(v_{\mathrm{i}})$ are probability vectors from the distributions of intensity values a wavenumber variable v_{i}. Probability vectors can be derived from various distributions, but only gaussian normal distributions were used in this study. The $\mathrm{D_{KL}}$ is related to cross-entropy as it is the additional entropy beyond the entropy of the data A. Since both distributions are already labeled and we are not interested in generating probability vectors, but rather supply them, the $\mathrm{D_{KL}}$ is described by Equation 3 below.

$$D_{KL}(\mathrm{A}||\mathrm{B}) = \sum_{i}^{n}(P_{A}(v_{i}) \log\left(\frac{p_{\mathrm{A}}(v_{\mathrm{i}})}{p_{\mathrm{B}}(v_{\mathrm{i}})}\right) \quad (3)$$

This divergence is made symmetric by equation 4 below. For this analysis probability vectors are of length 10^{7}

$$D_{KL} = \frac{D_{KL}(\mathrm{A}||\mathrm{B}) + D_{KL}(\mathrm{B}||\mathrm{A})}{2} \quad (4)$$

For multiclass situations in which the number of groups exceeds two, the average divergence is calculated following equation 5 (72, 73).

$$D(P_{1} \ldots P_{k}) = \frac{1}{k(k-1)} \sum_{i}^{k} \sum_{j}^{k} D_{KL}(P_{i}||P_{j}) \quad (5)$$

Relative Entropy is an excellent method for quantifying the relative importance of a wavenumber in the discrimination of spectral datasets because it is computationally fast and capitalizes on the tremendous chemometric potential afforded by the spectral resolution of modern spontaneous Raman systems. To date, this method has not yet been employed and reported in Raman spectroscopy applications. Second derivative barcode analysis may be coupled with this method, and improve it, as barcode analysis contains both amplitude and width information (74).

Image Analysis

Images were processed using MATLAB and ImageJ. 3D image stacks of lipid droplets underwent bandpass filtering to suppress horizontal noise artifacts from laser beam scanning, and smoothed. Lipid droplets received a sphericity score based on Euclidean distance from perfect spheres emanating from the center of mass of the lipid droplet to the surface of the lipid droplet. Those with low sphericity scores were discarded. Autofluorescence images underwent sliding paraboloid background subtraction before manual cell segmentation and measurement was conducted *via* ImageJ.

Statistical Analysis

All experiments were run in triplicate. LD spectra for display, ratiometric peak analysis, subtype clustering, and relative entropy comprise 5 LD spectra per cell, and 5 cells per experimental group per trial. SRS and TPF images used in multi-modal analysis consist of 3 or 4 ROI of approximately 5 cells per ROI in each experimental group per trial. SRS images used in 3D spatial analysis consist of 4 cells per experimental group per trial. 2-way balanced ANOVA results were consistent between trials, so to communicate the impact patterns of methionine and insulin, results were pooled such that each datum represents a measurement, not a trial mean.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

AF and LS conceived the idea and designed the study. AF conducted the experiments, analyzed the data with the input

from KH, DC, HZ, WZ and LS; AF and LS wrote and revised the manuscript with the input from all other authors. All authors contributed to the article and approved the submitted version.

FUNDING

This work is partially supported by UCSD Startup funds, U54 pilot grant 2U54CA132378-11A1, and Hellman Fellow Award.

ACKNOWLEDGMENTS

We thank Drs. Yajuan Li, Christian Metallo for helpful discussion. We thank Dr. Stephanie Fraley and Dr. Kun Zhang Lab for providing cell culture related reagents.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.858017/full#supplementary-material

REFERENCES

1. DeSantis CE, Ma J, Gaudet MM, Newman LA, Miller KD, Sauer AG, et al. Breast Cancer Statistics, 2019. *CA: A Cancer J Clin* (2019) 69:438–51. doi: 10.3322/caac.21583
2. Dai X, Li T, Bai Z, Yang Y, Liu X, Zhan J, et al. Breast Cancer Intrinsic Subtype Classification, Clinical Use and Future Trends. *Am J Cancer Res* (2015) 5:2929–43.
3. Olzmann JA, Carvalho P. Dynamics and Functions of Lipid Droplets. *Nat Rev Mol Cell Biol* (2019) 20:137–55. doi: 10.1038/s41580-018-0085-z
4. Cruz ALS, de Barreto QAE, Fazolini NPB, Viola JPB, Bozza PT. Lipid Droplets: Platforms With Multiple Functions in Cancer Hallmarks. *Cell Death Dis* (2020) 11:1–16. doi: 10.1038/s41419-020-2297-3
5. Daniëls VW, Smans K, Royaux I, Chypre M, Swinnen JV, Zaidi N. Cancer Cells Differentially Activate and Thrive on *De Novo* Lipid Synthesis Pathways in a Low-Lipid Environment. *PloS One* (2014) 9:e106913. doi: 10.1371/journal.pone.0106913
6. Franco D, Trusso S, Fazio E, Allegra A, Musolino C, Speciale A, et al. Raman Spectroscopy Differentiates Between Sensitive and Resistant Multiple Myeloma Cell Lines. *Spectrochim Acta Part A: Mol Biomol Spectros* (2017) 187:15–22. doi: 10.1016/j.saa.2017.06.020
7. Butler LM, Mah CY, Machiels J, Vincent AD, Irani S, Mutuku S, et al. Lipidomic Profiling of Clinical Prostate Cancer Reveals Targetable Alterations in Membrane Lipid Composition. *bioRxiv* (2020). doi: 10.1101/2020.10.27.356634
8. Dawaliby R, Trubbia C, Delporte C, Noyon C, Ruysschaert J-M, Van Antwerpen P, et al. Phosphatidylethanolamine Is a Key Regulator of Membrane Fluidity in Eukaryotic Cells. *J Biol Chem* (2016) 291:3658–67. doi: 10.1074/jbc.M115.706523
9. Bompard J, Rosso A, Brizuela L, Mebarek S, Blum LJ, Trunfio-Sfarghiu A-M, et al. Membrane Fluidity as a New Means to Selectively Target Cancer Cells With Fusogenic Lipid Carriers. *Langmuir* (2020) 36:5134–44. doi: 10.1021/acs.langmuir.0c00262
10. Rysman E, Brusselmans K, Scheys K, Timmermans L, Derua R, Munck S, et al. *De Novo* Lipogenesis Protects Cancer Cells From Free Radicals and Chemotherapeutics by Promoting Membrane Lipid Saturation. *Cancer Res* (2010) 70:8117–26. doi: 10.1158/0008-5472.CAN-09-3871
11. Schug ZT, Peck B, Jones DT, Zhang Q, Grosskurth S, Alam IS, et al. Acetyl-CoA Synthetase 2 Promotes Acetate Utilization and Maintains Cancer Cell Growth Under Metabolic Stress. *Cancer Cell* (2015) 27:57–71. doi: 10.1016/j.ccell.2014.12.002
12. Lisec J, Jaeger C, Zaidi N. Cancer Cell Lipid Class Homeostasis Is Altered Under Nutrient-Deprivation But Stable Under Hypoxia. *bioRxiv* (2018) 382457. doi: 10.1101/382457
13. Jarc E, Petan T. Lipid Droplets and the Management of Cellular Stress. *Yale J Biol Med* (2019) 92:435–52.
14. Wolins NE, Quaynor BK, Skinner JR, Schoenfish MJ, Tzekov A, Bickel PE. S3-12, Adipophilin, and TIP47 Package Lipid in Adipocytes. *J Biol Chem* (2005) 280:19146–55. doi: 10.1074/jbc.M500978200
15. Schott MB, Weller SG, Schulze RJ, Krueger EW, Drizyte-Miller K, Casey CA, et al. Lipid Droplet Size Directs Lipolysis and Lipophagy Catabolism in Hepatocytes. *J Cell Biol* (2019) 218:3320–35. doi: 10.1083/jcb.201803153
16. Abramczyk H, Surmacki J, Kopeć M, Olejnik AK, Lubecka-Pietruszewska K, Fabianowska-Majewska K. The Role of Lipid Droplets and Adipocytes in Cancer. Raman Imaging of Cell Cultures: MCF10A, MCF7, and MDA-MB-231 Compared to Adipocytes in Cancerous Human Breast Tissue. *Analyst* (2015) 140:2224–35. doi: 10.1039/C4AN01875C
17. Li X, Li Y, Jiang M, Wu W, He S, Chen C, et al. Quantitative Imaging of Lipid Synthesis and Lipolysis Dynamics in *Caenorhabditis Elegans* by Stimulated Raman Scattering Microscopy. *Anal Chem* (2019) 91:2279–87. doi: 10.1021/acs.analchem.8b04875
18. Paar M, Jüngst C, Steiner NA, Magnes C, Sinner F, Kolb D, et al. Remodeling of Lipid Droplets During Lipolysis and Growth in Adipocytes. *J Biol Chem* (2012) 287:11164–73. doi: 10.1074/jbc.M111.316794
19. Sun X, Wang M, Wang M, Yu X, Guo J, Sun T, et al. Metabolic Reprogramming in Triple-Negative Breast Cancer. *Front Oncol* (2020) 10:428. doi: 10.3389/fonc.2020.00428
20. Petővári G, Dankó T, Tőkés A-M, Vetlényi E, Krencz I, Raffay R, et al. *In Situ* Metabolic Characterisation of Breast Cancer and Its Potential Impact on Therapy. *Cancers* (2020) 12:2492. doi: 10.3390/cancers12092492
21. Lampa M, Arlt H, He T, Ospina B, Reeves J, Zhang B, et al. Glutaminase Is Essential for the Growth of Triple-Negative Breast Cancer Cells With a Deregulated Glutamine Metabolism Pathway and its Suppression Synergizes With mTOR Inhibition. *PloS One* (2017) 12:e0185092. doi: 10.1371/journal.pone.0185092
22. Jung SM, Hung C-M, Hildebrand SR, Sanchez-Gurmaches J, Martinez-Pastor B, Gengatharan JM, et al. Non-Canonical Mtorc2 Signaling Regulates Brown Adipocyte Lipid Catabolism Through SIRT6-Foxo1. *Mol Cell* (2019) 75:807–822.e8. doi: 10.1016/j.molcel.2019.07.023
23. Yee LD, Mortimer JE, Natarajan R, Dietze EC, Seewaldt VL, Health M. Insulin, and Breast Cancer: Why Oncologists Should Care About Insulin. *Front Endocrinol (Lausanne)* (2020) 11:58. doi: 10.3389/fendo.2020.00058
24. Sanderson SM, Gao X, Dai Z, Locasale JW. Methionine Metabolism in Health and Cancer: A Nexus of Diet and Precision Medicine. *Nat Rev Cancer* (2019) 19:625–37. doi: 10.1038/s41568-019-0187-8
25. Cai H, Dong L, Liu F. Recent Advances in Adipose mTOR Signaling and Function: Therapeutic Prospects. *Trends Pharmacol Sci* (2016) 37:303–17. doi: 10.1016/j.tips.2015.11.011
26. Yoon M-S. The Role of Mammalian Target of Rapamycin (mTOR) in Insulin Signaling. *Nutrients* (2017) 9(11):1176. doi: 10.3390/nu9111176
27. Hay N. Interplay Between FOXO, TOR, and Akt. *Biochim Biophys Acta (BBA) - Mol Cell Res* (2011) 1813:1965–70. doi: 10.1016/j.bbamcr.2011.03.013
28. Shi X, Wang J, Lei Y, Cong C, Tan D, Zhou X. Research Progress on the PI3K/AKT Signaling Pathway in Gynecological Cancer. *Mol Med Rep* (2019) 19:4529–35. doi: 10.3892/mmr.2019.10121
29. Kitada M, Xu J, Ogura Y, Monno I, Koya D. Mechanism of Activation of Mechanistic Target of Rapamycin Complex 1 by Methionine. *Front Cell Dev Biol* (2020) 8:715. doi: 10.3389/fcell.2020.00715
30. Zhou Y, Zhou Z, Peng J, Loor JJ. Methionine and Valine Activate the Mammalian Target of Rapamycin Complex 1 Pathway Through Heterodimeric Amino Acid Taste Receptor (TAS1R1/TAS1R3) and Intracellular Ca2+ in Bovine Mammary Epithelial Cells. *J Dairy Sci* (2018) 101:11354–63. doi: 10.3168/jds.2018-14461
31. Costantino A, Milazzo G, Giorgino F, Russo P, Goldfine ID, Vigneri R, et al. Insulin-Resistant MDA-MB231 Human Breast Cancer Cells Contain a

Tyrosine Kinase Inhibiting Activity. *Mol Endocrinol* (1993) 7:1667–76. doi: 10.1210/mend.7.12.8145772
32. Gupta C, Tikoo K. High Glucose and Insulin Differentially Modulates Proliferation in MCF-7 and MDA-MB-231 Cells. *J Mol Endocrinol* (2013) 51:119–29. doi: 10.1530/JME-13-0062
33. Wanders D, Hobson K, Ji X. Methionine Restriction and Cancer Biology. *Nutrients* (2020) 12(3):684. doi: 10.3390/nu12030684
34. Borrego SL, Fahrmann J, Datta R, Stringari C, Grapov D, Zeller M, et al. Metabolic Changes Associated With Methionine Stress Sensitivity in MDA-MB-468 Breast Cancer Cells. *Cancer Metab* (2016) 4:9. doi: 10.1186/s40170-016-0148-6
35. Jeon H, Kim JH, Lee E, Jang YJ, Son JE, Kwon JY, et al. Methionine Deprivation Suppresses Triple-Negative Breast Cancer Metastasis *In Vitro* and *In Vivo*. *Oncotarget* (2016) 7:67223–34. doi: 10.18632/oncotarget.11615
36. Morén B, Fryklund C, Stenkula K. Surface-Associated Lipid Droplets: An Intermediate Site for Lipid Transport in Human Adipocytes? *Adipocyte* (2020) 9:636–48. doi: 10.1080/21623945.2020.1838684
37. Covington JD, Johannsen DL, Coen PM, Burk DH, Obanda DN, Ebenezer PJ, et al. Intramyocellular Lipid Droplet Size Rather Than Total Lipid Content Is Related to Insulin Sensitivity After 8 Weeks of Overfeeding. *Obes (Silver Spring)* (2017) 25:2079–87. doi: 10.1002/oby.21980
38. DeBose-Boyd RA, Ye J. SREBPs in Lipid Metabolism, Insulin Signaling, and Beyond. *Trends Biochem Sci* (2018) 43:358–68. doi: 10.1016/j.tibs.2018.01.005
39. Borrego SL, Fahrmann J, Hou J, Lin D-W, Tromberg BJ, Fiehn O, et al. Lipid Remodeling in Response to Methionine Stress in MDA-MBA-468 Triple-Negative Breast Cancer Cells. *J Lipid Res* (2021) 62:100056. doi: 10.1016/j.jlr.2021.100056
40. Murata Y, Watanabe T, Sato M, Momose Y, Nakahara T, Oka S, et al. Dimethyl Sulfoxide Exposure Facilitates Phospholipid Biosynthesis and Cellular Membrane Proliferation in Yeast Cells. *J Biol Chem* (2003) 278:33185–93. doi: 10.1074/jbc.M300450200
41. Zou K, Rouskin S, Dervishi K, McCormick MA, Sasikumar A, Deng C, et al. Life Span Extension by Glucose Restriction Is Abrogated by Methionine Supplementation: Cross-Talk Between Glucose and Methionine and Implication of Methionine as a Key Regulator of Life Span. *Sci Adv* (2020) 6:eaba1306. doi: 10.1126/sciadv.aba1306
42. Weber FL, Veach GL, Friedman DW. Effects of Insulin and Glucagon on the Uptake of Amino Acids From Arterial Blood by Canine Ileum. *Digest Dis Sci* (1981) 26:113–8. doi: 10.1007/BF01312226
43. Hou J, Williams J, Botvinick EL, Potma EO, Tromberg BJ. Visualization of Breast Cancer Metabolism Using Multimodal Nonlinear Optical Microscopy of Cellular Lipids and Redox State. *Cancer Res* (2018) 78:2503–12. doi: 10.1158/0008-5472.CAN-17-2618
44. Souba WW. Glutamine and Cancer. *Ann Surg* (1993) 218:715–28. doi: 10.1097/00000658-199312000-00004
45. Wise DR, Thompson CB. Glutamine Addiction: A New Therapeutic Target in Cancer. *Trends Biochem Sci* (2010) 35:427–33. doi: 10.1016/j.tibs.2010.05.003
46. Charidemou E, Ashmore T, Li X, McNally BD, West JA, Liggi S, et al. A Randomized 3-Way Crossover Study Indicates That High-Protein Feeding Induces *De Novo* Lipogenesis in Healthy Humans. *JCI Insight* (2019) 4: e124819. doi: 10.1172/jci.insight.124819
47. Muthusamy T, Cordes T, Handzlik MK, You L, Lim EW, Gengatharan J, et al. Serine Restriction Alters Sphingolipid Diversity to Constrain Tumour Growth. *Nature* (2020) 586:790–5. doi: 10.1038/s41586-020-2609-x
48. Shi L, Zheng C, Shen Y, Chen Z, Silveira ES, Zhang L, et al. Optical Imaging of Metabolic Dynamics in Animals. *Nat Commun* (2018) 9:2995. doi: 10.1038/s41467-018-05401-3
49. *MDA-MB-231 (ATCC® HTB-26tm)* . Available at: https://www.atcc.org/products/all/htb-26.aspx#culturemethod.
50. Czamara K, Majzner K, Pacia MZ, Kochan K, Kaczor A, Baranska M. Raman Spectroscopy of Lipids: A Review. *J Raman Spectros* (2015) 46:4–20. doi: 10.1002/jrs.4607
51. Jamieson LE, Li A, Faulds K, Graham D. Ratiometric Analysis Using Raman Spectroscopy as a Powerful Predictor of Structural Properties of Fatty Acids. *R Soc Open Sci* (2018) 5:181483. doi: 10.1098/rsos.181483
52. Deevska GM, Nikolova-Karakashian MN. The Expanding Role of Sphingolipids in Lipid Droplet Biogenesis. *Biochim Biophys Acta (BBA) - Mol Cell Biol Lipids* (2017) 1862:1155–65. doi: 10.1016/j.bbalip.2017.07.008
53. Kullback S, Leibler RA. On Information and Sufficiency. *Ann Math Statistics* (1951) 22:79–86. doi: 10.1214/aoms/1177729694
54. Henne WM, Reese ML, Goodman JM. The Assembly of Lipid Droplets and Their Roles in Challenged Cells. *EMBO J* (2018) 37(12):e98947. doi: 10.15252/embj.201898947
55. Benador IY, Veliova M, Liesa M, Shirihai OS. Mitochondria Bound to Lipid Droplets: Where Mitochondrial Dynamics Regulate Lipid Storage and Utilization. *Cell Metab* (2019) 29:827–35. doi: 10.1016/j.cmet.2019.02.011
56. Cui L, Liu P. Two Types of Contact Between Lipid Droplets and Mitochondria. *Front Cell Dev Biol* (2020) 8:1589. doi: 10.3389/fcell.2020.618322
57. Kitt JP, Bryce DA, Minteer SD, Harris JM. Raman Spectroscopy Reveals Selective Interactions of Cytochrome C With Cardiolipin That Correlate With Membrane Permeability. *J Am Chem Soc* (2017) 139:3851–60. doi: 10.1021/jacs.7b00238
58. Sato ET, Martinho H. First-Principles Calculations of Raman Vibrational Modes in the Fingerprint Region for Connective Tissue. *Biomed Opt. Express* (2018) 9:1728. doi: 10.1364/BOE.9.001728
59. Farber C, Li J, Hager E, Chemelewski R, Mullet J, Yu. Rogachev A, et al. Complementarity of Raman and Infrared Spectroscopy for Structural Characterization of Plant Epicuticular Waxes. *ACS Omega* (2019) 4:3700–7. doi: 10.1021/acsomega.8b03675
60. Khalid M, Bora T, Ghaithi AA, Thukral S, Dutta J. Raman Spectroscopy Detects Changes in Bone Mineral Quality and Collagen Cross-Linkage in Staphylococcus Infected Human Bone. *Sci Rep* (2018) 8:9417. doi: 10.1038/s41598-018-27752-z
61. Da Silva E, Bresson S, Rousseau D. Characterization of the Three Major Polymorphic Forms and Liquid State of Tristearin by Raman Spectroscopy. *Chem Phys Lipids* (2009) 157:113–9. doi: 10.1016/j.chemphyslip.2008.11.002
62. Podsednik A, Jacob A, Li LZ, Xu HN. Relationship Between Optical Redox Status and Reactive Oxygen Species in Cancer Cells. *React Oxyg Species (Apex)* (2020) 9:95–108. doi: 10.20455/ros.2020.815
63. Ostrander JH, McMahon CM, Lem S, Millon SR, Brown JQ, Seewaldt VL, et al. Optical Redox Ratio Differentiates Breast Cancer Cell Lines Based on Estrogen Receptor Status. *Cancer Res* (2010) 70:4759–66. doi: 10.1158/0008-5472.CAN-09-2572
64. Yun Y-H, Bin J, Liu D-L, Xu L, Yan T-L, Cao D-S, et al. A Hybrid Variable Selection Strategy Based on Continuous Shrinkage of Variable Space in Multivariate Calibration. *Anal Chimica Acta* (2019) 1058:58–69. doi: 10.1016/j.aca.2019.01.022
65. Hotamisligil GS. Inflammation and Metabolic Disorders. *Nature* (2006) 444:860–7. doi: 10.1038/nature05485
66. Adams WR, Mehl B, Leiser E, Wang M, Patton S, Throckmorton GA, et al. "Multimodal Nonlinear Optical and Thermal Imaging Platform for Label-Free Characterization of Biological Tissue". *Biophysics* (2020). doi: 10.1101/2020.04.06.023820
67. Chen G, Deng X. Cell Synchronization by Double Thymidine Block. *Bio Protoc* (2018) 8:e2994. doi: 10.21769/BioProtoc.2994
68. O'Malley J, Kumar R, Kuzmin A, Pliss A, Yadav N, Balachandar S, et al. Lipid Quantification by Raman Microspectroscopy as a Potential Biomarker in Prostate Cancer. *Cancer Lett* (2017) 397:52–60. doi: 10.1016/j.canlet.2017.03.025
69. Zhang L, Li C, Peng D, Yi X, He S, Liu F, et al. Raman Spectroscopy and Machine Learning for the Classification of Breast Cancers. *Spectrochim Acta Part A: Mol Biomol Spectros* (2022) 264:120300. doi: 10.1016/j.saa.2021.120300
70. Kingma DP, Ba J. Adam: A Method for Stochastic Optimization. In: *Arxiv:1412.6980 [Cs]* (2017). Available at: http://arxiv.org/abs/1412.6980.
71. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-Learn: Machine Learning in Python. *J Mach Learn Res* (2011) 12:2825–30.
72. Sgarro A. Informational Divergence and the Dissimilarity of Probability Distributions. *Calcolo* (1981) 18:293–302. doi: 10.1007/BF02576360
73. Garcıa-Garcıa D, Williamson RC. *Divergences and Risks for Multiclass Experiments* JMLR: Workshop and Conference Proceedings 2012. 20:1–20.
74. Velioglu SD, Ercioglu E, Temiz HT, Velioglu HM, Topcu A, Boyaci IH. Raman Spectroscopic Barcode Use for Differentiation of Vegetable Oils and

Determination of Their Major Fatty Acid Composition. *J Am Oil Chem Soc* (2016) 93:627–35. doi: 10.1007/s11746-016-2808-7

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

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Oncology

MINI REVIEW
published: 11 March 2022
doi: 10.3389/fonc.2022.836058

Acetyl-CoA Carboxylases and Diseases

Yu Wang[1], *Weixing Yu*[1], *Sha Li*[1], *Dingyuan Guo*[1], *Jie He*[1,2] *and Yugang Wang*[1,3*]

[1] Department of Biochemistry and Molecular Biology, School of Basic Medicine, Tongji Medical College, Huazhong University of Science of Technology, Wuhan, China, [2] Department of Neurosurgery, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, [3] Cell Architecture Research Center, Huazhong University of Science and Technology, Wuhan, China

Acetyl-CoA carboxylases (ACCs) are enzymes that catalyze the carboxylation of acetyl-CoA to produce malonyl-CoA. In mammals, ACC1 and ACC2 are two members of ACCs. ACC1 localizes in the cytosol and acts as the first and rate-limiting enzyme in the *de novo* fatty acid synthesis pathway. ACC2 localizes on the outer membrane of mitochondria and produces malonyl-CoA to regulate the activity of carnitine palmitoyltransferase 1 (CPT1) that involves in the β-oxidation of fatty acid. Fatty acid synthesis is central in a myriad of physiological and pathological conditions. ACC1 is the major member of ACCs in mammalian, mountains of documents record the roles of ACC1 in various diseases, such as cancer, diabetes, obesity. Besides, acetyl-CoA and malonyl-CoA are cofactors in protein acetylation and malonylation, respectively, so that the manipulation of acetyl-CoA and malonyl-CoA by ACC1 can also markedly influence the profile of protein post-translational modifications, resulting in alternated biological processes in mammalian cells. In the review, we summarize our understandings of ACCs, including their structural features, regulatory mechanisms, and roles in diseases. ACC1 has emerged as a promising target for diseases treatment, so that the specific inhibitors of ACC1 for diseases treatment are also discussed.

Keywords: acetyl-CoA carboxylase, lipogenesis, cancer metabolism, tumorigenesis, metabolic diseases

OPEN ACCESS

Edited by:
Xuefei Li,
Institute of Synthetic Biology, Shenzhen Institutes of Advanced Technology (CAS), China

Reviewed by:
Kumar Pichumani,
Houston Methodist Research Institute, United States
De Huang,
Duke University, United States

***Correspondence:**
Yugang Wang
yugangw@hust.edu.cn

Specialty section:
This article was submitted to Cancer Metabolism, a section of the journal Frontiers in Oncology

Received: *15 December 2021*
Accepted: *10 February 2022*
Published: *11 March 2022*

Citation:
Wang Y, Yu W, Li S, Guo D, He J and Wang Y (2022) Acetyl-CoA Carboxylases and Diseases. Front. Oncol. 12:836058. doi: 10.3389/fonc.2022.836058

INTRODUCTION

In mammalian cells, acetyl-CoA is a global currency that can mediate the carbon transactions between metabolic pathways, including glycolysis, tricarboxylic acid cycle, amino acid metabolism, gluconeogenesis, and fatty acid synthesis. Lipid metabolism or fatty acid metabolism is the bank of acetyl-CoA. It can deposit extra acetyl-CoA in the form of fatty acids and regulate the intracellular availability of acetyl-CoA to the global metabolism network by controlling the conversion of acetyl-CoA into fatty acids. As such, fatty acid synthesis is a central pathway in harnessing a myriad of metabolic pathways and related physiologies in cells.

Acetyl-CoA carboxylases (ACCs) are enzymes that catalyze the carboxylation of acetyl-CoA to produce malonyl-CoA, which in turn is utilized by the fatty acid synthase (FASN) to produce long-chain saturated fatty acids (1). There are two members of ACCs in mammalian cells. ACC1 localizes in the cytosol and takes the major responsibility of converting cytoplasmic acetyl-CoA into

malonyl-CoA for fatty acid synthesis (2). Despite ACC2 also catalyzing the conversion of acetyl-CoA into malonyl-CoA, it localizes on the outer membrane of mitochondria that makes the downstream pathways of ACC2-produced malonyl-CoA different from ACC1. It is reported that the ACC2-produced malonyl-CoA can allosterically influence the activity of carnitine palmitoyltransferase 1 (CPT1) in the β-oxidation of fatty acid (3). More functional studies about ACC2 are expected in this field.

Fatty acid synthesis controls the storage and expenditure of carbon source and energy, which can regulate the activities of other metabolic pathways, such as amino acid metabolism and glucose metabolism, so that fatty acid synthesis is frequently alternated to harness the intracellular metabolism network to meet the requirement of materials and energy for diseases progressions, such as cancer and metabolic diseases (4–8). ACC1 is the first rate-limiting enzyme in the fatty acid synthesis that plays a central role in fatty acid synthesis, so that ACC1 is the hub of the fatty acid synthesis-related metabolism network. Its dysregulation in diseases is intensively studied, including the roles of ACC1 in regulating tumour cell proliferation, migration, and metabolic disease progression (9–12). In addition, because acetyl-CoA and malonyl-CoA are cofactors in protein acetylation and malonylation, respectively, the emerging non-metabolic functions of ACC1 in diseases are discussed in recent studies (11, 13, 14), which further expand the roles of ACC1 in physiologies and pathophysiologies. ACC1 is therefore considered as a promising therapeutic target for treating diseases, such as cancer and metabolic diseases.

This review summarizes our current knowledge about ACCs, including the structure of ACCs, the regulatory mechanisms, and the roles of ACCs in tumorigenesis and metabolic diseases. Besides, we briefly introduce ACCs inhibitors that are under investigation for cancer and metabolic diseases therapy.

STRUCTURE OF ACETYL-COA CARBOXYLASES

In mammals, ACCs have two isoforms: ACC1 and ACC2. Human ACC1 (ACCα, 265 kDa) is encoded by the *ACACA* gene on chromosome 17q12 while ACC2 (ACCβ, 275 kDa) is encoded by the *ACACB* gene on chromosome 12q23 (15). ACC1 and ACC2 share 75% identity in amino acid sequence and are composed of conservative domains for enzyme activity (16, 17).

ACC1 and ACC2 have similar structures and molecule bases in catalyzing carboxylation of acetyl-CoA to produce malonyl-CoA. ACC1 is discussed here in terms of ACCs' structure. ACC1 contains three major functional domains: a biotin carboxylase domain (BC domain), a carboxyl transferase domain (CT domain), and a biotin carboxyl carrier protein domain (BCCP domain) that links the BC domain and CT domain (**Figure 1**). To perform the catalytic activity, the BC domain of ACC1 firstly consumes ATP and bicarbonate to catalyze the carboxylation of biotin, in which bicarbonate serves as the donor of the carboxyl moiety (18). Then, the BCCP domain transfers the carboxyl moiety from the carboxylated biotin to the CT domain (19), where the carboxyl moiety is transferred to the acetyl-CoA to accomplish the carboxylation of acetyl-CoA, converting acetyl-CoA into malonyl-CoA (20) (**Figure 1**). Although the BC domain and CT domain are linked by the BCCP domain in a single ACC1 molecule, the spatial dimension of ACC1 is so broad that the functional domains are spatially separated, which makes the carboxylated biotin in the BC domain can hardly reach to the acetyl-CoA that bond by the CT domain of the same molecule of ACC1. To link the cascade reactions of acetyl-CoA carboxylation, ACC1 molecules form homodimers to enable the carboxylated biotin in the BC domain reaching to the acetyl-CoA in the CT domain of the other ACC1 molecule of

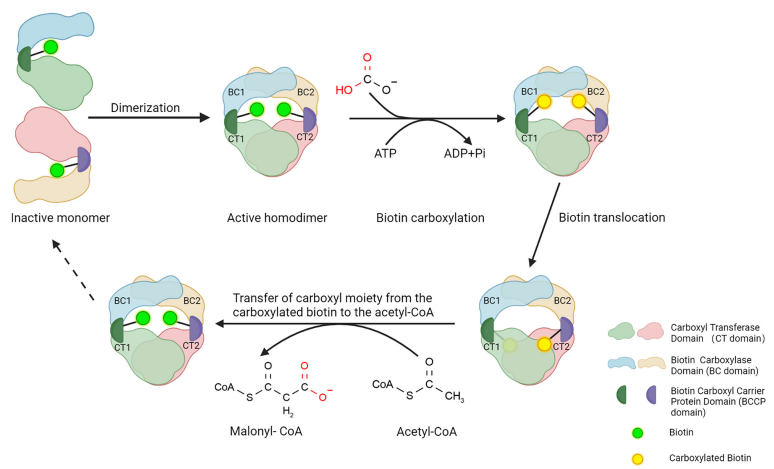


FIGURE 1 | Structure of ACC1 and function of three main domains. Three steps to a functional ACC1: First, the BC domain consumes ATP and catalyzes the carboxylation of biotin, in which bicarbonate serves as the donor of the carboxyl moiety. Subsequently, the BCCP domain transfers the bicarbonate moiety from carboxylated biotin to the CT domain of ACC1. Lastly, the CT domain catalyzes the carboxylation of acetyl-CoA carboxyl moiety, converting acetyl-CoA into malonyl-CoA. Created with BioRender.com.

the homodimer (19–23). Therefore, regulating the formation of ACC1 homodimers is considered as an important mechanism controlling the acetyl-CoA carboxylation activity of ACC1.

DISTRIBUTION AND FUNCTIONS OF ACETYL-COA CARBOXYLASES

ACC1 and ACC2 are widely distributed in organs and tissues in mammals. ACC1 is highly enriched in lipogenic tissues, such as liver and adipose tissue, while ACC2 is majorly expressed in oxidative tissues, such as cardiac and skeletal muscle (24), which are consistent with the functions of ACC1 in lipogenesis and ACC2 in regulating fatty acid β-oxidation. In mammalian cells, ACC1 and ACC2 are differently distributed (**Figure 2**). ACC1 is a cytoplasmic protein that catalyzes the conversion of acetyl-CoA into malonyl-CoA, which is utilized by the fatty acid synthetase (FASN) and subjected to the *de novo* fatty acid biosynthesis (2). It controls the synthesis of mid-and long-chain fatty acids that serve as building blocks for the cell biology process. Inhibiting ACC1 by 5-tetradecyloxy-2-furoic acid (TOFA) can completely inhibit hepatic *de novo* lipogenesis (DNL), which is considered a new strategy for non-alcoholic fatty liver disease (NAFLD) treatment (25). Soraphen A, another ACC1 inhibitor, can pharmacologically inhibit fatty acid synthesis in diet-induced obesity mice and significantly suppress weight gain, which sheds new light on controlling diet-induced obesity (26). Liver-specific ACC1 knockout (LACC1 KO) mice can survive but show dysregulated lipid metabolism and deficiency in triglycerides metabolism (27). In cancer cells, inhibition of ACC1 by Soraphen A significantly reduces saturated and mono-unsaturated phospholipid species and increases the proportion of poly-unsaturation, rendering cells vulnerable to oxidative stresses (28). Moreover, activity-impeded ACC1 reduces cytoplasmic membrane fluidity and impairs mobilities of transmembrane receptors, ultimately impairing cell membrane-dependent biological processes (29).

ACC2 contains a hydrophobic N-terminal region that leads ACC2 attaching to the outer membrane of mitochondria (25). The mitochondria-located ACC2 also catalyzes the conversion of acetyl-CoA into malonyl-CoA. However, instead of entering the *de novo* fatty acid biosynthesis, the ACC2-generated malonyl-CoA locally interacts with carnitine palmitoyltransferase 1 (CPT1) that also localizes on the outer mitochondrial membrane. CPT1 accounts for the first step of long-chain fatty acids β-oxidation in mitochondria. The binding of malonyl-CoA allosterically inhibits the activity of CPT1 and therefore influences the process of fatty acid β-oxidation in mitochondria (30). In animal experiments, inhibition of ACC2 can increase hepatic fat oxidation, reduce hepatic lipids, and improve hepatic insulin sensitivity in mice with NAFLD (31), which is further confirmed in mice with genetic depletion of ACC2 (3, 32, 33). In addition, the fatty acid oxidation rate in the soleus muscle of the ACC2-/- mice is 30% higher than that of wild-type mice and is not affected by the addition of insulin, leading to reduction of body weight under normal food intake and slower weight-gain with high-fat/high-carbohydrate diets (34).

In addition to the roles in metabolic flow, fatty acids, acetyl-CoA, and malonyl-CoA are effector molecules that participate in signaling pathways in cells (35–37). Correspondingly, ACCs, as the consumer of acetyl-CoA and the producer of malonyl-CoA that function as the rate-limiting enzyme in fatty acid synthesis, play intriguing roles in regulating cellular signaling networks. For example, polyunsaturated fatty acids (PUFAs) are the precursors of various signaling molecules, such as eicosanoids, that regulate the activity of sterol-regulatory-element-binding protein 1c (SREBP1c) in fatty acid metabolism in liver (38). Inhibition of

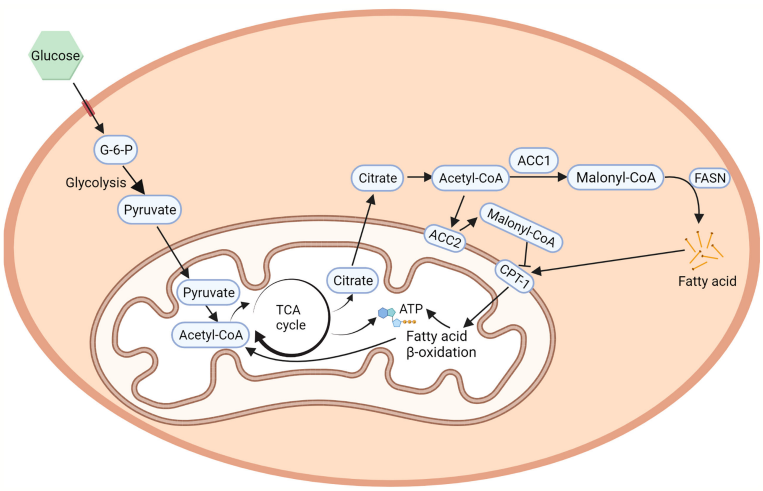


FIGURE 2 | ACCs in fatty acid metabolism. ACC1 is a cytoplasmic protein that catalyzes the conversing of acetyl-CoA to malonyl-CoA in the *de novo* fatty acid biosynthesis (2). On the other hand, the hydrophobicity of the N-terminal region of ACC2 allows its localization to the outer membrane of the mitochondria, regulates CPT1 which controls fatty acid β-oxidation (25). Created with BioRender.com.

ACCs is considered as a promising strategy for treating liver diseases (39). However, on the other hand, it leads to a decrease in malonyl-CoA and the synthesis of downstream PUFA, which in turn activates SREBP1c and upregulates the expression of glycerol-3-phosphate acyltransferase 1 (GPAT1) that catalyzes triglyceride synthesis, stimulating hepatic VLDL secretion and leading to hypertriglyceridemia (40). As such, hypertriglyceridemia is used to be accompanied with the ACCs-targeting therapies. Acetyl-CoA is another instance. It can regulate gene transcription by donating the acetyl-moiety in the acetyltransferases-mediated histone acetylation (41). Inhibition of ACCs can increase the intracellular acetyl-CoA level and stimulate the influx of calcium into the cells, which lead to the activation and translocation of NFAT (nuclear factor of activated T cells 1) into the nucleus to promote the transcription of adhesion and migration-related genes, promoting adhesion and migration of glioblastoma cells through Ca^{2+}– NFAT signaling. Malonyl-CoA plays roles in regulating dietary behavior and appetite (42). It is shown that mammalian neural tissue was able to rapidly convert administered acetate into acetyl-CoA, which subsequently entered the Krebs cycle to promote ATP production. Excessive ATP level, in turn, down-regulates AMPK activity and secures ACC2's enzymatic activity. As a result, malonyl-CoA was produced to a great extent, causing the downstream effector molecular pro-opiomelanocortin upregulation and neuropeptide Y downregulation, eventually leading to loss of appetite in mice (43, 44).

In addition to the metabolic functions, ACC1 and ACC2 regulate protein acetylation by manipulating the availability of acetyl-CoA in cells. In liver-specific ACC1 knockout mice, the amount of acetyl-CoA in the extra-mitochondrial space is substantially elevated, which can serve as the substrate cofactor for acetyltransferases and increase the acetylation of proteins to regulate the functionome, including metabolic enzymes that regulate the metabolism network in ACC1 KO mice (13). Another study also shows that attenuated expression of ACC1 leads to a substantial increase in histone acetylation and alters transcriptional regulation, resulting in increased histone acetylation that consequently regulates biological processes in cells *via* influencing gene transcription (14). While the causal relationship between ACCs' activities and protein acetylation is confirmed, the detailed mechanism underlying ACC1-related hyperacetylation remains elusive. Besides, ACC1 regulating protein acetylation by controlling the intracellular concentration of acetyl-CoA might also play a role in disease development. A supportive study reported that ACC inhibition regulates smad2 acetylation, which consequently affects the activity of smad2 and breast cancer metastasis (11). Malonyl-CoA is the product of ACCs' enzymatic reactions. It can also serve as the substrate cofactor in the enzyme-catalyzed protein malonylation. Increased intracellular malonyl-CoA can result in upregulation of protein malonylation, which might affect protein functions and biological activities in cells (9). Despite evidences are supporting the importance of the non-metabolic functions of ACC1 in regulating protein modifications and functions, it is premature to conclude the non-metabolic functions of ACC1 in diseases development and treatment.

Altogether, ACCs regulate the physiologies and pathophysiological processes of cells by executing metabolic and non-metabolic functions. It should sense the alternatives in cells and precisely translate the alternated signals into the responses of cells. As such, sophisticated regulation of ACCs is required to secure the metabolism network matching the physiologies of cells.

REGULATION OF ACETYL-COA CARBOXYLASES

The activities of ACCs in cells can be transcriptionally and post-transcriptionally regulated that are tightly associated with the metabolic status of cells. In general, the protein level and enzymatic activities of ACCs are upregulated in nutrient and energy abundant conditions, aiming to store the excess nutrient and energy in the form of fatty acids. Correspondingly, the protein level and enzymatic activities of ACCs are suppressed in nutrient and energy-deficient conditions, aiming to secure the limited energy and nutrients being utilized for survival (45, 46). The AMP-activated protein kinase (AMPK) is the most studied energy sensor that senses the nutrient and energy status of cells and is an important regulator of ACC1. When cells suffering metabolic stresses, such as glucose deprivation or hypoxia, AMPK is activated that can phosphorylate the serine-79 residues in ACC1 (equivalent to ACC2 Ser212) (47). Phosphorylation of the Ser-79 residue effectively blocks the formation of ACC1 homodimer, leaving ACC1 molecules as monomers that are unable to catalyze acetyl-CoA carboxylation (21). The fatty acid synthesis pathway is therefore suppressed. However, when cells return to a nutrient and energy-abundant environment, the phosphorylation of Ser-79 in ACC1 can be removed by the type 2A protein phosphatase (PP2A), allowing the reformation of ACC1 homodimer that is active in catalyzing acetyl-CoA carboxylation (48, 49). Besides nutrient and energy stresses, the Ser-79 residue in ACC1 can be phosphorylated and maintained to prevent lipogenesis in certain pathophysiological processes. For example, the susceptibility gene 1 (BRCA1) C-terminal (BRCT) domain binds to p-ACC1 to from BRCA1/p-ACC1 complex (50), which prevents dephosphorylation of p-ACC1 and constantly suppress the activity of ACC1 to reprogram the metabolism network in breast cancer. Insulin-like growth factor-1 (IGF-1) treatment can disrupt the interaction between BRCA1 and p-ACC1, which leads to the dephosphorylation and reactivation of ACC1 (51).

In addition to phosphorylation, metabolites that are associated with changes in metabolism can allosterically regulate the activities of ACCs. For example, citrate is an intermediate metabolite in the TCA cycle that can allosterically activate ACC1 to drive the fatty acid synthesis pathway in normal condition (52, 53). Intriguingly, opposite effects of citrate on ACCs' activities were reported (54), the underlying mechanism remains elusive. Glutamate can allosterically activate phosphatase that mediates dephosphorylation and activation of ACCs in cardiomyocytes, which may contribute to the cardioprotective effects of glutamine against lipolysis (55).. Fatty acyl-CoAs can induce the de-dimerization of ACC1 that inhibits the activities of

ACC1 and fatty acid synthesis in cells (54). By interplay with metabolites from different metabolic pathways, ACCs mediate the cross-talk between fatty acid synthesis and other metabolic pathways, forming a sophisticated regulation network to secure the metabolic status fit the physiologies of cells.

The protein levels of ACC1 and ACC2 are dynamically regulated in cells. The expression level of ACC1 can be regulated by certain transcription factors. SREBP1c is a well-studied instance. The *ACACA* gene has two distinct SREBP binding sites, which recruit SREBP1c to initiate RNA Polymerase II-dependent transcription. Carbohydrate response element (ChRE) -binding protein (ChREBP) is another transcription factor that regulates ACC1. It binds to the promoter regions and activates the transcription of *ACACA*, in response to the high-carbohydrate diet (56, 57). Besides transcription, the protein stability of ACC1 can also be regulated. In breast cancer, small interfering RNA-mediated Aldo-keto reductase family 1B10 (AKR1B10) knockdown induces ACC1 degradation *via* the ubiquitination-proteasome pathway, resulting in a markedly drop in fatty acid synthesis in RAO-3 cells (58). In prostate cancer, the expression of prolyl isomerase Pin1 positively correlates with the protein level of ACC1. It binds to ACC1 to prevent ACC1 from entering the lysosomal pathway, leading to the stabilization of ACC1 protein and resulting in enhanced activity of ACC1 in cells (59).

ACC1 activity can also be regulated by post-transcriptional and translational mechanisms. There are 64 exons included in the gene *ACACA* that result in 7 alternatively spliced minor exons (1A, 1B, 1C, 3, 5A', 5A, and 5B). The exon 5B can lead to transcriptional termination of the upstream exon 5 in two different transcripts, producing a short peptide that leads to the production of truncated ACC1 that affects the transcriptional efficiency and activity of ACC1. These studies suggest that ACC1 activity can be regulated by post-transcriptional and translational mechanisms and consequently result in suppression of fatty acid synthesis (60). The protein level of ACC1 can be post-transcriptionally reduced in calcium/calmodulin-dependent protein kinase kinase2 (CAMKK2) knock out cells, suppressing the proliferation of human prostate cancer cells (61).

Taken together, ACC1 and ACC2 are sophisticatedly regulated in cells to make the process of fatty acid synthesis, as well as its cross-talk metabolic networks, meet the physiologies of cells. There are myriad factors that regulate ACC1's activities, including nutrients, protein kinases, phosphatases, allosteric regulators, transcriptional factors et al. Dysregulation of these regulatory factors usually serves as causative signaling for the development of cancer and metabolic diseases (10, 54, 61–63). Dysregulation of ACCs in diseases is therefore intensively studied.

DYSREGULATION OF ACETYL-COA CARBOXYLASES IN DISEASES

Fatty acid synthesis is central in the cross-talk between multiple biological processes, including membrane biosynthesis, energy storage, and the generation of signaling molecules (64). Lipogenesis is dynamically regulated in response to the physiologies of cells. Correspondingly, dysregulation of fatty acid synthesis can induce or promote the development of diseases. ACCs is the first rate-limiting enzyme in fatty acid synthesis. It is therefore the focus of mountains of studies and be validated as a critical participant in diseases, especially cancer and metabolic diseases (11, 65–71).

Signaling regulators of lipid biosynthesis are major downstream targets of oncogenes and tumour suppressor pathways. Alternations of oncogenes and tumour suppressor pathways can manipulate *de novo* fatty acid synthesis. Dysregulation of fatty acid metabolism, in turn, influences the cellular processes that are linked to diseases, such as cancer. For example, the AMPK pathway is important in regulating cell growth, lipid and glucose metabolism, and autophagy (72). It senses the relative level of ADP to ATP and be activated when the ratio of ADP to ATP increased. When tumour cells suffering metabolic stresses, AMPK can be activated, which then phosphorylates ACCs to suppress the lipid biosynthesis pathway, resulting in metabolism reprogramming that influences the survival and growth of tumour cells. Mutagenic blockage of the AMPK phosphorylation site of ACC (ACC1 Ser76Ala and ACC2 Ser212Ala) increases liver DNL and accelerates the development of hepatocellular carcinoma (HCC). Liver-specific inhibitor ND-654, which mimics ACC phosphorylation, inhibits liver DNL and the progression of HCC, resulting in an improved prognosis for tumour-bearing mice (73). In head and neck squamous cell carcinoma cells (HNSCC), the AMPK activator cetuximab and 5-aminoimidazole-4-carboxamide-1-β-D-ribofuranoside (AICAR) can suppress tumour cell growth (74, 75). Abolishing the AMPK phosphorylation sites on ACC1 by mutagenesis can protect HNSCC from cetuximab-induced growth inhibition. Decreased AMPK activity in hereditary leiomyomatosis renal cell cancer (HLRCC) leads to the elevated activity of ACC1, which contributes to the oncogenic growth of HLRCC (76). Metformin is an agonist of AMPK that can promote phosphorylation of ACCs. Metformin treatment can effectively suppress lipogenesis and cancer cell proliferation (10). Because ACC1 can mediate the AMPK-sensed metabolic stress and the downstream of cancer metabolism reprogramming, it is considered a potential target for cancer therapy. However, some studies also show exceptional viewpoints on the AMPK/ACC signaling pathway for tumour growth (77). For example, under energy stress conditions, activated AMPK can phosphorylate and inhibit ACC1, which can suppress the NADPH-consuming fatty acid synthesis and maintain the NADPH homeostasis in tumour cells. Similarly, ACC1 depletion can also suppress the NADPH consumption by fatty acid synthesis, which in turn partially facilitates solid tumour survival under stress conditions (77). Thus, under special conditions, the AMPK/ACC signaling pathway that can alternatively regulate tumour cell proliferation by maintaining NADPH homeostasis.

The phosphatidylinositol-3 kinase (PI3K)/Akt/mammalian target of the rapamycin (mTOR) is another signaling pathway that senses the physiologies of cells and executes important functions by regulating ACC1 activities. In general, receptor tyrosine kinases (RTKs)-mediated activation of PI3K can activate

Akt. Hyperactivated Akt then activates mTOR, which processes the upstream signals and forms mTORC1 (78). PI3K/Akt/mTOR signaling pathway regulates tumour metabolism, growth, survival, and metastasis (79, 80). ACC1 is tightly associated with the PI3K/Akt/mTOR signaling pathway in cancer cells. For example, the melanoma antigen ganglioside GD3 is a downstream target of PI3K/Akt/mTOR signaling. In melanoma, GD3 can induce the activation of SREBP-1, which is a transcription factor that regulates the expression of ACC1 (81). In breast cancer, the HER2 oncogene can induce ACC1 expression through translational regulation of the mTOR signaling pathway (82). Correspondingly, inhibition of ACC1 by siRNA or chemical inhibitors can inhibit AKT-related pathways, which is detrimental for cancer, such as human HCC (83). It is therefore concluded that ACC1 protein level and activity can be regulated by various internal alterations, which in turn affects lipid synthesis in tumors. Dysregulated lipid metabolism impacts multiple intracellular processes, such as membrane synthesis and energy metabolism that may influence tumor development ultimately. However, the mechanisms underlying lipid metabolism influencing tumor progressions, such as proliferation and metastasis, have not been fully elucidated. How ACC1 cross-talk with other pathways remains open for discussion.

Metabolic diseases are also tightly associated with the dysregulation of ACCs. In mammals, the accumulation of lipid in tissues, such as muscle and liver, is closely related to insulin resistance that associates with a myriad of metabolic disorders (84, 85). Likewise, dysregulated lipogenesis may lead to metabolic diseases such as obesity, diabetes, and NAFLD (6–8). As a central player of lipogenesis, ACCs promptly participates in the progression of metabolic disease. For example, a high-fat diet leads to increased ACC1 activity and obesity in mice while inhibition of ACC1 antagonizes the high fat diet induced obesity. ACC2 plays roles in controlling diet-induced diabetic nephropathy (DN). High-glucose diets promote lipid deposition and reduce fatty acid β-oxidation in human podocytes. Depletion of ACC2 attenuates the high-glucose diet-promoted lipid deposition and podocyte injury. The expression of glucose transporter 4 (GLUT4) is also restored by ACC2 depletion, which hampers the insulin signaling pathway. Besides, the expression of SIRT1/PGC-1a, an important complex related to the insulin metabolic pathway is also restored in the cells with ACC2 depletion, leading to the reduction of cellular insulin resistance and ultimately alleviating DN-induced cell injury (86). ACCs knockout animal models are powerful tools to understand the roles of ACCs in the progression of metabolic diseases, with which, a study demonstrated that ACC1 is necessary to maintain functional pancreatic β cells and glucose homeostasis *in vivo*, which indicates that ACC1 might be used to improve insulin secretion during diabetes (71). Liver-specific ACC1-KO mice (LACC1 KO) accumulate 40%-70% lower triglycerides in livers than that of wide-type mice when overnutrition is provided. Similarly, ACC2 knockout (ACC2 KO) mice do not gain weight when fed with high-fat diet (HF) (34). It might be due to the hepatic peroxisome proliferator-activated receptor-γ (PPAR-γ) proteins that are significantly reduced in ACC2 KO mice that are fed with high-fat and high-carbohydrate diet (HFHC). In this case, lipid synthesis-related enzymes such as ACC1, FASN, and ATP citrate lyase (ACL) are decreased, which in turn reduced diet-induced obesity. Besides, ACC2 KO mice are able to alleviate the HFHC diet-induced insulin resistance. ACC2 KO mice with HF diet show reduced AKT level and increased phosphorylation of AKT, which is critical in the insulin signaling pathway that can protect the mice from diabetes (87). The above researches demonstrate that ACCs is responsible for metabolic disorders caused by dietary factors (27). Moreover, hyper-activation of ACC1 can also result in abnormal physiologies in metabolic disease. For instance, the enhanced activity of ACC1 accelerates lipogenesis and lipid accumulation when animal suffering overnutrition and obesity, which leads to the accumulation of triglycerides in hepatocytes and thus causing NAFLD (88).

In general, dysregulated lipogenesis leads to the development of tumorigenesis and metabolic diseases. The roles of ACCs in regulating metabolism reprogramming in cancer and metabolic diseases are revealed in accumulated studies, which shed bright light on diseases treatment. As such, ACCs are becoming a promising therapeutic target for discovering novel therapeutic strategies and therapeutics development.

ACETYL-COA CARBOXYLASES-TARGETING SMALL MOLECULES FOR THERAPEUTIC PROPOSES

With the evidence of ACCs participating in the progression of diseases and its structural information, countless screenings for ACCs antagonists are performed and several promising leading compounds are confirmed for further validations (65, 83, 89–94). The ACCs inhibitors mainly target its BC domain and CT domain.

The BC domain accounts for the biotin carboxylation and formation of the homodimer of ACCs molecules. The main mechanism of action (MOA) of the BC domain targeting inhibitors is allosterically inhibiting the dimerization of the BC domain, maintaining ACC1 molecules as inactive monomers that are unable to perform the catalytic activity (21). Soraphen A, AMPK activators, and ND-series inhibitors (ND-630, ND-646, ND-654) inhibit ACC1 belong to this category (73, 91, 95–98). These inhibitors can effectively inhibit ACCs activity and affect the process of lipid metabolism and the development of disease (10, 26, 28, 29, 90, 99, 100). It is worth noting that there are studies already confirmed that inhibiting ACCs in the liver by using ND-630 (GS-0976) can significantly reduce 29% liver fat, hepatic steatosis, and markers of liver injury in NAFLD patients (101, 102), which further encourage the finding of ACCs inhibitors for therapeutic proposes.

The CT domain catalyzes the transfer of carboxyl-moiety from the carboxylated biotin to the acetyl-CoA to produce malonyl-CoA. Competitive inhibitors targeting the binding of acetyl-CoA by CT domain are therefore another promising strategy for inhibiting ACCs. TOFA, CP-640186, piperidinyl derived analogs, and spiropiperidine derived compounds are antagonists that belong to this category (103–108). These

antagonists can reduce the mice's appetite and accelerate weight loss (93, 109) and lead to apoptosis in different cancer cell lines (92, 110, 111). Despite no relevant clinical trials of this class of antagonists are found, it keeps recruiting screenings for new leading compounds.

In conclusion, numbers of commercially available ACCs inhibitors have exhibited strong therapeutic effects on disease models *in vivo* and *in vitro*, supporting that ACCs are promising therapeutic targets for the treatment of tumour and metabolic diseases. However, no agonist can specifically inhibit one ACCs member and keep another member intact. This might lead to adverse effects, because ACC1 and ACC2, indeed, are different in physiologies and pathophysiologies. To this end, the development of agonists that are specifically against ACC1 or ACC2 might be a promising strategy to target ACCs for diseases treatment.

LIMITATIONS AND PROSPECTS

Antagonists that target ACCs are intensively studied in clinic but hampered by several side-effects. For example, inhibiting lipogenesis *via* suppressing the expression of ACCs can reduce hepatic steatosis, but it simultaneously results in hypertriglyceridemia due to the activation of SREBP-1c and increased VLDL secretion (40). Another instance is PF-05175157, the first-in-human clinical trials ACC inhibitor, contributes to DNL reduction in treatment for T2DM but with concomitant reductions in platelet count (112). Recently, an exciting result in phase II clinical trial shows that the co-administration of PF-0522134 (a new ACC1 inhibitor in clinical trial) and PF-6865571 (DGAT2 inhibitor) has a strong effect in treating NASH without the side effect of hypertriglyceridemia (113). However, there are several challenges to address the side-effects of ACCs inhibition in clinical practice.

The principal challenge is that the inhibitors can hardly distinguish ACC1 from ACC2. As described above, ACC1 and ACC2 share 75% identity in amino acid sequence and are similar in structures that are composed of conservative domains for enzyme activity. However, the ACCs antagonists, such as Soraphen A and TOFA, can target and influence the activity of both ACC1 and ACC2, which might lead to the side-effects caused by the inhibition of unwanted ACC isoform. In nutrient-abundant condition, fatty acid synthesis and breakdown are coordinately controlled, avoiding a wasteful cycle of metabolism. However, in cancer cells, both fatty acid synthesis and breakdown are boosted to support cancer growth. To this end, coordinately antagonizing the dysregulation of ACC1 and ACC2 in cancer cells would be a promising strategy for cancer treatment. So far, none of ACCs inhibitors is approved useful in clinic. This might be due to the fact that ACC inhibitors that are not isoform-specific only partially reverse cancer's preferences. Moreover, it is shown that the selectively inhibition of ACC2 may be ineffective in treating some metabolic diseases (114). A selective inhibitor targeting ACC1 that shows anti-NAFLD/NASH effects in pre-clinical models is reported in a recent study (115), which is expected to strengthen the efficacy.

Accumulating studies indicate the importance of ACCs in tumour cell growth which shows the great potential of ACCs in the treatment of cancer. However, studies on the role of ACCs in cancer have been attributed to their roles in fatty acid synthesis, the exact mechanism of which remains to be investigated. The role of fatty acid metabolism in cancer biology is not fully understood (116). More in-depth research about fatty acid metabolism in cancer will help examine and detail the roles of the ACCs, in cancer initiation, progression, and development.

AUTHOR CONTRIBUTIONS

YuW: draft the manuscript and iconography. WY, SL, DG, and JH: proofread the manuscript and iconography. YugW: conceptualized, supervised, and finalized the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China (31970577 and 91957110).

REFERENCES

1. Menendez JA, Lupu R. Fatty Acid Synthase and the Lipogenic Phenotype in Cancer Pathogenesis. *Nat Rev Cancer* (2007) 7:763–77. doi: 10.1038/nrc2222
2. Abu-Elheiga L, Matzuk MM, Kordari P, Oh W, Shaikenov T, Gu Z, et al. Mutant Mice Lacking Acetyl-CoA Carboxylase 1 Are Embryonically Lethal. *Proc Natl Acad Sci USA* (2005) 102:12011–6. doi: 10.1073/pnas.0505714102
3. Schreurs M, Kuipers F, van der Leij FR. Regulatory Enzymes of Mitochondrial Beta-Oxidation as Targets for Treatment of the Metabolic Syndrome. *Obes Rev* (2010) 11:380–8. doi: 10.1111/j.1467-789X.2009.00642.x
4. Currie E, Schulze A, Zechner R, Walther TC, Farese RVJr. Cellular Fatty Acid Metabolism and Cancer. *Cell Metab* (2013) 18:153–61. doi: 10.1016/j.cmet.2013.05.017
5. Hodson L, Gunn PJ. The Regulation of Hepatic Fatty Acid Synthesis and Partitioning: The Effect of Nutritional State. *Nat Rev Endocrinol* (2019) 15:689–700. doi: 10.1038/s41574-019-0256-9
6. Postic C, Girard J. Contribution of *De Novo* Fatty Acid Synthesis to Hepatic Steatosis and Insulin Resistance: Lessons From Genetically Engineered Mice. *J Clin Invest* (2008) 118:829–38. doi: 10.1172/JCI34275
7. Ronnett GV, Kim E-K, Landree LE, Tu Y. Fatty Acid Metabolism as a Target for Obesity Treatment. *Physiol Behav* (2005) 85:25–35. doi: 10.1016/j.physbeh.2005.04.014
8. Imamura F, Fretts AM, Marklund M, Ardisson Korat AV, Yang W-S, Lankinen M, et al. Fatty Acids in the *De Novo* Lipogenesis Pathway and Incidence of Type 2 Diabetes: A Pooled Analysis of Prospective Cohort Studies. *PloS Med* (2020) 17:e1003102. doi: 10.1371/journal.pmed.1003102
9. Bruning U, Morales-Rodriguez F, Kalucka J, Goveia J, Taverna F, Queiroz KC, et al. Impairment of Angiogenesis by Fatty Acid Synthase Inhibition

Involves mTOR Malonylation. *Cell Metab* (2018) 28:866–80.e15. doi: 10.1016/j.cmet.2018.07.019
10. Fullerton MD, Galic S, Marcinko K, Sikkema S, Pulinilkunnil T, Chen ZP, et al. Single Phosphorylation Sites in Acc1 and Acc2 Regulate Lipid Homeostasis and the Insulin-Sensitizing Effects of Metformin. *Nat Med* (2013) 19:1649–54. doi: 10.1038/nm.3372
11. Rios Garcia M, Steinbauer B, Srivastava K, Singhal M, Mattijssen F, Maida A, et al. Acetyl-CoA Carboxylase 1-Dependent Protein Acetylation Controls Breast Cancer Metastasis and Recurrence. *Cell Metab* (2017) 26:842–55.e5. doi: 10.1016/j.cmet.2017.09.018
12. Wakil SJ, Abu-Elheiga LA. Fatty Acid Metabolism: Target for Metabolic Syndrome. *J Lipid Res* (2009) 50:S138–S43. doi: 10.1194/jlr.R800079-JLR200
13. Chow JD, Lawrence RT, Healy ME, Dominy JE, Liao JA, Breen DS, et al. Genetic Inhibition of Hepatic Acetyl-CoA Carboxylase Activity Increases Liver Fat and Alters Global Protein Acetylation. *Mol Metab* (2014) 3:419–31. doi: 10.1016/j.molmet.2014.02.004
14. Galdieri L, Vancura A. Acetyl-CoA Carboxylase Regulates Global Histone Acetylation. *J Biol Chem* (2012) 287:23865–76. doi: 10.1074/jbc.M112.380519
15. Tong L. Acetyl-Coenzyme A Carboxylase: Crucial Metabolic Enzyme and Attractive Target for Drug Discovery. *Cell Mol Life Sci* (2005) 62:1784–803. doi: 10.1007/s00018-005-5121-4
16. Abu-Elheiga L, Almarza-Ortega DB, Baldini A, Wakil SJ. Human Acetyl-CoA Carboxylase 2. Molecular Cloning, Characterization, Chromosomal Mapping, and Evidence for Two Isoforms. *J Biol Chem* (1997) 272:10669–77. doi: 10.1074/jbc.272.16.10669
17. Madauss KP, Burkhart WA, Consler TG, Cowan DJ, Gottschalk WK, Miller AB, et al. The Human ACC2 CT-Domain C-Terminus Is Required for Full Functionality and Has a Novel Twist. *Acta Crystallogr D Biol Crystallogr* (2009) 65:449–61. doi: 10.1107/s0907444909008014
18. Lee CK, Cheong HK, Ryu KS, Lee JI, Lee W, Jeon YH, et al. Biotinoyl Domain of Human Acetyl-CoA Carboxylase: Structural Insights Into the Carboxyl Transfer Mechanism. *Proteins* (2008) 72:613–24. doi: 10.1002/prot.21952
19. Chen Y, Elizondo-Noriega A, Cantu DC, Reilly PJ. Structural Classification of Biotin Carboxyl Carrier Proteins. *Biotechnol Lett* (2012) 34:1869–75. doi: 10.1007/s10529-012-0978-4
20. Zhang H, Yang Z, Shen Y, Tong L. Crystal Structure of the Carboxyltransferase Domain of Acetyl-Coenzyme A Carboxylase. *Science* (2003) 299:2064–7. doi: 10.1126/science.1081366
21. Wei J, Tong L. Crystal Structure of the 500-kDa Yeast Acetyl-CoA Carboxylase Holoenzyme Dimer. *Nature* (2015) 526:723–7. doi: 10.1038/nature15375
22. Broussard TC, Kobe MJ, Pakhomova S, Neau DB, Price AE, Champion TS, et al. The Three-Dimensional Structure of the Biotin Carboxylase-Biotin Carboxyl Carrier Protein Complex of E. Coli Acetyl-CoA Carboxylase. *Structure* (2013) 21:650–7. doi: 10.1016/j.str.2013.02.001
23. Hunkeler M, Hagmann A, Stuttfeld E, Chami M, Guri Y, Stahlberg H, et al. Structural Basis for Regulation of Human Acetyl-CoA Carboxylase. *Nature* (2018) 558:470–4. doi: 10.1038/s41586-018-0201-4
24. Munday MR. Regulation of Mammalian Acetyl-CoA Carboxylase. *Biochem Soc Trans* (2002) 30:1059–64. doi: 10.1042/bst0301059
25. Harada N, Oda Z, Hara Y, Fujinami K, Okawa M, Ohbuchi K, et al. Hepatic *De Novo* Lipogenesis Is Present in Liver-Specific ACC1-Deficient Mice. *Mol Cell Biol* (2007) 27:1881–8. doi: 10.1128/mcb.01122-06
26. Schreurs M, van Dijk TH, Gerding A, Havinga R, Reijngoud DJ, Kuipers F. Soraphen, an Inhibitor of the Acetyl-CoA Carboxylase System, Improves Peripheral Insulin Sensitivity in Mice Fed a High-Fat Diet. *Diabetes Obes Metab* (2009) 11:987–91. doi: 10.1111/j.1463-1326.2009.01078.x
27. Mao J, DeMayo FJ, Li H, Abu-Elheiga L, Gu Z, Shaikenov TE, et al. Liver-Specific Deletion of Acetyl-CoA Carboxylase 1 Reduces Hepatic Triglyceride Accumulation Without Affecting Glucose Homeostasis. *Proc Natl Acad Sci USA* (2006) 103:8552–7. doi: 10.1073/pnas.0603115103
28. Rysman E, Brusselmans K, Scheys K, Timmermans L, Derua R, Munck S, et al. *De Novo* Lipogenesis Protects Cancer Cells From Free Radicals and Chemotherapeutics by Promoting Membrane Lipid Saturation. *Cancer Res* (2010) 70:8117–26. doi: 10.1158/0008-5472.Can-09-3871
29. Stoiber K, Nagło O, Pernpeintner C, Zhang S, Koeberle A, Ulrich M, et al. Targeting *De Novo* Lipogenesis as a Novel Approach in Anti-Cancer Therapy. *Br J Cancer* (2018) 118:43–51. doi: 10.1038/bjc.2017.374
30. Abu-Elheiga L, Brinkley WR, Zhong L, Chirala SS, Woldegiorgis G, Wakil SJ. The Subcellular Localization of Acetyl-CoA Carboxylase 2. *Proc Natl Acad Sci USA* (2000) 97:1444–9. doi: 10.1073/pnas.97.4.1444
31. Savage DB, Choi CS, Samuel VT, Liu ZX, Zhang D, Wang A, et al. Reversal of Diet-Induced Hepatic Steatosis and Hepatic Insulin Resistance by Antisense Oligonucleotide Inhibitors of Acetyl-CoA Carboxylases 1 and 2. *J Clin Invest* (2006) 116:817–24. doi: 10.1172/jci27300
32. Essop MF, Camp HS, Choi CS, Sharma S, Fryer RM, Reinhart GA, et al. Reduced Heart Size and Increased Myocardial Fuel Substrate Oxidation in ACC2 Mutant Mice. *Am J Physiol Heart Circ Physiol* (2008) 295:H256–65. doi: 10.1152/ajpheart.91489.2007
33. Eaton S. Control of Mitochondrial Beta-Oxidation Flux. *Prog Lipid Res* (2002) 41:197–239. doi: 10.1016/s0163-7827(01)00024-8
34. Abu-Elheiga L, Matzuk MM, Abo-Hashema KA, Wakil SJ. Continuous Fatty Acid Oxidation and Reduced Fat Storage in Mice Lacking Acetyl-CoA Carboxylase 2. *Science* (2001) 291:2613–6. doi: 10.1126/science.1056843
35. Pietrocola F, Galluzzi L, Bravo-San Pedro JM, Madeo F, Kroemer G. Acetyl Coenzyme A: A Central Metabolite and Second Messenger. *Cell Metab* (2015) 21:805–21. doi: 10.1016/j.cmet.2015.05.014
36. Snaebjornsson MT, Janaki-Raman S, Schulze A. Greasing the Wheels of the Cancer Machine: The Role of Lipid Metabolism in Cancer. *Cell Metab* (2020) 31:62–76. doi: 10.1016/j.cmet.2019.11.010
37. Saggerson D. Malonyl-CoA, a Key Signaling Molecule in Mammalian Cells. *Annu Rev Nutr* (2008) 28:253–72. doi: 10.1146/annurev.nutr.28.061807.155434
38. Wang M, Ma LJ, Yang Y, Xiao Z, Wan JB. N-3 Polyunsaturated Fatty Acids for the Management of Alcoholic Liver Disease: A Critical Review. *Crit Rev Food Sci Nutr* (2019) 59:S116–s29. doi: 10.1080/10408398.2018.1544542
39. Zhang XJ, Cai J, Li H. Targeting ACC for NASH Resolution. *Trends Mol Med* (2022) 28:5–7. doi: 10.1016/j.molmed.2021.11.002
40. Kim CW, Addy C, Kusunoki J, Anderson NN, Deja S, Fu X, et al. Acetyl CoA Carboxylase Inhibition Reduces Hepatic Steatosis But Elevates Plasma Triglycerides in Mice and Humans: A Bedside to Bench Investigation. *Cell Metab* (2017) 26:394–406.e6. doi: 10.1016/j.cmet.2017.07.009
41. Lee JV, Berry CT, Kim K, Sen P, Kim T, Carrer A, et al. Acetyl-CoA Promotes Glioblastoma Cell Adhesion and Migration Through Ca(2+)-NFAT Signaling. *Genes Dev* (2018) 32:497–511. doi: 10.1101/gad.311027.117
42. Fadó R, Rodríguez-Rodríguez R, Casals N. The Return of Malonyl-CoA to the Brain: Cognition and Other Stories. *Prog Lipid Res* (2021) 81:101071. doi: 10.1016/j.plipres.2020.101071
43. Frost G, Sleeth ML, Sahuri-Arisoylu M, Lizarbe B, Cerdan S, Brody L, et al. The Short-Chain Fatty Acid Acetate Reduces Appetite *via* a Central Homeostatic Mechanism. *Nat Commun* (2014) 5:3611. doi: 10.1038/ncomms4611
44. Wyss MT, Magistretti PJ, Buck A, Weber B. Labeled Acetate as a Marker of Astrocytic Metabolism. *J Cereb Blood Flow Metab* (2011) 31:1668–74. doi: 10.1038/jcbfm.2011.84
45. Swinnen JV, Ulrix W, Heyns W, Verhoeven G. Coordinate Regulation of Lipogenic Gene Expression by Androgens: Evidence for a Cascade Mechanism Involving Sterol Regulatory Element Binding Proteins. *Proc Natl Acad Sci USA* (1997) 94:12975–80. doi: 10.1073/pnas.94.24.12975
46. Atkinson LL, Fischer MA, Lopaschuk GD. Leptin Activates Cardiac Fatty Acid Oxidation Independent of Changes in the AMP-Activated Protein Kinase-Acetyl-CoA Carboxylase-Malonyl-CoA Axis. *J Biol Chem* (2002) 277:29424–30. doi: 10.1074/jbc.M203813200
47. Steinberg GR, Kemp BE. AMPK in Health and Disease. *Physiol Rev* (2009) 89:1025–78. doi: 10.1152/physrev.00011.2008
48. Sanders FW, Griffin JL. *De Novo* Lipogenesis in the Liver in Health and Disease: More Than Just a Shunting Yard for Glucose. *Biol Rev* (2016) 91:452–68. doi: 10.1111/brv.12178
49. Auger C, Knuth CM, Abdullahi A, Samadi O, Parousis A, Jeschke MG. Metformin Prevents the Pathological Browning of Subcutaneous White Adipose Tissue. *Mol Metab* (2019) 29:12–23. doi: 10.1016/j.molmet.2019.08.011
50. Moreau K, Dizin E, Ray H, Luquain C, Lefai E, Foufelle F, et al. BRCA1 Affects Lipid Synthesis Through Its Interaction With Acetyl-CoA Carboxylase. *J Biol Chem* (2006) 281:3172–81. doi: 10.1074/jbc.M504652200
51. Koobotse M, Holly J, Perks C. Elucidating the Novel BRCA1 Function as a Non-Genomic Metabolic Restraint in ER-Positive Breast Cancer Cell Lines. *Oncotarget* (2018) 9:33562–76. doi: 10.18632/oncotarget.26093

52. Beaty NB, Lane MD. Kinetics of Activation of Acetyl-CoA Carboxylase by Citrate. Relationship to the Rate of Polymerization of the Enzyme. *J Biol Chem* (1983) 258:13043–50. doi: 10.1016/S0021-9258(17)44077-4
53. Vagelos PR, Alberts AW, Martin DB. Studies on the Mechnism of Activation of Acetyl Coenzyme A Carboxylase by Citrate. *J Biol Chem* (1963) 238:533–40 https://pubmed.ncbi.nlm.nih.gov/13995702/.
54. Wakil SJ, Stoops JK, Joshi VC. Fatty Acid Synthesis and Its Regulation. *Annu Rev Biochem* (1983) 52:537–79. doi: 10.1146/annurev.bi.52.070183.002541
55. Boone AN, Chan A, Kulpa JE, Brownsey RW. Bimodal Activation of Acetyl-CoA Carboxylase by Glutamate. *J Biol Chem* (2000) 275:10819–25. doi: 10.1074/jbc.275.15.10819
56. Ishii S, Iizuka K, Miller BC, Uyeda K. Carbohydrate Response Element Binding Protein Directly Promotes Lipogenic Enzyme Gene Transcription. *Proc Natl Acad Sci USA* (2004) 101:15597–602. doi: 10.1073/pnas.0405238101
57. Iizuka K, Horikawa Y. ChREBP: A Glucose-Activated Transcription Factor Involved in the Development of Metabolic Syndrome. *Endocr J* (2008) 55:617–24. doi: 10.1507/endocrj.k07e-110
58. Ma J, Yan R, Zu X, Cheng JM, Rao K, Liao DF, et al. Aldo-Keto Reductase Family 1 B10 Affects Fatty Acid Synthesis by Regulating the Stability of Acetyl-CoA Carboxylase-Alpha in Breast Cancer Cells. *J Biol Chem* (2008) 283:3418–23. doi: 10.1074/jbc.M707650200
59. Ueda K, Nakatsu Y, Yamamotoya T, Ono H, Inoue Y, Inoue MK, et al. Prolyl Isomerase Pin1 Binds to and Stabilizes Acetyl CoA Carboxylase 1 Protein, Thereby Supporting Cancer Cell Proliferation. *Oncotarget* (2019) 10:1637–48. doi: 10.18632/oncotarget.26691
60. Mao J, Chirala SS, Wakil SJ. Human Acetyl-CoA Carboxylase 1 Gene: Presence of Three Promoters and Heterogeneity at the 5'-Untranslated mRNA Region. *Proc Natl Acad Sci USA* (2003) 100:7515–20. doi: 10.1073/pnas.1332670100
61. Penfold L, Woods A, Muckett P, Nikitin AY, Kent TR, Zhang S, et al. CAMKK2 Promotes Prostate Cancer Independently of AMPK *via* Increased Lipogenesis. *Cancer Res* (2018) 78:6747–61. doi: 10.1158/0008-5472.Can-18-0585
62. Karami KJ, Coppola J, Krishnamurthy K, Llanos DJ, Mukherjee A, Venkatachalam KV. Effect of Food Deprivation and Hormones of Glucose Homeostasis on the Acetyl CoA Carboxylase Activity in Mouse Brain: A Potential Role of Acc in the Regulation of Energy Balance. *Nutr Metab (Lond)* (2006) 3:15. doi: 10.1186/1743-7075-3-15
63. Kaulage MH, Bhattacharya S, Muniyappa K. Structural Characterization of I-Motif Structure in the Human Acetyl-CoA Carboxylase 1 Gene Promoters and Their Role in the Regulation of Gene Expression. *Chembiochem* (2018) 19:1078–87. doi: 10.1002/cbic.201800021
64. Röhrig F, Schulze A. The Multifaceted Roles of Fatty Acid Synthesis in Cancer. *Nat Rev Cancer* (2016) 16:732–49. doi: 10.1038/nrc.2016.89
65. Luo DX, Tong DJ, Rajput S, Wang C, Liao DF, Cao D, et al. Targeting Acetyl-CoA Carboxylases: Small Molecular Inhibitors and Their Therapeutic Potential. *Recent Pat Anticancer Drug Discov* (2012) 7:168–84. doi: 10.2174/157489212799972918
66. Wang MD, Wu H, Fu GB, Zhang HL, Zhou X, Tang L, et al. Acetyl-Coenzyme A Carboxylase Alpha Promotion of Glucose-Mediated Fatty Acid Synthesis Enhances Survival of Hepatocellular Carcinoma in Mice and Patients. *Hepatology* (2016) 63:1272–86. doi: 10.1002/hep.28415
67. Milgraum LZ, Witters LA, Pasternack GR, Kuhajda FP. Enzymes of the Fatty Acid Synthesis Pathway Are Highly Expressed in *In Situ* Breast Carcinoma. *Clin Cancer Res* (1997) 3:2115–20 https://pubmed.ncbi.nlm.nih.gov/9815604/.
68. Yang JH, Kim NH, Yun JS, Cho ES, Cha YH, Cho SB, et al. Snail Augments Fatty Acid Oxidation by Suppression of Mitochondrial ACC2 During Cancer Progression. *Life Sci Alliance* (2020) 3:e202000683. doi: 10.26508/lsa.202000683
69. Brusselmans K, De Schrijver E, Verhoeven G, Swinnen JV. RNA Interference-Mediated Silencing of the Acetyl-CoA-Carboxylase-Alpha Gene Induces Growth Inhibition and Apoptosis of Prostate Cancer Cells. *Cancer Res* (2005) 65:6719–25. doi: 10.1158/0008-5472.Can-05-0571
70. Li K, Zhang C, Chen L, Wang P, Fang Y, Zhu J, et al. The Role of acetyl-coA Carboxylase2 in Head and Neck Squamous Cell Carcinoma. *PeerJ* (2019) 7:e7037. doi: 10.7717/peerj.7037
71. Cantley J, Davenport A, Vetterli L, Nemes NJ, Whitworth PT, Boslem E, et al. Disruption of Beta Cell Acetyl-CoA Carboxylase-1 in Mice Impairs Insulin Secretion and Beta Cell Mass. *Diabetologia* (2019) 62:99–111. doi: 10.1007/s00125-018-4743-7
72. Herzig S, Shaw RJ. AMPK: Guardian of Metabolism and Mitochondrial Homeostasis. *Nat Rev Mol Cell Biol* (2018) 19:121–35. doi: 10.1038/nrm.2017.95
73. Lally JSV, Ghoshal S, DePeralta DK, Moaven O, Wei L, Masia R, et al. Inhibition of Acetyl-CoA Carboxylase by Phosphorylation or the Inhibitor ND-654 Suppresses Lipogenesis and Hepatocellular Carcinoma. *Cell Metab* (2019) 29:174–82.e5. doi: 10.1016/j.cmet.2018.08.020
74. Luo J, Hong Y, Lu Y, Qiu S, Chaganty BK, Zhang L, et al. Acetyl-CoA Carboxylase Rewires Cancer Metabolism to Allow Cancer Cells to Survive Inhibition of the Warburg Effect by Cetuximab. *Cancer Lett* (2017) 384:39–49. doi: 10.1016/j.canlet.2016.09.020
75. Li K, Chen L, Lin Z, Zhu J, Fang Y, Du J, et al. Role of the AMPK/ACC Signaling Pathway in TRPP2-Mediated Head and Neck Cancer Cell Proliferation. *BioMed Res Int* (2020) 2020:4375075. doi: 10.1155/2020/4375075
76. Tong WH, Sourbier C, Kovtunovych G, Jeong SY, Vira M, Ghosh M, et al. The Glycolytic Shift in Fumarate-Hydratase-Deficient Kidney Cancer Lowers AMPK Levels, Increases Anabolic Propensities and Lowers Cellular Iron Levels. *Cancer Cell* (2011) 20:315–27. doi: 10.1016/j.ccr.2011.07.018
77. Jeon SM, Chandel NS, Hay N. AMPK Regulates NADPH Homeostasis to Promote Tumour Cell Survival During Energy Stress. *Nature* (2012) 485:661–5. doi: 10.1038/nature11066
78. Schlessinger J. Cell Signaling by Receptor Tyrosine Kinases. *Cell* (2000) 103:211–25. doi: 10.1016/s0092-8674(00)00114-8
79. Alzahrani AS. PI3K/Akt/mTOR Inhibitors in Cancer: At the Bench and Bedside. *Semin Cancer Biol* (2019) 59:125–32. doi: 10.1016/j.semcancer.2019.07.009
80. Porta C, Paglino C, Mosca A. Targeting PI3K/Akt/mTOR Signaling in Cancer. *Front Oncol* (2014) 4:64. doi: 10.3389/fonc.2014.00064
81. Yamauchi Y, Furukawa K, Hamamura K, Furukawa K. Positive Feedback Loop Between PI3K-Akt-Mtorc1 Signaling and the Lipogenic Pathway Boosts Akt Signaling: Induction of the Lipogenic Pathway by a Melanoma Antigen. *Cancer Res* (2011) 71:4989–97. doi: 10.1158/0008-5472.Can-10-4108
82. Yoon S, Lee MY, Park SW, Moon JS, Koh YK, Ahn YH, et al. Up-Regulation of Acetyl-CoA Carboxylase Alpha and Fatty Acid Synthase by Human Epidermal Growth Factor Receptor 2 at the Translational Level in Breast Cancer Cells. *J Biol Chem* (2007) 282:26122–31. doi: 10.1074/jbc.M702854200
83. Calvisi DF, Wang C, Ho C, Ladu S, Lee SA, Mattu S, et al. Increased Lipogenesis, Induced by AKT-Mtorc1-RPS6 Signaling, Promotes Development of Human Hepatocellular Carcinoma. *Gastroenterology* (2011) 140:1071–83. doi: 10.1053/j.gastro.2010.12.006
84. Engin A. The Definition and Prevalence of Obesity and Metabolic Syndrome. *Adv Exp Med Biol* (2017) 960:1–17. doi: 10.1007/978-3-319-48382-5_1
85. Santoleri D, Titchenell PM. Resolving the Paradox of Hepatic Insulin Resistance. *Cell Mol Gastroenterol Hepatol* (2019) 7:447–56. doi: 10.1016/j.jcmgh.2018.10.016
86. Wang Q, Zhao B, Zhang J, Sun J, Wang S, Zhang X, et al. Faster Lipid β-Oxidation Rate by Acetyl-CoA Carboxylase 2 Inhibition Alleviates High-Glucose-Induced Insulin Resistance *via* SIRT1/PGC-1α in Human Podocytes. *J Biochem Mol Toxicol* (2021) 35:e22797. doi: 10.1002/jbt.22797
87. Abu-Elheiga L, Wu H, Gu Z, Bressler R, Wakil SJ. Acetyl-CoA Carboxylase 2-/- Mutant Mice are Protected Against Fatty Liver Under High-Fat, High-Carbohydrate Dietary and *De Novo* Lipogenic Conditions. *J Biol Chem* (2012) 287:12578–88. doi: 10.1074/jbc.M111.309559
88. Alves-Bezerra M, Cohen DE. Triglyceride Metabolism in the Liver. *Compr Physiol* (2017) 8:1–8. doi: 10.1002/cphy.c170012
89. Sugimoto Y, Naniwa Y, Nakamura T, Kato H, Yamamoto M, Tanabe H, et al. A Novel Acetyl-CoA Carboxylase Inhibitor Reduces *De Novo* Fatty Acid Synthesis in HepG2 Cells and Rat Primary Hepatocytes. *Arch Biochem Biophys* (2007) 468:44–8. doi: 10.1016/j.abb.2007.09.012
90. Corominas-Faja B, Cuyàs E, Gumuzio J, Bosch-Barrera J, Leis O, Martin ÁG, et al. Chemical Inhibition of Acetyl-CoA Carboxylase Suppresses Self-

Renewal Growth of Cancer Stem Cells. *Oncotarget* (2014) 5:8306–16. doi: 10.18632/oncotarget.2059
91. Svensson RU, Parker SJ, Eichner LJ, Kolar MJ, Wallace M, Brun SN, et al. Inhibition of Acetyl-CoA Carboxylase Suppresses Fatty Acid Synthesis and Tumor Growth of Non-Small-Cell Lung Cancer in Preclinical Models. *Nat Med* (2016) 22:1108–19. doi: 10.1038/nm.4181
92. He D, Sun X, Yang H, Li X, Yang D. TOFA Induces Cell Cycle Arrest and Apoptosis in ACHN and 786-O Cells Through Inhibiting PI3K/Akt/mTOR Pathway. *J Cancer* (2018) 9:2734–42. doi: 10.7150/jca.26374
93. Loftus TM, Jaworsky DE, Frehywot GL, Townsend CA, Ronnett GV, Lane MD, et al. Reduced Food Intake and Body Weight in Mice Treated With Fatty Acid Synthase Inhibitors. *Science* (2000) 288:2379–81. doi: 10.1126/science.288.5475.2379
94. Gao YS, Qian MY, Wei QQ, Duan XB, Wang SL, Hu HY, et al. WZ66, a Novel Acetyl-CoA Carboxylase Inhibitor, Alleviates Nonalcoholic Steatohepatitis (NASH) in Mice. *Acta Pharmacol Sin* (2020) 41:336–47. doi: 10.1038/s41401-019-0310-0
95. Shen Y, Volrath SL, Weatherly SC, Elich TD, Tong L. A Mechanism for the Potent Inhibition of Eukaryotic Acetyl-Coenzyme A Carboxylase by Soraphen A, a Macrocyclic Polyketide Natural Product. *Mol Cell* (2004) 16:881–91. doi: 10.1016/j.molcel.2004.11.034
96. Jang S, Gornicki P, Marjanovic J, Bass E, PI T, Rodriguez P, et al. Activity and Structure of Human Acetyl-CoA Carboxylase Targeted by a Specific Inhibitor. *FEBS Lett* (2018) 592:2048–58. doi: 10.1002/1873-3468.13097
97. Cho YS, Lee JI, Shin D, Kim HT, Jung HY, Lee TG, et al. Molecular Mechanism for the Regulation of Human ACC2 Through Phosphorylation by AMPK. *Biochem Biophys Res Commun* (2010) 391:187–92. doi: 10.1016/j.bbrc.2009.11.029
98. Harriman G, Greenwood J, Bhat S, Huang X, Wang R, Paul D, et al. Acetyl-CoA Carboxylase Inhibition by ND-630 Reduces Hepatic Steatosis, Improves Insulin Sensitivity, and Modulates Dyslipidemia in Rats. *Proc Natl Acad Sci USA* (2016) 113:E1796–805. doi: 10.1073/pnas.1520686113
99. Beckers A, Organe S, Timmermans L, Scheys K, Peeters A, Brusselmans K, et al. Chemical Inhibition of Acetyl-CoA Carboxylase Induces Growth Arrest and Cytotoxicity Selectively in Cancer Cells. *Cancer Res* (2007) 67:8180–7. doi: 10.1158/0008-5472.Can-07-0389
100. Scaglia N, Chisholm JW, Igal RA. Inhibition of Stearoylcoa Desaturase-1 Inactivates Acetyl-CoA Carboxylase and Impairs Proliferation in Cancer Cells: Role of AMPK. *PloS One* (2009) 4:e6812. doi: 10.1371/journal.pone.0006812
101. Lawitz EJ, Coste A, Poordad F, Alkhouri N, Loo N, McColgan BJ, et al. Acetyl-CoA Carboxylase Inhibitor GS-0976 for 12 Weeks Reduces Hepatic *De Novo* Lipogenesis and Steatosis in Patients With Nonalcoholic Steatohepatitis. *Clin Gastroenterol Hepatol* (2018) 16:1983–91.e3. doi: 10.1016/j.cgh.2018.04.042
102. Alkhouri N, Lawitz E, Noureddin M, DeFronzo R, Shulman GI. GS-0976 (Firsocostat): An Investigational Liver-Directed Acetyl-CoA Carboxylase (ACC) Inhibitor for the Treatment of Non-Alcoholic Steatohepatitis (NASH). *Expert Opin Investig Drugs* (2020) 29:135–41. doi: 10.1080/13543784.2020.1668374
103. Chen L, Duan Y, Wei H, Ning H, Bi C, Zhao Y, et al. Acetyl-CoA Carboxylase (ACC) as a Therapeutic Target for Metabolic Syndrome and Recent Developments in ACC1/2 Inhibitors. *Expert Opin Investig Drugs* (2019) 28:917–30. doi: 10.1080/13543784.2019.1657825
104. McCune SA, Harris RA. Mechanism Responsible for 5-(Tetradecyloxy)-2-Furoic Acid Inhibition of Hepatic Lipogenesis. *J Biol Chem* (1979) 254:10095–101.
105. Halvorson DL, McCune SA. Inhibition of Fatty Acid Synthesis in Isolated Adipocytes by 5-(Tetradecyloxy)-2-Furoic Acid. *Lipids* (1984) 19:851–6. doi: 10.1007/bf02534514
106. Zhang H, Tweel B, Li J, Tong L. Crystal Structure of the Carboxyltransferase Domain of Acetyl-Coenzyme A Carboxylase in Complex With CP-640186. *Structure* (2004) 12:1683–91. doi: 10.1016/j.str.2004.07.009
107. Harwood HJJr., Petras SF, Shelly LD, Zaccaro LM, Perry DA, Makowski MR, et al. Isozyme-Nonselective N-Substituted Bipiperidylcarboxamide Acetyl-CoA Carboxylase Inhibitors Reduce Tissue Malonyl-CoA Concentrations, Inhibit Fatty Acid Synthesis, and Increase Fatty Acid Oxidation in Cultured Cells and in Experimental Animals. *J Biol Chem* (2003) 278:37099–111. doi: 10.1074/jbc.M304481200
108. Chonan T, Oi T, Yamamoto D, Yashiro M, Wakasugi D, Tanaka H, et al. (4-Piperidinyl)-Piperazine: A New Platform for Acetyl-CoA Carboxylase Inhibitors. *Bioorg Med Chem Lett* (2009) 19:6645–8. doi: 10.1016/j.bmcl.2009.10.012
109. Bengtsson C, Blaho S, Saitton DB, Brickmann K, Broddefalk J, Davidsson O, et al. Design of Small Molecule Inhibitors of Acetyl-CoA Carboxylase 1 and 2 Showing Reduction of Hepatic Malonyl-CoA Levels *In Vivo* in Obese Zucker Rats. *Bioorg Med Chem* (2011) 19:3039–53. doi: 10.1016/j.bmc.2011.04.014
110. Nishi K, Suzuki K, Sawamoto J, Tokizawa Y, Iwase Y, Yumita N, et al. Inhibition of Fatty Acid Synthesis Induces Apoptosis of Human Pancreatic Cancer Cells. *Anticancer Res* (2016) 36:4655–60. doi: 10.21873/anticanres.11016
111. Guseva NV, Rokhlin OW, Glover RA, Cohen MB. TOFA (5-Tetradecyl-Oxy-2-Furoic Acid) Reduces Fatty Acid Synthesis, Inhibits Expression of AR, Neuropilin-1 and Mcl-1 and Kills Prostate Cancer Cells Independent of P53 Status. *Cancer Biol Ther* (2011) 12:80–5. doi: 10.4161/cbt.12.1.15721
112. Griffith DA, Kung DW, Esler WP, Amor PA, Bagley SW, Beysen C, et al. Decreasing the Rate of Metabolic Ketone Reduction in the Discovery of a Clinical Acetyl-CoA Carboxylase Inhibitor for the Treatment of Diabetes. *J Med Chem* (2014) 57:10512–26. doi: 10.1021/jm5016022
113. Calle RA, Amin NB, Carvajal-Gonzalez S, Ross TT, Bergman A, Aggarwal S, et al. ACC Inhibitor Alone or Co-Administered With a DGAT2 Inhibitor in Patients With Non-Alcoholic Fatty Liver Disease: Two Parallel, Placebo-Controlled, Randomized Phase 2a Trials. *Nat Med* (2021) 27(10):1836–48. doi: 10.1038/s41591-021-01489-1
114. Olson DP, Pulinilkunnil T, Cline GW, Shulman GI, Lowell BB. Gene Knockout of Acc2 has Little Effect on Body Weight, Fat Mass, or Food Intake. *Proc Natl Acad Sci USA* (2010) 107:7598–603. doi: 10.1073/pnas.0913492107
115. Tamura Y, Sugama J, Iwasaki S, Sasaki M, Yasuno H, Aoyama K, et al. Selective Acetyl-CoA Carboxylase 1 Inhibitor Improves Hepatic Steatosis and Hepatic Fibrosis in a Pre-Clinical NASH Model. *J Pharmacol Exp Ther* (2021) 379(3):280–9. doi: 10.1124/jpet.121.000786
116. Hanahan D, Weinberg RA. Hallmarks of Cancer: The Next Generation. *Cell* (2011) 144:646–74. doi: 10.1016/j.cell.2011.02.013

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

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Oncology

ORIGINAL RESEARCH
published: 14 April 2022
doi: 10.3389/fonc.2022.857572

Computational Model of Heterogeneity in Melanoma: Designing Therapies and Predicting Outcomes

Arran Hodgkinson[1], *Dumitru Trucu*[2], *Matthieu Lacroix*[3,4], *Laurent Le Cam*[3,4] *and Ovidiu Radulescu*[5*]

[1] Living Systems Institute, University of Exeter, Exeter, United Kingdom, [2] Division of Mathematics, University of Dundee, Dundee, United Kingdom, [3] IRCM, Institut de Recherche en Cancérologie de Montpellier, INSERM U1194, Univ Montpellier, Institut régional du Cancer de Montpellier, Montpellier, France, [4] Equipe Labélisée Ligue contre le cancer, Paris, France, [5] LPHI, University of Montpellier and CNRS UMR 5235, Montpellier, France

OPEN ACCESS

Edited by:
Dongya Jia,
Center for Cancer Research, National Cancer Institute (NIH), United States

Reviewed by:
Bishal Paudel,
University of Virginia, United States
Chunmei Liu,
Institute for Systems Biology (ISB), United States

***Correspondence:**
Ovidiu Radulescu
ovidiu.radulescu@umontpellier.fr

Specialty section:
This article was submitted to Cancer Metabolism, a section of the journal Frontiers in Oncology

Received: *18 January 2022*
Accepted: *07 March 2022*
Published: *14 April 2022*

Citation:
Hodgkinson A, Trucu D, Lacroix M, Le Cam L and Radulescu O (2022) Computational Model of Heterogeneity in Melanoma: Designing Therapies and Predicting Outcomes. Front. Oncol. 12:857572. doi: 10.3389/fonc.2022.857572

Cutaneous melanoma is a highly invasive tumor and, despite the development of recent therapies, most patients with advanced metastatic melanoma have a poor clinical outcome. The most frequent mutations in melanoma affect the BRAF oncogene, a protein kinase of the MAPK signaling pathway. Therapies targeting both BRAF and MEK are effective for only 50% of patients and, almost systematically, generate drug resistance. Genetic and non-genetic mechanisms associated with the strong heterogeneity and plasticity of melanoma cells have been suggested to favor drug resistance but are still poorly understood. Recently, we have introduced a novel mathematical formalism allowing the representation of the relation between tumor heterogeneity and drug resistance and proposed several models for the development of resistance of melanoma treated with BRAF/MEK inhibitors. In this paper, we further investigate this relationship by using a new computational model that copes with multiple cell states identified by single cell mRNA sequencing data in melanoma treated with BRAF/MEK inhibitors. We use this model to predict the outcome of different therapeutic strategies. The reference therapy, referred to as "continuous" consists in applying one or several drugs without disruption. In "combination therapy", several drugs are used sequentially. In "adaptive therapy" drug application is interrupted when the tumor size is below a lower threshold and resumed when the size goes over an upper threshold. We show that, counter-intuitively, the optimal protocol in combination therapy of BRAF/MEK inhibitors with a hypothetical drug targeting cell states that develop later during the tumor response to kinase inhibitors, is to treat first with this hypothetical drug. Also, even though there is little difference in the timing of emergence of the resistance between continuous and adaptive therapies, the spatial distribution of the different melanoma subpopulations is more zonated in the case of adaptive therapy.

Keywords: cancer heterogeneity, melanoma, targeted treatment, single cell data, mathematical modeling

INTRODUCTION

More than one half of melanomas carry mutations of the gene coding the BRAF kinase, a key upstream component of the MAPK signaling pathway, which is involved in cell growth and proliferation. In this pathway, BRAF phosphorylates and activates MEK that in turn phosphorylates and activates ERK, a potent effector that induces the transcription of many important genes that play a dominant role in survival and development of tumor cells. In melanoma, targeted therapies based on BRAF inhibitors (vemurafenib, dabrafenib, encorafenib) and MEK inhibitors (trametinib, cobimetinib, binimetinib) aim at reducing the activity of this key signaling cascade (1–4). BRAF inhibitors act differentially on cancer and healthy cells. Indeed, elevated MEK and ERK activity is induced mainly by BRAF dimers, and less by monomers. In BRAF-mutated melanomas, RAS-GTP levels are insufficient to promote BRAF dimerization, therefore the inhibition of BRAF monomers is sufficient for ERK inactivation. This specificity reduces the toxicity of this type of treatment (5). Although the treatment based on these kinase inhibitors initially leads to efficient tumor regression, resistance appears almost systematically. Several mechanisms have been associated to acquired resistance, such as RAS mutation, receptor tyrosine kinase activation that either compromise ERK inactivation or induce other survival pathways such as PI3K/AKT (5).

We focus here on a non-exclusive, but different cause of resistance, that involves the development of several drug tolerant cell states by non-genetic mechanisms. The non-genetic nature of adaptive resistance in melanoma was first suggested by the reversibility of this process: resistant tumors can re-sensitize upon a drug holiday (6, 7). Coexistence of sensitive and resistant cells with anti-correlated fitness in treated and untreated conditions can also explain apparent tumor re-sensitization in the absence of drug by positive selection of sensitive cells and negative selection of resistant cells, without the need for transitions between different cell states (8). Moreover, single cell RNA analysis has demonstrated plastic transitions between distinct cellular phenotypes in cell lines (9–11) and in patient-derived xenograft (PDX) mouse models (12, 13) submitted to BRAF/MEK inhibitors. Treatment-induced transitions between cell states have robust features, common to many patient-derived cultures and different cell lines (11). Between the melanocytic and mesenchymal-like states which represent the sensitive and resistant extremes there are intermediate states resembling nutrient-starved cells and evolving *via* several trajectories towards mensenchymal-like states. The intermediate states and the trajectories originating therein show intrinsic variability of the gene expression, which suggests that the transitions between states are continuous rather than discrete (9, 11, 12).

These fundamental findings could be used to design new therapeutic strategies to avoid resistance. The re-sensitization, either real or apparent, arising when resistant cells are slowly growing in untreated conditions, suggest that a discontinuous adaptive treatment, alternating 'on' and 'off' drug periods, may be able to control tumor size, at least for some time. Combination therapy may also depend on one's capacity to predict the changes induced by the primary tumor treatment, in space and time. For instance, drug tolerant neural crest stem cells, which are enriched upon treatment with BRAF/MEK inhibitors, display an RXR-driven signature, suggesting that these cells could eventually be targeted pharmacologically by using RXR-inhibitors (12). Besides anti-BRAF/MEK targeted therapies, the recent discovery that immune checkpoint inhibitors, targeting regulatory molecules on T lymphocytes (anti-CTLA4, anti-PD-1, and anti-PD-L1), are highly efficient in melanoma patients has revolutionized the treatment of metastatic melanoma. However, each treatment modality has limitations. While treatment with targeted therapies is associated with a strong beneficial short-term response but is followed by systematic resistance, treatment with immune checkpoint inhibitors has a lower response rate but associates with better long-term responses on a subset of melanoma patients. Thus, despite these considerable improvements in melanoma treatment, the development of new clinical strategies remains necessary and a better understanding of melanoma biology is likely to provide additional therapeutic options to patients with resistant cancers (14, 15).

In this paper we use a computational framework to study the heterogeneity of melanoma and develop a predictive model for various therapeutic outcomes. We base our model on data obtained in MEL06 patient-derived melanoma cells, which were demonstrated to develop non-genetic resistance to BRAF/MEK inhibitors (12). Given the complexity of the resistance mechanisms in melanoma, our conclusions may not hold true for all melanomas, which may evolve during treatments through multiple mechanisms of resistance.

RESULTS

Multidimensional, Data Driven Model of Heterogeneity

Our main assumption is that under treatment melanoma cells undergo a series of non-genetic transitions, leading to drug tolerant and resistant cell states. Contrary to more traditional models of heterogeneity that consider a finite number of discrete cell states (16), our model can cope with a continuous spectrum of states. In this model, cell populations are represented as distributions (heatmaps) over many dimensions; spatial, coping with cell motility and cell interactions with extracellular matrix, diffusive drug and signaling molecules, but also structural, representing internal cell-state variables such as gene expression, signaling, and metabolic activities (**Figures 1A, B**). An interesting possibility is to use single cell data and feature extraction methods such as t-distributed stochastic neighbor embedding (t-SNE) in order to define reduced structural dimensions (**Figures 1C, D**). In this case, the distributions (heatmaps) predicted by the model (**Figure 1B**) can be directly compared to the empirical single cell distributions. We call this approach 'mesoscopic' as it is intermediate between a microscopic approach, which simulates each cell individually, and a macroscopic one, in which the cell-to-cell variability is averaged out. Even though this method can be applied to any

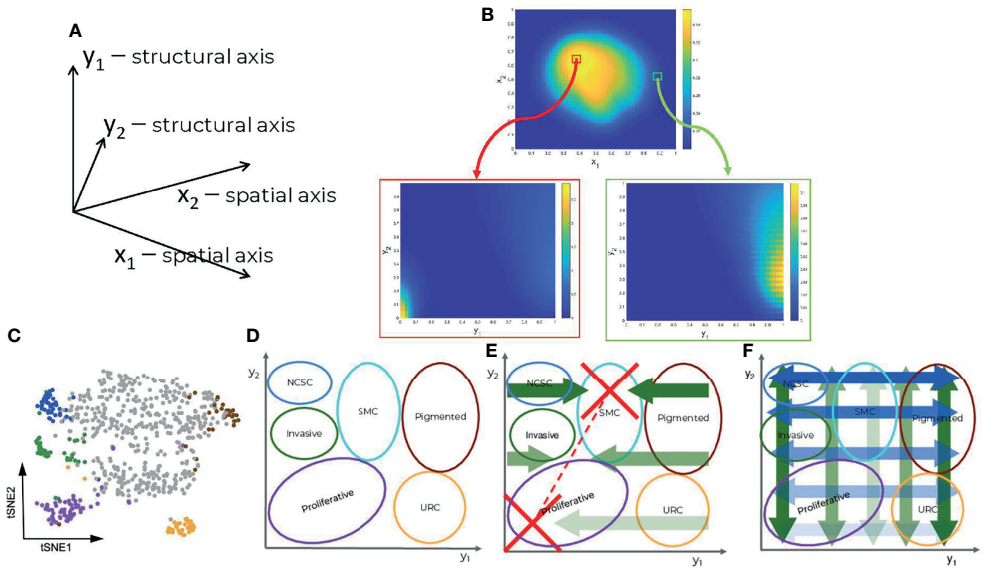


FIGURE 1 | Components of the data driven heterogeneity model: **(A)** Dimensions of the model. **(B)** Multidimensional cell distributions predicted by the model emphasizing treatment induced zoning. **(C)** Reduced representation of single cell expression data, from (12). **(D)** Cell states represented as domains in structural/gene expression space. **(E)** Directed (advective) structural fluxes. **(F)** Undirected (diffusive) structural fluxes.

type of single cell data (transcriptomic, proteomic or metabolomic), our analysis is based on the single cell mRNA sequencing data from (12).

In order to clarify the structural dimensions of our tumor, we distinguished six different melanoma subpopulations; namely proliferative, invasive, pigmented cells, neural-crest stem cells (NCSCs), starved-like melanoma cells (SMCs), and uncharacterized resistant cells (URCs). These are schematically represented in the structural plane in **Figure 1D** and are in line with previous experimental results (12). It should be recalled that individual cells may concurrently occupy several states, existing in a continuum of gene expression profiles across the structural domain.

The model predicts dynamic heterogeneity, meaning that the multidimensional distributions depend also on time. The evolution of these distributions is driven by spatial fluxes, involving undirected (diffusive) and directed (advective) spatial cell motility, and by structural fluxes, corresponding to changes of the cell state. The undirected spatial diffusion fluxes describe a cellular spatial random walk process, whereas directed spatial fluxes describe controlled cell migration mediated by adhesive extracellular matrix substrates or to sites of more elevated nutritional content (see *Methods* and **Supplementary Methods**). The undirected structural fluxes (structural diffusion) correspond to random changes of the cell state leading to the spread of the cell distributions (increased heterogeneity) without changes of modal positions in the structural dimensions. The directed structural fluxes (structural advection) correspond to deterministic changes of the cell state, leading to shifts of the distribution modes. The cell distribution dynamics, represented as one 4D (2 spatial and 2 structural dimensions) partial differential equation (PDE), is coupled to five other 2D (2 spatial dimensions) PDEs coping with the spatial distribution of other variables such as extracellular nutritional environment (ECNE), chemo-attractant molecules (surrogate for mediated cell-cell communications), and drug concentrations. The effect of the drugs on the cells' distribution is taken into account in the negative (degradative) source terms that depend on their position within the structural domain, i.e., on the cell state. For details, see *Methods*.

Targeted Treatment Exacerbates Heterogeneity

The model recapitulates the dynamics of the cell heterogeneity observed in (12) (see **Movies S1**, **S2**). Starting with a naïve tumor containing a population of sensitive melanocytes, several cell subpopulations are induced by the therapy. In our simulations, this is seen by the multimodality of the cell population's structural distribution, with positions of the modes depending on time. As shown in **Figure 1D**, each sub-population is characterized by the position of the mode and by its spread in the structural domain. For a more quantitative approach, we use the variance in the cell structural distribution as a metric and show that heterogeneity increases with time upon drug administration (**Figure S1**).

The model predicts the typical three phase tumor growth curve under kinase inhibitors; a first phase wherein the tumor responds and shrinks, a second phase wherein the tumor is no longer visible corresponding to the minimal residual disease (MRD), and a third phase during which tumor growth resumes after the emergence of resistance. During the MRD phase, heterogeneity strongly increases through continuous spreading of the cell distributions in the structural dimensions

and, moreover, by development of co-existing, drug-tolerant, intermediate states between sensitivity and resistance (multi-modality, see **Movie S2**).

Order in Combination Therapy Matters

We have tested, in our computational setting, combination therapies by successively applying two differing types of treatments: (1) using BRAF/MEK inhibitors (BRAF/MEKi) as in (12), and (2) a hypothetical cancer treatment (HCT). We have considered that the tumor has the same intrinsic dynamics, defined by the same diffusion and advection terms, for the two treatments. In particular, the cell states and their transitions will be the same for the two treatments. However, the two treatments eliminate cells differently, depending on their states. This difference between treatments was modeled by using a drug response function, defining how the drug dependent cell degradation changes with the cell state. This function peaks in the modal position of the primary tumor, in the case of BRAF/MEKi, or in the positions of the BRAF/MEKi resistant states, typically invasive and URC cell populations, in the case of HCT (see *Methods* and **Figure S2**). Applied alone, the BRAF/MEKi treatment induces immediate and drastic tumor reduction, followed by MRD and development of resistance after approximately four months of continuous administration of the drugs. The HCT treatment leads to a mild response initially, but like BRAF/MEKi treatment, induces tumor adaptation. However, the representation of invasive and URC cell states is only moderate because they are now more effectively eliminated.

Treatments using BRAF/MEKi (**Figure 2A** and **Movie S3**) or HCT (**Figure 2B** and **Movie S4**), alone, resulted in a re-establishment of initial tumor volume, prior to the end of the study period, with HCT inducing resistance far earlier than BRAF/MEKi. For the combination therapy, BRAF/MEKi then HCT, we observe a later time-point for the re-establishment of the initial tumor volume, in comparison to BRAF/MEKi only, but still resulted in a significant increased tumor growth rate (**Figure 2C** and **Movie S5**). Starting first with HCT and then using BRAF/MEKi, however, was a better strategy that significantly delayed resistance and also reduced the tumor load by combining the advantages of the two treatments (**Figure 2D** and **Movie S6**).

Output in Terms of Heterogeneity Depends on the Therapeutic Strategy

The dynamics of melanoma cells submitted to kinase inhibitors is typically robust. In the case of adaptive therapy, although the intermediate dynamics is modified by allowing the tumor to grow before re-applying treatment, our model predicted that resistance development cannot be avoided (**Movies S7**, **S8**). However, in terms of spatial heterogeneity, the outcome is much more variable. In **Figure 3** we have represented the spatial distribution of different cell states at the end of MRD and beginning of resistance, for various treatments. In all cases cell states depend on position, a phenomenon called zoning. The details of this phenomena depend on the type of therapy. Our model predicted that the adaptive therapy generates more pronounced zoning, with steeper and mutually exclusive patterns (**Figure 3**) than those predicted under continuous therapies.

DISCUSSION AND CONCLUSION

Treatment by kinase inhibitors leads to a heterogeneity upsurge in melanoma. At least part of this heterogeneity is generated by non-genetic mechanisms and involves continuous modifications of gene expression programs which lead to transitions between cell states. Our mathematical model captures the essential features of non-genetic transitions and explains the heterogeneous dynamics by diffusive and advective spatial and structural fluxes. The increased heterogeneity results from the multiplicity of drug tolerant and resistant states induced by the treatment, and from cooperative strategies in a spatially heterogeneous tumor where resistant cells protect sensitive cells from elevated drug exposure.

Moreover, our model predicts *in silico* the outcomes of various therapies.

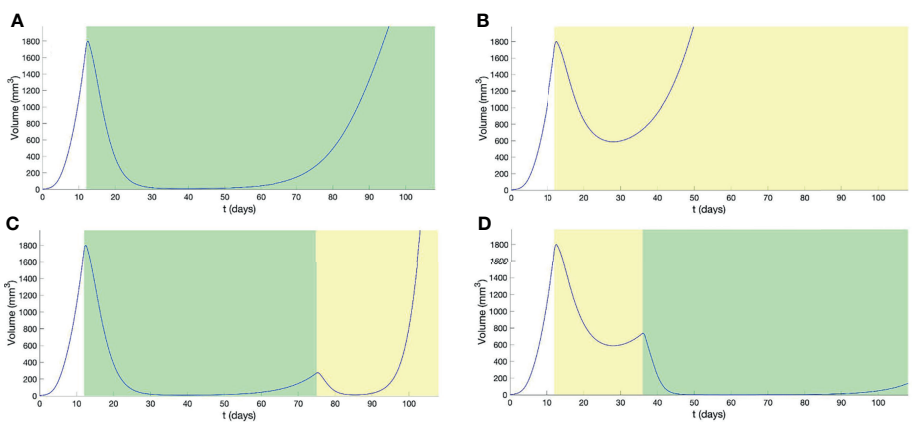


FIGURE 2 | Outcomes from combination therapy, with treatment intervals indicated by graphical shading for BRAF/MEKi (*green*) and HCT (*yellow*). Panels show the outcomes from **(A)** continuous BRAF/MEKi, **(B)** continuous HCT, **(C)** combination BRAF/MEKi then HCT, and **(D)** combination HCT then BRAF/MEKi treatment regimes.

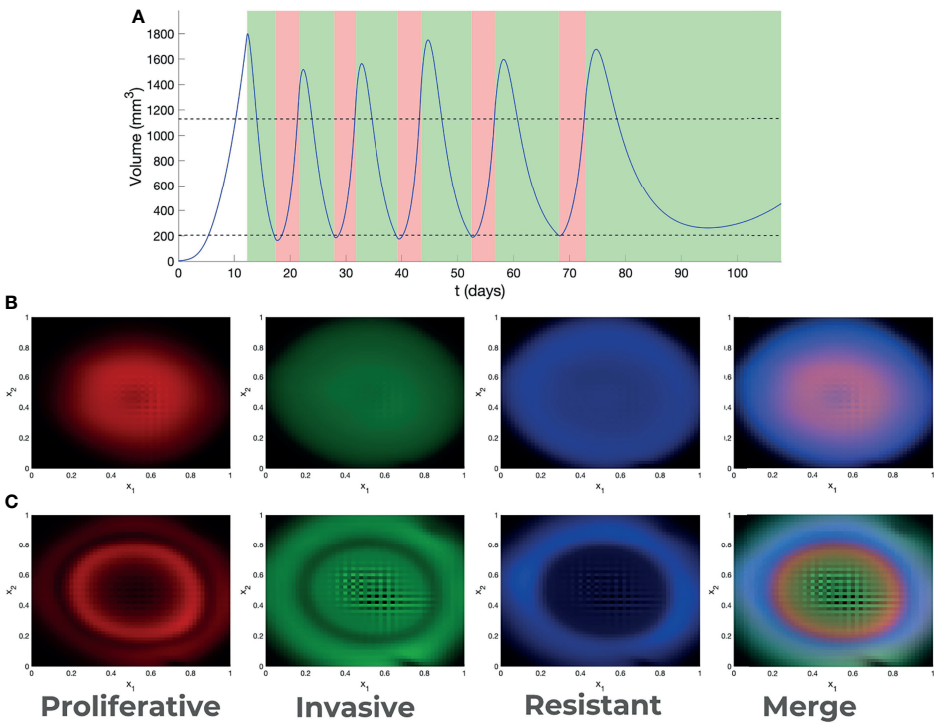


FIGURE 3 | Adaptive therapy. **(A)** Outcome of adaptive therapy, with BRAF/MEKi treatment intervals indicated in *green* and drug hollidays in *pink*. Decision about treatment is taken every day. Treatment is applied if the volume is higher than the upper threshold, stopped if the volume is lower than the lower threshold (thresholds are indicated as dotted lines). Spatial heterogeneity (zoning) generated by continuous **(B)** and adaptive **(C)** therapy.

Here, we show that, considering combination therapies, it is better to treat first with a less effective hypothetical drug, targeting sub-populations that develop during tumor resistance phases, before treating with BRAF/MEKi. The explanation of this rather counter-intuitive result can be found in the cell population dynamics. We supposed that the intermediate response of the tumor to any of the treatments results in increased heterogeneity by non-genetic processes. If the first applied treatment is BRAF/MEKi, this acts mainly on cells belonging to early modes and there is a non-negligible probability that some cells escape treatment and become resistant. The same probability is very small when the first applied treatment is HCT that acts preferentially on cells belonging to late modes; the role of HCT initial application is to avoid starting the BRAF/MEKi treatment with some cells that are not sensitive. Then, the application of BRAF/MEKi kills practically all the remaining cells and resistance takes much longer time to develop (see **Movies S5**, **S6**). One should note that, due to structural diffusion, any cell state can, in theory, give rise to all other cell states. Therefore, in order to confine cells to BRAF/MEKi-sensitive modes, the drug has to act on a large domain of cell states, not only on a single intermediate drug tolerant sub-population. This is difficult to perform using targeted therapies.

A possible candidate for the hypothetical cancer treatment (HCT) is the drug family of immune checkpoint inhibitors (ICIs). Although this treatment does not act directly on melanoma cells, it can have a differential indirect effect on melanoma sub-populations, and acts more generally than targeted treatments. Very recent Phase III trials combining kinase inhibitors and ICIs show that starting a first line treatment with ICIs leads to better results in terms of survival time and duration of response than starting with kinase inhibitors (17). This is explained if checkpoint inhibitors induce effective prior elimination of resistant stage sub-populations. There are, however, other interpretations of the interplay between kinase inhibitors and immunotherapy. Obenauf et al. (18) showed that kinase inhibitors induce changes in the stroma and cell secretome and hypothesized changes of immune cells infiltration. Other authors suggested that treatment with BRAFi leads to favorable changes in the tumor microenvironment in synergy with immune checkpoint inhibitors (see (14) for a review). The interactions between the immune cells and the various melanoma sub-populations are still poorly defined. We hope that future experimental and modelling work in the field, will elucidate the mechanistic aspects of these interactions.

Simple models of adaptive therapy were based on the idea that, in the absence of drugs, resistant cells grow more slowly than sensitive cells (10). It is believed that this fitness advantage allows sensitive cells to recover at least partially during a drug holiday. Although this effect is present in our model, it is compensated by structural and spatial diffusion that led to increased heterogeneity and delay only moderately the time to resistance. The resulting tumor, however, depends on the type, continuous or adaptive, of treatment. Irrespective of the treatment, zoning is a population-scale

strategy to increase the mean fitness by cooperative protection of sensitive cells by resistant cells. In the adaptive treatment, the growth of sensitive cells is also favored by drug holidays, which lead to a more pronounced zoning.

From a theoretical perspective, our model shows the interplay between directed and undirected structural fluxes for the development of plasticity and heterogeneity. Undirected fluxes correspond to diffusion and random changes of cell states. As well known in physics, or in neutral theory of molecular evolution, free diffusion can reach any state from any other state if one waits a time proportional to the square of the state change. In the presence of treatment, diffusion is not free and has to cross barriers generated by the drugs action. In this case, the escape transition time is exponential. The escape transition time and the proportion of escaping cells depend on the position, height, and width of the barriers, which are different for different treatments. This dependence further explains why order matters in combination therapy and why heterogeneity may differ when employing adaptive strategies, since the barrier is time-transient. Another important theoretical aspect is the symmetry breaking induced by the treatment. Although a barrier can be crossed in both directions, the transition probability is asymmetric if one of the states is more stable than the others. This leads to the notion of metastable states hierarchy, in which states are distinguished by the time that cells spend in each one of these; this time can be very long for highly stable states. Adaptive therapies favor the stabilization of one metastable state by alternating treatment and holiday periods. The success of this strategy depends on conditions that may be difficult to guarantee, especially in a multidimensional context and for a spatially heterogeneous drug distribution.

We should nevertheless emphasize that our model is mostly phenomenological with structural dimensions representing nonlinear functions of the gene expression data. As several findings point towards the role of BRAFi in metabolic remodeling (19), it would be very useful to interpret the structure variables in terms of metabolic changes. This is possible within our formalism as metabolic ODE models [see for instance (20)] are transposable into structural advection fluxes, where metabolic stochasticity or uncertainty would translate to diffusive fluxes. This possibility will be investigated in future work. Furthermore, the distribution of blood vessels that are sources of nutrition and drug compounds is an important variable for understanding zoning aspects of cancer adaptation to treatment [see also (21)]. Like in (21), we expect that the spatial distribution of sensitive and resistant cells depends on the distance to these sources. Blood vessels distribution can be reconstructed from *ex vivo* tumor sections (22) and we will use these distributions to increase the realism of future models.

METHODS

General Formalism

Mesoscale models of cancer heterogeneity are based on partial differential equations and can be generically obtained from the Liouville continuity equation (8, 23, 24). Let us consider that there are n types of cells. In this model cells are distinguished by two types of variables, a discrete one representing the type $i \in \{1,\ldots, n\}$ and a continuous one $\boldsymbol{y} = (y_1, \ldots, y_m)$ representing the internal state (vector of concentrations of biochemical species, for instance). Then $\boldsymbol{c} = (c_1,\ldots, c_n)$ represents a vector of cell distributions satisfying the equation

$$\frac{\partial \boldsymbol{c}(\boldsymbol{x},\boldsymbol{y},t)}{\partial t} = -\nabla_{\boldsymbol{x}} \cdot \boldsymbol{F}_{\boldsymbol{x}}(\boldsymbol{x},\boldsymbol{y},t) - \nabla_{\boldsymbol{y}} \cdot \boldsymbol{F}_{\boldsymbol{y}}(\boldsymbol{x},\boldsymbol{y},t) + \boldsymbol{S}(\boldsymbol{x},\boldsymbol{y},t), \quad (1)$$

where $\boldsymbol{x}$ is the spatial position, $\boldsymbol{y}$ is the cell's internal state (structure variable), $\boldsymbol{F_x}$ is the spatial flux, $\boldsymbol{F_y}$ is the structural flux and $\boldsymbol{S}$ is the source term. If the cell's internal state follows ODEs $\frac{dy}{dt} = \boldsymbol{\Phi}(y)$, then the structural flux is advective $\boldsymbol{F_y} = \boldsymbol{c\Phi}$. If the cell's state follows random Brownian motion in the structural space, then the structural flux is diffusive resulting from Fick's law $\boldsymbol{F_y} = -\boldsymbol{D}_y\, \nabla_{\boldsymbol{y}}\boldsymbol{c}$, where $\boldsymbol{D_y}$ is the structural diffusion coefficient matrix. The spatial flux contains terms related to cell motility: undirected (diffusion), or directed (chemotaxis, haptotaxis) (25). The source term can integrate cell proliferation, death, and discrete stochastic changes (finite jumps) of the cell state $\boldsymbol{y}$, other than those included in the continuous flux $\boldsymbol{F_y}$.

Model Derived From Single Cell Expression Data

In this section we present only the broad lines of the model construction. The details can be found in the **Supplementary Methods**.

Model Components

Our melanoma progression model has two main components: the cancer cell population density $c(t, x, y)$ and the extra-cellular nutritional environment (ECNE) density $v(t, x)$. In our minimal melanoma model, space positions x, and structural positions y are two dimensional (in space we consider a 2D tumor section, and in structure we use a 2D t-SNE representation of the tumor transcriptome). We also consider spatial gradients of three types of diffusible molecules, namely 1) the nutritional molecular species, provided by the ECNE and consumed by cells, 2) the acidic molecular species, produced by cells and degrading the ECNE as in (26), and 3) drugs.

The fluxes defining the model dynamics have been derived using the following assumptions:

Spatial Variables and Fluxes

We assume, consistently with previous mathematical studies of spatial cancer dynamics (25), that the spatial dynamics of melanoma cells are governed by both random (diffusive) and deterministic (advective) components. The random (diffusive) component is assumed to occur as a result of tissue-scale reorientation and volume-filling processes. The deterministic (advective) component is assumed to result from directed cell motility and is driven by cell-environment interactions. In particular, we assume that cells exhibit controlled migration to sites of chemically elevated nutritional content (chemotaxis), as well as to sites of higher ECNE density (haptotaxis).

Structural Variables and Fluxes

The definition of the structural variables follows from the data analysis in (12). Unsupervised clustering of single cell mRNA-seq

data identified several types of cell sub-populations with distinct transcription states. The high-dimensional transcriptome was compressed to a 2D map using t-distributed stochastic neighbor embedding (t-SNE). The support of this 2D map is our structure space domain. The different transcription states represent sub-domains in this representation (see **Figure 1D**). The cell-state transitions observed experimentally can be represented as diffusion and advective flow in this domain. The flow changes the positions of the cells in the 2D structure domain, moving them from one state to another. Thus, rather than considering a number of distinct cell types, we have built a model with only one cell type whose state can change continuously by the structural fluxes. A cell is added to a sub-population if its state enters the corresponding structural sub-domain and is subtracted if it dies or if it leaves the sub-domain. In order to define the structural fluxes, we start by identifying the sub-domains corresponding to different sub-populations inside the tumor at different times. Although seven transcriptional signatures were identified [**Table S1** of (12)], we focus upon the description of six primary states important for resistance. For their localization in the structural domain, we use cardinal points, as follows:

i. Melanoma cells with a "proliferative" signature are predominant in naive tumors, localized south-west (SW).
ii. Invasive cells are also present in naive tumors. They are localized east (E).
iii. Pigmented cells expressing markers of differentiation are induced by the treatment. They are localized north-west (NW).
iv. Neural crest stem cells (NCSC) are enriched by the treatment, have a maximum during the minimal residual disease and are diluted out during the development of resistance. They are localized north-east (NE).
v. Starved-like melanoma cells (SMC) are rapidly induced by the treatment, and become predominant during MRD. They are localized north (N).
vi. Uncharacterized resistant cells (URCs) were not thoroughly biologically investigated, though the model predicts they may have a biological interest. They are localized south-west (SW).

The structural fluxes describe the metabolic adaptation within the structural domain and diffusion-like exchanges between cell populations (**Figures 1E, F**). In order to define these fluxes, we use the following dynamical assumptions:

- Horizontal advection is assumed to stabilize the proliferative (SW) state, since there is no known emergence of URCs prior to resistance;
- SMC states (N) are also stabilized by horizontal advection fluxes that converge towards this state, allowing cells to populate this minimally mitotic state;
- advection is assumed to interpolate linearly at intermediate phenotypes between proliferative cells and SMCs;
- horizontal diffusion is assumed to be maximal in the northern regions of the structural plane and decreases in southern regions, illustrating rare, stochastic transitions between proliferative and URC states;
- vertical diffusion is maximal towards the western and eastern regions, allowing transitions between proliferative, invasive, and NCSC or pigmented and URC populations, but lower transition rates between SMC and southern states.

In principle, by diffusion any cell state can give rise to all cell states. However, advection maintains a certain degree of cellular hierarchy. These assumptions have been made upon a reasoned analysis of the figures presented in (12), as a minimal set of functional assumptions to reproduce observed patterns, but do not necessarily represent an optimal or biologically motivated set of assumptions.

Source Terms and Degradation

The source and degradation terms describe cell proliferation and death, respectively. Like in (12) we consider that proliferation is significantly reduced among SMC cells and increased among proliferative cells. We consider that treatment is the only cause of active cell death. Due to the nature of our modelling framework, drugs may target cells with a spectrum of specific expression markers as would be the case in the clinical scenario. In this case, we assume the existence of two particular treatments. Firstly, BRAF and MEK inhibitors were employed within the study conducted by (12) and, as such, are assumed to primarily target a distribution centered around the proliferative population, stretching into the invasive population but with diminished success among cells in the NW of the structural domain (**Figure S2**). Secondly, a hypothetical cancer therapeutic (HCT) has been used for the sake of illustration and targets primarily the invasive and URC cell populations, with an expansive effectiveness span E and SW (**Figure S2**).

Spatial Dynamics of Other Components

Given the complexity of the dynamics in the primary cancer cell population, the dynamics of other components have been kept as simple as possible. It is assumed that the ECNE exhibits only a natural restorative growth process, as well as acidic species-induced and natural degradation kinetics. Nutritional and acidic species exhibit diffusion, as well as controlled production, and natural degradation. Finally, the drug species also exhibit diffusion, time dependent administration, as well as natural and cell-based degradation.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The data can be found here: https://github.com/oradules/Melanoma2D_2021/.

AUTHOR CONTRIBUTIONS

AH and OR conceived the project. OR wrote the paper with help from AH, ML and LC. AH designed the model, numerical scheme, and performed the simulations. DT contributed to the

general formalism and numerical scheme. All authors contributed to the article and approved the submitted version.

FUNDING

We acknowledge financial support from Itmo Cancer on funds administered by INSERM (project MALMO) and from I-SITE MUSE (project MEL-ECO). AH has been funded by the Wellcome Trust Institutional Strategic Support Fund. ML and LLC are financially supported by the Ligue Nationale Contre le Cancer, the Institut National du Cancer (INCa) and the cancéropôle Grand Sud Ouest.

ACKNOWLEDGMENTS

The authors would like to thank Jean-Christophe Marine for very helpful discussions.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.857572/full#supplementary-material

REFERENCES

1. Gross S, Rahal R, Stransky N, Lengauer C, Hoeflich KP. Targeting Cancer With Kinase Inhibitors. *J Clin Invest* (2015) 125:1780–9. doi: 10.1172/JCI76094
2. Zhang C, Spevak W, Zhang Y, Burton EA, Ma Y, Habets G, et al. RAF Inhibitors That Evade Paradoxical MAPK Pathway Activation. *Nature* (2015) 526:583–6. doi: 10.1038/nature14982
3. McClure E, Patel A, Carr MJ, Sun J, Zager JS. The Combination of Encorafenib and Binimetinib for the Treatment of Patients With BRAF-Mutated Advanced, Unresectable, or Metastatic Melanoma: An Update. *Expert Rev Precis Med Drug Dev* (2021) 6:19–29. doi: 10.1080/23808993.2021.1847639
4. Zhang C, Bollag G. Triple Therapy to Outwit the Braf Oncogene. *Cancer Discovery* (2021) 11:1620–2. doi: 10.1158/2159-8290.CD-21-0378
5. Poulikakos PI, Rosen N. Mutant BRAF Melanomas–Dependence and Resistance. *Cancer Cell* (2011) 19:11–5. doi: 10.1016/j.ccr.2011.01.008
6. Das Thakur M, Salangsang F, Landman AS, Sellers WR, Pryer NK, Levesque MP, et al. Modelling Vemurafenib Resistance in Melanoma Reveals a Strategy to Forestall Drug Resistance. *Nature* (2013) 494:251–5. doi: 10.1038/nature11814
7. Sun C, Wang L, Huang S, Heynen GJ, Prahallad A, Robert C, et al. Reversible and Adaptive Resistance to Braf (V600e) Inhibition in Melanoma. *Nature* (2014) 508:118–22. doi: 10.1038/nature13121
8. Hodgkinson A, Le Cam L, Trucu D, Radulescu O. Spatio-Genetic and Phenotypic Modelling Elucidates Resistance and Re-Sensitisation to Treatment in Heterogeneous Melanoma. *J Theor Biol* (2019) 466:84–105. doi: 10.1016/j.jtbi.2018.11.037
9. Fallahi-Sichani M, Becker V, Izar B, Baker GJ, Lin JR, Boswell SA, et al. Adaptive Resistance of Melanoma Cells to RAF Inhibition *via* Reversible Induction of a Slowly Dividing De-Differentiated State. *Mol Syst Biol* (2017) 13:905. doi: 10.15252/msb.20166796
10. Smalley I, Kim E, Li J, Spence P, Wyatt CJ, Eroglu Z, et al. Leveraging Transcriptional Dynamics to Improve BRAF Inhibitor Responses in Melanoma. *EBioMedicine* (2019) 48:178–90. doi: 10.1016/j.ebiom.2019.09.023
11. Wouters J, Kalender-Atak Z, Minnoye L, Spanier KI, De Waegeneer M, González-Blas CB, et al. Robust Gene Expression Programs Underlie Recurrent Cell States and Phenotype Switching in Melanoma. *Nat Cell Biol* (2020) 22:986–98. doi: 10.1038/s41556-020-0547-3
12. Rambow F, Rogiers A, Marin-Bejar O, Aibar S, Femel J, Dewaele M, et al. Toward Minimal Residual Disease-Directed Therapy in Melanoma. *Cell* (2018) 174:843–55. doi: 10.1016/j.cell.2018.06.025
13. Marine JC, Dawson SJ, Dawson MA. Non-Genetic Mechanisms of Therapeutic Resistance in Cancer. *Nat Rev Cancer* (2020) 20:743–56. doi: 10.1038/s41568-020-00302-4
14. Naderi-Azad S, Sullivan R. The Potential of BRAF-Targeted Therapy Combined With Immunotherapy in Melanoma. *Expert Rev Anticancer Ther* (2020) 20:131–6. doi: 10.1080/14737140.2020.1724097
15. Ma VT, Daignault-Newton S, Waninger JJ, Journey S, Chopra Z, Tezel A, et al. The Impact of BRAF Mutation Status on Clinical Outcomes With Anti-PD-1 Monotherapy Versus Combination Ipilimumab/Nivolumab in Treatment-Naïve Advanced Stage Melanoma. *Pigment Cell Melanoma Res* (2021) 34:629–40. doi: 10.1111/pcmr.12944
16. Delitala M, Lorenzi T. A Mathematical Model for Progression and Heterogeneity in Colorectal Cancer Dynamics. *Theor Population Biol* (2011) 79:130–8. doi: 10.1016/j.tpb.2011.01.001
17. Atkins MB, Lee SJ, Chmielowski B, Ribas A, Tarhini AA, Truong TG, et al. Dreamseq (Doublet, Randomized Evaluation in Advanced Melanoma Sequencing): A Phase III Trial–ECOG-ACRIN Ea6134. *J Clin Onco* (2021)39(36_suppl):356154. doi: 10.1200/JCO.2021.39.36_suppl.356154
18. Obenauf AC, Zou Y, Ji AL, Vanharanta S, Shu W, Shi H, et al. Therapy-Induced Tumor Secretomes Promote Resistance and Tumor Progression. *Nature* (2015) 520:368–72. doi: 10.1038/nature14336
19. Corazao-Rozas P, Guerreschi P, Jendoubi M, André F, Jonneaux A, Scalbert C, et al. Mitochondrial Oxidative Stress Is the Achille's Heel of Melanoma Cells Resistant to BRAF-Mutant Inhibitor. *Oncotarget* (2013) 4:1986. doi: 10.18632/oncotarget.1420
20. Jia D, Lu M, Jung KH, Park JH, Yu L, Onuchic JN, et al. Elucidating Cancer Metabolic Plasticity by Coupling Gene Regulation With Metabolic Pathways. *Proc Natl Acad Sci* (2019) 116:3909–18. doi: 10.1073/pnas.1816391116
21. Kumar S, Sharife H, Kreisel T, Mogilevsky M, B+ar-Lev L, Grunewald M, et al. Intra-Tumoral Metabolic Zonation and Resultant Phenotypic Diversification Are Dictated by Blood Vessel Proximity. *Cell Metab* (2019) 30:201–11. doi: 10.1016/j.cmet.2019.04.003
22. Kiemen A, Braxton AM, Grahn MP, Han KS, Babu JM, Reichel R, et al. *In Situ* Characterization of the 3d Microanatomy of the Pancreas and Pancreatic Cancer at Single Cell Resolution. *bioRxiv* (2020). doi: 10.1101/2020.12.08.416909
23. Hodgkinson A, Chaplain MAJ, Domschke P, Trucu D. Computational Approaches and Analysis for a Spatio-Structural-Temporal Invasive Carcinoma Model. *Bull Math Biol* (2018) 80:701–37. doi: 10.1007/s11538-018-0396-4
24. Hodgkinson A, Radulescu O, Uzé G, Trucu D. Signal Propagation in Sensing and Reciprocating Cellular Systems With Spatial and Structural Heterogeneity. *Bull Math Biol* (2018) 80:1900–36. doi: 10.1007/s11538-018-0439-x
25. Chaplain MAJ, Lolas G. Mathematical Modelling of Cancer Invasion of Tissue: Dynamic Heterogeneity. *Networks Heterog Media* (2006) 1:399–439. doi: 10.3934/nhm.2006.1.399
26. Gatenby RA, Gawlinski ET. A Reaction-Diffusion Model of Cancer Invasion. *Cancer Res* (1996) 56:5745–53.

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

this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Oncology

ORIGINAL RESEARCH
published: 03 May 2022
doi: 10.3389/fonc.2022.853768

Quantification of the Landscape for Revealing the Underlying Mechanism of Intestinal-Type Gastric Cancer

Chong Yu[1] *and Jin Wang*[2]*

[1] Department of Statistics, Jilin University of Finance and Economics, Changchun, Jilin, China, [2] Department of Chemistry and of Physics and Astronomy, State University of New York at Stony Brook, Stony Brook, NY, United States

Gastric cancer is a daunting disease with a tragic impact on global health. It is the fourth most common cancer and has become the second most frequent cause of cancer death in recent times. According to the Lauren classification, gastric cancer can be classified into two types: intestinal and diffuse. Intestinal-type gastric cancer (IGC) is more common in elderly people, and atrophic gastritis (AG) and intestinal metaplasia (IM) have been proven to be the main premalignant causes of intestinal-type gastric cancer. In turn, *Helicobacter pylori* infection has been identified as the most significant cause of AG and IM. In this study, we determine the mechanism of IGC progression and how *H. pylori* infection induces IGC. Through researching the relevant literature, we identified the key genes associated with gastric cancer and the specific genes associated with IGC. We then use hese genes to build up a gene regulatory network for IGC. Based on this gene regulatory network, we quantify the IGC landscape. Within this landscape, there are three stable states, which are classified as the normal, AG, and gastric cancer states. Through landscape topography, we can determine the biological features and progression process of IGC. To investigate the influence of *H. pylori* infection on IGC, we simulated different degrees of *H. pylori* infection. As the *H. pylori* infection becomes more serious, the landscape topography changes accordingly. A fourth state, named the intestinal metaplasia (IM) state, emerges on the landscape and is associated with a very high risk of developing gastric cancer. The emergence of this state is due to the interactions/regulations among genes. Through variations in the landscape topography, we can determine the influence of *H. pylori* infection on IGC. Finally, we use global sensitivity analysis to research the regulations most sensitive to IGC prevention or therapies. This study presents a new approach and a novel model with which to explore the mechanism of IGC. The simulations of different degrees of *H. pylori* infection can provide us with a systematic view of IGC progression. The key regulations found can give us some insight and guidance for clinical trials and experimental studies.

Keywords: gastric cancer, *Helicobacter pylori*, intestinal type, gene regulatory network, landscape, flux

OPEN ACCESS

Edited by:
Dongya Jia,
National Cancer Institute (NIH), United States

Reviewed by:
Mohit Kumar Jolly,
Indian Institute of Science (IISc), India
Wentao Dong,
Stanford University, United States

***Correspondence:**
Jin Wang
jin.wang.1@stonybrook.edu

Specialty section:
This article was submitted to Cancer Metabolism, a section of the journal Frontiers in Oncology

Received: *13 January 2022*
Accepted: *15 March 2022*
Published: *03 May 2022*

Citation:
Yu C and Wang J (2022) Quantification of the Landscape for Revealing the Underlying Mechanism of Intestinal-Type Gastric Cancer. Front. Oncol. 12:853768. doi: 10.3389/fonc.2022.853768

1 INTRODUCTION

Cancer has long been considered the most daunting disease, and gastric cancer is the second most aggressive cancer, having a tremendous, large-scale impact on global health. Despite a huge amount of research, gastric cancer remains the fourth most common cause of cancer-related deaths worldwide (1). Despite a decline in incidence in the last several decades, the prognosis for gastric cancer is still very poor. The five-year survival rates for gastric cancer are less than 20% (2). Early-stage gastric cancer has a better prognosis, with five-year survival rates of up to 95% (3). According to the Lauren classification, gastric cancer can be divided into two types: intestinal and diffuse. The intestinal type occurs more frequently, in about 54% of cases, and more commonly in men and elderly people (4, 5). Atrophic gastritis (AG) and intestinal metaplasia (IM) have been proven to be the main premalignant factors in the intestinal type of gastric cancer (2).

Intestinal-type gastric cancer (IGC) is caused mainly by environmental factors such as salty food, alcohol, and cigarette smoking. These factors may contribute to AG, which is considered one of the main precursor lesions of IGC (6). Moreover, *Helicobacter pylori* infection can increase the risk of IGC developed from AG. The stomach is the natural reservoir of *H. pylori*. Studies show that about 50% of the world's human population is chronically colonized by *H. pylori* and about 15% of infected people develop gastric cancer from AG and IM (7). *H. pylori* infection may cause epithelial damage, which can trigger a multistep progression to gastric cancer from AG, gastric atrophy, and IM (8, 9). These changes are mainly caused by epigenetic alterations (10). Epigenetic modifications such as DNA methylation and histone modifications can alter cell cycles. Aberrant DNA methylation can also induce IGC formation.

However, the oncogene gene c-met is related to the development of about 20% of IGC cases, and alterations in c-met have also been associated with many types of diseases, particularly diseases of the digestive system (11). IM, dysplasia, and invasive carcinomas are associated with K-Ras mutations (12). Abnormal expressions of tumor suppressor genes, such as TP53 and APC, are found in many IGC subjects (13). Therefore, the development of IGC is genetic and epigenetic, and neither of these two factors can be ignored. In the study of tumor biology, network-based models have received more and more attention recently. Many studies have shown that molecular targeted therapy can help predict cancer biomarkers, design network-based anti-cancer therapies, and provide clinical strategies for cancer studies (14–17). This is because gene regulatory networks can help resolve key issues in cancer research by reflecting not just information at the genetic level but also epigenetic information embedded in gene regulation strengths.

In this study, we investigate IGC formation and mechanisms at both the genetic and epigenetic levels. From literature research, we built an IGC-related gene regulatory network. Some other network-based methods, such as the correlation network (18), do not contain information on gene regulations. The regulations in networks built using regression methods do not contain regulatory directions (19). Networks built using other machine learning methods, such as dynamic Bayesian networks, contain regulatory directions and feedback loops, but researchers must consider biases in the algorithm's accuracy (20). The regulations in the gene regulatory network we build are all from the experimental literature and contain the regulation directions and regulatory types (activation or repression) from experiments that are more reliable compared to high-throughput data mining. Based on the IGC gene regulatory network, we quantify the corresponding landscape. There are three stable states on the landscape—the normal, AG, and gastric cancer (IGC) states. The landscape can give us a better understanding of IGC formation through molecular mechanisms and epigenetic information. The dominant paths between state attractors can be quantified and used to understand the development and progression of IGC.

To investigate how *H. pylori* infection can increase the risk of IGC developed from AG and IM, we simulated different degrees of *H. pylori* infection to provide a global perspective on IGC development. Finally, we use global sensitivity analysis to determine which regulations are more sensitive to IGC prevention and treatment strategies. Three regulations are found—RAS → HIF-1α, ZEB → TGF-β, and HIF-1α→ RAS. These results may guide clinical treatment and the design of drugs based on network strategies.

2 DEVELOPMENT OF INTESTINAL GASTRIC CANCER MODEL

We researched the literature to collect information on genes related to gastric cancer and then used these gastric genes to build the gastric-cancer-related gene regulatory network shown in **Figure 1**. In **Figure 1**, there are 17 genes and 82 regulations. The activating regulations are represented by arrows, and the repression regulations are represented by blunted arrows. The 17 genes were identified by mining the gastric cancer literature. We first collected the genes that are highlighted in publications about gastric cancer, then those related to the 10 hallmarks of cancer and mentioned in the gastric cancer literature. We put these genes into the EVEX database to do text mining and search for interactions. The regulations were collected and the results with low confidence were removed. Very high or high confidence regulations were kept. Moreover, we examined the literature to make sure the regulations identified were correct. All of these results are from the experimental literature. Such literature can clarify how one gene or protein influences another gene or protein. We then identified two genes and determined the regulated relationship between them. These interactions are listed in the **Supplementary Material**. The genes Bcl-2, c-erBb2, and K-ras are the specific genes for IGC. Another 14 genes are crucial for gastric cancer. The genes DCC and Beta-caterin are also specific genes associated with IGC. However, regulation of these two genes in the gene regulatory network does not have feedback loops. Therefore, we remove them from the gene regulatory network. Feedback loops are important because they ensure interactions are non-trivial. Genes with feedback

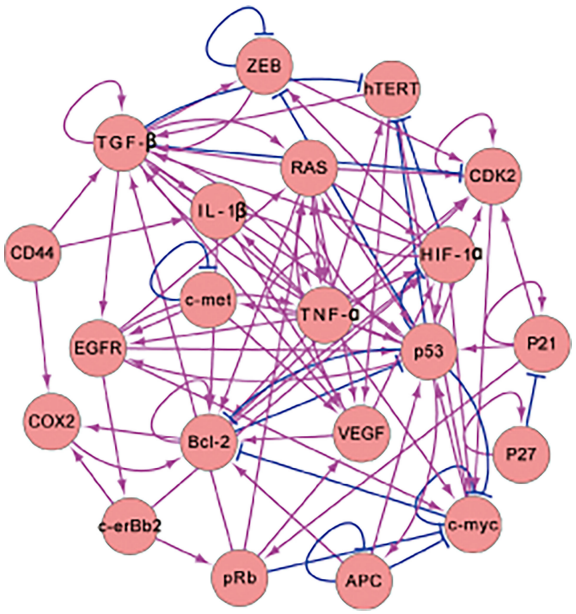


FIGURE 1 | The 17-node gene regulatory network for intestinal-gastric cancer. The network contains 17 genes and 82 regulations (There are 67 activating and 15 repressing regulators. The arrows and the blunted arrows denote the activating and repressing regulations, respectively).

loops must be included in the network. The behavior of genes without feedback loops is relatively straightforward to understand. To simplify the computational process, we reduce the dimensionality by taking these genes as one effective node instead of many until feedback occurs. When the gene does not have incoming edges (there is no gene activating or repressing it), gene expression or protein concentration will be determined by its own self-generation and degradation. Even if a gene has no feedback edges, it can regulate downstream genes. It can influence other genes or proteins through its regulation of others, which depends on its expression or concentration. The regulation of the downstream genes can be manifested by a regulation parameter that is kept constant in the process and determined from the ODEs of the model. Due to the constant regulation, the influence of the gene on others is kept constant. Therefore, we can simplify the model by using the network without it to avoid redundant computations. We list the entire network in the **Supplementary Material**.

In the gene regulatory network, P53 and APC are vital for gastric disease (13, 21) and are both tumor suppressor genes. Abnormalities in TP53 expression have also been observed in *H. pylori*-related AG, IM, dysplasia, and others (22). C-myc participates in cell proliferation and apoptosis, which is significant for gastric cancer and other digestive-system-related cancers (23). HIF-1α is involved in glycolysis pathways for hypoxia (24) and is a critical prognosis element for gastric cancer (25). The RAS gene participates in certain cellular functions such as cell proliferation, differentiation, survival, and apoptosis (26). TGF-β plays a major role in cancer metastasis and participates in the transduction of self-sufficiency growth signals (27). TNF-α is involved in gastric cancer progression, such as in invasion and metastasis (28). The gene c-erbB2 encodes a type of kinase that shows a response to prognosis and is associated with IGC therapy (29). ZEB is a key gene for epithelial–mesenchymal transitions (EMT) that promotes cancer metastasis (30). EGFR is a vital prognostic factor of IGC that is related to the transduction of proliferative signals (31). The VEGF gene is often highlighted in IGC prognosis and has a vital gene response to angiogenesis (32, 33). The c-met gene is a prominent drug target of IGC (34). Bcl-2 plays a key role in apoptosis, and the dysfunction of Bcl-2 is the basis of carcinogenesis (35). COX2 is a key player in IGC development and is associated with risks for numerous types of cancer (36). hTERT is a potent part of IGC and is related to unlimited DNA replication (37). CDK2 is known as an evading growth suppressor and is indispensable in gastric cancer therapy (38). IL-1β is a cytokine associated with lesions, inflammation, and wound healing (39).

Once the gene regulatory network has been developed, we can use ordinary differential equations to describe the dynamics of the related network, with the equations as shown below:

$$\frac{dX_i}{dt} = F_i = g_i \prod_{j=1}^{n_i} H_{ji} - k_i X_i \quad (1)$$

In Eq. (1), $\frac{dX_i}{dt}$ represents the gene expression (protein concentrations), which changes with respect to time. The parameters g and k are used to illustrate the protein generation rate and the protein self-degradation rate, respectively. X_i is used to represent the gene expression or the amount of protein that causes the transcript of the gene i. The subscript j represents the gene regulating the gene i. n_i is the amount of gene regulating the gene i. H_{ji} quantifies the regulations among genes through a Hill function (40), which can be defined as the following:

$$H_{ji} = \frac{S_{ji}^n}{S_{ji}^n + X_j^n} + \lambda_{ji}^r \frac{X_j^n}{S_{ji}^n + X_j^n} \quad (2)$$

Here, the parameter S denotes the threshold, which is the half-maximum value of the sigmoid function. When the value of S is very large, the regulation strength will tend toward a fixed value of 1. When the value of S is very small, the regulation strength will tend toward another fixed value, λ_{ji}^r. To keep the regulation strength in an appropriate range, S is set to be 2.5. The meanings of the subscripts i and j are equivalent to those in Eq. (1). The parameter n denotes the steepness of the sigmoid function and demonstrates the protein cooperatives. In a biochemical system, a protein binding complex can be a monomer, dimer, trimer, tetramer, etc. In our system, since tetramers are more frequent, the parameter n is set to 4. The parameter λ_{ji} is defined to be greater than 1, and denotes the regulation strength of X_j in regulating X_i. The parameter r denotes the regulation type ($r = +1$ represents the activation type and $r = -1$ represents the inhibition type). The parameters g and k denote the protein generation rate and the protein self-degradation rate, respectively. To simplify the calculation of the whole system, we set them to be comparable with each other, with $g = 1$ and $k = 1$.

The parameter λ is a matrix representing the weight of the gene regulatory network (see **Table S1** of the **Supplementary Material**). The gene regulatory network contains 17 genes.

Therefore, 17 ODEs were used to describe the dynamics of the whole biological system. The weights of the network are different, and related to the gene-gene regulation strengths. We determined the weights of the network from the biological functions and gene expression in different stages of gastric cancer. For example, P53 is a tumor suppressor gene, and the gene expression of P53 will be high in the normal state and low in the gastric cancer state. Another 16 gene expressions are consistent with the literature results. If the behavior of gene expressions is inconsistent with the results in the literature, we modify the parameters of the network accordingly until the simulated gene expressions are consistent with experiments. As there are ranges of parameters that can produce similar behavior, we vary the parameters by about 10% and ensure that the gene expressions of the normal and cancer states, as well as the associated landscape topography, do not change significantly. We believe this gives a range of parameters that lead to behavior consistent with the literature.

3 RESULTS AND DISCUSSION

3.1 The IGC Landscape and Related Kinetic Paths

The gene regulatory network (**Figure 1**) of IGC contains 17 genes. Through the collection of our simulation trajectories and associated statistics, we can quantify the landscape on a 17-dimensional probability distribution. The related potential landscape U can be defined as $U = -lnP_{ss}$ (41, 42). P_{ss} is the probability of the steady-states, with'ss' being an abbreviation of steady state. It is difficult to visually display a 17-dimensional landscape, so we take two genes or dimensions (*HIF* - 1α and COX2, two genes very crucial to IGC) to visualize the landscape clearly. In **Figure 2**, the X-axis shows the expression level of hypoxia-inducible factor-1α (HIF-1α), which plays the role of an 'angiogenic switch' in the hypoxia microenvironment in many types of tumors, including IGC (43). The Y-axis shows the expression level of Cyclooxygenase-2 (COX2), which plays a crucial role in cancer development and clinical metastasis (44). We can also choose two other genes (such as EGFR and VEGF), as shown in the **Supplementary Material**. Here, there are still three stable states. There are some changes in the landscape topography, as different genes have different values for the three stable points.

On the IGC landscape, there are three stable state attractors, which are the normal, AG, and cancer (IGC) states, respectively. The definition of these three stable states is based on the biological functions and gene expression levels of the 17 genes in the gene regulatory network. The parameters we set in our model depend on the simulated gene expression levels of the 17 genes, and are all in agreement with the trends seen in the clinical data at different cancer stages (different states). P53 and APC are tumor suppressor genes that have high expression levels in the normal state but low expression levels in cancer cells. The other 15 genes have high expression levels in cancer and low expression levels in normal cells. The simulated gene expression levels of the 17 genes are all consistent with the experimental literature (details can be seen in **Supplementary Material Table S3**).

In **Figure 2**, the genes HIF-1α and COX2 have high expression levels in the cancer state, low expression levels in the normal state, and intermediate expression levels in the intermediate state (AG state) (45). The barrier height between the normal and AG states is relatively low, which indicates that AG infection and recovery is relatively achievable. The barrier height between AG and gastric cancer states is much higher, which indicates that gastric cancer formation and recovery is much more difficult. The barrier height can quantify the difficulty of attractor transfer from one state to another. The barrier height from gastric cancer to the AG state is very high, which shows why gastric cancer is so difficult to cure (or reverse).

To describe the gastric cancer progression process quantitatively, we quantified the dominant paths from normal to AG to gastric cancer states using previously-explored approaches (path integral approaches) (46, 47). In **Figure 2B**, we can see the dominant paths colored in red, blue, yellow, and purple, respectively, as the dominant paths from the normal to the AG state, from the AG to the normal state, from the AG to the gastric cancer state, and from the gastric cancer to the AG state, respectively. These dominant paths are separated and irreversible (48) as the rotational flux force (as part of the driving force) in addition to the gradient force makes the dominant path separate from the gradient direction of the potential. In our model, the driving force can be mathematically decomposed into two directions, the rotational flux force and the gradient force of the potential landscape. The green arrows represent the rotational flux force and the white arrows represent the gradient force direction. The dominant paths through the normal state to the AG and gastric cancer states are irreversible, which can help us understand why the processes of IGC formation and IGC treatment are separate and irreversible biological processes.

3.2 Simulations of the Effect of *H. pylori* Infection on IGC

To investigate the influence of *H. pylori* on IGC, we performed simulations to observe the cancer progression on the landscape. We used a term in the ODEs to simulate different degrees of *H. pylori* infection. The term $F(x_i)$ can be rewritten as $F'(x_i) = F(x_i) + H_i (i = 1,2\ldots,17)$. The term H_i is used to denote the degree of Helicobacter pylori infection on the related gene expression level. The value of H is set according to the experiments. If the gene expression level is increased from the *H. pylori* infection, the value of H will be >0. In the opposite case, $H < 0$.

For **Figure 3**, we chose genes hTERT and MYC to show the landscape layers with variations associated with the development of IGC under the effects of *H. pylori* infection. *H. pylori* infection can result in the gene expression levels of both genes being increased. From **Figure 3**, we can see that when the term $H = 0$, the normal, AG, and gastric cancer states are visible on the first layer of the landscape. A value of $H = 0.05$ indicates infection. There are normal, AG, IM, and gastric cancer states on the second layer of the landscape. The probability of the AG state is dominant, indicating that the *H. pylori* infection accelerates the development of AG. When the term $H = 0.1$, there are normal, IM, and gastric cancer states on the third layer of the landscape. The AG state has disappeared and the IM state is dominant,

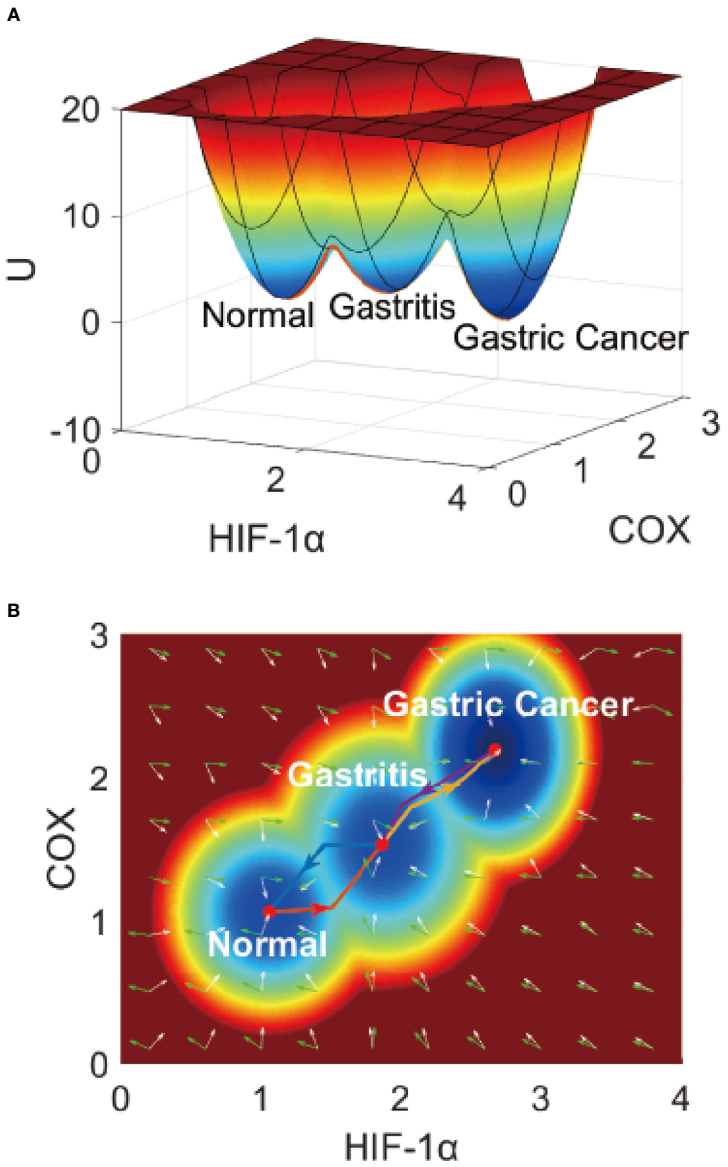


FIGURE 2 | The IGC landscape, which contains three stable states. **(A)** The 3-dimensional landscape of IGC. **(B)** The 2-dimensional landscape of IGC. The lines in red, blue, yellow, and purple represent the dominant kinetic path from the normal to the AG state, from the AG to the normal state, from the AG to the gastric cancer state, and from the gastric cancer to the AG state, respectively. The white arrows denote the negative gradient of the potential landscape, and green arrows denote the curl flux force of the potential landscape.

which indicates that the *H. pylori* infection worsens gastritis and causes further change into intestinal metaplasia. When the term $H = 0.2$, there are normal and gastric cancer states on the fourth layer of the landscape. The IM and gastric cancer states gradually converge and merge into one. The cancer state becomes the dominant state. When the term $H = 0.6$, there is only one gastric cancer state on the fifth layer of the landscape. The normal state disappears and the cancer state is dominant. It is impossible for a patient to recover to their normal state while suffering from *H. pylori* infection can lead to the aggravation of AG, and then the appearance of the IM state. The IM state can be considered very close to the cancerous state during IGC development. When the *H. pylori* infection becomes more and more serious, the AG state disappears and the IM state becomes dominant, finally leaving only one gastric cancer state.

This simulation is on the epigenetic level to illustrate the progression and development of IGC when one gets infected with *H. pylori*. The regulations of the network do not change with this series of variations in the landscape. The effect of *H. pylori* infection results in variations in the IGC landscape. Depending on the degree of *H. pylori* infection, which is becoming more and more serious, landscape development is moving on the cancer direction. As AG and IM are seen in the development of the landscape, which demonstrates the dynamic

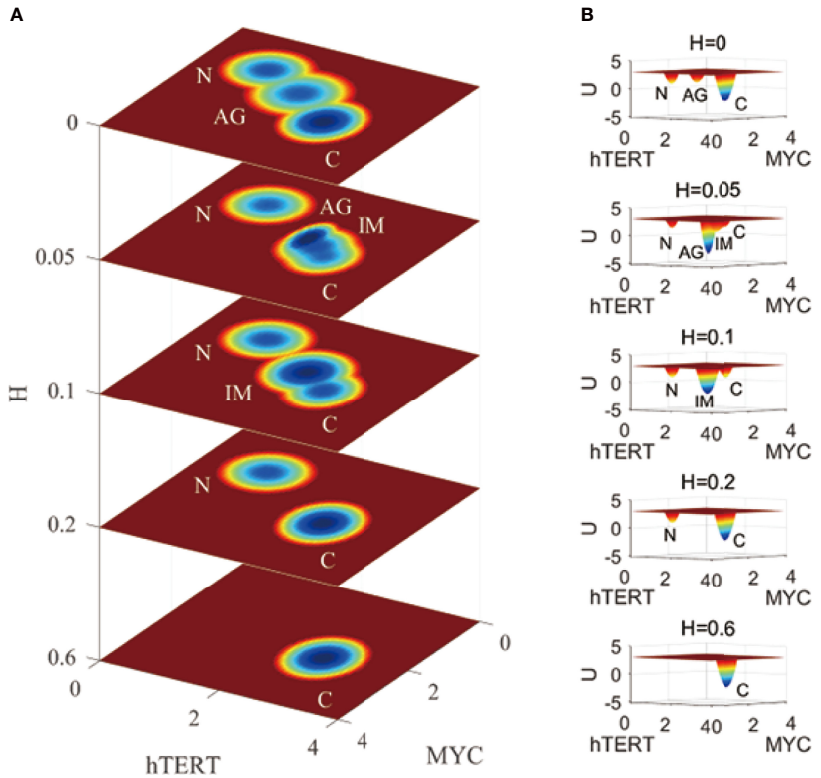


FIGURE 3 | A comparison of the landscape topography variations fir IGC with Helicobacter pylori infection. The X and Y-axes represent the gene expression levels of hTERT and MYC, respectively, while the Z-axis represents the variations in H. H denotes the degree of Helicobacter pylori infection. N, AG, IM, and C represent the normal, atrophic gastritis, intestinal metaplasia, and gastric cancer states, respectively. **(A)** is the comparison of the 2-dimensional landscape. **(B)** is the comparision of the 3-dimensional landscape.

process of IGC development as the degree of *H. pylori* infection changes.

3.3 Identifying Key Regulations of IGC Through Global Sensitivity Analysis

To further investigate the key regulations crucial to IGC therapy or prevention, we apply a global sensitivity analysis method to the landscape model. Each gene regulation or protein concentration can contribute to system dynamics. A small change in the regulatory strength of one gene in the gene regulatory network can lead to the whole landscape topography varying accordingly. As the barrier heights between the biological states can quantify the difficulty associated with transferring between the states, we calculate the variation in the barrier height when regulation strength is changed. The greater the variation, the more sensitive the regulation is.

Figure 4 shows the global sensitivity analysis for IGC. We selected the top 10 most sensitive regulations, which are shown in **Figure 4A**. When we changed the regulation strength to 0.9 of the original regulation strength, these regulations showed the most significant variations in ΔU_{ng} and ΔU_{gn}. ΔU_{ng} is the variation of the barrier height from the normal state to the AG state. ΔU_{gn} is the variation of the barrier height from the AG state to the normal state. From **Figure 4A**, we can see that the values of ΔU_{ng} changed most significantly are for regulating RAS → HIF-1α and ZEB → TGF-β. The higher value of ΔU_{ng} indicates that it is more difficult for the cells to transform from the normal state to the AG state than before. This is because the barrier height between the normal and the AG state is much higher than before. The variations of ΔU_{ng} are most significant when the regulation strengths of the RAS → HIF-1α and ZEB → TGF-β are varied. This type of variation can be used in gastritis prevention as the cell transformation to the AG state becomes more difficult. When the regulation strength is reduced to 0.9 of the original value, the activation of the genes HIF-1α and TGF-β decreases and the concentrations of HIF-1α and TGF-β decrease accordingly. Experiments show that the expression levels of HIF-1α in gastric cancer patients are higher than those in healthy subjects (25). HIF-1α participates in the activation of numerous target genes to adapt to the hypoxic environment (49), which leads to gastric cancer development. We changed the regulation strength, which can cause the concentration of HIF-1α to decrease. This will inhibit the transcription of those target genes and reduce the ability of the cells to adapt to the hypoxic environment, which leads to higher ΔU_{ng}. Therefore, reducing the regulation strength to 0.9 times the original regulation

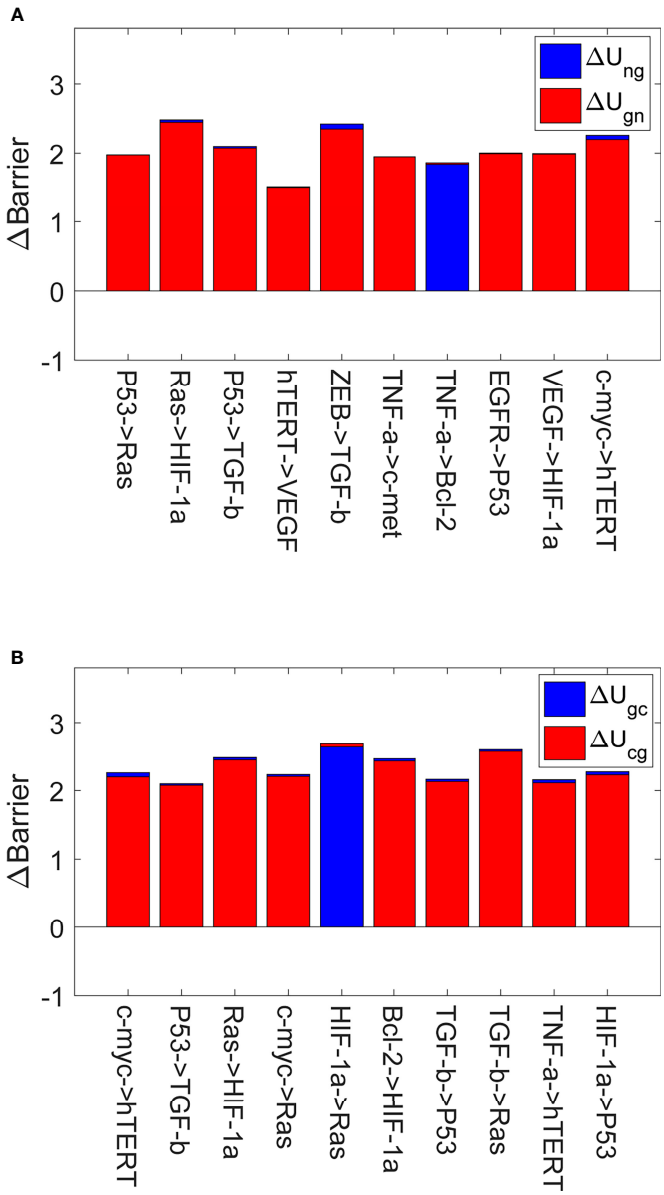


FIGURE 4 | The global sensitivity analysis for IGC. The X-axis represents the top 10 most sensitive regulations, and the Y-axis represents variations in the barrier height (ΔBarrier). **(A)** The variations in the barrier height between the normal and AG states. ΔUng is the variation in the barrier height between the normal and AG states. ΔUgn is the variation in the barrier height between the AG and normal states. **(B)** The variations in the barrier height between the AG and cancer states. ΔUgc is the variation in the barrier height between the AG and cancer states. ΔUcg is the variation in the barrier height between the cancer and AG states.

strength can inhibit the cells from adapting to the hypoxic environment, which can prevent the gastritis cells from developing into gastric cancer. Studies show that inhibiting TGF-β expression can serve as a potential therapeutic target or a biomarker for gastric cancer treatment (50, 51). Therefore, the decreased regulation strength can reduce the transcription of TGF-β, which increases in ΔU_{ng} and inhibits gastritis from developing into gastric cancer.

Figure 4B shows the top 10 most sensitive regulations when we changed the regulation strengths to 0.9 times their original values. These regulations show the most significant variations of ΔU_{gc} and ΔU_{cg}. A higher value of ΔU_{gc} compared to before indicates that switching from the AG state to the gastric cancer state becomes more difficult, because the barrier height between the AG and gastric cancer states is much higher. Reducing the regulation strength can inhibit the expression level of RAS, which increases in ΔU_{gc} and helps inhibit gastric cancer progression and development. Such variation can be used in gastric cancer prevention, as the transformation of the cell from the AG state to the gastric cancer state becomes more difficult. When the

regulation strength is reduced to 0.9 times the original value, the activation of the RAS gene decreases and the concentration of RAS also decreases accordingly. There are many studies showing that K-ras is associated with the development and progression of gastric cancer. Overexpression of K-ras can increase the risk of gastric cancer development (52) (K-ras is a protein of the RAS family). Reducing the regulation strength can inhibit the expression level of RAS, which increases in ΔU_{gc} and helps inhibit gastric cancer progression and development.

Figure 4 displays the most sensitive regulations on global topography in terms of the barrier height between normal and cancer states. When the regulation strengths of P53 → RAS, c-myc→ RAS, HIF-1α → RAS and TGF-β → RAS are reduced to 0.9 of the original values, the expression of RAS will decrease accordingly. Studies show that K-ras regulates cell survival, motility, proliferation, angiogenesis, and metastasis (53). Therefore, this participates in gastritis and gastric cancer formation and metastasis. When the regulation strengths of RAS → HIF-1α, VEGF → HIF-1α, and Bcl-2 → HIF-1α are reduced to 0.9 of the original values, the expression of HIF-1α is reduced accordingly. The gene HIF-1α can activate the transcription of many target genes to adapt to the hypoxic environment of cancer cells (54). The overexpression of the gene HIF-1α can induce cancer cell development. When the regulation strengths ofP53 → TGF-β and ZEB → TGF-β are reduced to 0.9 of the original values, TGF-β expression decreases. TGF-β promotes cancer-related characteristics in most gastric cancer cell lines (55). In TGF-β and HIF-1α gene expressions appear twice as that of the target genes in **Figure 4A**. Ras expression appears thrice in **Figure 4**, and the expressions of HIF-1α, P53, and hTERT genes appear twice as that of the target genes. We can pay more attention to these genes in designing strategies in clinical experiments or trials to prevent gastritis or gastric cancer formation.

3.4 Identifying Key Regulations of IGC With *H. pylori* Infection

In **Figure 3**, we can see that when a patient is infected with *H. pylori* (term $H = 0.05$), four stable states emerge, which are the normal, AG, IM, and gastric cancer states. To figure out which regulations are more sensitive to IGC with *H. pylori* infection, we performed a global sensitivity analysis on this condition.

Figure 5 shows global sensitivity analysis for IGC with *H. pylori* infection. We reduced the regulation strength to 0.9 of the original value. **Figure 5A** displays the top 10 regulations most sensitive to variations in ΔU_{gm} and ΔU_{mg}. ΔU_{gm} is the variation in the barrier height from the AG state to the IM state. ΔU_{mg} is the variation in the barrier height from the IM state to the AG state. The regulations for HIF-1α→c-myc and CDK2→c-myc cause the most significant changes in ΔU_{gm}. The value of ΔU_{gm} becoming higher indicates that it becomes more difficult for cells to transform from the AG state to the IM state as the barrier height is higher. This type of variation can be used to prevent cell transformation from the AG state to the IM state. As the activation strengths of HIF-1α→c-myc and CDK2→c-myc are decreased, the expression of c-myc will be reduced accordingly. Gene c-myc has been studied as a biomarker with which to identify *H. pylori* infection (56). Gastric cancer treatment and gastric cancer progression are complicated by aberrant expressions of c-myc (57). Therefore, inhibiting the expression of c-myc will benefit the treatment of IGC with *H. pylori* infection and prevent transformations from AG to an IM state.

Figure 5B displays the top 10 regulations most sensitive to variations in ΔU_{mc} and ΔU_{cm}. ΔU_{mc} is the variation of the barrier height from the IM state to the gastric cancer state. ΔU_{cm} is the variation of the barrier height from the gastric cancer state to the IM state. The regulation RAS→ TGF-β caused the most significant changes in ΔU_{mc}. The value of ΔU_{mc} becoming higher indicates that it becomes more difficult for cells to transform from the IM state to the gastric cancer state because the barrier height between the two states is higher. As the activation strength of RAS→ TGF-β is decreased, the expression of TGF-β will be reduced accordingly. TGF-β can trigger epithelial–mesenchymal transition (EMT) markers, which are crucial for canceration and metastasis. *H. pylori* infection can induce TGF- β to trigger the EMT pathway. When *H. pylori* is eradicated, TGF- β is inhibited from triggering the EMT pathway (58). TGF- β is a key gene in gastric cancer prevention and treatment, which we have identified in this study. C-myc and TGF- β are vital for the treatment and prevention of IGC from *H. pylori* infection.

From **Figure 4** and **Figure 5** we can see that the key regulations are different depending on whether the *H. pylori* is infected or uninfected. When infected with *H. pylori*, another state (IM state) emerged on the landscape, which is different from the IGC landscape without *H. pylori* infection. The c-myc gene is essential as it appears in the two key regulations (HIF-1α→c-myc and CDK2→c-myc). The c-myc gene is a biomarker to identify *H. pylori* infection in clinical trials (56). We can take other regulations such as TGF-β→ IL-1β and TGF-β→ ZEB, which are more sensitive in clinical experiments. The gene TGF-β is vital for both *H. pylori* infected and uninfected as it is sensitive to the two conditions when the cell states switch from AG (or IM) state to cancer state. TGF-β plays a critical role in cancer metastasis (58). TGF-β appears thrice in the top 10 regulations in **Figure 5B**, while TNF-α and c-myc genes appear twice. We should take these genes into consideration in designing strategies in clinical experiments for preventing of gastric cancer with *H. pylori* infection.

4 CONCLUSIONS

In this work, we have studied the formation and development of IGC in a systematic and quantitative way. We have built a gene regulatory network for IGC. The genes and gene regulations were collected through experimental literature research. The gene regulatory network reflects both genetic and epigenetic level information. After the construction of the gene regulatory network, we used ODEs to describe the dynamics of IGC. We then obtained a systematic landscape for IGC. There are normal, AG, and gastric cancer states on the IGC landscape. The landscape can provide us with a global overview of IGC progression and development, which can help us understand IGC formation

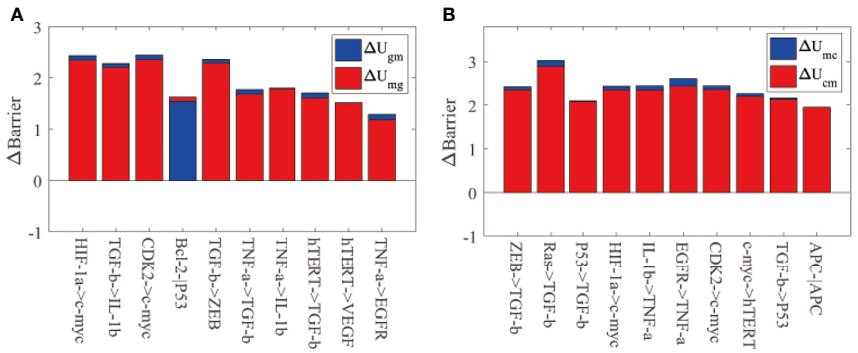


FIGURE 5 | The global sensitivity analysis for IGC with Helicobacter pylori infection. The meanings of the X-axis and Y-axis are the same as **Figure 3**. **(A)** The barrier height variations between the AG and IM states. ΔUgm represents the barrier height variation between the AG and IM states. ΔUmg is the barrier height variation between the IM and AG states. **(B)** The barrier height variation between the IM and cancer states. ΔUmc is the barrier height variation between the IM and cancer states. ΔUcm is the barrier height variation between the cancer and IM states.

systematically. The dominant paths can describe the IGC progression and dynamical transitions can help us understand the IGC development quantitatively. The dominant paths between neighboring states (the normal and AG states or the AG and gastric cancer states) are separate and irreversible. The irreversibility of the dominant paths explains why the IGC formation and recovery processes are complex and independent.

To investigate the effect of *H. pylori* infection on IGC formation, we simulated different degrees of *H. pylori* infection, resulting in variations in the landscape topography. When one is infected with *H. pylori* (H = 0.05), a state called intestinal metaplasia IM appears in the landscape, and the atrophic gastritis (AG) state becomes dominant. When the degree of *H. pylori* infection becomes serious, the AG state disappears and the IM state becomes dominant. As the degree of *H. pylori* infection increases, the normal state disappears, eventually leaving only one gastric cancer state. This demonstrates that *H. pylori* infection leads to gastric cancer progression and illustrates how *H. pylori* infection can increase the risk of gastric cancer development.

To further highlight the key regulations associated with IGC therapy and treatment, we performed a global sensitivity analysis and found three key regulations to be more sensitive than the others as the landscape topography varies. The three regulations are RAS $\rightarrow$ HIF-1α, ZEB $\rightarrow$ TGF-β, and HIF-1$\alpha\rightarrow$ RAS. We predicted that these regulations would serve as a guide for developing network-based anti-cancer drug targets.

This study provides a new approach and a novel yet simple model to analyze IGC in a global and systematic way. This model can help us understand the formation and development of IGC, not only from genetic variations but also from epigenetic modifications. Furthermore, *H. pylori* infection can be simulated and investigated with the landscape model. Global sensitivity analysis can help us determine which regulations are more sensitive for gastric cancer prevention or therapy. The results can help us develop clinical strategies by designing polygenic drugs to fight cancer.

5 SUPPORT MATERIALS

5.1 Landscape-Flux Decomposition of the Driving Force

A gene regulatory network consists of genes and gene regulatory relationships, represented by nodes and edges in the network, respectively. We use an n-component vector $\boldsymbol{x} = (x_1, x_2,\ldots, x_n)$ to quantify the genes in the network. Here n is the number of genes in the network and x_i (i =1,2,…, n) denote the expression levels (or protein concentrations) of the corresponding genes. A system of ordinary differential equations, written in the compact form $\dot{x} = F(x)$, can be employed to study the deterministic dynamics of the network, where $\boldsymbol{F}(\boldsymbol{x})$ denotes the driving force of the deterministic dynamics.

In biological systems, stochastic fluctuations of internal or external origins are ubiquitous and may have a significant impact on the dynamics of the system. To incorporate the effects of stochastic fluctuations, a stochastic force $\xi(t)$ may be attached to the ordinary differential equations$\dot{x} = F(x)$. This leads to a stochastic differential equation of the form$\dot{x} = F(x) + \xi(t)$, also known as the Langevin equation. The stochastic force $\xi(t)$ modeling random fluctuations is assumed to be Gaussian white noise in time, with the mean$\langle\ \xi(t)\ \rangle$=0 and the correlation $\langle\ \xi(t)\xi^T(t')\ \rangle$=$2D\delta(t-t')$. Here $\boldsymbol{D}$ is the diffusion matrix characterizing the fluctuation strength.

An equivalent description of the Langevin dynamics is in terms of the probability distribution $P(\boldsymbol{x}, t)$, whose time evolution is governed by the corresponding Fokker-Planck equation: $\partial P/\partial t = -\nabla \cdot [\boldsymbol{F}P - \nabla\cdot (\boldsymbol{D}P)]$. It can also be written as $\partial P/\partial t = -\nabla \cdot \boldsymbol{J}$, with $\boldsymbol{J}$ denoting the probability flux. The steady state characterized by $\partial P_{ss}/\partial t = \nabla \cdot \boldsymbol{J}_{ss} = 0$ is of particular interest. In an equilibrium system, the probability flux at the steady state vanishes, i.e. $\boldsymbol{J}_{ss} = 0$. In a non-equilibrium system, there is in general a nonvanishing probability flux at the steady state, i.e. $\boldsymbol{J}_{ss} \neq 0$, which signifies the time-irreversible nature of the non-equilibrium steady state. From the expression $\boldsymbol{J}_{ss} = \boldsymbol{F}P_{ss} - \boldsymbol{D}\cdot\nabla P_{ss}$, the driving force $\boldsymbol{F}$ can be written in

the landscape-flux decomposition form (59): $\boldsymbol{F} = -\boldsymbol{D} \cdot \nabla U + \boldsymbol{J}_{ss}/P_{ss}$. Here $-\boldsymbol{D} \cdot \nabla U$ is the part of the driving force contributed by the gradient of the potential landscape $U = -\ln P_{ss}$, and $\boldsymbol{J}_{ss}/P_{ss}$ is the other part contributed by the probability flux that is associated with the nonequilibrium nature of the system.

5.2 Self-Consistent Mean Field Approach

Self-consistent mean field (60) serves as an effective approximation method of solving Fokker-Planck equations with a large number of variables. In this approximation, the joint probability distribution of all the variables is substituted by the product of marginal probability distributions of each variable, namely, $P(x_1, x_2, \ldots, x_n, t) \sim \Pi_i P(x_i, t)$, so that the latter can be solved in a self-consistent manner. The dimensionality of the problem is reduced significantly from m^n to $m \times n$, where m is the number of possible values each variable may take. This makes the computations much more feasible.

A further approximation is invoked to simplify the problem, which postulates $P(x_1, x_2, \ldots, x_n, t)$ with the form of a multivariate Gaussian distribution. When the magnitude of the diffusion matrix $\boldsymbol{D}$ is small, the equations governing the mean vector $\bar{x}(t)$ and the covariance matrix $\boldsymbol{\sigma}(t)$ of the Gaussian distribution are given by:

$$\dot{x}(t) = F(\bar{x}(t)) \quad (3)$$

$$\dot{\sigma}(t) = A(t)\sigma(t) + \sigma(t)A^T(t) + 2D. \quad (4)$$

Here the matrix $\boldsymbol{A}$ has elements $A_{ij}(t) = \frac{\partial F_i(\bar{x}(t))}{\partial \bar{x}_j(t)}$. Given the self-consistent mean field approximation, only the diagonal elements of $\sigma(t)$ need to be considered. The combination of the self-consistent mean field approach and the Gaussian distribution approximation leads to the following form of probability distribution evolution for each x_i:

$$P(x_i, t) = \frac{1}{\sqrt{2\pi\sigma_i(t)}} \exp\left\{-\frac{[x_i - \bar{x}_i(t)]^2}{2\sigma_i(t)}\right\}. \quad (5)$$

The steady-state probability distribution of a monostable system with one fixed point can be easily obtained on the basis of Eq.(5) as a single Gaussian distribution. In a multistable system with more than one fixed point, the steady-state distribution may be constructed as a combination of multiple Gaussian distributions with the form $P_{ss}(\boldsymbol{x}) = \Sigma_k \omega_k P_k(\boldsymbol{x})$, where k labels different fixed points, ω_k represents the weight of each fixed point, and $P_k(\boldsymbol{x})$ is the Gaussian distribution corresponding to each fixed point.

5.3 The Path Integral Approach

Based on the Onsager-Machlup functional, the transition probability of the Fokker-Planck equation has the following path-integral formulation (46):

$$\begin{aligned} P(x_f, t_f; x_0, t_0) &= \int D\ [x(t)] exp\{-S[x(t)]\} \\ &= \int D\ [x(t)] exp\left\{-\int L(x(t))dt\right\}. \end{aligned} \quad (6)$$

In the above, $\boldsymbol{x}_0$ denotes the initial state at time t_0 and $\boldsymbol{x}_f$ represents final state at time t_f. $P(\boldsymbol{x}_f, t_f; \boldsymbol{x}_0, t_0)$ is the transition probability from the initial state to the final state. The notation $\int D\ [\boldsymbol{x}(t)]$ represents an integral over the all possible paths starting from the initial state $\boldsymbol{x}_0$ at time t_0 and ending at the final state $\boldsymbol{x}_f$ at time t_f. $L(\boldsymbol{x}(t))$ is the Lagrangian with the expression $L(x(t)) = \frac{1}{4}(\dot{x} - F(x)) \cdot D^{-1} \cdot (\dot{x} - F(x)) + \frac{1}{2} \nabla \cdot F(x)$. Its time integration gives the action $S[\boldsymbol{x}(t)] = \int L(\boldsymbol{x}(t))dt$ associated with each path as in classical mechanics. The action $S[\boldsymbol{x}(t)]$ determines the probability weight $e^{-S[\boldsymbol{x}(t)]}$ contributed by the corresponding path. The summation (or integration) of these probability weights over all the paths gives the transition probability. Since the contribution of each path has the exponential form $e^{-S[\boldsymbol{x}(t)]}$, the dominant path with maximum probability is the path with minimum action, which can be determined by the variational principle $\delta S[\boldsymbol{x}(t)] = 0$ and the resulting Euler-Lagrange equation. For non-equilibrium systems the existence of nonvanishing probability flux $\boldsymbol{J}_{ss}$ cannot be ignored. As a consequence, the dominant kinetic paths in non-equilibrium systems are separated and irreversible (46).

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

Conception and design: CY and JW. Development of methodology: CY and JW. Acquisition of data: CY. Analysis and interpretation of data: CY and JW. Writing the manuscript: CY and JW. Study supervision: JW. All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.853768/full#supplementary-material

REFERENCES

1. Miller KD, Nogueira L, Mariotto AB, Rowland JH, Yabroff KR, Alfano CM, et al. Cancer Treatment and Survivorship Statistics. *CA A CA Cancer J Clin* (2019) 2019(69):363–85. doi: 10.3322/caac.21565
2. Park YH, Kim N. Review of Atrophic Gastritis and Intestinal Metaplasia as a Premalignant Lesion of Gastric Cancer. *J Cancer Prev* (2015) 20:25–40. doi: 10.15430/jcp.2015.20.1.25
3. Park JM, Ryu WS, Kim JH, Park SS, Kim SJ, Kim CS, et al. Prognostic Factors for Advanced Gastric Cancer: Stage-Stratified Analysis of Patients Who Underwent Curative Resection. *Cancer Res Treat* (2006) 38:13. doi: 10.4143/crt.2006.38.1.13
4. Werner M, Becker KF, Keller G, Höfler H. Gastric Adenocarcinoma: Pathomorphology and Molecular Pathology. *J Cancer Res Clin Oncol* (2001) 127(4):207–16. doi: 10.1007/s004320000195
5. Cisło M, Filip AA, Offerhaus GJA, Ciseł B, Rawicz-Pruszyński K, Skierucha M, et al. Distinct Molecular Subtypes of Gastric Cancer: From Laurén to

Molecular Pathology. *Oncotarget* (2018) 9:19427–42. doi: 10.18632/oncotarget.24827
6. Carcas L. Gastric Cancer Review. *J Carcinogene* (2014) 13:14. doi: 10.4103/1477-3163.146506
7. Bravo D, Hoare A, Soto C, Valenzuela MA, Quest AF. Helicobacter Pyloriin Human Health and Disease: Mechanisms for Local Gastric and Systemic Effects. *World J Gastroenterol* (2018) 24:3071–89. doi: 10.3748/wjg.v24.i28.3071
8. Huang JQ, Sridhar S, Chen Y, Hunt RH. Meta-Analysis of the Relationship Between Helicobacter Pylori Seropositivity and Gastric Cancer. *Gastroenterology* (1998) 114:1169–79. doi: 10.1016/s0016-5085(98)70422-6
9. Correa P. Gastric Cancer. *Gastroenterol Clin North Am* (2013) 42:211–7. doi: 10.1016/j.gtc.2013.01.002
10. Carrasco G, Corvalan AH. Helicobacter Pylori-Induced Chronic Gastritis and Assessing Risks for Gastric Cancer. *Gastroenterol Res Pract* (2013) 2013:1–8. doi: 10.1155/2013/393015
11. Smith MG. Cellular and Molecular Aspects of Gastric Cancer. *World J Gastroenterol* (2006) 12:2979. doi: 10.3748/wjg.v12.i19.2979
12. Yasui W, Oue N, Kuniyasu H, Ito R, Tahara E, Yokozaki H. Molecular Diagnosis of Gastric Cancer: Present and Future. *Gastric Cancer* (2001) 4:113–21. doi: 10.1007/pl00011733
13. Yasui W, Sentani K, Motoshita J, Nakayama H. Molecular Pathobiology of Gastric Cancer. *Scandina J Surg* (2006) 95:225–31. doi: 10.1177/145749690609500403
14. Yu C, Liu Q, Chen C, Wang J. Quantification of the Underlying Mechanisms and Relationships Among Cancer, Metastasis, and Differentiation and Development. *Front Genet* (2020) 10:1388. doi: 10.3389/fgene.2019.01388
15. Yu C, Xu H, Wang J. A Global and Physical Mechanism of Gastric Cancer Formation and Progression. *J Theor Biol* (2021) 520:110643. doi: 10.1016/j.jtbi.2021.110643
16. Yu C, Wang J. A Physical Mechanism and Global Quantification of Breast Cancer. *PloS One* (2016) 11:e0157422. doi: 10.1371/journal.pone.0157422
17. Li C, Wang J. Quantifying the Underlying Landscape and Paths of Cancer. *J R Soc Interface* (2014) 11:20140774. doi: 10.1098/rsif.2014.0774
18. Schwanhäusser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, et al. Global Quantification of Mammalian Gene Expression Control. *Nature* (2011) 473:337–42. doi: 10.1038/nature10098
19. Haury AC, Mordelet F, Vera-Licona P, Vert JP. TIGRESS: Trustful Inference of Gene REgulation Using Stability Selection. *BMC Syst Biol* (2012) 6:145. doi: 10.1186/1752-0509-6-145
20. Xuan N, Chetty M, Coppel R, Wangikar PP. Gene Regulatory Network Modeling via Global Optimization of High-Order Dynamic Bayesian Network. *BMC Bioinf* (2012) 13:131. doi: 10.1186/1471-2105-13-131
21. Wang JY, Hsieh JS, Chen CC, Tzou WS, Cheng TL, Chen FM, et al. Alterations of APC, C-Met, and P53 Genes in Tumor Tissue and Serum of Patients With Gastric Cancers. *J Surg Res* (2004) 120:242–8. doi: 10.1016/j.jss.2003.12.018
22. Morgan C, Jenkins GJS, Ashton T, Griffiths AP, Baxter JN, Parry EM, et al. Detection of P53 Mutations in Precancerous Gastric Tissue. *Br J Cancer* (2003) 89:1314–9. doi: 10.1038/sj.bjc.6601302
23. Saha P, Kumar YP, Das T, Müller D, Bessi I, Schwalbe H, et al. G-Quadruplex-Specific Cell-Permeable Guanosine–Anthracene Conjugate Inhibits Telomere Elongation and Induces Apoptosis by Repressing the C-MYC Gene. *Bioconjug Chem* (2019) 30:3038–45. doi: 10.1021/acs.bioconjchem.9b00655
24. Chi PI, Huang WR, Chiu HC, Li JY, Nielsen BL, Liu HJ. Avian Reovirus Sigma a-Modulated Suppression of Lactate Dehydrogenase and Upregulation of Glutaminolysis and the Mtoc1/Eif4e/HIF-1alpha Pathway to Enhance Glycolysis and the TCA Cycle for Virus Replication. *Cell Microbiol* (2018) 20:e12946. doi: 10.1111/cmi.12946
25. Xiong Q. Diagnostic Value of Serum Mir-378, Angptl2, Hif-1 and Cea Alone or in Combination in Patients With Gastric Cancer. *Acta Med Mediterr* (2019) 35:3159–63. doi: 10.19193/0393-6384
26. Gonzalez-Hormazabal P, Musleh M, Bustamante M, Stambuk J, Pisano R, Valladares H, et al. Polymorphisms in RAS/RAF/MEK/ERK Pathway are Associated With Gastric Cancer. *Genes* (2018) 10:20. doi: 10.3390/genes10010020
27. Sarkar A, Rahaman A, Biswas I, Mukherjee G, Chatterjee S, Bhattacharjee S, et al. TGF*beta* Mediated LINC00273 Upregulation Sponges Mir200a-3p and Promotes Invasion and Metastasis by Activating Zeb1. *J Cell Physiol* (2020) 235(10):7159–72. doi: 10.1002/jcp.29614
28. Cui X, Zhang H, Na Cao A, Cao L, Hu X. Cytokine TNF-*alpha* Promotes Invasion and Metastasis of Gastric Cancer by Down-Regulating Pentraxin3. *J Cancer* (2020) 11:1800–7. doi: 10.7150/jca.39562
29. Bayrak M, Olmez OF, Kurt E, Cubukcu E, Evrensel T, Kanat O, et al. Prognostic Significance of C-Erbb2 Overexpression in Patients With Metastatic Gastric Cancer. *Clin Trans Oncol* (2012) 15:307–12. doi: 10.1007/s12094-012-0921-0
30. Wellner U, Schubert J, Burk UC, Schmalhofer O, Zhu F, Sonntag A, et al. The Emt-Activator Zeb1 Promotes Tumorigenicity by Repressing Stemness-Inhibiting Micrornas. *Nat Cell Biol* (2009) 11:1487–95. doi: 10.1038/ncb1998
31. Parker MI, Nikonova AS, Sun D, Golemis EA. Proliferative Signaling by ERBB Proteins and RAF/MEK/ERK Effectors in Polycystic Kidney Disease. *Cell Signal* (2020) 67:109497. doi: 10.1016/j.cellsig.2019.109497
32. Lieto E, Ferraraccio F, Orditura M, Castellano P, Mura AL, Pinto M, et al. Expression of Vascular Endothelial Growth Factor (VEGF) and Epidermal Growth Factor Receptor (EGFR) is an Independent Prognostic Indicator of Worse Outcome in Gastric Cancer Patients. *Ann Surg Oncol* (2007) 15:69–79. doi: 10.1245/s10434-007-9596-0
33. Hanahan D, Weinberg RA. Hallmarks of Cancer: The Next Generation. *Cell* (2011) 144:646–74. doi: 10.1016/j.cell.2011.02.013
34. Gymnopoulos M, Betancourt O, Blot V, Fujita R, Galvan D, Lieuw V, et al. TR1801-ADC: A Highly Potent Cmet Antibody–Drug Conjugate With High Activity in Patient-Derived Xenograft Models of Solid Tumors. *Mol Oncol* (2019) 14:54–68. doi: 10.1002/1878-0261.12600
35. Gryko M, Pryczynicz A, Zareba K, Kkedra B, Kemona A, Guzinska-Ustymowicz K. The Expression of Bcl-2 and BID in Gastric Cancer Cells.*J Immunol Res* (2014) 2014:1–5. doi: 10.1155/2014/953203
36. Hao Q, Zhang C, Gao Y, Wang S, Li J, Li M, et al. FOXP3 Inhibits NF-B Activity and Hence COX2 Expression in Gastric Cancer Cells. *Cell Signal* (2014) 26:564–9. doi: 10.1016/j.cellsig.2013.11.030
37. Ohira T, Kojima H, Kuroda Y, Aoki S, Inaoka D, Osaki M, et al. PITX1 Protein Interacts With ZCCHC10 to Regulate hTERT mRNA Transcription. *PloS One* (2019) 14:e0217605. doi: 10.1371/journal.pone.0217605
38. Tang Z, Li L, Tang Y, Xie D, Wu K, Wei W, et al. CDK2 Positively Regulates Aerobic Glycolysis by Suppressing SIRT5 in Gastric Cancer. *Cancer Sci* (2018) 109:2590–8. doi: 10.1111/cas.13691
39. Chang YW, Jang JY, Kim NH, Lee JW, Lee HJ, Jung WW, et al. Interleukin-1b (IL-1b) Polymorphisms and Gastric Mucosal Levels of IL-1? Cytokine in Korean Patients With Gastric Cancer. *Int J Cancer* (2005) 114:465–71. doi: 10.1002/ijc.20724
40. Yu L, Lu M, Jia D, Ma J, Ben-Jacob E, Levine H, et al. Modeling the Genetic Regulation of Cancer Metabolism: Interplay Between Glycolysis and Oxidative Phosphorylation. *Cancer Res* (2017) 77:1564–74. doi: 10.1158/0008-5472.can-16-2074
41. Wang J, Li C, Wang E. Potential and Flux Landscapes Quantify the Stability and Robustness of Budding Yeast Cell Cycle Network. *Proc Natl Acad Sci* (2010) 107:8195–200. doi: 10.1073/pnas.0910331107
42. Li C, Wang E, Wang J. Landscape Topography Determines Global Stability and Robustness of a Metabolic Network. *ACS Synthet Biol* (2012) 1:229–39. doi: 10.1021/sb300020f
43. Wu Y, Meng D, You Y, Sun R, Fu M, Yan Q, et al. Hypoxia Inducible Factor-1alpha (HIF-1a) Plays Different Roles in Gallbladder Cancer and Normal Gallbladder Tissues. *J Cancer* (2021) 12:827–39. doi: 10.7150/jca.46749
44. Lucci A, Krishnamurthy S, Singh B, Bedrosian I, Meric-Bernstam F, Reuben J, et al. Cyclooxygenase-2 Expression in Primary Breast Cancers Predicts Dissemination of Cancer Cells to the Bone Marrow. *Breast Cancer Res Treat* (2008) 117:61–8. doi: 10.1007/s10549-008-0135-x
45. Rohwer N, Lobitz S, Daskalow K, Jöns T, Vieth M, Schlag PM, et al. HIF-1 Determines the Metastatic Potential of Gastric Cancer Cells. *Br J Cancer* (2009) 100:772–81. doi: 10.1038/sj.bjc.6604919
46. Wang J, Zhang K, Xu L, Wang E. Quantifying the Waddington Landscape and Biological Paths for Development and Differentiation. *Proc Natl Acad Sci* (2011) 108:8257–62. doi: 10.1073/pnas.1017017108
47. Li C, Wang J. Quantifying Cell Fate Decisions for Differentiation and Reprogramming of a Human Stem Cell Network: Landscape and Biological Paths. *PloS Comput Biol* (2013) 9:e1003165. doi: 10.1371/journal.pcbi.1003165

48. Maier RS, Stein DL. Escape Problem for Irreversible Systems. *Phys Rev E* (1993) 48:931–8. doi: 10.1103/physreve.48.931
49. Kitajima Y, Miyazaki K. The Critical Impact of HIF-1a on Gastric Cancer Biology. *Cancers* (2013) 5:15–26. doi: 10.3390/cancers5010015
50. Xu G, Chen Y, Fu M, Zang X, Cang M, Niu Y, et al. Circular RNA CCDC66 Promotes Gastric Cancer Progression by Regulating C-Myc and TGF-Signaling Pathways. *J Cancer* (2020) 11:2759–68. doi: 10.7150/jca.37718
51. Liu F, Cao QH, Lu DJ, Luo B, Lu XF, Luo RC, et al. TMEM16a Overexpression Contributes to Tumor Invasion and Poor Prognosis of Human Gastric Cancer Through TGF- Signaling. *Oncotarget* (2015) 6:11585–99. doi: 10.18632/oncotarget.3412
52. Hiyama T, Haruma K, Kitadai Y, Masuda H, Miyamoto M, Tanaka S, et al. K-Ras Mutation Inhelicobacter Pylori-Associated Chronic Gastritis in Patients With and Without Gastric Cancer. *Int J Cancer* (2002) 97:562–6. doi: 10.1002/ijc.1644
53. Peng N, Zhao X. Comparison of K-Ras Mutations in Lung, Colorectal and Gastric Cancer. *Oncol Lett* (2014) 8:561–5. doi: 10.3892/ol.2014.2205
54. Lee DY, Jung DE, Yu SS, Lee YS, Choi BK, Lee YC. Regulation of SIRT3 Signal Related Metabolic Reprogramming in Gastric Cancer by Helicobacter Pylori Oncoprotein CagA. *Oncotarget* (2017) 34(8):78365–78. doi: 10.18632/oncotarget.18695
55. Veen LM, Skrabanja TL, Derks S, de Gruijl TD, Bijlsma MF, van Laarhoven HW. The Role of Transforming Growth Factor in Upper Gastrointestinal Cancers: A Systematic Review. *Cancer Treat Rev* (2021) 100:102285. doi: 10.1016/j.ctrv.2021.102285
56. George S, Lucero Y, Torres JP, Lagomarcino AJ, O'Ryan M. Gastric Damage and Cancer-Associated Biomarkers in Helicobacter Pylori-Infected Children. *Front Microbiol* (2020) 11:90. doi: 10.3389/fmicb.2020.00090
57. Hu Y, Yu K, Wang G, Zhang D, Shi C, Ding Y, et al. Lanatoside C Inhibits Cell Proliferation and Induces Apoptosis Through Attenuating Wnt/-Catenin/C-Myc Signaling Pathway in Human Gastric Cancer Cell. *Biochem Pharmacol* (2018) 150:280–92. doi: 10.1016/j.bcp.2018.02.023
58. Kim N. Chemoprevention of Gastric Cancer Byhelicobacter Pylorieradication and its Underlying Mechanism. *J Gastroenterol Hepatol* (2019) 34(8):1287–95. doi: 10.1111/jgh.14646
59. Wang J, Xu L, Wang EK. Potential Landscape and Flux Framework of Nonequilibrium Networks: Robustness, Dissipation, and Coherence of Biochemical Oscillations. *Proc Natl Acad Sci USA* (2008) 105:12271–6. doi: 10.1073/pnas.0800579105
60. Sasai M, Wolynes PG. Stochastic Gene Expression as a Many-Body Problem. *Proc Natl Acad Sci USA* (2003) 100:2374–9. doi: 10.1073/pnas.2627987100

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

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Oncology

REVIEW
published: 07 July 2022
doi: 10.3389/fonc.2022.914594

Constraint-Based Reconstruction and Analyses of Metabolic Models: Open-Source Python Tools and Applications to Cancer

Rachel H. Ng[1,2], *Jihoon W. Lee*[3,4], *Priyanka Baloni*[1], *Christian Diener*[1], *James R. Heath*[1,2*] *and Yapeng Su*[4,5*]

[1] Institute for Systems Biology, Seattle, WA, United States, [2] Department of Bioengineering, University of Washington, Seattle, WA, United States, [3] Medical Scientist Training Program, University of Washington, Seattle, WA, United States, [4] Program in Immunology, Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, WA, United States, [5] Herbold Computational Biology Program, Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, WA, United States

OPEN ACCESS

Edited by:
Miguel Rocha,
University of Minho, Portugal

Reviewed by:
Ovidiu Radulescu,
Université de Montpellier, France
Frederic CADET,
Dynamique des Structures et Interactions des Macromolécules Biologiques (DSIMB) (INSERM), France

***Correspondence:**
James R. Heath
jheath@isbscience.org
Yapeng Su
suyapeng.tju@gmail.com

Specialty section:
This article was submitted to Cancer Metabolism, a section of the journal Frontiers in Oncology

Received: 07 April 2022
Accepted: 30 May 2022
Published: 07 July 2022

Citation:
Ng RH, Lee JW, Baloni P, Diener C, Heath JR and Su Y (2022) Constraint-Based Reconstruction and Analyses of Metabolic Models: Open-Source Python Tools and Applications to Cancer. Front. Oncol. 12:914594. doi: 10.3389/fonc.2022.914594

The influence of metabolism on signaling, epigenetic markers, and transcription is highly complex yet important for understanding cancer physiology. Despite the development of high-resolution multi-omics technologies, it is difficult to infer metabolic activity from these indirect measurements. Fortunately, genome-scale metabolic models and constraint-based modeling provide a systems biology framework to investigate the metabolic states and define the genotype-phenotype associations by integrations of multi-omics data. Constraint-Based Reconstruction and Analysis (COBRA) methods are used to build and simulate metabolic networks using mathematical representations of biochemical reactions, gene-protein reaction associations, and physiological and biochemical constraints. These methods have led to advancements in metabolic reconstruction, network analysis, perturbation studies as well as prediction of metabolic state. Most computational tools for performing these analyses are written for MATLAB, a proprietary software. In order to increase accessibility and handle more complex datasets and models, community efforts have started to develop similar open-source tools in Python. To date there is a comprehensive set of tools in Python to perform various flux analyses and visualizations; however, there are still missing algorithms in some key areas. This review summarizes the availability of Python software for several components of COBRA methods and their applications in cancer metabolism. These tools are evolving rapidly and should offer a readily accessible, versatile way to model the intricacies of cancer metabolism for identifying cancer-specific metabolic features that constitute potential drug targets.

Keywords: cancer, metabolism, constraint-based modeling, genome-scale metabolic models, systems biology, omics, python, single-cell analysis

INTRODUCTION

Cancer involves a complex set of dysregulations in multiple biomolecular layers including metabolism. Metabolic changes in cancer result from and lead to profound changes in the behavior of cancer cells and their surrounding environment. Although extensively studied, these metabolic changes are difficult to accurately measure and model in an unbiased manner due to the need to consider a heterogeneous tumor environment encompassing different cell types, many difficult-to-measure metabolites, and lack of standardization of models (1). While recent years have yielded a wealth of methods to measure and analyze biological systems at multiple omics layers (genomic (2, 3), epigenomic (4), proteomic (5–8), and metabolomic (9–11), often extending to single-cell resolution (12), metabolic systems are difficult to systematically assess because gene expression or protein levels may not directly translate into metabolic activity (1).

Genome-scale metabolic models (GEMs) can provide a compelling approach towards understanding cellular metabolism. GEMs are curated computational descriptions of entire cellular metabolic networks. Derived from genome annotations and experimental data, GEMs are composed of mass-balanced metabolic reactions and gene-protein associations that map the relationship of genes to proteins involved in each reaction (**Figure 1**). The accumulation of high-throughput data has contributed to the reconstruction of GEMs for hundreds of organisms, from microbes and model organisms to animals and humans (13). Whole-organism GEMs can further be reduced into context-specific and cell type-specific models for analyzing specific tissue phenotypic states performing different cellular functions. Metabolic flux analyses of GEMs have led to various model-guided applications, such as hypothesis generation, strain design, drug target discovery, multicellular interactions modeling, and disease etiology (14–16). With the rapidly increasing availability of high-resolution multi-omics datasets, there is an increasing need for tools to interpret data using a mathematical framework that also integrates existing vast and complex biological knowledge. In particular, dysregulated metabolic systems in cancer interact heavily with the surrounding environment, and metabolic flux analysis may prove especially beneficial to modeling these systems.

Compared to omics analysis, cancer metabolism may be more accurately modeled by combination of GEMs and a family of methods called Constraint-Based Reconstruction and Analysis (COBRA). COBRA methods perform systems-level analyses on metabolic networks to uncover how genetic and environmental factors affect phenotype on a biomolecular basis. COBRA framework utilizes a stoichiometric matrix that transcribes mass-balanced metabolic reactions of a cellular system, including the system's uptake and secretion rates, into a matrix that represents the change in levels of reactants and products for each reaction (**Figure 1**). While there are many allowable states of reaction fluxes through a metabolic network, COBRA reduces this solution space of feasible flux distributions by adding constraints. Some basic constraints are mass conservation (stoichiometry of reaction and products in a reaction), steady-state assumption (input and output fluxes are balanced), and reaction flux bounds (inequalities of upper and lower bounds).

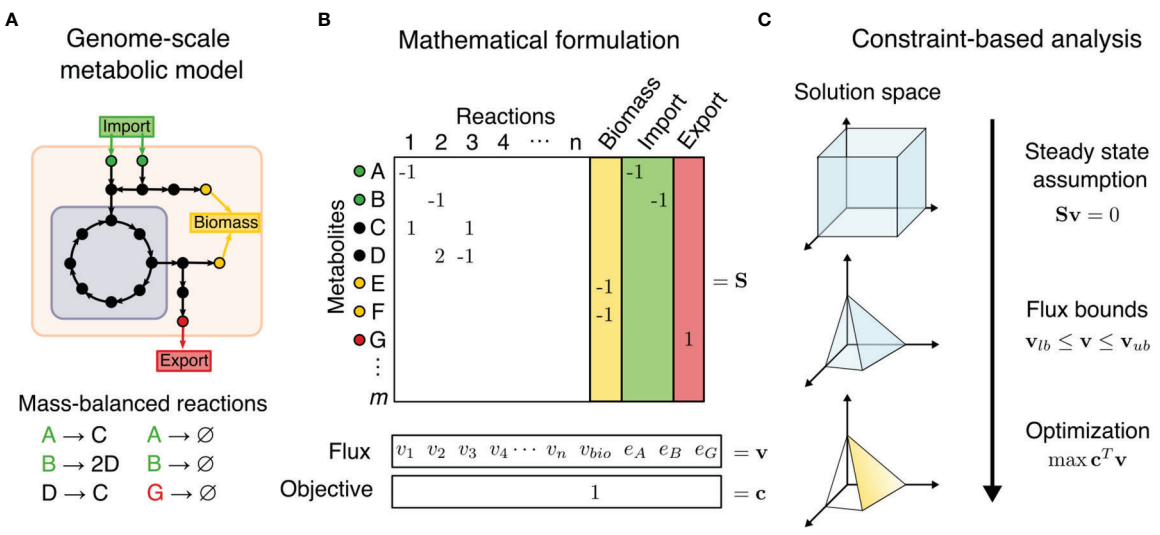


FIGURE 1 | Constraint-based metabolic modeling. **(A)** A genome-scale metabolic model is a compartmentalized network of mass-balanced reactions that convert products to reactants, and boundary pseudo-reactions that import or export metabolites. Biological objectives, such as biomass production, require activity through a subset of internal reactions. **(B)** The metabolic model is converted into a stoichiometric matrix (S) of size $m \times n$, with rows representing m metabolites and columns n reactions. Reaction flux through all internal reaction (v_i) and exchange reactions (e_i) is represented by vector v of length n. Objective function $Z = c^Tv$ is formulated as a linear combination of desired fluxes, weighted by vector c. **(C)** At steady state, the rate of production and consumption of a metabolite must be zero, which is described by the system of equations $Sv = 0$. There are many solutions to this system of equations, but the solution space can be constrained by imposing flux bounds ($v_{lb} \leq v \leq v_{ub}$) and optimization such as maximization of objective function.

Additional constraints can be determined by metabolite and enzyme levels, thermodynamics directionality, enzyme capacities, spatial compartmentalization, and genome regulatory mechanisms (15, 17). This induces a space of feasible fluxes which fulfill the used balance equations and constraints, often called the "flux cone". Constraint-based analysis methods then aim to find biologically relevant flux distributions within the flux cone.

COBRA methods for metabolic network analysis are now incorporated into many software packages across several programming languages like MATLAB and Python (15). Of these, MATLAB packages such as COBRA Toolbox, Raven Toolbox, and CellNetAnalyzer have been the leading standard platforms that integrate with many existing COBRA methods (18–20). However, the reliance on MATLAB, a proprietary and closed-source software, reduces the accessibility of metabolic flux analysis, especially for teaching and reproducibility purposes. Recent open-source community efforts have promoted the development of a similar ecosystem of COBRA software in Python, starting with the development of COBRApy (21) under the openCOBRA Project (22) and PySCeS CBMPy (23). As an open-source language, Python opens COBRA methods to greater possibilities by enabling deployment on machines without a proprietary license, which is especially convenient for cloud computing. Due to Python being widely adopted for data science and computation, it provides state-of-the-art scientific tools for accessing databases, integrating various data modalities, and interfacing with computational tools like parallel computing, machine learning, visualizations, and web applications.

This review will summarize the set of packages currently available in Python for various COBRA methods. We identify the advantages and shortcomings of the Python ecosystem to guide users' decisions on their choice of a software platform and inspire future research ideas. We focus on the application of COBRA methods to cancer metabolism. Finally, we will explore the future directions of COBRA methods development and their importance in cancer modeling.

COBRA METHODS IN PYTHON

To make COBRA open-source and accessible, multiple Python packages have been developed by the scientific community to perform the different analyses within COBRA. Here we describe the major components of COBRA and list their associated packages (**Figure 2**; **Table 1**), and assess their strengths and weaknesses (**Table 2**). First, we start with the core package COBRApy, which handles the details of metabolic models and basic simulations. We then describe methods for determining metabolic flux, such as flux balance analysis, flux variability analysis, and *in silico* perturbation. Next, we summarize various methods for adding biological constraints like multi-omics and biophysics. In addition, we review methods for unbiased pathway analysis and sampling methods. We also summarize the development of COBRA methods for models at the single-cell and population level. Finally, we touch upon packages for visualization and interactive web applications.

Modeling Framework

COBRA for Python (COBRApy) uses an object-oriented programming approach to represent models, metabolites, reactions, and genes as class objects with accessible attributes. Using this design, COBRApy recapitulates functions for standard metabolic flux analyses of its MATLAB counterpart while being extendible and accessible. First, it has the capabilities to read and write models in various formats such as MAT-file (for storing MATLAB variables), JSON, YAML, and Systems Biology Markup Language (SBML) (93), the current community-accepted standard for computational systems biology. SBML incorporates the Flux Balance Constraints (FBC) version 2 package (94), which supports constraint-based modeling by encoding objective functions, flux bounds, model components, and gene-protein associations, whose usage will be discussed below. COBRApy can also load SBML models from web databases such as BiGG and BioModels (95, 96). The quality of such metabolic models can be assessed using a Python test suite called MEMOTE that integrates version control of models *via* GitHub and checks for correct annotation, model components, and stoichiometry (24). To use these models for various optimization problems, COBRApy interfaces with either commercial or open-source solvers that implement linear programming algorithms. We will detail additional built-in or integrated functionalities for various COBRA methods (**Figure 2**).

Flux Balance Analysis

The most common COBRA method is flux balance analysis (FBA), which assumes the system is at steady state, follows mass-balance described in the stoichiometric matrix, and restricts reaction fluxes by bounds. Furthermore, FBA searches for sets of steady-state reaction fluxes that maximize or minimize an objective function representing a biological function, such as using biomass production objective to model cellular growth (29). The objective function is an artificial reaction formulated by linear combinations of reactions that would contribute to the desired biological function. For example, the biomass production can be represented by the consumption of biomass precursors in different proportions. Components of the biomass production may include amino acids, lipids, nucleotides, carbohydrates, cofactors, and other molecules based stoichiometrically on the macromolecular composition of a cell measured as weight fractions under specific experimental conditions, typically during exponential growth. Although the biomass equation is the de facto choice for the objective function and macromolecular compositions are more similar across related species, certain components such as fatty acids are sensitive to environmental and genetic conditions (97). Therefore, caution is required when choosing an appropriate objective function that reflects the system's experimental condition. Sensitivity analysis of FBA could be performed using different objectives (or ensemble of objectives) accounting for the natural variation in biomass equation across different conditions (97). Assessment of bias introduced by the objective function would require experimental validation of growth dynamics or knockout simulations discussed below.

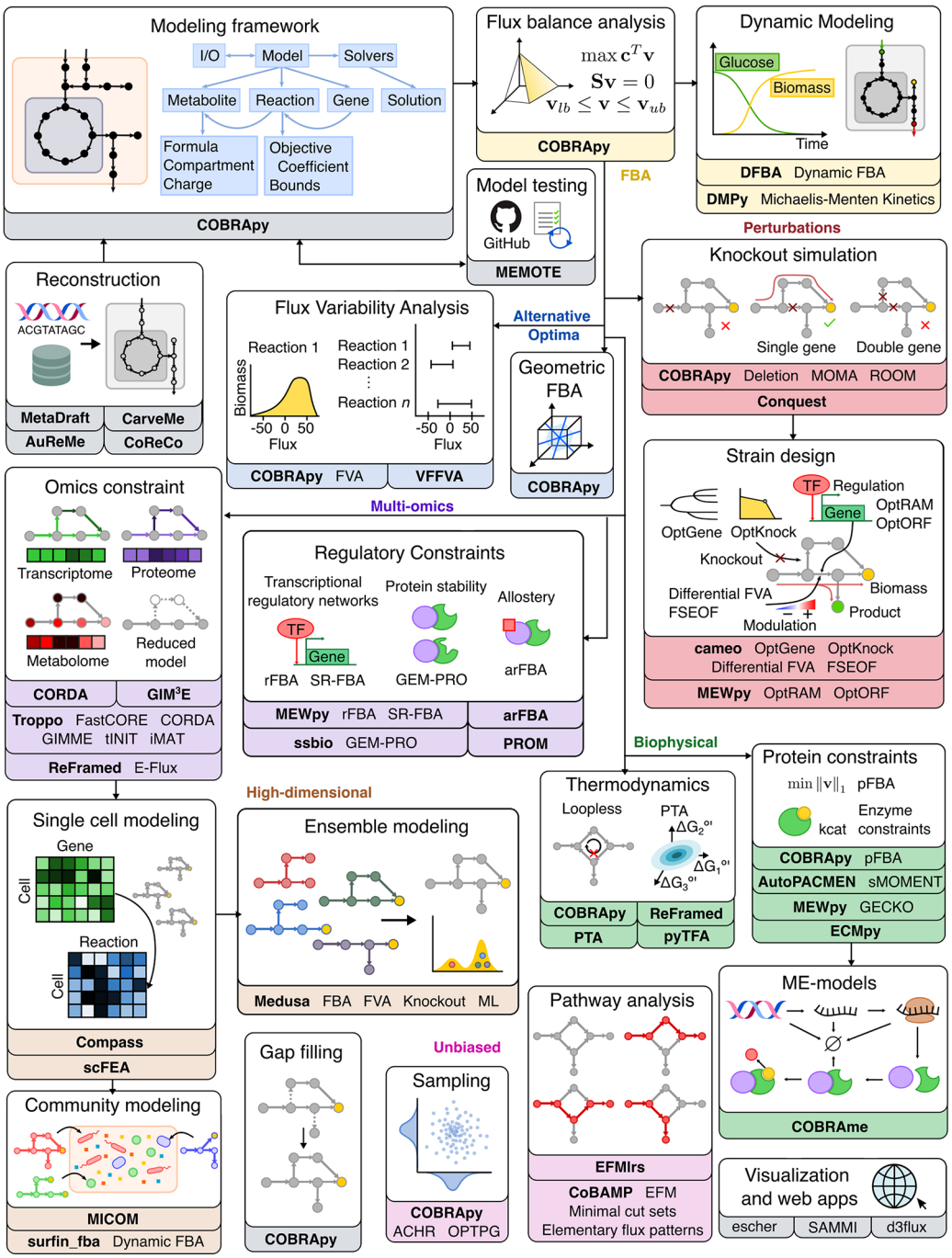


FIGURE 2 | Overview of Python software for major components of COBRA methods. Constraint-based metabolic modeling first requires loading a metabolic model into software that handles the various parts of the modeling framework (grey), such as metabolites, reactions, genes, stoichiometric matrix, and flux solutions. New metabolic models can be reconstructed from genome sequences and database, quality-checked by model testing software, made consistent using gap-filling tools, and visualized using web-based packages. Using the metabolic model, FBA (yellow) finds an optimal flux distribution that follows stoichiometry under steady state and can further be extended to dynamic systems. Since there are alternative optima (blue) to FBA, FVA and geometric FBA can be used to characterize the solution space. We can perturb (red) the system to predict the effect of knockouts and use such predictions to design an optimal system ('strain'). To improve FBA predictions, we can add biophysical (green) constraints based on thermodynamics, proteins, and macromolecular expression. Metabolic modeling can be further enhanced by integration of multi-omics (purple) data, such as extracting reduced models based on omics data and adding regulatory constraints. Using omics data, metabolic modeling can become high-dimensional (brown), through single cell modeling and community modeling. Multiple metabolic models can be reduced into ensemble objects. In contrast to FBA, unbiased (pink) approaches do not require an objective function. These include methods for sampling flux distributions and pathway analyses. Names of software packages are in bold.

TABLE 1 | Python tools for constraint-based modeling.

Category	Method	Software	URL	Doc.
Modeling framework	Object-oriented programming	COBRApy (21)	https://cobrapy.readthedocs.io https://github.com/opencobra/cobrapy	✓
	Testing	MEMOTE (24)	https://memote.readthedocs.io https://github.com/opencobra/memote	✓
Reconstruction	Template-based	AuReMe (25)	https://aureme.readthedocs.io http://aureme.genouest.org	✓
	Template-based, gap-filling	CarveMe (26)	https://carveme.readthedocs.io https://github.com/cdanielmachado/carveme	✓
	Template-based	MetaDraft (27)	https://systemsbioinformatics.github.io/cbmpy-metadraft/	✓
	Homology-based, multi-species, gap-filling	CoReCo (28)	https://github.com/esaskar/CoReCo	✓
FBA	FBA (29)	COBRApy	See above	✓
Dynamic metabolic modeling	Dynamic FBA (30)	dfba (31)	https://dynamic-fba.readthedocs.io https://gitlab.com/davidtourigny/dynamic-fba	✓
	Michaelis-Menten kinetics	DMPy (32)	https://gitlab.com/wurssb/DMPy	✓
Alternative optima	Geometric FBA (33) FVA (34)	COBRApy	See above	✓
	VFFVA	VFFVA (35)	https://vffva.readthedocs.io https://github.com/marouenbg/VFFVA	✓
Knockout Simulation	Single/Double deletions (36) MOMA (37) ROOM (38)	COBRApy	See above	✓
	Flux- and graph-based	Conquest (39)	https://github.com/laniauj/conquests	✓
Strain Design	OptGene (40) OptKnock (42) Differential FVA FSEOF (43)	Cameo (41)	https://cameo.bio/ https://github.com/biosustain/cameo	✓
	OptRAM (44) OptORF (46)	MEWpy (45)	https://mewpy.readthedocs.io https://github.com/BioSystemsUM/mewpy	✓
Omics constraints	E-flux (47)	ReFramed (48)	https://reframed.readthedocs.io https://github.com/cdanielmachado/reframed	✓
	CORDA	CORDA (49)	https://github.com/resendislab/corda	✓
	GIM^{3}E	GIM^{3}E (50)	https://github.com/brianjamesschmidt/gim3e	✓
	FASTCORE (51) CORDA (49) GIMME (53) tINIT (54) iMAT (55)	Troppo (52)	https://github.com/BioSystemsUM/troppo	✗
Regulatory constraints	rFBA (56) SR-FBA (57)	MEWpy	See above	✓
	PROM	PROM (58)	https://github.com/jseidel5/Python-Probabilistic-Regulation-of-Metabolism	✓
	GEM-PRO (59)	ssbio (60)	https://ssbio.readthedocs.io	✓
	arFBA	arFBA (61)	https://github.com/cdanielmachado/arfba	✗
Thermodynamics	ll-FBA (62) CycleFreeFlux (63)	COBRApy	See above	✓
	PTA	PTA (64)	https://probabilistic-thermodynamic-analysis.readthedocs.io https://gitlab.com/csb.ethz/pta	✓
	TFA, TVA (65)	ReFramed	See above	✓
	TFA, TVA (65)	pyTFA (66)	https://pytfa.readthedocs.io https://github.com/EPFL-LCSB/pytfa	✓
Protein constraints	pFBA (67)	COBRApy	See above	✓
	GECKO (68)	MEWpy	See above	✓
	sMOMENT	AutoPACMEN (69)	https://github.com/klamt-lab/autopacmen	✓
	ECMpy	ECMpy (70)	https://github.com/tibbdc/ECMpy	✓
ME-modeling	COBRAme	COBRAme (71)	https://cobrame.readthedocs.io	✓
Gap filling	MILP	COBRApy	See above	✓
Ensemble modeling	FBA FVA Deletion ML	Medusa (72)	https://medusa.readthedocs.io/ https://github.com/opencobra/Medusa	✓

(Continued)

TABLE 1 | Continued

Category	Method	Software	URL	Doc.
Single cell modeling	Compass	Compass (73)	https://yoseflab.github.io/Compass/ https://github.com/YosefLab/Compass	✓
	scFEA	scFEA (74)	https://github.com/changwn/scFEA	✓
Community modeling	MICOM	MICOM (75)	https://micom-dev.github.io/micom/ https://github.com/micom-dev/micom	✓
	Dynamic FBA	surfin_fba (76)	https://github.com/jdbrunner/surfin_fba	✓
Sampling	ACHR (77) OPTPG (78)	COBRApy	See above	✓
Pathway analysis	EFM	EFMlrs (79)	https://github.com/BeeAnka/EFMlrs	✓
	EFM (80) Minimal cut sets (82) Elementary flux patterns (83)	CoBAMP (81)	https://cobamp.readthedocs.io https://github.com/BioSystemsUM/cobamp	✓
Visualization and web apps	Plug-in, website	Escher (84)	https://escher.readthedocs.io https://escher.github.io	✓
	Plug-in, website	SAMMIpy (85)	https://sammipy.readthedocs.io www.SammiTool.com	✓
	Plug-in	d3flux (86)	https://pstjohn.github.io/d3flux/ https://github.com/pstjohn/d3flux	✓

Methods and their associated software packages as illustrated in ***Figure 2****, organized by their general function. Weblinks to each software's official website and documentation are provided if available. Each software is assessed for availability of documentation (Doc.) or any form of demonstrative examples.*

Under the above mathematical constraints, FBA is an optimization problem involving a system of equations that can be solved by linear programming, as initially proposed in 1984 (98). Functions for FBA and customization of objective functions are included in COBRApy. With these basic constraints, FBA is the foundation from which many forms of COBRA methods evolved.

Dynamic Metabolic Modeling

Although FBA assumes that a system is unchanging at steady state, these pseudo-steady states can be coupled to a dynamical system with changing environmental variables using dynamic FBA (DFBA) (99). There are several approaches to DFBA: 1) dynamical optimization approach (DOA) that uses ordinary differential equations (ODEs) to describe an optimization problem of entire time profiles of metabolites, 2) statistic optimization approach (SOA) that divides the time period into time intervals to perform instantaneous optimization (LP) per time interval with flux rate-of-change constraints, 3) direct approach (DA) that resolves the LP of the right-hand side of ODEs, and 4) reformulation of the ODEs as differential-algebraic equation (DAE) system (30, 89, 99). The fourth approach via DAE is implemented in Python package dfba (31), while the second approach *via* SOA can be implemented using COBRApy and SciPy. Alternatively, a very different approach to dynamical metabolic modeling was proposed by DMPy, which translates a GEM into a dynamic reaction equation model using Michaelis-Menten approximations and infers missing kinetic constants using Bayesian parameter estimation with thermodynamics constraints (32). However, this method requires extensive measurements of reaction rates to accurately parameterize a large-scale model. All these constraint-based methods for dynamical metabolic modeling enable the utilization of high-throughput and longitudinal data to interrogate changes in metabolism.

Alternative Optimal Solutions

Flux distributions, even under an optimal objective, are usually not unique as many alternative fluxes can yield a maximum biomass production. The most representative solution can be found using geometric FBA in COBRApy, which looks for a unique flux distribution that is central to the entire solution space (33). To better characterize all alternative optima that satisfy the constraints of FBA, flux variability analysis (FVA) finds the range of alternative fluxes for a reaction that maintains optimization of the objective function within a margin of error (34). The search for alternate optimal solutions is time-intensive, but COBRApy has addressed this problem in FVA by implementing parallel computing. For example, Very Fast Flux Variability Analysis (VFFVA) is available in Python and its implementation of FVA is much faster and more memory-efficient than its analog in MATLAB, fastFVA (35).

System Perturbations, *In Silico* Knockout, and Strain Design

Quantitative flux predictions are useful to experimentalists because of their potential to explain or even predict the effect of environmental and genetic changes. For investigating the relationship between the external environment and the modeled system, COBRApy provides tools for specifying the growth medium and exchange rates of a model. Instead of extracellular conditions, intracellular changes such as genetic mutations and gene modulation can be interrogated as well. To identify essential genes and reactions for biological functions, FBA is performed with gene knockout simulations to assess the effects of the knockouts on objective functions (36). Similar to COBRA Toolbox, COBRApy includes functions for knocking out single or double genes and reactions by restricting the flux through associated reactions. Another algorithm for assessing the effect of a perturbation is minimization of metabolic adjustment (MOMA), which determines the post-perturbation

TABLE 2 | Pros and cons of COBRA methods.

Category	Method/Tool	Pros	Cons
Reconstruction	AuReMe	- Support for eukaryotes model - Good traceability - Automatic integration of experimental data	- Some manual refinement assistance - Not FBA-ready
	CarveMe	- GEMs ready for FBA - Fast - Customizable for large number of genomes	- No manual refinement assistance - Some support for eukaryotes model
	MetaDraft	-Support for eukaryotes model - Fast	- No manual refinement assistance - Not FBA-ready
	CoReCo	- Support for eukaryotes model - GEMs nearly ready for FBA - Simultaneous reconstruction for multiple species (parallelizable)	- Requires KEGG license - No manual refinement assistance
FBA	FBA	- Does not require kinetic parameters	- Requires objective function - Requires reaction bounds (especially exchange flux)
Dynamic modeling	Dynamic FBA (SOA and DAE)	- Couples pseudo-steady states to dynamical systems - Does not require kinetic parameters	- SOA requires small steps and thus more computation
	DMPy	- Infers missing kinetic parameters using thermodynamics constraints	- Requires >80% of kinetic parameters for accuracy
Alternative optima	Geometric FBA	- Gives single representative solution – Reproducible typical solution (avoids randomly picking one solution from flux cone)	- Weak correlation with protein levels (without omics constraint)
	FVA/VFFVA	- Determines min and max flux for a reaction would achieve optimal objective state - (VFFVA) Increased speed and reduced memory usage	- Varies one reaction at a time
	Sampling	- Estimates probability distribution of feasible fluxes - Can be unbiased (not using an objective function)	- Computationally intensive
Omics constraints	E-flux	- Constraints reaction bounds only - No discretization of data	- May over-constrain model based on noisy data - Poor growth rate prediction
	GIMME	- LP problem (fast) - Ensures operability of required metabolic function - Predicts growth rate, uptake/secretion rates, essential genes, and oncogenes	- Discretizes data - Models have high fractions of blocked reactions, moderate resolution power, poor robustness to missing data/noise
	GIM^{3}E	- Ensures operability of required metabolic function - Integrates metabolomics data	- Discretizes data - MILP problem (slow)
	(t)INIT	- Ensures operability of required metabolic functions - (INIT) predicts oncogenes and tumor suppressor genes, consistent model, good resolution power, robust to noise/missing data	- MILP problem (slow) - (INIT) Poor predictions of growth rate, uptake/secretion rates, and essential genes
	iMAT	- No objective required - Consistent model, good resolution power, robust to noise/missing data - Predicts oncogenes	- Discretizes data - MILP problem (slow) - Weak predictions of growth rate, uptake/secretion rates, and essential genes
	FASTCORE	- LP problem (fast) - Obtains minimal consistent model - Predicts oncogenes and loss of function mutations - Moderately consistent model, good resolution power, robust to noise	- Requires specification of core reactions - Poor predictions of growth rate, uptake/secretion rates, and essential genes
	CORDA	- LP problem (fast) - Non-parsimonious pruning - Predicts oncogenes and loss of function mutations	- Requires specification of core reactions - Weak predictions of growth rate and essential genes - Poor predictions of uptake/secretion rates
Regulatory constraints	rFBA	- Predicts flux over time intervals - Models transcriptional regulation	- Uses boolean TRN - Stepwise calculation of metabolic and regulatory states - Chooses only one solution per time interval
	SR-FBA	- Combined calculation using metabolic and regulatory constraints - Models transcriptional regulation	- Uses boolean TRN - Calculates flux for one time step (steady-state) - Does not account for metabolic transitions and feedback loops
	PROM	- Uses continuous TRN - Models transcriptional regulation	- Requires TF-target gene relationships
	GEM-PRO	- Models protein instability	- Requires protein structures
	arFBA	- Models allosteric regulation	- Requires regulation matrix defining effector-reaction relationship - Small-scale applications
Thermodynamics	ll-FBA	- Does not require metabolite concentrations or free energies	- MILP problem (slow)
	CycleFreeFlux	- Post-process using LP problem (fast) - Can be applied to any flux distribution including sampled solutions - Does not require metabolite concentrations or free energies	- Biased towards solutions with small total flux and those with same direction as their overlapping internal cycles

(Continued)

TABLE 2 | Continued

Category	Method/Tool	Pros	Cons
	TFA, TVA	- Explicitly models thermodynamics	- Requires metabolite concentrations and free energies - Over-approximates uncertainty
	PTA	- Explicitly models thermodynamics for optimization and sampling - Models uncertainty of free energies and metabolite concentrations	- Requires metabolite concentrations and free energies - Computationally intensive
Protein constraints	pFBA	- Predicts growth rate, uptake/secretion rates, and essential genes	- Assumes that flux distribution with smallest magnitude minimizes protein costs
	Enzymatic constraints (GECKO, sMOMENT, ECMpy)	- Model proteome limitation at enzyme resolution - (sMOMENT) Automates enzyme database query - (ECMpy) Automates enzyme parameters calibration - (ECMpy) Does not increase model size	- Requires experimentally measured enzyme turnover numbers - (GECKO) Increases model size - (sMOMENT) Moderately increases model size - (ECMpy) Manually obtains protein subunit composition data
ME-modeling	COBRAme	- Modeling proteome composition improves predictive accuracy - Framework for building ME-models for new organisms	- Large model size and complexity - No standardized SBML format for ME-models - Only applied to bacteria so far
Ensemble modeling	Medusa	- Compresses multiple models into compact ensemble objects - Reduces memory usage of storing ensembles - Interfaces with machine learning	- No standardized SBML format for ensemble objects
Single cell modeling	Compass	- Genome-scale modeling - Maximizes agreement with gene expression - Handles sparsity by sharing information across neighbors - Uses multiple objective functions	- Map gene expression to reaction expression using boolean relationships (GPR)
	scFEA	- Minimizes flux imbalance of all cells to simulate exchange of metabolites - Less stringent flux balance and steady-state assumption - Uses neural net to model nonlinear relationship between gene expression and reaction rates	- Not easily scalable due to large memory usage - Applied to small-scale models
Community modeling	MICOM	- Models exchanges and interactions between communities and environment - Automates building community models from a model database - Predicts replication rates in human gut microbiome	- Assumes trade-offs between individual and community growth rate (gut microbiome specific) - Metabolic models may not be accurate (labratory vs. gut conditions, species differences)
	Dynamic FBA (surfin_fba)	- Reduces optimizations problems (and parameter space) required for dynamic FBA for communities	- Non-biological approach to choosing between non-unique optima
Pathway Analysis	EFM	- Unbiased characterization of models (no objective function required) - (EFMlrs) Pre- and post-process models for EFM calculations	- (EFMlrs) EFM calculation performed by other tools not included in program - EFM calculations are memory intensive and not scalable

Some method comparisons extracted from literature for reconstruction (87, 88), dynamic modeling (89), omics constraints (90, 91), and regulatory constraints (92). Growth rate, uptake/secretion rates, and cancer essential gene prediction performances from Jamialahmadi et al. are based on human metabolic models and available only for GIMME, INIT, iMAT, FASTCORE, CORDA, and pFBA (91).

flux vector that is closest to a reference flux vector (e.g., FBA solution before change) (37). Currently, COBRApy implementation of MOMA is the only one that does not require a commercial quadratic programming solver but instead uses OSQP, which is an open-source solver (100). Another method, called Regulatory-on-off minimization (ROOM), finds the new flux distribution with minimal reaction changes compared to a reference state (38). Available in COBRApy, these methods characterize the effects of gene deletion relative to a wild-type reference. Adding to flux-based determination of essentiality, a new metabolite essentiality analysis combining graph-based and flux-based analysis was proposed by Conquests (Crossroad in metabOlic Networks from Stoichiometric and Topologic Studies) (39).

The iterative testing of gene or reaction deletions was initially developed for *in silico* strain design, which determines optimal genetic changes that would maximize production of desired metabolites. Straight maximization of only the desired reaction is problematic, since it ignores the drainage of cellular resources needed for cellular growth. Therefore, strain design methods couple product yields with cellular objectives to optimize for fast-growing cells that have high productivity. Such metabolic engineering tools are available in a COBRApy-derived package called cameo (41). It provides efficient, parallelized implementations of standard *in silico* strain design methods for predicting gene knockout strategies (OptGene [evolutionary algorithm] (40), OptKnock [linear programming] (42) and for predicting gene expression modulation targets (Differential FVA, Flux Scanning based on Enforced Objective Flux [FSEOF] (43). Instead of modulating genes, there are algorithms that optimize at the regulatory level by changing transcription factors, such as OptRAM (44) and OptORF (46) in MEWpy (Metabolic Engineering Workbench in python) (45). These simulation tools for strain design and *in silico* knockouts/perturbations can be easily adapted to study metabolism in the context of physiology and disease, especially cancer. For example, we will later discuss studies that use *in silico* knockout to screen for cancer drug targets. Other studies integrated genetic variants by

simulating knock out of enzymes with loss of function mutations (101–103).

Integrating Multi-Omics Data With GEMs

Integration of omics data into metabolic models is now critical to standard analysis of GEMs to improve flux predictions and interpret multi-omics data. Prior to applying constraints, gene-level data must first be processed to reflect reaction-level data. This involves calculating a reaction expression matrix that evaluates gene-protein-associations (GPR, nested logic rules representing gene essentiality and redundancy). For example, we take the minimum expression of required subunits, but take the sum of isozyme expression. This calculation can be performed in Python packages like CORDA (Cost Optimization Reaction Dependency Assessment) (49) and MEWpy (45). Marín de Mas et al. further improved GPR evaluation in their Python implementation of stoichiometric GPR (S-GPR) that considers the stoichiometry of protein subunits (104).

The resulting reaction expression levels are used subsequently to extract a context-specific metabolic model of active reactions from the whole-organism GEM to reflect a phenotypic state specific to cell type and condition, such as disease state or nutrient level. The simplest transcriptome constraints can be applied by setting associated expression levels as the reaction upper bound, as demonstrated in E-flux and other studies (47, 105, 106). Instead of constraining all genes, PRIME is method that adjusts reaction upper bounds of phenotype-associated genes that are correlated with phenotypic data such as growth rate (90). Additional methods for extraction of context-specific models from transcriptome, metabolome, and proteome have been reviewed previously and can be summarized into three main families of approaches (107): 1) GIMME-like (GIMME (53), GIM^{3}E (50), tINIT (54)), which aims to maximize the correspondence of flux phenotype to data while maintaining required metabolic functions; 2) iMAT-like (iMAT (55), INIT (108), Lee-12 (109), which only maximizes similarity of flux phenotype to data; and 3) MBA-like (MBA (110), mCADRE (111), FASTCORE (51), FASTCORMICS (112), CORDA (49), which removes non-core reactions while ensuring consistency of the model. Currently, integration of these methods with COBRApy is still in development within the DRIVEN project (113). Fortunately, some of these reconstruction methods have been reimplemented in other Python packages (**Table 1**). For example, ReFramed implemented E-flux (48), CORDA and GIM^{3}E have standalone Python packages, and Troppo implemented FASTCORE, CORDA, GIMME, tINIT, and iMAT (52). Nonetheless, the Python ecosystem has shortcomings in reconstruction methods, such as the unavailability of some methods (INIT, MBA, mCADRE, FASTCORMICS, and PRIME), and the lack of documentation and usage examples for the Troppo package.

Reconstruction methods could result in incomplete and infeasible networks, partly due to errors in experimental data and curated knowledge, and partly due to parsimonious approaches when pruning reactions. To make reconstructed models feasible, one can use the gap-filling functionality in COBRApy to infer missing pathways using mixed-integer linear program (MILP). However, due to stochasticity and existence of alternative optima, GEM reconstruction and gap-filling of the same network can give rise to multiple GEMs that could yield different flux predictions. To account for the uncertainty in network structure, ensemble modeling compresses such a set of alternative models into an ensemble object to reduce redundancy while capturing variation. Ensemble modeling can be performed through Medusa, a Python package for generating ensembles, performing ensemble simulations, and coupling ensembles with machine learning (ML) (72).

Despite reconstruction of context-specific GEMs, GEMs are still flawed in flux prediction due to their inability to account for cellular mechanisms that regulate metabolic activity. A recent review has outlined the major methods for integrating regulatory mechanisms into metabolic models as the following: transcriptional regulatory networks (TRNs), post-translational modifications, epigenetics, protein–protein interactions and protein stability, allostery, and signaling networks (92). Several methods using TRNs have been translated from MATLAB to Python (**Table 1**), including boolean TRN methods like regulatory FBA (rFBA) (56) and steady-state regulatory FBA (SR-FBA) (57) available *via* MEWpy, and a continuous TRN method called probabilistic regulation of metabolism (PROM) (58, 114). Other regulatory mechanisms are also available: 1) GEM-PRO (59) integrates protein structure information, and 2) arFBA (61) integrates allosteric interactions respectively. However, methods for integrating post-translational modifications, epigenetics, and signaling networks are not yet available in Python. Future development is needed to account for the complex cellular regulatory activity.

Extraction of context-specific GEMs requires a reference GEM that is often manually curated. To automate the laborious process of GEM reconstruction, several tools were developed to reconstruct microbial GEMs from genome sequences (87). Several examples of Python-based software are AuReMe (25), CarveMe (26), MetaDraft (27), and CoReCo (28). Among these, Mendoza et al. (87) reviewed the first three and found them all to generate GEMs that have high reaction sets similarity to manually curated models, but only CarveMe generates GEMs ready-to-use for FBA (**Table 2**). A more recent tool called gapseq (88) was shown to outperform CarveMe, but it is written in shell-script and R.

Biophysical Constraints

To ensure that reaction directionalities in computational results agree with biological findings, COBRA methods include addition of thermodynamic constraints *via* removal of thermodynamically infeasible pathways or calculations of Gibbs free energy. The vastness of solution space can also be attributed to thermodynamically infeasible loops where metabolites are cycled infinitely. COBRApy includes two implementations for removing such loops: one method ll-FBA (add_loopless) utilizes mixed-integer linear programming (62), and another faster method CycleFreeFlux (loopless_solution) uses postprocessing of solutions (63). Additionally, there are other Python packages that interface with COBRApy to

implement thermodynamics analysis. For example, probabilistic thermodynamics analysis (PTA) models use joint probability distributions of free energies and concentrations for stream optimization and sampling flux analysis (64). Earlier methods such as thermodynamic flux analysis (TFA) and thermodynamic variability analysis (TVA) (65) were implemented in ReFramed (48). Another Python package for thermodynamic-based flux analysis (pyTFA) couples thermodynamics feasibility into FBA calculations (66). Thermodynamics constraints ensure physiological flux predictions and help to reduce the solution space.

Another theme of biophysical constraints involves modeling the proteome limitation of a cell due to molecular crowding in a cell. A simple method within this theme is parsimonious FBA (pFBA), which assumes that minimizing overall total flux approximately finds efficient pathways that minimizes the total enzyme mass (67). Available in COBRApy, pFBA first determines the maximum value of the objective function, then adds it as a model constraint and solves for the flux distribution with the smallest magnitude, minimizing protein costs (67). However, this assumption may not always hold for all conditions and complex cellular networks. Another way to limit proteins is to add constraints based on enzyme parameters such as turnover number (k_{cat}) and molecular weight. These protein allocation constraints are applied by Python package MEWpy using a method called GECKO (Genome-scale model enhancement with Enzymatic Constraints accounting for Kinetic and Omics data), which adds many pseudo-metabolites and pseudo-reactions to represent enzymes (68). Another package for protein allocation constraints is AutoPACMEN (Automatic integration of Protein Allocation Constraints in MEtabolic Networks) (69). AutoPACMEN can automate database query and creation of models using sMOMENT (short metabolic modeling with enzyme kinetics), which introduces only one pseudo-reaction and pseudo-metabolite. Further improving upon these methods, ECMpy adds enzyme constraints without increasing model size (70). Studies have shown that adding protein constraints improves the accuracy of flux predictions by explaining suboptimal overflow metabolism and metabolic switches (69, 70). Instead of high-level protein constraints, the machinery cost of protein expression can be explicitly modeled using genome-scale models of metabolism and macromolecular expression (ME-models). ME-models extend GEMs by computing optimal composition of macromolecules like proteins, nucleotides, and cofactors, to model the entire process from transcription and translation, to complex formation and metabolic reaction. Software for building and simulating ME-models is currently only available in Python *via* COBRAme (71) and was extended to dynamic systems *via* dynamicME (115). All packages for protein constraints mentioned above are compatible with COBRApy.

Unbiased Characterization of Solution Space

There are unbiased methods for analyzing distribution of steady-state flux through a metabolic model. One set of unbiased methods performs network-based pathway analysis without knowledge of traditional pathway annotations: elementary flux mode (EFM) analysis finds the minimum reaction sets (i.e., pathways) that can maintain steady state. Different variations of EFM have been implemented in Python. For example, EFMlrs is a Python package that performs EFM enumeration *via* lexicographic reverse search, an implementation that significantly improves performance and memory usage (79). In addition, CoBAMP is another package that has implemented EFM (80), minimal cut sets (82), and elementary flux patterns (81, 116). Extreme pathway (ExPa) analysis is another method for identifying reaction sets but it is not currently available in Python (83).

Another set of unbiased methods is Markov chain Monte Carlo (MCMC) sampling methods, which can characterize the solution space by estimating the probability distribution of feasible fluxes. This could be performed with or without constraining by an objective function. Currently, COBRApy integrated MCMC methods such as artificial centering hit-and-run (ACHR) (77) and optimized general parallel (OPTPG) (78) samplers, but not coordinate hit-and-run with round (CHRR) (117) that was found to be the best performing (118).

Single-Cell Metabolic Modeling

Our ability to interrogate the heterogeneity of cell populations has grown rapidly due to advances in single-cell technologies that can measure the transcriptome, proteome, epigenome, and even metabolome at the single-cell level (2–7, 11, 12, 119–126). While single cell multi-omics data can be analyzed by pathway enrichment, clustering, and correlation methods (16, 122, 123), recent studies have developed algorithms in Python to calculate metabolic flux from single-cell transcriptome (119, 127). Zhang et al. demonstrated the usage of CORDA for the reconstruction of cell type-specific metabolic models from murine single-cell transcriptome and their subsequent FBA simulations of NAD^+ biosynthesis using COBRApy (128). Instead of optimizing for a specified objective function, Compass is an FBA-based method that scores the ability of cell transcriptome to maintain high flux through each reaction (73). Rather than using linear programming to solve for flux distribution, scFEA first reconstructs a metabolic model into a directed factor graph, then trains a deep neural network to learn metabolic flux distributions by minimizing flux imbalance across all cells and maximizing correspondence with gene expression (74). Due to drop-outs in single-cell RNA-seq, these algorithms took different approaches to handle the sparsity of expression data: 1) Zhang et al. calculated mean expression profiles per tissue and cell ontology class, 2) Compass allows information sharing between cells that are similar in transcriptional space, and 3) scFEA trains the model on all cells and removes metabolic modules only if they are entirely composed of significantly unexpressed genes. These methods allow metabolic flux interpretation of single-cell transcriptome at the single-cell resolution; however, not all flux estimation methods account for the interaction of cells *via* uptake and secretion of metabolites into the environment.

Multicellular Metabolic Modeling

To account for metabolic interactions, multicellular modeling was devised to model interplay between multiple metabolic networks coming from different species or tissues, with applications from microbiology to human physiology (129). Community modeling of the human gut microbiome reveals community-level function and cross-feeding interactions, as demonstrated by Python package MICOM (75). Community models are further extended using dynamic FBA of microbial communities, which can be efficiently calculated using Python package called surfin_fba that reduces the number of optimization timesteps when modeling communities (76). Early attempts to model human cell populations were explored using MATLAB, beginning with popFBA that simulated clones of cancer cells with identical stoichiometry and capacity constraints while allowing extracellular fluxes (130). PopFBA searched for combinations of individual metabolic flux distributions that would maximize a population object, e.g., total biomass, to explore metabolic heterogeneity and cooperation between single cells. However, this method gives many possible solutions and ignores the differences in metabolic requirements, functions, and proliferation rates of heterogeneous populations. To address both issues, single-cell FBA (scFBA) in MATLAB optimizes individual objective functions within a multi-scale model constrained by single-cell transcriptome and bulk extracellular fluxes to reduce the solution space (131). Overall, the added complexity of multicellular modeling can improve our interpretation of omics data and provide insights into cell-cell interactions important to many biological systems.

Visualization and Web Application

While algorithm development for COBRA is important, the utility of COBRA methods also depends on the usability and dissemination of scientific results. Python libraries have enabled the development of more interactive, user-friendly applications for analysis and visualization of metabolic networks. For example, Escher is a web application for visualizing metabolic models and also a Python package with interactive widgets for Jupyter Notebooks that can visualize COBRApy models (84). Escher has been integrated into other Python COBRA packages such as cameo to visualize flux analysis results. Additional interactive visualization packages include SAMMI for semi-automated visualization and d3flux for d3.js based plots (85, 86). Due to open-source nature of Python packages, future COBRA web applications can be deployed for public use without licensing limitations.

GENOME-SCALE MODELING OF CANCER METABOLISM WITH COBRA TOOLS

Cancer cells undergo metabolic reprogramming to promote proliferation and invasion, and in turn alter the nutrient-levels and cell types within the tumor microenvironment (TME). We here summarize these metabolic changes and provide the rationale for using COBRA methods to analyze cancer metabolism and TME. Indeed, COBRA methods have been utilized for various applications in cancer research in the past decades. We describe how the analyses begin with building cancer-specific metabolic models, from which one can infer metabolic dysregulation through pathway and network analyses. Next, we showed how these models were used for quantitative prediction of cancer metabolic activity and drug targets. Finally, we highlight the frontiers of modeling the TME using multicellular or single-cell COBRA methods.

Metabolism of Cancer and the Tumor Microenvironment

The dramatic functional and environmental changes that occur during cancer formation and progression are accompanied by accordingly dramatic metabolic reprogramming in cancer cells (**Figure 3**). These changes canonically include the Warburg effect (132, 133), the switch from predominantly mitochondrial oxidative phosphorylation to aerobic glycolysis, potentially done to increase biomass production critical to maintain high proliferation (133); this leads to increased glucose uptake and lactate secretion by cancer cells. Increased energy and biomass production in cancer cells is also associated with increased uptake and synthesis of amino acids (134), fatty acids (135), and nucleotides (136). The TME is also quite distinct from normal physiology as it espouses a different set of spatial structures, nutrient/metabolite compositions, and cellular heterogeneities, and thus the metabolism of cancer cells is further perturbed just as the cancer cells metabolically influence the TME in turn (137). In the TME, tumor cells also inhibit immune cells by outcompeting them for critical nutrients with finite supply, such as glucose and amino acids, thereby limiting immune anti-tumor activity. The manifold metabolic changes that occur in cancer pose a challenging question to faithfully model. However, overcoming this challenge to establish an accurate model of this complicated metabolic reprogramming may prove useful for identifying potential targets, such as cell-cell metabolic interactions between tumor and immune cells, for cancer therapy.

COBRA methods offer a way to computationally achieve this goal, such as inferring metabolic state *via* FBA which requires an objective function. While designing an objective function for tissue-specific eukaryotic cells is usually challenging, cancer cells can be reasonably modeled by biomass objective function, because cancer is mainly characterized by cellular growth (138). This makes flux predictions better suited for modeling cancer than healthy tissues, which do not actively proliferate. To simulate flux through cancer GEMs, studies have used objective functions representing growth as consumption of biomass precursors (139), or individual required metabolic tasks such as energy and redox, internal conversions, substrate utilization, biosynthesis, and biomass growth (54, 108). Some studies found gene-essentiality predictions from GEMs to be robust to definition of biomass composition (139) and capable of predicting growth kinetics in small-scale model (140), suggesting that the biomass equation is not significantly biased.

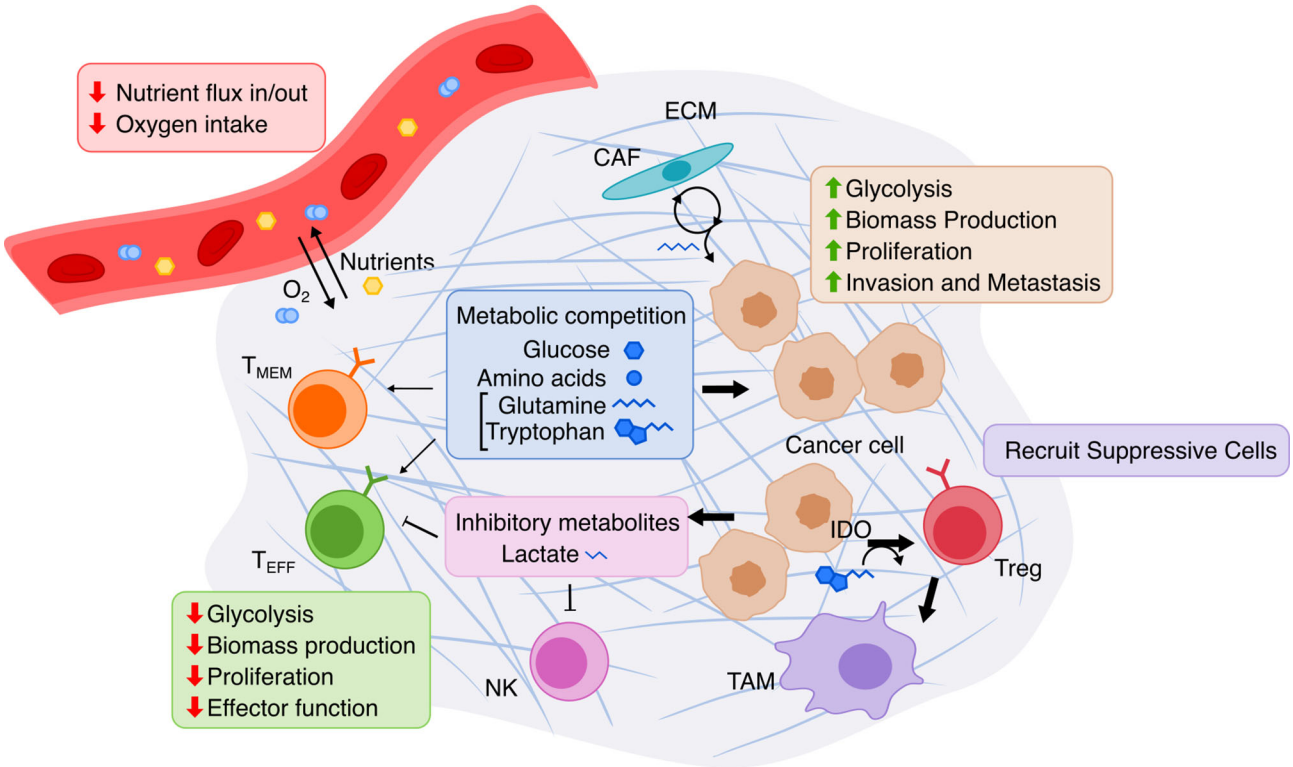


FIGURE 3 | Overview of metabolic interactions within the tumor microenvironment. The TME is composed of cancer cells, immune cells, and stromal cells embedded in extracellular matrix (ECM). Limited nutrients and oxygen lead to metabolic competition between cancer and various lymphocytes, especially hampering anti-tumor activity of effector T cells (T_{EFF}). Cancer cells adapts *via* upregulating nutrient transport and altering cancer-associated fibroblasts (CAF) to replenish metabolites. T cell immunity is further suppressed by cancer cells' release of lactate produced by glycolysis and by recruitment of immune-suppressive cells due to Indoleamine 2,3-dioxygenase (IDO) activity. T_{MEM}, memory T cell; NK, natural killer cell; Treg, regulatory T cell; TAM, tumor-associated macrophage.

However, another small-scale model claimed that elemental mode flux predictions using lactate objective is better than biomass objective at predicting experimental fluxes (141). These differences emphasize the importance of experimental validation to look for bias and sensitivity analysis to see if our biological insights are heavily affected by objective function definition and other system assumptions. Furthermore, the assumption that cancer cells optimize for cell growth may not always hold as tumors adapt, especially under selective pressure from therapies and immune system to adopt a quiescent state (138). Even if a proper objective is used, there are many optimal FBA solutions, and some may not be biologically viable due to inaccurate reaction bounds, violation of steady-state assumption, regulatory processes, and other limitations to our biological knowledge. Despite these limitations, past cancer applications of COBRA methods strived to improve our understanding of the disease and identify drug targets *via* comparative analysis, network analyses, quantitative flux simulations, and TME modeling. These studies have been reviewed multiple times (13, 14, 138, 142–144), and we have compiled the collection of these studies in **Table 3** and summarized their applications below (**Figure 4**).

Reconstruction of Cancer Metabolic Models

To date, numerous efforts have iteratively improved reconstruction of the human metabolic network within the Recon series (Recon 1, 2, 3D) (103, 183, 184), the Human Metabolic Reaction (HMR) series (HMR 1 and 2) (185, 186), and their derived unified model Human1 (187) (**Table 4**). From these generic human GEMs, cancer-specific metabolic models were generated by integrating multi-omics data to reduce the number of reactions to reflect cancer-specific activity. To extract multiple healthy and cancerous tissue-specific GEMs, studies utilized protein levels from Human Protein Atlas along with INIT algorithm (108) or CORDA algorithm (170). Other studies constructed cancer GEMs using transcriptomic data from 1) cancer cell lines in combination with different integration algorithms such as MBA (139), tINIT (150), a likelihood-based method (156), PRIME (90), and FASTCORMICS (112), or 2) transcriptomic data from tissue samples in combination with mCADRE algorithm (111). While transcriptome measurements can capture more genes, its data is noisy and does not correlate well with protein levels (190). In contrast, proteomic data more directly corresponds to enzymatic activity, but was previously

TABLE 3 | List of cancer metabolic modeling studies.

Ref.	Cancer	Purpose					Method				
		Drug design	Predict drug targets	Explore Cancer biology	Explain Warburg effect	Patient classification	Integration	Model	Analysis	Constraints	Objective
(145)	Breast	x		x			Lee-12	HMR 1	FBA, Comparative, Topological	Transcriptome, Fluxomic	Data Correlation
(146)	Colorectal	x		x			tINIT	Human1	TFA, TFVA, pTFVA	Thermodynamic, Transcriptome, Biomass	Biomass
(147)	Eye	x		x			iMAT	Recon 2	Gap filling, FBA, FVA, Knockout	Transcriptome	Biomass, Tasks
(148)	Head and Neck	x		x			Upper bound	Recon 2	FBA, Sampling, Knockout	Thermodynamics, Enzyme kinetics, Transcriptome, Metabolome	ATP, NADPH
(149)	Liver	x		x			iMAT-like	Recon 1	Comparative, FBA, Sampling	Transcriptome	Data Similarity
(150)	Multiple	x		x			tINIT	HMR 2	Comparative, Knockout	Transcriptome	Tasks
(151)	Multiple	x		x			tINIT	HMR 2	FBA, Knockout	Transcriptome	Biomass, Tasks
(140)	Generic	x			x			Small-scale	FBA, DFBA, FVA, Knockout, Sampling		Biomass
(152)	Kidney	x				x	tINIT	iCancer-Core	FBA, Knockout	Transcriptome	Biomass
(153)	Brain	x					tINIT	HMR 2	Comparative, FBA, Knockout	Transcriptome	Biomass, Tasks
(105)	Breast, Lung, Multiple	x					Upper bound	HMR 1	FBA, Sampling, Knockout	Transcriptome	Biomass
(139)	Generic	x					MBA	Recon 1	FBA, Knockout	Transcriptome	Biomass
(154)	Kidney	x					MBA	Recon 1	FBA, Knockout	Transcriptome	Biomass
(54)	Liver	x					tINIT	HMR 2	Comparative, FBA, Knockout	Proteome	Biomass, Tasks
(155)	Liver	x					tINIT	HMR 2	FBA, Knockout, Topological	Transcriptome	Biomass, Tasks
(156)	Multiple	x					iMAT	Recon 1	FBA, ML, Topological	Transcriptome	Data Similarity
(102)	Multiple	x					GIMME	Recon 2	FBA, Sampling, Knockout	Mutations, Transcriptome	Biomass
(157)	Prostate	x					tINIT	iCancer-Core	FBA, Knockout, Sampling	Transcriptome, Proteome	Biomass, Tasks
(158)	Breast, Kidney, Liver, Prostate		x	x				KEGG	Network Propagation, Knockout, ML	Transcriptome	
(159)	Colorectal		x	x			tINIT	HMR 2	Comparative	Transcriptome	
(160)	Multiple		x	x				Recon 2	Regulatory, Topological, ML	Transcriptome, Metabolome	
(111)	Multiple		x	x			mCADRE	Recon 1	Comparative	Transcriptome	Tasks, Biomass
(103)	Multiple, Brain, Lung, Breast, Leukemia, Prostate		x	x			tINIT	Recon 3D	Comparative, Knockout, ML	Mutations, Protein Structures	Biomass
(161)	Prostate		x	x			iMAT	Recon 2	FBA, FVA	Transcriptome	Data Similarity
(162)	Kidney, Prostate		x				INIT	HMR 1	Knockout	Proteome, Fluxomic	Biomass
(108)	Multiple		x				INIT	HMR 1	Comparative	Proteome	
(163)	Multiple		x						Topological		
(164)	Generic			x	x			C2M2N	FBA		Biosynthesis, Biomass

(Continued)

TABLE 3 | Continued

Ref.	Cancer	Purpose					Method				
		Drug design	Predict drug targets	Explore Cancer biology	Explain Warburg effect	Patient classification	Integration	Model	Analysis	Constraints	Objective
(165)	Breast, Colorectal			x		x		Recon 2	FVA, ML	Metabolome	Metabolite Change
(166)	Liver			x		x	tINIT	HMR 2	Comparative	Transcriptome, Proteome	Biomass, Tasks
(167)	Brain			x			GIMME, MADE	iMS570	pFBA, Sampling	Transcriptome	Biomass
(168)	Breast			x			E-Flux	Recon2	FBA	Proteome	Biomass
(169)	Colorectal			x				Recon 2.2	Comparative	Transcriptome	
(170)	Colorectal			x			CORDA	Recon 2.2	FBA, FVA, Topological	Proteome	Biomass, ATP
(130)	Generic			x				HMRcore	popFBA, Sampling	Loopless	Biomass
(171)	Kidney			x				Recon 1	pFBA	Flux measurements	Biomass
(172)	Kidney			x			INIT	HMR 1	Comparative	Proteome	
(173)	Liver			x			tINIT	HMR 2	Comparative, Gap filling, Regulatory, FBA	Transcriptome, Metabolome	Biomass
(174)	Lung			x					13C flux analysis	Flux measurements, Labeling measurements	
(141)	Lung			x				Central Carbon, Recon 2	Elementary modes, Structural fluxes, pFBA	Protein efficiency	Biomass, Biosynthesis
(131)	Lung, Breast			x			E-Flux	HMRcore	scFBA	scRNA-seq, metabolomics	Biomass
(175)	Lung, Prostate			x			E-Flux	Recon 1	FVA	Transcriptome	Biomass
(176)	Multiple			x			tINIT	HMR 2	Comparative	Transcriptome	
(104)	Prostate			x			IMAT, GIMME, Gonçalves, MADE	HMR 2	FBA	Transcriptome	Data Similarity
(177)	Generic				x			Recon 1	FBA, FVA, Sampling	Protein efficiency, Enzyme kinetics	Biomass
(178)	Generic				x			ATP	FBA	Protein efficiency	ATP
(179)	Generic				x			ATP, BiGG	FBA	Protein efficiency	ATP, Nutrient cost
(180)	Liver				x		MADE	Recon 2	Comparative	Transcriptome	Data Similarity
(106)	Liver				x		Upper bound	Recon 3D	FBA, FVA	Transcriptome, Nutrient availability	Biomass
(181)	Liver				x		Bounds	Recon 2	FBA	Protein efficiency, Transcriptome	ATP
(182)	Multiple				x		E-Flux	Recon 1	FBA	Transcriptome	Biomass

limited by antibody or spectrometry methods that are low-throughput and less quantitative. Emerging evidence shows that newly developed quantitative proteome may better explain genetic disease and metabolism (191), emphasizing the advantage of using proteome evidence for metabolic model reconstruction. However, accuracy of proteome-based reconstructions is still limited due to various regulatory mechanisms such as protein modifications that have yet to be integrated into cancer metabolic models.

In the past decade, many more cancer-specific models have been reconstructed for liver (106, 149, 155, 166, 173, 181), kidney (152, 154, 162, 172, 172), breast (103, 105, 131, 145, 168), prostate (103, 104, 157, 161, 162, 175), brain (103, 153, 167), colorectal (159, 170, 192), head and neck (148), eye (147), and lung (103, 105, 131, 175) cancer to generate cancer-specific hypotheses. To compare these various methods for reconstruction of cancer metabolic models, a study benchmarked their predictive performance and consistency (91), with relevant findings summarized in **Table 2**. In pursuit of personalized medicine to find optimal treatment based on patient' genetic factors, researchers have also built personalized cancer GEMs from patient sample data to identify metabolic features that are commonly-shared or patient-specific (54). Furthermore, patient genetic variants were integrated in the

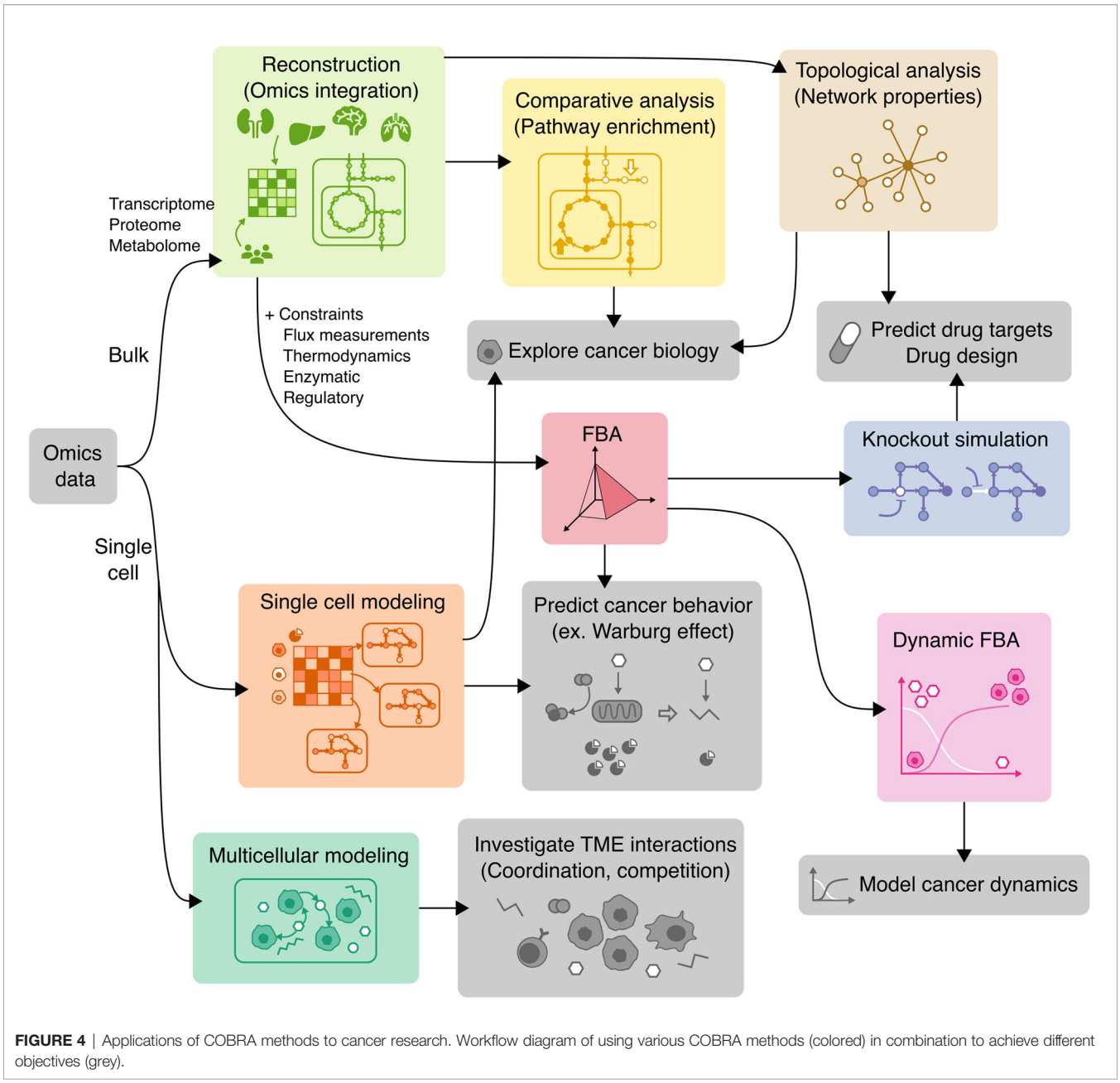


FIGURE 4 | Applications of COBRA methods to cancer research. Workflow diagram of using various COBRA methods (colored) in combination to achieve different objectives (grey).

Recon3D model with protein structures to look for cancer mutation hotspots in glioblastoma patients (103). Nam et al. modeled loss of function mutations *via* knockouts and analyzed potential gain of function mutations by adding promiscuous reactions predicted by chemoinformatics (102). Overall, these various reconstructions of cancer metabolic models aim to capture the heterogeneity of cancer.

Pathway and Network Analyses of Cancer GEM

To find metabolic differences between cancer and healthy cell types and between patients, these reconstructed metabolic networks are analyzed for enrichment of biological features, generating biologically relevant hypotheses that can guide mechanistic interpretation, biomarker discovery, and drug development. Comparative analysis involves statistical testing for the enrichment of reactions, genes, and metabolites to identify differentially activated pathways. Comparing networks of healthy and cancer cell types using hypergeometric test identified enrichment of not only well-known drug targets (polyamines, isoprenoid biosynthesis, prostaglandins and leukotrienes), but also new drug targets explained by protection against oxidative stress and methylglyoxal toxicity (108). Another study that used Wilcoxon rank sum test to

TABLE 4 | Human metabolic generic models and cancer models.

Model	Scale	No. of Reactions	No. of Metabolite	No. of Genes	Web link
HMR 1 (185)	Genome	8174	6006	3674	https://metabolicatlas.org/gems/repository/366
HMR 2 (186)	Genome	8181	6007	3765	https://metabolicatlas.org/gems/repository/367
Recon 1 (183)	Genome	3741	2766	1905	http://bigg.ucsd.edu/models/RECON1
Recon 2 (184)	Genome	7440	5063	2194	https://www.ebi.ac.uk/biomodels/MODEL1109130000
Recon 3D (103)	Genome	10600	5835	2248	https://www.vmh.life/#downloadview http://bigg.ucsd.edu/models/Recon3D
Human1 (187)	Genome	13069	8366	3067	https://github.com/SysBioChalmers/Human-GEM
Cancer central metabolism (140)	Small	80	66	46	https://doi.org/10.1371/journal.pone.0012383
iCancer-Core (iHCC2578) (151, 166)	Genome	7762	5566	2892	https://github.com/sysmedicine/phd2020/tree/master/GEM/data
C2M2N (164)	Small	77	54	–	https://doi.org/10.3390/metabo9050081
HMRcore (131, 188)	Intermediate	315	256	418	https://github.com/BIMIB-DISCo/scFBA
Central Carbon (141)	Small	114	120	–	https://doi.org/10.1042/bst20150149
iMS570 (brain) (189)	Genome	630	524	570	http://dx.doi.org/10.1016/j.fob.2014.05.006

This table describes various human reconstructions that are used as the starting reference models in various cancer applications listed in ***Table 3****. The most updated online links to these models may be different than previously described in their original manuscripts.*

compare tumor and normal metabolic models also found enrichment of leukotriene synthesis in addition to other tumor supporting pathways such as folate metabolism, eicosanoid metabolism, fatty acid synthesis, and nucleotide metabolism (111). Of note, these pathways were not statistically significant from pathway analysis of gene expression data alone, emphasizing the importance of systems-level network analysis to extract biological signal. In addition, the presence and absence of active genes, metabolites, and reactions can be characterized by clustering to validate similarity of related cell types (108), and calculating Hamming distance or pairwise comparisons to find the most different cancer GEMs (150). Comparing cancer-specific GEMs can reveal cancer types with more severe metabolic dysfunction. For example, clear cell renal cell carcinoma (ccRCC) GEM showed loss of redundant genes in key metabolic pathways (162, 172), suggesting that ccRCC might be more responsive to metabolic anticancer drugs due to reduced capacity to evade drug inhibition *via* alterative enzymes and pathways.

While the presence of pathways is indicative of activity, analyzing the pattern of how these pathways connect could provide additional insights. For this purpose, topological analysis is a network-based analysis that characterizes metabolic models based on network properties that describes the degree and patterns of connection between metabolites, genes, and reactions. The same models from Agren et al. (108) were converted to enzyme-enzyme networks and re-analyzed using topological analysis, which revealed that most approved cancer drugs do not correlate with centrality (measure of importance) of individual enzymes, but do belong to a specific cluster in a cancer enzyme-centric networks (163). Furthermore, the analysis found that certain network motifs, such as feed-forward loop, are enriched in cancer networks compared to healthy cell type. Utilized in several other cancer studies (**Table 3**), topological analyses reveal insights about cancer based on the structure of cancer-specific metabolic networks without using flux simulations. Topological analyses emphasize the importance of system-oriented cancer drug design to find therapy that change the entire metabolic state instead of a single drug target that can be easily compensated by alternative pathways.

Quantitative Prediction of Cancer Behavior

To better understand metabolic reprogramming within cancer cells, cancer-specific metabolic models were used to simulate flux distributions to illustrate their metabolic state. Initial efforts built generic small-scale cancer models that only included the major pathways in cancer such as ATP and biomass production (140, 178) to demonstrate the usefulness of standard COBRA methods as such FBA, FVA, and *in silico* knockouts (140). Performing dynamical FBA on such model was able to predict the growth rates of HeLa cells, validating the use of biomass objective with FBA for cancer predictions (140). While constraints on glucose uptake and solvent capacity initially predicted the Warburg effect (178), later implementations of protein constraints in these small-scale (179) and genome-scale (177) cancer models explained the Warburg effect as a result of maximizing enzyme efficiency. Another protein efficiency constraint, flux minimization with FBA, predicted the Warburg effect in liver-specific GEMs and agreed with metabolic profiling of *Mir122a* knockout mice (181). Another cancer metabolic adaptation that bypass mutation of enzymes from the TCA cycle was recapitulated by adding upper flux bounds during flux simulations (154). In addition to these methods for modeling intracellular constraints, it is also important to account for cell-extrinsic factors imposed by the tumor microenvironment. Approaches to impose nutrient constraints include constraining exchange reaction bounds by experimentally measured flux (145, 162), transporter expression (105), concentration and membrane potential-dependent free energy calculations (148), and concentration gradient over time (106). These quantitative predictions of cancer metabolic reprogramming further demonstrate the applicability of COBRA methods to model cancer metabolic programs.

In Silico Drug Discovery

Furthermore, quantitative flux predictions can guide drug therapy design by simulating the effect of enzyme inhibition on

cellular metabolic function in both cancer and healthy GEMs to maximize therapeutic effect while minimizing toxicity. *In silico* knockouts are performed by constraining one or more reactions' flux to zero, setting an objective function that represents growth or other metabolic tasks, and finally performing FBA to calculate the change in maximum objective. One approach aims to find drug targets based on gene essentiality–knock out of enzymes that inhibit cancer growth. *In silico* gene knockout simulations of genome-scale cancer model identified drug targets and combination drug strategies (double gene knockout) that could reduce cancer growth (139). These candidates include known drugs and are validated *via* shRNA gene silencing data and cancer somatic mutations. Another study found that gene essentiality by FBA using biomass objective is better than chance but has limited accuracy depending on cancer type, especially after adding exchange flux constraints (162). A second approach based on metabolite essentiality screens for antimetabolites (metabolite analogs), which would compete with endogenous metabolites to inhibit their associated enzymes. By simulating *in silico* knockout of all enzymes acting on each metabolite, studies have identified antimetabolite drug candidates that could selectively disable critical metabolic task in cancer cell line-specific GEMs (150) and personalized hepatocellular carcinoma (HCC) patient GEMs (54). Out of 101 antimetabolite candidates, many were already used (22%) or proposed as anticancer drug targets (60%), and some targets were shown to be highly patient-specific, supporting the use of flux predictions of cancer GEMs for both general and personalized drug discovery (54). Many more studies applying *in silico* knockouts are listed in **Table 3**. While using FBA for *in silico* drug design is well established, the predictions maybe inaccurate due to bias introduced by the choice of objective function and reaction bounds, such as those for cell-specific exchange fluxes that are not always experimentally determined (91, 162). Furthermore, simulations based on cell line measurements and culturing conditions cannot faithfully reflect multi-cellular tissues and physiological environments *in vivo*.

Multicellular and Single-Cell Modeling of TME

To analyze cell-heterogeneous systems like the TME, it is important to investigate metabolic programs within a multi-scale population model and at the single-cell level. To model interactions between multiple cells, multicellular modeling accounts for metabolite exchange between single cells within the environment. This was attempted by popFBA (130), which simulated a spatial model of identical cancer cells that adapted heterogeneously and cooperatively to maximize growth of the entire tumor mass. To account for tumor heterogeneity, a population model can be constrained by single-cell RNA-seq (scRNA-seq) data containing different tissue subpopulations in the scFBA method (131). When applied to lung adenocarcinoma and breast cancer cells, scFBA reveals metabolically defined subpopulations, some of which have coordinated metabolic fluxes (e.g., uptake or secretion of opposite sets of metabolites) suggesting potential cell-cell metabolic interactions. Other methods, such as scFEA or Compass, calculates cell-wise metabolic flux from scRNA-seq data to interpret cellular metabolic activity. Compass revealed metabolic states associated with functional states of T helper 17 (Th17) cells, in particular an increase in arginine and polyamine metabolism that resulted in a regulatory T cell (Treg)-like, dysfunctional cell state (73). The other single-cell method, scFEA, applied to patient-derived pancreatic cancer cells with metabolic perturbations (gene knockout, hypoxia), predicted flux variation that correlates with measured metabolomics. These methods could be applied to infer metabolic states of tumor and immune cells from existing scRNA-seq datasets of tumor samples. In future studies, algorithms for microbial community-modeling can be repurposed to investigate the interactions of cancer and immune cells in the TME (MICOM) and model the dynamics of immunosurveillance and tumor resistance (surfin_fba).

DISCUSSION

COBRA methods have proved useful for systems-level inference of metabolic activity under a mathematical framework built upon biomolecular knowledge. The accessibility and algorithms of COBRA methods have been improved with the development of open-source COBRA Python packages. We have identified Python packages available to handle the major areas of COBRA methods: FBA, FVA, gene knockout, strain design, omics integration, regulatory constraints, reconstruction, gap filling, ensemble modeling, thermodynamics, enzymatic constraints, EFM, sampling, single-cell modeling, multicellular modeling, and visualization. However, the Python COBRA ecosystem is currently missing some methods for constraining models by regulatory mechanisms and reconstruction of context-specific GEMs. However, these gaps are only due to limitations of time and effort, not limitations of the Python programming language. In fact, many features involving complex models, parallelization, and efficient memory management are available in Python instead of MATLAB. For example, ME-models, a set of multi-scale problems describing multiple biological processes across different space and time scales such as transcription, translation, and protein interactions, are handled by Python packages only for now. Integration of protein structure into the Recon3D human GEM was facilitated by Python packages ssbio and GEM-PRO (103). GEMs interface with machine learning in Medusa and scFEA. Likewise, upcoming COBRA packages will likely integrate with existing Python tools for statistical learning and analysis of single-cell multi-omics data. As models and omics datasets increase in complexity, COBRA methods will thrive in the open-source Python environment. While we improve our modeling techniques, it is also important to validate flux predictions using experimental techniques such as metabolomics profile and label tracing experiments. To interpret isotope tracing data, ^{13}C-Metabolic Flux Analysis was developed to infer intracellular fluxes. While ^{13}C-MFA allows direct

measurement of metabolic flux, the method is limited to small-scale models (central metabolism) and requires more expertise than the typical omics measurements for constraining COBRA methods. Python packages for modeling label tracing data are available *via* FluxPyt and mfapy (193, 194). While these experimental techniques are outside the scope of this review, they have been reviewed previously for bulk, single-cell, and cancer applications (119, 195, 196). Another alternative computational metabolic modeling approach is parametric kinetic modeling, which mathematically describes enzyme activity involving regulatory mechanisms (17). While this paradigm may offer accurate prediction of perturbation outcomes, systems emergent properties (e.g., switches, oscillations, bistability), and non-steady state concentrations, scaling kinetic models to genome-scale metabolic models is a challenge due to the requirement for intracellular concentrations, kinetic parameters, and rate laws. DMPy attempts to overcome the challenge by incorporating thermodynamics constraints to infer missing kinetic parameters. Hybrid approaches combining kinetic modeling with constraints-based models may bring kinetic modeling closer to genome-scale.

Applications of GEMs and COBRA methods to cancer research have improved our understanding of how molecular mechanisms translate to cancer phenotype, aiding interpretation of multi-omics data and guiding drug designs that target cell metabolism at the systems-level. Metabolic models of cancer have evolved from small-scale models of essential pathways to genome-scale cancer-specific models, and they are now expanding to the realm of single-cell modeling. The computational resources required for numerous single-cell reconstructions and optimizations can be costly. Single-cell methods reduce complexity by pooling of reactions and similar cells and could benefit from ensemble modeling techniques that reduce a large number of models into ensemble objects. As demonstrated by bulk-level modeling, future single-cell modeling can improve prediction accuracy by incorporating constraints determined by multi-omics, thermodynamics, protein crowding and kinetics, genotype, and regulatory mechanisms. Furthermore, single-cell methods that estimate the metabolic flux of individual cells can be improved by integration of spatial information and inter-cell metabolic exchange to model crosstalk between cancer, immune, and stromal cells within the TME. By understanding the cancer-immune metabolic competition, we can design drugs that disrupt pathophysiologic interactions to enhance antitumor immune response and prevent evasion of immunosurveillance.

AUTHOR CONTRIBUTIONS

RN conceived the review, wrote the manuscript, and created visualizations. JL wrote the manuscript and created visualizations. YS and JH supervised and edited the manuscript. PB and CD edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Andy Hill Cancer Research Endowment Fund (JH) and the Parker Institute for Cancer Immunotherapy (JH). YS is a Damon Runyon Quantitative Biology Fellow supported by the Damon Runyon Cancer Research Foundation (DRQ-13-22). YS was additionally supported by the Mahan Fellowship at the Herbold Computational Biology Program of Fred Hutch Cancer Research Center and the Translational Data Science Integrated Research Center New Collaboration Award at Fred Hutch Cancer Research Center and in part through a pilot fund from the NIH/NCI Cancer Center Support Grant P30 CA015704. PB acknowledges the support of 5U01AG061359-02 and 5U01AG061359-03 from NIA and 5R01HD091527-06 from NICHD. JL was funded by the Brotman Baty Institute Catalytic Collaborations Trainee Grant program.

ACKNOWLEDGMENTS

We thank Herbert Sauro for comments on initial draft for Computation Systems Biology class at University of Washington.

REFERENCES

1. Kumar A, Misra BB. Challenges and Opportunities in Cancer Metabolomics. *Proteomics* (2019) 19:1900042. doi: 10.1002/pmic.201900042
2. Chen C, Xing D, Tan L, Li H, Zhou G, Huang L, et al. Single-Cell Whole-Genome Analyses by Linear Amplification *via* Transposon Insertion (LIANTI). *Science* (2017) 356:189–94. doi: 10.1126/science.aak9787
3. Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, et al. Highly Parallel Genome-Wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* (2015) 161:1202–14. doi: 10.1016/j.cell.2015.05.002
4. Shema E, Bernstein BE, Buenrostro JD. Single-Cell and Single-Molecule Epigenomics to Uncover Genome Regulation at Unprecedented Resolution. *Nat Genet* (2019) 51:19–25. doi: 10.1038/s41588-018-0290-x
5. Bendall SC, Simonds EF, Qiu P, Amir ED, Krutzik PO, Finck R, et al. Single-Cell Mass Cytometry of Differential Immune and Drug Responses Across a Human Hematopoietic Continuum. *Science* (2011) 332:687–96. doi: 10.1126/science.1198704
6. Ma C, Fan R, Ahmad H, Shi Q, Comin-Anduix B, Chodon T, et al. A Clinical Microchip for Evaluation of Single Immune Cells Reveals High Functional Heterogeneity in Phenotypically Similar T Cells. *Nat Med* (2011) 17:738–43. doi: 10.1038/nm.2375
7. Shi Q, Qin L, Wei W, Geng F, Fan R, Shik Shin Y, et al. Single-Cell Proteomic Chip for Profiling Intracellular Signaling Pathways in Single Tumor Cells. *Proc Natl Acad Sci* (2012) 109:419–24. doi: 10.1073/pnas.1110865109
8. Stoeckius M, Hafemeister C, Stephenson W, Houck-Loomis B, Chattopadhyay PK, Swerdlow H, et al. Simultaneous Epitope and Transcriptome Measurement in Single Cells. *Nat Methods* (2017) 14:865–8. doi: 10.1038/nmeth.4380
9. Xue M, Wei W, Su Y, Kim J, Shin YS, Mai WX, et al. Chemical Methods for the Simultaneous Quantitation of Metabolites and Proteins From Single Cells. *J Am Chem Soc* (2015) 137:4066–9. doi: 10.1021/jacs.5b00944

10. Xue M, Wei W, Su Y, Johnson D, Heath JR. Supramolecular Probes for Assessing Glutamine Uptake Enable Semi-Quantitative Metabolic Models in Single Cells. *J Am Chem Soc* (2016) 138:3085–93. doi: 10.1021/jacs.5b12187
11. Du J, Su Y, Qian C, Yuan D, Miao K, Lee D, et al. Raman-Guided Subcellular Pharmaco-Metabolomics for Metastatic Melanoma Cells. *Nat Commun* (2020) 11:4830. doi: 10.1038/s41467-020-18376-x
12. Chappell L, Russell AJC, Voet T. Single-Cell (Multi)omics Technologies. *Annu Rev Genomics Hum Genet* (2018) 19:15–41. doi: 10.1146/annurev-genom-091416-035324
13. Gu C, Kim GB, Kim WJ, Kim HU, Lee SY. Current Status and Applications of Genome-Scale Metabolic Models. *Genome Biol* (2019) 20:121. doi: 10.1186/s13059-019-1730-3
14. Chowdhury S, Fong SS. Leveraging Genome-Scale Metabolic Models for Human Health Applications. *Curr Opin Biotechnol* (2020) 66:267–76. doi: 10.1016/j.copbio.2020.08.017
15. Lewis NE, Nagarajan H, Palsson BO. Constraining the Metabolic Genotype–Phenotype Relationship Using a Phylogeny of *In Silico* Methods. *Nat Rev Microbiol* (2012) 10:291–305. doi: 10.1038/nrmicro2737
16. Lee JW, Su Y, Baloni P, Chen D, Pavlovitch-Bedzyk AJ, Yuan D, et al. Integrated Analysis of Plasma and Single Immune Cells Uncovers Metabolic Changes in Individuals With COVID-19. *Nat Biotechnol* (2022), 40:110–120. doi: 10.1038/s41587-021-01020-4
17. Yasemi M, Jolicoeur M. Modelling Cell Metabolism: A Review on Constraint-Based Steady-State and Kinetic Approaches. *Processes* (2021) 9:322. doi: 10.3390/pr9020322
18. Heirendt L, Arreckx S, Pfau T, Mendoza SN, Richelle A, Heinken A, et al. Creation and Analysis of Biochemical Constraint-Based Models Using the COBRA Toolbox V.3.0. *Nat Protoc* (2019) 14:639–702. doi: 10.1038/s41596-018-0098-2
19. Agren R, Liu L, Shoaie S, Vongsangnak W, Nookaew I, Nielsen J. The RAVEN Toolbox and Its Use for Generating a Genome-Scale Metabolic Model for Penicillium Chrysogenum. *PLOS Comput Biol* (2013) 9:e1002980. doi: 10.1371/journal.pcbi.1002980
20. von Kamp A, Thiele S, Hädicke O, Klamt S. Use of CellNetAnalyzer in Biotechnology and Metabolic Engineering. *J Biotechnol* (2017) 261:221–8. doi: 10.1016/j.jbiotec.2017.05.001
21. Ebrahim A, Lerman JA, Palsson BO, Hyduke DR. COBRApy: COnstraints-Based Reconstruction and Analysis for Python. *BMC Syst Biol* (2013) 7:74. doi: 10.1186/1752-0509-7-74
22. The openCOBRA Project. *SourceForge* . Available at: https://sourceforge.net/projects/opencobra/ (Accessed December 14, 2021).
23. Olivier B, Gottstein W, Molenaar D, Teusink B. *CBMPy Release 0.8.2*. Zenodo (2021). doi: 10.5281/ZENODO.5546608
24. Lieven C, Beber ME, Olivier BG, Bergmann FT, Ataman M, Babaei P, et al. MEMOTE for Standardized Genome-Scale Metabolic Model Testing. *Nat Biotechnol* (2020) 38:272–6. doi: 10.1038/s41587-020-0446-y
25. Aite M, Chevallier M, Frioux C, Trottier C, Got J, Cortés MP, et al. Traceability, Reproducibility and Wiki-Exploration for "À-La-Carte" Reconstructions of Genome-Scale Metabolic Models. *PLOS Comput Biol* (2018) 14:e1006146. doi: 10.1371/journal.pcbi.1006146
26. Machado D, Andrejev S, Tramontano M, Patil KR. Fast Automated Reconstruction of Genome-Scale Metabolic Models for Microbial Species and Communities. *Nucleic Acids Res* (2018) 46:7542–53. doi: 10.1093/nar/gky537
27. Hanemaaijer M, Olivier BG, Röling WFM, Bruggeman FJ, Teusink B. Model-Based Quantification of Metabolic Interactions From Dynamic Microbial-Community Data. *PLOS One* (2017) 12:e0173183. doi: 10.1371/journal.pone.0173183
28. Pitkänen E, Jouhten P, Hou J, Syed MF, Blomberg P, Kludas J, et al. Comparative Genome-Scale Reconstruction of Gapless Metabolic Networks for Present and Ancestral Species. *PLOS Comput Biol* (2014) 10: e1003465. doi: 10.1371/journal.pcbi.1003465
29. Orth JD, Thiele I, Palsson BØ. What Is Flux Balance Analysis? *Nat Biotechnol* (2010) 28:245–8. doi: 10.1038/nbt.1614
30. Harwood SM, Höffner K, Barton PI. Efficient Solution of Ordinary Differential Equations With a Parametric Lexicographic Linear Program Embedded. *Numer Math* (2016) 133:623–53. doi: 10.1007/s00211-015-0760-3
31. Tourigny DS, Muriel JC, Beber ME. Dfba: Software for Efficient Simulation of Dynamic Flux-Balance Analysis Models in Python. *J Open Source Softw* (2020) 5:2342. doi: 10.21105/joss.02342
32. Smith RW, van Rosmalen RP, Martins dos Santos VAP, Fleck C. DMPy: A Python Package for Automated Mathematical Model Construction of Large-Scale Metabolic Systems. *BMC Syst Biol* (2018) 12:72. doi: 10.1186/s12918-018-0584-8
33. Smallbone K, Simeonidis E. Flux Balance Analysis: A Geometric Perspective. *J Theor Biol* (2009) 258:311–5. doi: 10.1016/j.jtbi.2009.01.027
34. Mahadevan R, Schilling CH. The Effects of Alternate Optimal Solutions in Constraint-Based Genome-Scale Metabolic Models. *Metab Eng* (2003) 5:264–76. doi: 10.1016/j.ymben.2003.09.002
35. Guebila MB. VFFVA: Dynamic Load Balancing Enables Large-Scale Flux Variability Analysis. *BMC Bioinf* (2020) 21:424. doi: 10.1186/s12859-020-03711-2
36. Edwards JS, Palsson BO. The *Escherichia Coli* MG1655 *In Silico* Metabolic Genotype: Its Definition, Characteristics, and Capabilities. *Proc Natl Acad Sci* (2000) 97:5528–33. doi: 10.1073/pnas.97.10.5528
37. Segrè D, Vitkup D, Church GM. Analysis of Optimality in Natural and Perturbed Metabolic Networks. *Proc Natl Acad Sci* (2002) 99:15112–7. doi: 10.1073/pnas.232349399
38. Shlomi T, Berkman O, Ruppin E. Regulatory on/Off Minimization of Metabolic Flux Changes After Genetic Perturbations. *Proc Natl Acad Sci* (2005) 102:7695–700. doi: 10.1073/pnas.0406346102
39. Laniau J, Frioux C, Nicolas J, Baroukh C, Cortes M-P, Got J, et al. Combining Graph and Flux-Based Structures to Decipher Phenotypic Essential Metabolites Within Metabolic Networks. *PeerJ* (2017) 5:e3860. doi: 10.7717/peerj.3860
40. Patil KR, Rocha I, Förster J, Nielsen J. Evolutionary Programming as a Platform for in Silico Metabolic Engineering. *BMC Bioinf* (2005) 6:308. doi: 10.1186/1471-2105-6-308
41. Cardoso JGR, Jensen K, Lieven C, Lærke Hansen AS, Galkina S, Beber M, et al. Cameo: A Python Library for Computer Aided Metabolic Engineering and Optimization of Cell Factories. *ACS Synth Biol* (2018) 7:1163–6. doi: 10.1021/acssynbio.7b00423
42. Burgard AP, Pharkya P, Maranas CD. Optknock: A Bilevel Programming Framework for Identifying Gene Knockout Strategies for Microbial Strain Optimization. *Biotechnol Bioeng* (2003) 84:647–57. doi: 10.1002/bit.10803
43. Choi HS, Lee SY, Kim TY, Woo HM. *In Silico* Identification of Gene Amplification Targets for Improvement of Lycopene Production. *Appl Environ Microbiol* (2010) 76:3097–105. doi: 10.1128/AEM.00115-10
44. Shen F, Sun R, Yao J, Li J, Liu Q, Price ND, et al. OptRAM: *In Silico* Strain Design *via* Integrative Regulatory-Metabolic Network Modeling. *PLOS Comput Biol* (2019) 15:e1006835. doi: 10.1371/journal.pcbi.1006835
45. Pereira V, Cruz F, Rocha M. MEWpy: A Computational Strain Optimization Workbench in Python. *Bioinformatics* (2021) 37:2494–6. doi: 10.1093/bioinformatics/btab013
46. Kim J, Reed JL. OptORF: Optimal Metabolic and Regulatory Perturbations for Metabolic Engineering of Microbial Strains. *BMC Syst Biol* (2010) 4:53. doi: 10.1186/1752-0509-4-53
47. Colijn C, Brandes A, Zucker J, Lun DS, Weiner B, Farhat MR, et al. Interpreting Expression Data With Metabolic Flux Models: Predicting Mycobacterium Tuberculosis Mycolic Acid Production. *PLOS Comput Biol* (2009) 5:e1000489. doi: 10.1371/journal.pcbi.1000489
48. Machado D. *ReFramed: Metabolic Modeling Package* (2021). Available at: https://github.com/cdanielmachado/reframed (Accessed February 22, 2022).
49. Schultz A, Qutub AA. Reconstruction of Tissue-Specific Metabolic Networks Using CORDA. *PLOS Comput Biol* (2016) 12:e1004808. doi: 10.1371/journal.pcbi.1004808
50. Schmidt BJ, Ebrahim A, Metz TO, Adkins JN, Palsson BØ, Hyduke DR. GIM3E: Condition-Specific Models of Cellular Metabolism Developed From Metabolomics and Expression Data. *Bioinformatics* (2013) 29:2900–8. doi: 10.1093/bioinformatics/btt493
51. Vlassis N, Pacheco MP, Sauter T. Fast Reconstruction of Compact Context-Specific Metabolic Network Models. *PLOS Comput Biol* (2014) 10:e1003424. doi: 10.1371/journal.pcbi.1003424
52. Ferreira J, Vieira V, Gomes J, Correia S, Rocha M. "Troppo - A Python Framework for the Reconstruction of Context-Specific Metabolic Models.,".

In: F Fdez-Riverola, M Rocha, MS Mohamad, N Zaki and JA Castellanos-Garzón, editors. *Practical Applications of Computational Biology and Bioinformatics, 13th International Conference. Advances in Intelligent Systems and Computing*. Cham: Springer International Publishing (2020). p. 146–53. doi: 10.1007/978-3-030-23873-5_18

53. Becker SA, Palsson BO. Context-Specific Metabolic Networks Are Consistent With Experiments. *PLOS Comput Biol* (2008) 4:e1000082. doi: 10.1371/journal.pcbi.1000082
54. Agren R, Mardinoglu A, Asplund A, Kampf C, Uhlen M, Nielsen J. Identification of Anticancer Drugs for Hepatocellular Carcinoma Through Personalized Genome-Scale Metabolic Modeling. *Mol Syst Biol* (2014) 10:721. doi: 10.1002/msb.145122
55. Shlomi T, Cabili MN, Herrgård MJ, Palsson BØ, Ruppin E. Network-Based Prediction of Human Tissue-Specific Metabolism. *Nat Biotechnol* (2008) 26:1003–10. doi: 10.1038/nbt.1487
56. Covert MW, Palsson BØ. Transcriptional Regulation in Constraints-Based Metabolic Models of Escherichia Coli. *J Biol Chem* (2002) 277:28058–64. doi: 10.1074/jbc.M201691200
57. Shlomi T, Eisenberg Y, Sharan R, Ruppin E. A Genome-Scale Computational Study of the Interplay Between Transcriptional Regulation and Metabolism. *Mol Syst Biol* (2007) 3:101. doi: 10.1038/msb4100141
58. jseidel5. *Python Implementation of Probabilistic Regulation of Metabolism (PROM)* (2020). Available at: https://github.com/jseidel5/Python-Probabilistic-Regulation-of-Metabolism (Accessed February 22, 2022).
59. Brunk E, Mih N, Monk J, Zhang Z, O'Brien EJ, Bliven SE, et al. Systems Biology of the Structural Proteome. *BMC Syst Biol* (2016) 10:26. doi: 10.1186/s12918-016-0271-6
60. Mih N, Brunk E, Chen K, Catoiu E, Sastry A, Kavvas E, et al. Ssbio: A Python Framework for Structural Systems Biology. *Bioinformatics* (2018) 34:2155–7. doi: 10.1093/bioinformatics/bty077
61. Machado D, Herrgård MJ, Rocha I. Modeling the Contribution of Allosteric Regulation for Flux Control in the Central Carbon Metabolism of E. Coli (2015) (Accessed February 9, 2022).
62. Schellenberger J, Lewis NE, Palsson BØ. Elimination of Thermodynamically Infeasible Loops in Steady-State Metabolic Models. *Biophys J* (2011) 100:544–53. doi: 10.1016/j.bpj.2010.12.3707
63. Desouki AA, Jarre F, Gelius-Dietrich G, Lercher MJ. CycleFreeFlux: Efficient Removal of Thermodynamically Infeasible Loops From Flux Distributions. *Bioinformatics* (2015) 31:2159–65. doi: 10.1093/bioinformatics/btv096
64. Gollub MG, Kaltenbach H-M, Stelling J. Probabilistic Thermodynamic Analysis of Metabolic Networks. *Bioinforma Oxf Engl* (2021) 37:2938–2945, btab194. doi: 10.1093/bioinformatics/btab194
65. Henry CS, Broadbelt LJ, Hatzimanikatis V. Thermodynamics-Based Metabolic Flux Analysis. *Biophys J* (2007) 92:1792–805. doi: 10.1529/biophysj.106.093138
66. Salvy P, Fengos G, Ataman M, Pathier T, Soh KC, Hatzimanikatis V. pyTFA and matTFA: A Python Package and a Matlab Toolbox for Thermodynamics-Based Flux Analysis. *Bioinforma Oxf Engl* (2019) 35:167–9. doi: 10.1093/bioinformatics/bty499
67. Lewis NE, Hixson KK, Conrad TM, Lerman JA, Charusanti P, Polpitiya AD, et al. Omic Data From Evolved E. Coli Are Consistent With Computed Optimal Growth From Genome-Scale Models. *Mol Syst Biol* (2010) 6:390. doi: 10.1038/msb.2010.47
68. Sánchez BJ, Zhang C, Nilsson A, Lahtvee P, Kerkhoven EJ, Nielsen J. Improving the Phenotype Predictions of a Yeast Genome-Scale Metabolic Model by Incorporating Enzymatic Constraints. *Mol Syst Biol* (2017) 13:935. doi: 10.15252/msb.20167411
69. Bekiaris PS, Klamt S. Automatic Construction of Metabolic Models With Enzyme Constraints. *BMC Bioinf* (2020) 21:19. doi: 10.1186/s12859-019-3329-9
70. Mao Z, Zhao X, Yang X, Zhang P, Du J, Yuan Q, et al. ECMpy, a Simplified Workflow for Constructing Enzymatic Constrained Metabolic Network Model. *Biomolecules* (2022) 12:65. doi: 10.3390/biom12010065
71. Lloyd CJ, Ebrahim A, Yang L, King ZA, Catoiu E, O'Brien EJ, et al. COBRAme: A Computational Framework for Genome-Scale Models of Metabolism and Gene Expression. *PLOS Comput Biol* (2018) 14:e1006302. doi: 10.1371/journal.pcbi.1006302
72. Medlock GL, Moutinho TJ, Papin JA. Medusa: Software to Build and Analyze Ensembles of Genome-Scale Metabolic Network Reconstructions. *PLOS Comput Biol* (2020) 16:e1007847. doi: 10.1371/journal.pcbi.1007847
73. Wagner A, Wang C, Fessler J, DeTomaso D, Avila-Pacheco J, Kaminski J, et al. Metabolic Modeling of Single Th17 Cells Reveals Regulators of Autoimmunity. *Cell* (2021) 184:4168–85.e21. doi: 10.1016/j.cell.2021.05.045
74. Alghamdi N, Chang W, Dang P, Lu X, Wan C, Gampala S, et al. A Graph Neural Network Model to Estimate Cell-Wise Metabolic Flux Using Single-Cell RNA-Seq Data. *Genome Res* (2021) 31:1867–84. doi: 10.1101/gr.271205.120
75. Diener C, Gibbons SM, Resendis-Antonio O. MICOM: Metagenome-Scale Modeling To Infer Metabolic Interactions in the Gut Microbiota. *mSystems* (2020) 5:e00606–19. doi: 10.1128/mSystems.00606-19
76. Brunner JD, Chia N. Minimizing the Number of Optimizations for Efficient Community Dynamic Flux Balance Analysis. *PLOS Comput Biol* (2020) 16:e1007786. doi: 10.1371/journal.pcbi.1007786
77. Kaufman DE, Smith RL. Direction Choice for Accelerated Convergence in Hit-And-Run Sampling(1998) (Accessed May 19, 2022).
78. Megchelenbrink W, Huynen M, Marchiori E. Optgpsampler: An Improved Tool for Uniformly Sampling the Solution-Space of Genome-Scale Metabolic Networks. *PLOS ONE* (2014) 9:e86587. doi: 10.1371/journal.pone.0086587
79. Buchner BA, Zanghellini J. EFMlrs: A Python Package for Elementary Flux Mode Enumeration *via* Lexicographic Reverse Search. *BMC Bioinf* (2021) 22:547. doi: 10.1186/s12859-021-04417-9
80. Schuster S, Hilgetag C. On Elementary Flux Modes in Biochemical Reaction Systems at Steady State. *J Biol Syst* (1994) 02:165–82. doi: 10.1142/S0218339094000131
81. Vieira V, Rocha M. CoBAMP: A Python Framework for Metabolic Pathway Analysis in Constraint-Based Models. *Bioinformatics* (2019) 35:5361–2. doi: 10.1093/bioinformatics/btz598
82. Klamt S, Gilles ED. Minimal Cut Sets in Biochemical Reaction Networks. *Bioinformatics* (2004) 20:226–34. doi: 10.1093/bioinformatics/btg395
83. Schilling CH, Letscher D, Palsson BØ. Theory for the Systemic Definition of Metabolic Pathways and Their Use in Interpreting Metabolic Function From a Pathway-Oriented Perspective. *J Theor Biol* (2000) 203:229–48. doi: 10.1006/jtbi.2000.1073
84. King ZA, Dräger A, Ebrahim A, Sonnenschein N, Lewis NE, Palsson BO. Escher: A Web Application for Building, Sharing, and Embedding Data-Rich Visualizations of Biological Pathways. *PLOS Comput Biol* (2015) 11:e1004321. doi: 10.1371/journal.pcbi.1004321
85. Schultz A, Akbani R. SAMMI: A Semi-Automated Tool for the Visualization of Metabolic Networks. *Bioinformatics* (2020) 36:2616–7. doi: 10.1093/bioinformatics/btz927
86. John PS. *D3flux* (2021). Available at: https://github.com/pstjohn/d3flux (Accessed December 14, 2021).
87. Mendoza SN, Olivier BG, Molenaar D, Teusink B. A Systematic Assessment of Current Genome-Scale Metabolic Reconstruction Tools. *Genome Biol* (2019) 20:158. doi: 10.1186/s13059-019-1769-1
88. Zimmermann J, Kaleta C, Waschina S. Gapseq: Informed Prediction of Bacterial Metabolic Pathways and Reconstruction of Accurate Metabolic Models. *Genome Biol* (2021) 22:81. doi: 10.1186/s13059-021-02295-1
89. Höffner K, Harwood SM, Barton PI. A Reliable Simulator for Dynamic Flux Balance Analysis. *Biotechnol Bioeng* (2013) 110:792–802. doi: 10.1002/bit.24748
90. Yizhak K, Gaude E, Le Dévédec S, Waldman YY, Stein GY, van de Water B, et al. Phenotype-Based Cell-Specific Metabolic Modeling Reveals Metabolic Liabilities of Cancer. *eLife* (2014) 3:e03641. doi: 10.7554/eLife.03641
91. Jamialahmadi O, Hashemi-Najafabadi S, Motamedian E, Romeo S, Bagheri F. A Benchmark-Driven Approach to Reconstruct Metabolic Networks for Studying Cancer Metabolism. *PLOS Comput Biol* (2019) 15:e1006936. doi: 10.1371/journal.pcbi.1006936
92. Chung CH, Lin D-W, Eames A, Chandrasekaran S. Next-Generation Genome-Scale Metabolic Modeling Through Integration of Regulatory Mechanisms. *Metabolites* (2021) 11:606. doi: 10.3390/metabo11090606
93. Hucka M, Finney A, Bornstein BJ, Keating SM, Shapiro BE, Matthews J, et al. Evolving a Lingua Franca and Associated Software Infrastructure for

Computational Systems Biology: The Systems Biology Markup Language (SBML) Project. *Syst Biol* (2004) 1:41–53. doi: 10.1049/sb:20045008
94. Olivier BG, Bergmann FT. SBML Level 3 Package: Flux Balance Constraints Version 2. *J Integr Bioinforma* (2018) 15:20170082. doi: 10.1515/jib-2017-0082
95. King ZA, Lu J, Dräger A, Miller P, Federowicz S, Lerman JA, et al. BiGG Models: A Platform for Integrating, Standardizing and Sharing Genome-Scale Models. *Nucleic Acids Res* (2016) 44:D515–22. doi: 10.1093/nar/gkv1049
96. Li C, Donizelli M, Rodriguez N, Dharuri H, Endler L, Chelliah V, et al. BioModels Database: An Enhanced, Curated and Annotated Resource for Published Quantitative Kinetic Models. *BMC Syst Biol* (2010) 4:92. doi: 10.1186/1752-0509-4-92
97. Choi Y-M, Choi D-H, Lee YQ, Koduru L, Lewis NE, Lakshmanan M, et al. Mitigating Biomass Composition Uncertainties in Flux Balance Analysis Using Ensemble Representations. *bioRxiv* (2022), 652040. doi: 10.1101/652040
98. Watson MR. Metabolic Maps for the Apple II. *Biochem Soc Trans* (1984) 12:1093–4. doi: 10.1042/bst0121093
99. Mahadevan R, Edwards JS, Doyle FJ. Dynamic Flux Balance Analysis of Diauxic Growth in Escherichia Coli. *Biophys J* (2002) 83:1331–40. doi: 10.1016/S0006-3495(02)73903-9
100. Stellato B, Banjac G, Goulart P, Bemporad A, Boyd S. OSQP: An Operator Splitting Solver for Quadratic Programs. *Math Program Comput* (2020) 12:637–72. doi: 10.1007/s12532-020-00179-2
101. Cardoso JGR, Andersen MR, Herrgård MJ, Sonnenschein N. Analysis of Genetic Variation and Potential Applications in Genome-Scale Metabolic Modeling(2015) (Accessed February 9, 2022).
102. Nam H, Campodonico M, Bordbar A, Hyduke DR, Kim S, Zielinski DC, et al. A Systems Approach to Predict Oncometabolites *via* Context-Specific Genome-Scale Metabolic Networks. *PLOS Comput Biol* (2014) 10: e1003837. doi: 10.1371/journal.pcbi.1003837
103. Brunk E, Sahoo S, Zielinski DC, Altunkaya A, Dräger A, Mih N, et al. Recon3D Enables a Three-Dimensional View of Gene Variation in Human Metabolism. *Nat Biotechnol* (2018) 36:272–81. doi: 10.1038/nbt.4072
104. Marín de Mas I, Torrents L, Bedia C, Nielsen LK, Cascante M, Tauler R. Stoichiometric Gene-to-Reaction Associations Enhance Model-Driven Analysis Performance: Metabolic Response to Chronic Exposure to Aldrin in Prostate Cancer. *BMC Genomics* (2019) 20:652. doi: 10.1186/s12864-019-5979-4
105. Raškevičius V, Mikalayeva V, Antanavičiūtė I, Ceslevičienė I, Skeberdis VA, Kairys V, et al. Genome Scale Metabolic Models as Tools for Drug Design and Personalized Medicine. *PLOS ONE* (2018) 13:e0190636. doi: 10.1371/journal.pone.0190636
106. Weglarz-Tomczak E, Mondeel TDGA, Piebes DGE, Westerhoff HV. Simultaneous Integration of Gene Expression and Nutrient Availability for Studying the Metabolism of Hepatocellular Carcinoma Cell Lines. *Biomolecules* (2021) 11:490. doi: 10.3390/biom11040490
107. Robaina Estévez S, Nikoloski Z. Generalized Framework for Context-Specific Metabolic Model Extraction Methods. *Front Plant Sci* (2014) 5:491. doi: 10.3389/fpls.2014.00491
108. Agren R, Bordel S, Mardinoglu A, Pornputtapong N, Nookaew I, Nielsen J. Reconstruction of Genome-Scale Active Metabolic Networks for 69 Human Cell Types and 16 Cancer Types Using INIT. *PLOS Comput Biol* (2012) 8: e1002518. doi: 10.1371/journal.pcbi.1002518
109. Lee D, Smallbone K, Dunn WB, Murabito E, Winder CL, Kell DB, et al. Improving Metabolic Flux Predictions Using Absolute Gene Expression Data. *BMC Syst Biol* (2012) 6:73. doi: 10.1186/1752-0509-6-73
110. Jerby L, Shlomi T, Ruppin E. Computational Reconstruction of Tissue-Specific Metabolic Models: Application to Human Liver Metabolism. *Mol Syst Biol* (2010) 6:401. doi: 10.1038/msb.2010.56
111. Wang Y, Eddy JA, Price ND. Reconstruction of Genome-Scale Metabolic Models for 126 Human Tissues Using mCADRE. *BMC Syst Biol* (2012) 6:153. doi: 10.1186/1752-0509-6-153
112. Pacheco MP, John E, Kaoma T, Heinäniemi M, Nicot N, Vallar L, et al. Integrated Metabolic Modelling Reveals Cell-Type Specific Epigenetic Control Points of the Macrophage Metabolic Network. *BMC Genomics* (2015) 16:809. doi: 10.1186/s12864-015-1984-4
113. Cardoso J, Redestig H, Galkina S, Sonnenschein N. *Driven* (2021). openCOBRA. Available at: https://github.com/opencobra/driven (Accessed February 9, 2022).
114. Chandrasekaran S, Price ND. Probabilistic Integrative Modeling of Genome-Scale Metabolic and Regulatory Networks in *Escherichia Coli* and *Mycobacterium Tuberculosis*. *Proc Natl Acad Sci* (2010) 107:17845–50. doi: 10.1073/pnas.1005139107
115. Yang L, Ebrahim A, Lloyd CJ, Saunders MA, Palsson BO. DynamicME: Dynamic Simulation and Refinement of Integrated Models of Metabolism and Protein Expression. *BMC Syst Biol* (2019) 13:2. doi: 10.1186/s12918-018-0675-6
116. Kaleta C, de Figueiredo LF, Schuster S. Can the Whole be Less Than the Sum of its Parts? Pathway Analysis in Genome-Scale Metabolic Networks Using Elementary Flux Patterns. *Genome Res* (2009) 19:1872–83. doi: 10.1101/gr.090639.108
117. Haraldsdóttir HS, Cousins B, Thiele I, Fleming RMT, Vempala S. CHRR: Coordinate Hit-and-Run With Rounding for Uniform Sampling of Constraint-Based Models. *Bioinformatics* (2017) 33:1741–3. doi: 10.1093/bioinformatics/btx052
118. Fallahi S, Skaug HJ, Alendal G. A Comparison of Monte Carlo Sampling Methods for Metabolic Network Models. *PLOS ONE* (2020) 15:e0235393. doi: 10.1371/journal.pone.0235393
119. Artyomov MN, Van den Bossche J. Immunometabolism in the Single-Cell Era. *Cell Metab* (2020) 32:710–25. doi: 10.1016/j.cmet.2020.09.013
120. Su Y, Wei W, Robert L, Xue M, Tsoi J, Garcia-Diaz A, et al. Single-Cell Analysis Resolves the Cell State Transition and Signaling Dynamics Associated With Melanoma Drug-Induced Resistance. *Proc Natl Acad Sci* (2017) 114:13679–84. doi: 10.1073/pnas.1712064115
121. Su Y, Ko ME, Cheng H, Zhu R, Xue M, Wang J, et al. Multi-Omic Single-Cell Snapshots Reveal Multiple Independent Trajectories to Drug Tolerance in a Melanoma Cell Line. *Nat Commun* (2020) 11:2345. doi: 10.1038/s41467-020-15956-9
122. Su Y, Chen D, Yuan D, Lausted C, Choi J, Dai CL, et al. Multi-Omics Resolves a Sharp Disease-State Shift Between Mild and Moderate COVID-19. *Cell* (2020) 183:1479–95.e20. doi: 10.1016/j.cell.2020.10.037
123. Su Y, Yuan D, Chen DG, Ng RH, Wang K, Choi J, et al. Multiple Early Factors Anticipate Post-Acute COVID-19 Sequelae. *Cell* (2022) 185:881–95.e20. doi: 10.1016/j.cell.2022.01.014
124. Wei W, Shin YS, Xue M, Matsutani T, Masui K, Yang H, et al. Single-Cell Phosphoproteomics Resolves Adaptive Signaling Dynamics and Informs Targeted Combination Therapy in Glioblastoma. *Cancer Cell* (2016) 29:563–73. doi: 10.1016/j.ccell.2016.03.012
125. Su Y, Shi Q, Wei W. Single Cell Proteomics in Biomedicine: High-Dimensional Data Acquisition, Visualization, and Analysis. *Proteomics* (2017) 17:1600267. doi: 10.1002/pmic.201600267
126. Heath JR, Ribas A, Mischel PS. Single-Cell Analysis Tools for Drug Discovery and Development. *Nat Rev Drug Discov* (2016) 15:204–16. doi: 10.1038/nrd.2015.16
127. Hrovatin K, Fischer DS, Theis FJ. Toward Modeling Metabolic State From Single-Cell Transcriptomics. *Mol Metab* (2022) 57:101396. doi: 10.1016/j.molmet.2021.101396
128. Zhang Y, Kim MS, Nguyen E, Taylor DM. Modeling Metabolic Variation With Single-Cell Expression Data. *bioRxiv [Preprint]* (2020). doi: 10.1101/2020.01.28.923680
129. Martins Conde P do R, Sauter T, Pfau T. Constraint Based Modeling Going Multicellular. *Front Mol Biosci* (2016) 3:3. doi: 10.3389/fmolb.2016.00003
130. Damiani C, Di Filippo M, Pescini D, Maspero D, Colombo R, Mauri G. popFBA: Tackling Intratumour Heterogeneity With Flux Balance Analysis. *Bioinforma Oxf Engl* (2017) 33:i311–8. doi: 10.1093/bioinformatics/btx251
131. Damiani C, Maspero D, Filippo MD, Colombo R, Pescini D, Graudenzi A, et al. Integration of Single-Cell RNA-Seq Data Into Population Models to Characterize Cancer Metabolism. *PLOS ONE Comput Biol* (2019) 15: e1006733. doi: 10.1371/journal.pcbi.1006733
132. Warburg O, Wind F, Negelein E. The Metabolism Of Tumors In The Body. *J Gen Physiol* (1927) 8:519–30. doi: 10.1085/jgp.8.6.519
133. Heiden MGV, Cantley LC, Thompson CB. Understanding the Warburg Effect: The Metabolic Requirements of Cell Proliferation. *Science* (2009) 324:1029–1033. doi: 10.1126/science.1160809

134. Dolfi SC, Chan LL-Y, Qiu J, Tedeschi PM, Bertino JR, Hirshfield KM, et al. The Metabolic Demands of Cancer Cells are Coupled to Their Size and Protein Synthesis Rates. *Cancer Metab* (2013) 1:20. doi: 10.1186/2049-3002-1-20
135. Koundouros N, Poulogiannis G. Reprogramming of Fatty Acid Metabolism in Cancer. *Br J Cancer* (2020) 122:4–22. doi: 10.1038/s41416-019-0650-z
136. Villa E, Ali ES, Sahu U, Ben-Sahra I. Cancer Cells Tune the Signaling Pathways to Empower De Novo Synthesis of Nucleotides. *Cancers* (2019) 11:688. doi: 10.3390/cancers11050688
137. Anderson KG, Stromnes IM, Greenberg PD. Obstacles Posed by the Tumor Microenvironment to T Cell Activity: A Case for Synergistic Therapies. *Cancer Cell* (2017) 31:311–25. doi: 10.1016/j.ccell.2017.02.008
138. Nilsson A, Nielsen J. Genome Scale Metabolic Modeling of Cancer. *Metab Eng* (2017) 43:103–12. doi: 10.1016/j.ymben.2016.10.022
139. Folger O, Jerby L, Frezza C, Gottlieb E, Ruppin E, Shlomi T. Predicting Selective Drug Targets in Cancer Through Metabolic Networks. *Mol Syst Biol* (2011) 7:501. doi: 10.1038/msb.2011.35
140. Resendis-Antonio O, Checa A, Encarnación S. Modeling Core Metabolism in Cancer Cells: Surveying the Topology Underlying the Warburg Effect. *PLOS One* (2010) 5:e12383. doi: 10.1371/journal.pone.0012383
141. Schwartz J-M, Barber M, Soons Z. Metabolic Flux Prediction in Cancer Cells With Altered Substrate Uptake. *Biochem Soc Trans* (2015) 43:1177–81. doi: 10.1042/BST20150149
142. Frades I, Foguet C, Cascante M, Araúzo-Bravo MJ. Genome Scale Modeling to Study the Metabolic Competition Between Cells in the Tumor Microenvironment. *Cancers* (2021) 13:4609. doi: 10.3390/cancers13184609
143. Lewis N, Abdel-Haleem A. The Evolution of Genome-Scale Models of Cancer Metabolism. *Front Physiol* (2013) 4:237. doi: 10.3389/fphys.2013.00237
144. Bordbar A, Palsson BO. Using the Reconstructed Genome-Scale Human Metabolic Network to Study Physiology and Pathology. *J Intern Med* (2012) 271:131–41. doi: 10.1111/j.1365-2796.2011.02494.x
145. Granata I, Troiano E, Sangiovanni M, Guarracino MR. Integration of Transcriptomic Data in a Genome-Scale Metabolic Model to Investigate the Link Between Obesity and Breast Cancer. *BMC Bioinf* (2019) 20:162. doi: 10.1186/s12859-019-2685-9
146. Herrmann HA, Rusz M, Baier D, Jakupec MA, Keppler BK, Berger W, et al. Thermodynamic Genome-Scale Metabolic Modeling of Metallodrug Resistance in Colorectal Cancer. *Cancers* (2021) 13:4130. doi: 10.3390/cancers13164130
147. Sahoo S, Ravi Kumar RK, Nicolay B, Mohite O, Sivaraman K, Khetan V, et al. Metabolite Systems Profiling Identifies Exploitable Weaknesses in Retinoblastoma. *FEBS Lett* (2019) 593:23–41. doi: 10.1002/1873-3468.13294
148. Lewis JE, Costantini F, Mims J, Chen X, Furdui CM, Boothman DA, et al. Genome-Scale Modeling of NADPH-Driven β-Lapachone Sensitization in Head and Neck Squamous Cell Carcinoma. *Antioxid Redox Signal* (2018) 29:937–52. doi: 10.1089/ars.2017.7048
149. Goldstein I, Yizhak K, Madar S, Goldfinger N, Ruppin E, Rotter V. P53 Promotes the Expression of Gluconeogenesis-Related Genes and Enhances Hepatic Glucose Production. *Cancer Metab* (2013) 1:1–6. doi: 10.1186/2049-3002-1-9
150. Ghaffari P, Mardinoglu A, Asplund A, Shoaie S, Kampf C, Uhlen M, et al. Identifying Anti-Growth Factors for Human Cancer Cell Lines Through Genome-Scale Metabolic Modeling. *Sci Rep* (2015) 5:8183. doi: 10.1038/srep08183
151. Uhlen M, Zhang C, Lee S, Sjöstedt E, Fagerberg L, Bidkhori G, et al. A Pathology Atlas of the Human Cancer Transcriptome. *Science* (2017) 357:eaan2507. doi: 10.1126/science.aan2507
152. Li X, Kim W, Juszczak K, Arif M, Sato Y, Kume H, et al. Stratification of Patients With Clear Cell Renal Cell Carcinoma to Facilitate Drug Repositioning. *iScience* (2021) 24:102722. doi: 10.1016/j.isci.2021.102722
153. Larsson I, Uhlén M, Zhang C, Mardinoglu A. Genome-Scale Metabolic Modeling of Glioblastoma Reveals Promising Targets for Drug Development (2020) (Accessed March 17, 2022).
154. Frezza C, Zheng L, Folger O, Rajagopalan KN, MacKenzie ED, Jerby L, et al. Haem Oxygenase is Synthetically Lethal With the Tumour Suppressor Fumarate Hydratase. *Nature* (2011) 477:225–8. doi: 10.1038/nature10363
155. Bidkhori G, Benfeitas R, Elmas E, Kararoudi MN, Arif M, Uhlen M, et al. Metabolic Network-Based Identification and Prioritization of Anticancer Targets Based on Expression Data in Hepatocellular Carcinoma. *Front Physiol* (2018) 9:916. doi: 10.3389/fphys.2018.00916
156. Li L, Zhou X, Ching W-K, Wang P. Predicting Enzyme Targets for Cancer Drugs by Profiling Human Metabolic Reactions in NCI-60 Cell Lines. *BMC Bioinf* (2010) 11:501. doi: 10.1186/1471-2105-11-501
157. Turanli B, Zhang C, Kim W, Benfeitas R, Uhlen M, Arga KY, et al. Discovery of Therapeutic Agents for Prostate Cancer Using Genome-Scale Metabolic Modeling and Drug Repositioning. *eBioMedicine* (2019) 42:386–96. doi: 10.1016/j.ebiom.2019.03.009
158. Çubuk C, Hidalgo MR, Amadoz A, Rian K, Salavert F, Pujana MA, et al. Differential Metabolic Activity and Discovery of Therapeutic Targets Using Summarized Metabolic Pathway Models. *NPJ Syst Biol Appl* (2019) 5:1–11. doi: 10.1038/s41540-019-0087-2
159. Zhang C, Aldrees M, Arif M, Li X, Mardinoglu A, Aziz MA. Elucidating the Reprograming of Colorectal Cancer Metabolism Using Genome-Scale Metabolic Modeling(2019) (Accessed February 6, 2022).
160. Ortmayr K, Dubuis S, Zampieri M. Metabolic Profiling of Cancer Cells Reveals Genome-Wide Crosstalk Between Transcriptional Regulators and Metabolism. *Nat Commun* (2019) 10:1841. doi: 10.1038/s41467-019-09695-9
161. Marin de Mas I, Aguilar E, Zodda E, Balcells C, Marin S, Dallmann G, et al. Model-Driven Discovery of Long-Chain Fatty Acid Metabolic Reprogramming in Heterogeneous Prostate Cancer Cells. *PLOS Comput Biol* (2018) 14:e1005914. doi: 10.1371/journal.pcbi.1005914
162. Gatto F, Miess H, Schulze A, Nielsen J. Flux Balance Analysis Predicts Essential Genes in Clear Cell Renal Cell Carcinoma Metabolism. *Sci Rep* (2015) 5:10738. doi: 10.1038/srep10738
163. Asgari Y, Salehzadeh-Yazdi A, Schreiber F, Masoudi-Nejad A. Controllability in Cancer Metabolic Networks According to Drug Targets as Driver Nodes. *PLOS ONE* (2013) 8:e79397. doi: 10.1371/journal.pone.0079397
164. Mazat J-P, Ransac S. The Fate of Glutamine in Human Metabolism. The Interplay With Glucose in Proliferating Cells. *Metabolites* (2019) 9:81. doi: 10.3390/metabo9050081
165. Cakmak A, Celik MH. Personalized Metabolic Analysis of Diseases. *IEEE/ACM Trans Comput Biol Bioinform* (2021) 18:1014–25. doi: 10.1109/TCBB.2020.3008196
166. Björnson E, Mukhopadhyay B, Asplund A, Pristovsek N, Cinar R, Romeo S, et al. Stratification of Hepatocellular Carcinoma Patients Based on Acetate Utilization. *Cell Rep* (2015) 13:2014–26. doi: 10.1016/j.celrep.2015.10.045
167. Özcan E, Çakır T. Reconstructed Metabolic Network Models Predict Flux-Level Metabolic Reprogramming in Glioblastoma. *Front Neurosci* (2016) 10:156. doi: 10.3389/fnins.2016.00156
168. Gámez-Pozo A, Trilla-Fuertes L, Berges-Soria J, Selevsek N, López-Vacas R, Díaz-Almirón M, et al. Functional Proteomics Outlines the Complexity of Breast Cancer Molecular Subtypes. *Sci Rep* (2017) 7:10100. doi: 10.1038/s41598-017-10493-w
169. Fuhr L, El-Athman R, Scrima R, Cela O, Carbone A, Knoop H, et al. The Circadian Clock Regulates Metabolic Phenotype Rewiring *Via* HKDC1 and Modulates Tumor Progression and Drug Response in Colorectal Cancer. *EBioMedicine* (2018) 33:105–21. doi: 10.1016/j.ebiom.2018.07.002
170. Wang F-S, Wu W-H, Hsiu W-S, Liu Y-J, Chuang K-W. Genome-Scale Metabolic Modeling With Protein Expressions of Normal and Cancerous Colorectal Tissues for Oncogene Inference. *Metabolites* (2020) 10:16. doi: 10.3390/metabo10010016
171. Fan J, Ye J, Kamphorst JJ, Shlomi T, Thompson CB, Rabinowitz JD. Quantitative Flux Analysis Reveals Folate-Dependent NADPH Production. *Nature* (2014) 510:298–302. doi: 10.1038/nature13236
172. Gatto F, Nookaew I, Nielsen J. Chromosome 3p Loss of Heterozygosity is Associated With a Unique Metabolic Network in Clear Cell Renal Carcinoma. *Proc Natl Acad Sci* (2014) 111:E866–75. doi: 10.1073/pnas.1319196111
173. Hur W, Ryu JY, Kim HU, Hong SW, Lee EB, Lee SY, et al. Systems Approach to Characterize the Metabolism of Liver Cancer Stem Cells Expressing CD133. *Sci Rep* (2017) 7:45557. doi: 10.1038/srep45557

174. Faubert B, Li KY, Cai L, Hensley CT, Kim J, Zacharias LG, et al. Lactate Metabolism in Human Lung Tumors. *Cell* (2017) 171:358–71.e9. doi: 10.1016/j.cell.2017.09.019
175. Asgari Y, Khosravi P, Zabihinpour Z, Habibi M. Exploring Candidate Biomarkers for Lung and Prostate Cancers Using Gene Expression and Flux Variability Analysis. *Integr Biol* (2018) 10:113–20. doi: 10.1039/c7ib00135e
176. Gatto F, Ferreira R, Nielsen J. Pan-Cancer Analysis of the Metabolic Reaction Network. *Metab Eng* (2020) 57:51–62. doi: 10.1016/j.ymben.2019.09.006
177. Shlomi T, Benyamini T, Gottlieb E, Sharan R, Ruppin E. Genome-Scale Metabolic Modeling Elucidates the Role of Proliferative Adaptation in Causing the Warburg Effect. *PLOS Comput Biol* (2011) 7:e1002018. doi: 10.1371/journal.pcbi.1002018
178. Vazquez A, Liu J, Zhou Y, Oltvai ZN. Catabolic Efficiency of Aerobic Glycolysis: The Warburg Effect Revisited. *BMC Syst Biol* (2010) 4:58. doi: 10.1186/1752-0509-4-58
179. Vazquez A, Oltvai ZN. Molecular Crowding Defines a Common Origin for the Warburg Effect in Proliferating Cells and the Lactate Threshold in Muscle Physiology. *PLOS ONE* (2011) 6:e19538. doi: 10.1371/journal.pone.0019538
180. Steenbergen R, Oti M, ter Horst R, Tat W, Neufeldt C, Belovodskiy A, et al. Establishing Normal Metabolism and Differentiation in Hepatocellular Carcinoma Cells by Culturing in Adult Human Serum. *Sci Rep* (2018) 8:11685. doi: 10.1038/s41598-018-29763-2
181. Wu H-Q, Cheng M-L, Lai J-M, Wu H-H, Chen M-C, Liu W-H, et al. Flux Balance Analysis Predicts Warburg-Like Effects of Mouse Hepatocyte Deficient in miR-122a. *PLOS Comput Biol* (2017) 13:e1005618. doi: 10.1371/journal.pcbi.1005618
182. Asgari Y, Zabihinpour Z, Salehzadeh-Yazdi A, Schreiber F, Masoudi-Nejad A. Alterations in Cancer Cell Metabolism: The Warburg Effect and Metabolic Adaptation. *Genomics* (2015) 105:275–81. doi: 10.1016/j.ygeno.2015.03.001
183. Duarte NC, Becker SA, Jamshidi N, Thiele I, Mo ML, Vo TD, et al. Global Reconstruction of the Human Metabolic Network Based on Genomic and Bibliomic Data. *Proc Natl Acad Sci* (2007) 104:1777–82. doi: 10.1073/pnas.0610772104
184. Thiele I, Swainston N, Fleming RMT, Hoppe A, Sahoo S, Aurich MK, et al. A Community-Driven Global Reconstruction of Human Metabolism. *Nat Biotechnol* (2013) 31:419–25. doi: 10.1038/nbt.2488
185. Mardinoglu A, Agren R, Kampf C, Asplund A, Nookaew I, Jacobson P, et al. Integration of Clinical Data With a Genome-Scale Metabolic Model of the Human Adipocyte. *Mol Syst Biol* (2013) 9:649. doi: 10.1038/msb.2013.5
186. Mardinoglu A, Agren R, Kampf C, Asplund A, Uhlen M, Nielsen J. Genome-Scale Metabolic Modelling of Hepatocytes Reveals Serine Deficiency in Patients With non-Alcoholic Fatty Liver Disease. *Nat Commun* (2014) 5:3083. doi: 10.1038/ncomms4083
187. Robinson JL, Kocabaş P, Wang H, Cholley P-E, Cook D, Nilsson A, et al. An Atlas of Human Metabolism. *Sci Signal* (2020) 13:eaaz1482. doi: 10.1126/scisignal.aaz1482
188. Di Filippo M, Colombo R, Damiani C, Pescini D, Gaglio D, Vanoni M, et al. Zooming-In on Cancer Metabolic Rewiring With Tissue Specific Constraint-Based Models. *Comput Biol Chem* (2016) 62:60–9. doi: 10.1016/j.compbiolchem.2016.03.002
189. Sertbaş M, Ülgen K, Çakır T. Systematic Analysis of Transcription-Level Effects of Neurodegenerative Diseases on Human Brain Metabolism by a Newly Reconstructed Brain-Specific Metabolic Network. *FEBS Open Bio* (2014) 4:542–53. doi: 10.1016/j.fob.2014.05.006
190. Liu Y, Beyer A, Aebersold R. On the Dependency of Cellular Protein Levels on mRNA Abundance. *Cell* (2016) 165:535–50. doi: 10.1016/j.cell.2016.03.014
191. Jiang L, Wang M, Lin S, Jian R, Li X, Chan J, et al. A Quantitative Proteome Map of the Human Body. *Cell* (2020) 183:269–83.e19. doi: 10.1016/j.cell.2020.08.036
192. Herrmann HA, Dyson BC, Vass L, Johnson GN, Schwartz J-M. Flux Sampling Is a Powerful Tool to Study Metabolism Under Changing Environmental Conditions. *NPJ Syst Biol Appl* (2019) 5:1–8. doi: 10.1038/s41540-019-0109-0
193. Desai TS, Srivastava S. FluxPyt: A Python-Based Free and Open-Source Software for 13C-Metabolic Flux Analyses. *PeerJ* (2018) 6:e4716. doi: 10.7717/peerj.4716
194. Matsuda F, Maeda K, Taniguchi T, Kondo Y, Yatabe F, Okahashi N, et al. Mfapy: An Open-Source Python Package for 13C-Based Metabolic Flux Analysis. *Metab Eng Commun* (2021) 13:e00177. doi: 10.1016/j.mec.2021.e00177
195. Antoniewicz MR. A Guide to 13C Metabolic Flux Analysis for the Cancer Biologist. *Exp Mol Med* (2018) 50:1–13. doi: 10.1038/s12276-018-0060-y
196. Lagziel S, Lee WD, Shlomi T. Studying Metabolic Flux Adaptations in Cancer Through Integrated Experimental-Computational Approaches. *BMC Biol* (2019) 17:51. doi: 10.1186/s12915-019-0669-x

Conflict of Interest: JH is a board member of PACT Pharma and Isoplexis and receives support from Gilead, Regeneron and Merck.

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

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Oncology

REVIEW
published: 11 July 2022
doi: 10.3389/fonc.2022.938847

Targeted Arginine Metabolism Therapy: A Dilemma in Glioma Treatment

Xiaoshuang Hou, Sui Chen, Po Zhang, Dongsheng Guo and Baofeng Wang**

Department of Neurosurgery, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

Efforts in the treatment of glioma which is the most common primary malignant tumor of the central nervous system, have not shown satisfactory results despite a comprehensive treatment model that combines various treatment methods, including immunotherapy. Cellular metabolism is a determinant of the viability and function of cancer cells as well as immune cells, and the interplay of immune regulation and metabolic reprogramming in tumors has become an active area of research in recent years. From the perspective of metabolism and immunity in the glioma microenvironment, we elaborated on arginine metabolic reprogramming in glioma cells, which leads to a decrease in arginine levels in the tumor microenvironment. Reduced arginine availability significantly inhibits the proliferation, activation, and function of T cells, thereby promoting the establishment of an immunosuppressive microenvironment. Therefore, replenishment of arginine levels to enhance the anti-tumor activity of T cells is a promising strategy for the treatment of glioma. However, due to the lack of expression of argininosuccinate synthase, gliomas are unable to synthesize arginine; thus, they are highly dependent on the availability of arginine in the extracellular environment. This metabolic weakness of glioma has been utilized by researchers to develop arginine deprivation therapy, which 'starves' tumor cells by consuming large amounts of arginine in circulation. Although it has shown good results, this treatment modality that targets arginine metabolism in glioma is controversial. Exploiting a suitable strategy that can not only enhance the antitumor immune response, but also "starve" tumor cells by regulating arginine metabolism to cure glioma will be promising.

Keywords: glioma, arginine metabolism, T lymphocytes, tumor microenvironment, metabolic reprogramming

OPEN ACCESS

Edited by:
Xuefei Li,
Shenzhen Institutes of Advanced Technology (CAS), China

Reviewed by:
Guohui Qin,
The First Affiliated Hospital of Zhengzhou University, China
Xu Qian,
Nanjing Medical University, China

****Correspondence:***
Dongsheng Guo
tjguodongsheng@163.com
Baofeng Wang
wbf620@163.com

Specialty section:
This article was submitted to Cancer Metabolism, a section of the journal Frontiers in Oncology

Received: *08 May 2022*
Accepted: *20 June 2022*
Published: *11 July 2022*

Citation:
Hou X, Chen S, Zhang P, Guo D and Wang B (2022) Targeted Arginine Metabolism Therapy: A Dilemma in Glioma Treatment. Front. Oncol. 12:938847. doi: 10.3389/fonc.2022.938847

INTRODUCTION

Glioma is the most common primary malignant tumor of the central nervous system (CNS), accounting for 48% of all primary malignant CNS tumors (1); the most malignant type of glioma is glioblastoma (GBM). Although various treatment modalities including surgery, radiotherapy, chemotherapy, tumor treatment fields, molecular targeted therapy as well as supportive care have been employed in the treatment of GBM, the median survival time of the patients is less than two years, and the 5-year survival rate is less than 10% (2). The main reasons for the poor prognosis of

patients with GBM are tumor infiltration, recurrence, and resistance to conventional therapy, which are closely related to intra-tumoral heterogeneity and phenotypic plasticity in GBM (3).

The CNS was considered an immune-privileged organ. However, this dogma was broken with the discovery that lymphatic vessels exist in the CNS (4) and that immune cells can cross the blood-brain barrier (BBB) (5). Many innate and acquired immune cells reside in the boundary zones of the CNS (6, 7). Different from the brain parenchyma, there is a large amount of lymphocyte infiltration that mediates the immune response (8). The lymphatic system in the meninges, and the cerebrospinal fluid, and the lymphocytes present in the meninges form a relatively mature network. This network allows antigens in the cerebrospinal fluid to enter the lymphatic system through the cervical lymph nodes, thereby initiating the activation of T cells (9, 10). This process has been confirmed in various diseases, including GBM. However, in pathological conditions such as GBM, the blood-brain barrier is destroyed; this results in increased permeability. Leukocytes, including antigen-presenting cells, enter the CNS through the choroid plexus, meningeal barrier, and postcapillary venules (5, 11), leading to infiltration of immune cells into tumor tissues (12, 13). GBM is a "cold" tumor owing to a lack of lymphocyte infiltration (14). The immune cells that infiltrate GBM are mainly macrophages and lymphocytes, such as CD4+ and CD8+ T cells; the concentration of T lymphocytes is positively correlated with the survival time of patients (15).

Advances in immunotherapy, such as the use of immune checkpoint inhibitors, have revolutionized cancer therapy. Unfortunately, these have been unsuccessful in the treatment of GBM (14, 16). The main obstacle in the treatment of GBM is the heterogenous and immunosuppressed tumor microenvironment, which results partly due to altered cellular metabolism (17). Cellular metabolism has become a determinant of the viability and function of cancer cells as well as immune cells. Tumors are metabolically reprogrammed to maintain enormous anabolic demands, which leads to the development of a microenvironment that is acidic, hypoxic, and devoid of the key nutrients required by immune cells. In this context, tumor metabolism is a checkpoint because it mediates tumor immune escape (18). The interplay between immune regulation and metabolic reprogramming in GBM is an active and stimulating area of research (18, 19). For example, enhanced glycolysis results in a glucose-starved microenvironment that makes tumors more aggressive. Glucose is a key nutrient that supports the rapid and dynamic transition of immune cells from the naïve state to an activated state (20). Reprogramming of amino acid metabolism in tumors often involves nutritional competition between cancer and immune cells. A large number of basic and clinical studies have shown that the use of new drugs that target tumor-dependent amino acid metabolism can effectively inhibit tumor growth. We noticed that arginine in the GBM microenvironment may be associated with the antitumor function of T lymphocytes.

Arginine promotes a series of metabolic reactions, including the synthesis of nitric oxide, polyamines, glutamine, and proline, all of which are important regulators of cell growth and survival (21). Arginine also exerts an essential regulatory effect on the immune system. Arginine-deficient T cells exhibit cell cycle arrest, impaired proliferation, reduced activation, and reduced antitumor activity (22–25). The reprogramming of arginine metabolism in GBM includes upregulation of the expression of amino acid transporters for intake of arginine, upregulation of the expression of arginase to decompose arginine, and downregulation of the expression of key enzymes involved in the endogenous arginine synthesis pathway. The former causes a deficit of arginine in the microenvironment, thereby inhibiting the function of T lymphocytes and promoting the formation of an immunosuppressive microenvironment. The latter represents a defect in cancer cell metabolism, and targeting this metabolic defect is a strategy used for treating tumors. Since the rate of proliferation of cancer cells is much higher, they require more nutrients, which exceeds their ability to synthesize amino acids (26). Cancer cells are dependent on extracellular arginine because of the decreased expression of arginine-synthesizing enzymes, argininosuccinate synthase (ASS1) and argininosuccinate lyase (ASL). In the absence of extracellular arginine, cancer cells become arginine dystrophic, or "arginine auxotrophic" (27). This strategy has been successfully used to treat acute lymphoblastic leukemia, in which asparaginase combined with chemotherapy has become the standard treatment (28). Mycoplasma infection was initially found to kill cancer cells (29). It was subsequently found that this is due to arginine deaminase (ADI), which degrades arginine in Mycoplasma (30, 31). Researchers then began using arginine deaminase and another enzyme, arginase (ARG), to break down arginine for the treatment of various tumors, including gliomas. Extensive preclinical and clinical research is being conducted on arginine deprivation therapy (32).

In this review, we describe how the unique metabolism of arginine in the glioma microenvironment leads to the suppression of the antitumor activity of T lymphocytes, thereby leading to tumor immune escape. We also discuss how targeting arginine metabolism in gliomas not only inhibits tumor growth, but also promotes effective and durable antitumor immunity.

METABOLISM OF ARGININE

Arginine Metabolism in Humans

Arginine is a semi-essential amino acid that is found in adults. The humans can synthesize arginine, but under certain physiological stresses, such as burns or severe immune challenges, the humans needs to supplement dietary arginine (33–36). Arginine in adult circulation has a short half-life (37). Plasma arginine concentration is regulated by dietary arginine intake, endogenous arginine synthesis, arginine catabolism, hepatic urea cycle, and protein synthesis. It is important to note that changes in the dietary intake of arginine do not alter the rate of its endogenous synthesis, which lays the foundation for targeting arginine metabolism for the treatment of some specific diseases (38). Endogenous arginine is mainly synthesized

through the intestinal–renal axis (39, 40). Although the urea cycle in the liver can synthesize arginine, there is very little net arginine synthesis in the liver (41, 42).

CNS tumors, such as gliomas, require more arginine; however, the CNS cannot increase the synthesis of arginine to meet the needs of cancer cells, and it can only increase arginine intake from the blood (43). For infiltrating immune cells, macrophages can express both ASS1 and ASL to synthesize arginine from citrulline (44, 45), which may be related to the fact that macrophages can account for 30%-50% of cells in the glioma microenvironment (46). However, not all immune cells simultaneously express all the enzymes required for *de novo* synthesis of arginine. For example, T cells rely only on a circulating supply of arginine or its immediate precursor.

In addition to protein synthesis, arginine has multiple functions such as vasodilation, neurotransmission, cell proliferation, and immune regulation (47, 48). The effect of arginine on the immune system has been gradually discovered in the last century. In 1968, the inhibitory effect of arginine deficiency on T lymphocyte activation *in vitro* was first described (49). Clinically, arginine is required for wound healing (50–52). Immune-enhancing diets (IED) use dietary arginine to stimulate the immune system (53, 54). These diets contain two to six times the arginine content of a normal diet. IEDs can boost immunity in trauma patients and reduce infection risk in surgical patients (55–57). It is important to note that IEDs do not benefit all patients (58). Determining whether arginine metabolism modulates immune cell function in specific diseases will undoubtedly lead to the development of more efficient individualized treatments.

Metabolism of Arginine in Cells

The intracellular arginine concentration is much higher than the extracellular or plasma arginine concentration. The arginine transporter in most cells is CAT-1, which transports arginine into cells to form the arginine pool. Several enzymes can break down arginine, including arginase, nitric oxide synthase (NOS), arginine decarboxylase, and arginine: glycine amidinotransferase (**Figure 1**) (33, 59).

Quantitatively, arginase is the most important enzyme for arginine decomposition in the body (60). Intracellular arginase hydrolyzes arginine to urea and ornithine. There are two arginase isoenzymes in humans, arginase 1 (Arg1) and arginase 2 (Arg2). Arginase 1 is located in the cytoplasm, its expression is restricted to specific cell types. Moreover, it is transcriptionally regulated by cytokines. Arginase 2 is primarily located in the mitochondria and exhibits a more ubiquitous and constitutive expression pattern, independent of cytokine regulation (61, 62). Ornithine is a metabolite of arginine. Ornithine can enter the urea cycle and is converted to citrulline by ornithine transcarbamylase (OTC). Citrulline synthesizes argininosuccinate through ASS1, which in turn synthesizes arginine through ASL, thus repeating the urea cycle. Ornithine can also generate polyamines *via* ornithine decarboxylase (ODC). Polyamines, including putrescine, spermine, and spermidine, are important products of the arginase metabolic pathway and have tumor-promoting effects (60, 63).

NOS is another important enzyme that breaks down arginine. It breaks down arginine to produce nitric oxide (NO) and citrulline. Notably, arginine is the only substrate for NO production (64). Intracellular arginine increases NO production in a dose-dependent manner (65). There are three distinct

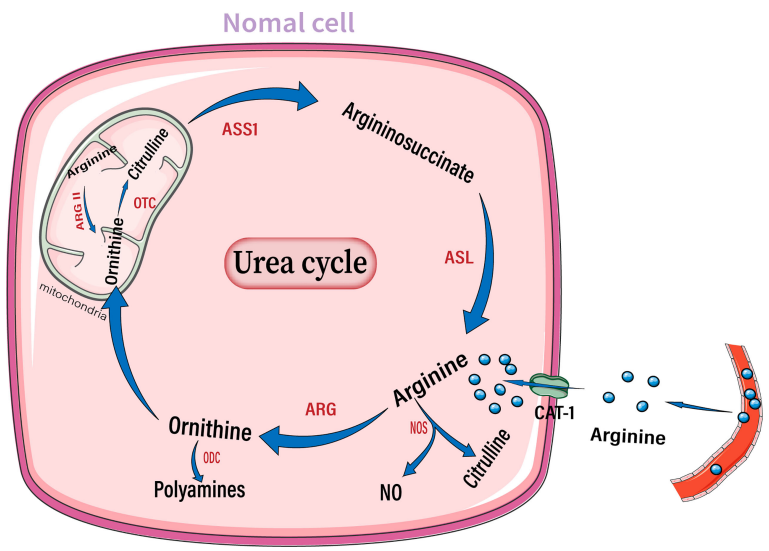


FIGURE 1 | Diagram of the arginine metabolism pattern in normal cells. After entering the extracellular matrix from the circulation, arginine enters the cell through the CAT-1 transporter on the cell membrane. Arginine can be broken down by NOS into NO and citrulline, or be broken down by arginase into ornithine, thus entering the urea cycle. Ornithine can also generate polyamines through ODC. Arginase II in mitochondria is also involved in the degradation of arginine. NO, Nitric oxide; NOS, nitric oxide synthase; ARG, arginase; ODC, ornithine decarboxylase; OTC, ornithine transcarbamylase; ASS1, argininosuccinate synthase; ASL, argininosuccinate lyase.

isoforms of nitric oxide synthase in the body: NOS1, NOS2, and NOS3, which are encoded by different genes. NOS1 and NOS3 are constitutively expressed in neural and endothelial cells, respectively. NOS2 is a ubiquitous isoform in immune cells, but is not constitutively expressed. Instead, its expression is induced by lipopolysaccharide and inflammatory cytokines; thus, it is called inducible NOS. The roles of NO in tumors are conflicting and may depend on the concentration of NO, type of effector cells, and duration of exposure (66). In general, low concentrations of NO may promote carcinogenesis, cancer cell proliferation, and tumor angiogenesis (67). However, high concentrations of NO can exert cytotoxic effects on tumor cells by inducing DNA damage (68). The complex role of NO in tumors suggests that a comprehensive evaluation of the effect of NO on tumors *in vivo* is essential when targeting arginine metabolism for the treatment of gliomas.

REPROGRAMMING OF ARGININE METABOLISM IN GLIOMA

Healthy adults obtain arginine primarily through dietary intake and intracellular protein degradation but can also synthesize arginine when needed. This is sufficient to meet the body's general arginine requirements (69). However, owing to metabolic reprogramming, cancer cells have a greater demand for arginine and rely on the extracellular pool of arginine to sustain their growth (70, 71). Moreover, ASS1 and ASL are downregulated in cancer cells, resulting in the inability to synthesize endogenous arginine, which makes cancer cells more dependent on the extracellular arginine pool (21, 72). This has laid the foundation for arginine deprivation therapy. The expression of ASS1 is varied in different types of tumors; further, the expression of ASS1 is heterogenous even within the same tumor, reflecting tumor heterogeneity (**Figure 2**). In the case of gliomas, 30% of GBM cell lines lack ASS1 expression (**Figure 3**) (73). In general, the downregulation of ASS1 is mediated by promoter methylation or hypoxia-inducible factor (HIF) 1α in multiple cancers. ASS1 levels in cancer are differentially regulated under various environmental conditions to metabolically benefit cancer progression. For example, ASS1 is downregulated under acidic conditions, and ASS1-depleted cancer cells maintain a higher intracellular pH, depend less on extracellular glutamine, and display higher glutathione levels. Cancer cells in an acidic or hypoxic environment downregulate the expression of the urea cycle enzyme ASS1, which provides them with redox and pH advantages, resulting in better survival (74). In response to genotoxic stress, p53 directly promotes ASS1 expression, resulting in increased ASS1 activity. P53-mediated ASS1 induction is a systemic response to genotoxic stress, which can lead to the rearrangement of arginine metabolism at the organism level, as seen in mice (75). Additionally, proline, creatine, and metabolites related to the arginine synthesis pathway were upregulated in ASS1-positive GBM cells compared to ASS1-negative cells. Pyruvic acid, citric acid, and α-ketoglutaric acid are metabolites in the initial phase of the citric acid cycle and are decreased in ASS1 positive cell lines (32). Similarly, tumor cells resistant to the arginine deprivation agent ADI-PEG20, which had upregulated ASS1 expression compared with sensitive cells, showed enhanced expression of glucose transporter-1 and lactate dehydrogenase-A, reduced expression of pyruvate dehydrogenase, and elevated sensitivity to the glycolytic inhibitors, 2-deoxyglucose and 3-bromopyruvate, consistent with the enhanced glycolytic pathway (the Warburg effect). Simultaneously, these cells showed higher glutamine dehydrogenase and glutaminase expression (76). Furthermore, activity-based proteomic profiling and phosphoproteomic profiling were performed before and after ADI-PEG20 treatment of ADI-PEG20-sensitive and -resistant sarcoma cells. Proteomic changes that facilitate oxaloacetate production by enhancing glutamine and pyruvate anaplerosis and altering lipid metabolism to recycle citrate for oxidative glutaminolysis have been elucidated (77). However, whether alterations in these metabolites affect the biological characteristics of gliomas is unclear. However, there is evidence that ASS1 may act as a tumor suppressor gene. For example, patients with GBM lacking ASS1 expression have worse prognosis than ASS1-

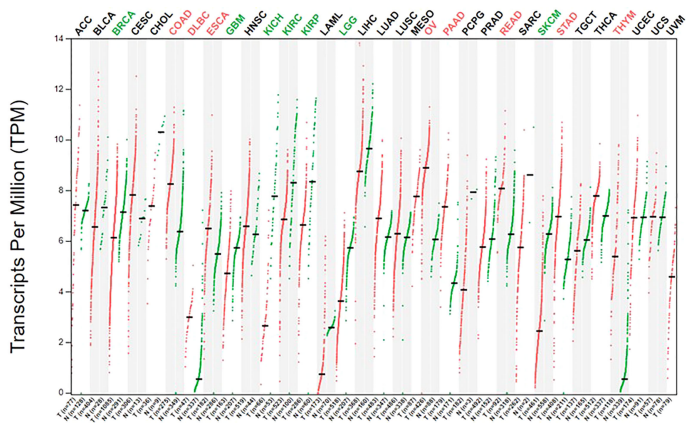


FIGURE 2 | The expression of ASS1 in human normal tissue and cancer cells.The gene expression profile across all tumor samples and paired normal tissues. The figure was excerpted from GEPIA2 (http://gepia2.cancer-pku.cn/#index).

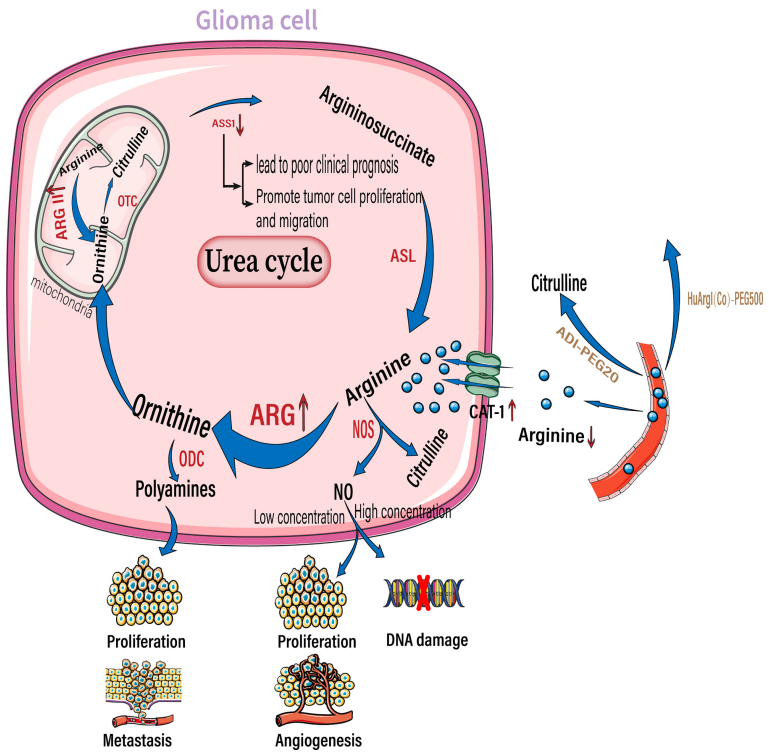


FIGURE 3 | Arginine reprogramming in glioma cells. In glioma cells, ASS1 expression was downregulated while CAT-1 and arginase were upregulated. The upregulated arginase is mainly arginase II located in the mitochondria. The use of ADI-PEG20 and HuArgI(CO)-PEG5000 to break down Circulating arginine results in a significant decrease in arginine concentration in the extracellular environment. Among the downstream metabolites of arginine, polyamines can promote tumor proliferation and metastasis, low concentrations of NO promote tumor proliferation and angiogenesis, and high concentrations of NO cause DNA damage. NO, nitric oxide; NOS, nitric oxide synthase; ARG, arginase; ODC, ornithine decarboxylase; OTC, ornithine transcarbamylase. ASS1,, argininosuccinate synthase; ASL, argininosuccinate lyase.;ADI-PEG20, pegylated arginine deaminase; HuArgI(CO)-PEG5000 ,Pegylated recombinant human arginase I.

positive patients (32). Consistent with this finding, decreased ASS1 levels were also significantly associated with postoperative lung metastases and poor clinical outcomes in patients with osteosarcoma. In preclinical studies, overexpression of ASS1 inhibited tumor growth (78). Epigenetic silencing of ASS1 can stimulate tumor cell proliferation and migration (79). These results suggest that ASS1 is a tumor suppressor gene (80). Interestingly, ASS1 may have opposite effects on other tumors. For example, the expression of ASS1 in gastric cancer can promote the invasion of cancer cells, resulting in poor prognosis in patients with gastric cancer (81, 82). Additionally, high ASS1 levels are an indicator of poor disease-free survival in patients with head and neck cancer (83). The dual role of ASS1 in tumors is not fully understood. However, these findings indicate that it is essential to fully understand the expression of ASS1 and its role, before using arginine deprivation therapy for the treatment of specific tumors. The influence of individual differences and tumor heterogeneity should also be considered. The mechanism of ASS1 downregulation, even though not fully elucidated, is undoubtedly beneficial for tumors if ASS1 acts as a tumor suppressor gene. Recent studies have shown that epigenetic changes in two genes involved in arginine biosynthesis in gliomas, namely CpG island methylation of ASS1 and ASL, lead to decreased protein expression. This results in glioma sensitivity to arginine deprivation therapy (84).

Reprogramming of arginine metabolism in gliomas provides a new approach for targeted therapy. But the downside is that this reprogramming also profoundly affects the infiltrating T lymphocytes. This has often been overlooked by researchers who use arginine deprivation therapy to treat gliomas. However, it is not clear whether adaptive changes in T lymphocytes in an arginine-deficient environment can cause glioma tolerance to arginine deprivation therapy. In the following discourse, we explain how gliomas cause a deficit of arginine in the tumor microenvironment and subsequent immunosuppression.

GLIOMA LEADS TO AN ARGININE-DEFICIENT IMMUNOSUPPRESSIVE MICROENVIRONMENT

Glioma Leads to an Arginine-Deficient Microenvironment

Solid tumors reside in harsh tumor microenvironments together with various stromal cell types. Tumor cells metabolically coordinate or compete with their “neighbors” to meet

biosynthetic and bioenergetic demands, while escaping immunosurveillance or therapeutic interventions. The consumption of essential nutrients by cancer cells directly limits the availability of nutrition to the tumor-killing immune cells; this is observed especially with cytotoxic T cells, leading to impaired antitumor immunity. In addition to rapid proliferation, cancer cells outcompete cancer cells by overexpressing transporters for nutrient uptake, and enzymes for nutrient catabolism (85). By upregulating amino acid transporters, glioma cells take up more arginine from the extracellular environment to meet their own proliferation and metabolism requirements. As their requirement of arginine is more than what they are capable of synthesizing, they are highly dependent on arginine availability in the extracellular environment. Therefore, arginine is an essential amino acid (86, 87). Elevated arginine catabolism is a common feature of the tumor microenvironment. The most important enzyme involved in arginine catabolism is arginase, which converts arginine into urea and ornithine. Arginase expression and activity are increased in patients with cancers including glioma, colon cancer, lung cancer, breast cancer, thyroid cancer, prostate cancer, compared to the surrounding healthy tissues in these patients (**Figure 3**) (88, 89). Arginase II is a major subtype expressed by tumor cells (90, 91). Moreover, arginase II is released from tumor cells, such as acute myeloblastoma, and is present in patient plasma at high concentrations (91). Whether arginase II is released outside the cell depends on the type of tumor, as neuroblastomas do not release free arginase II (90). It is unclear whether glioma cells that highly express arginase II release this enzyme. However, regardless of whether tumor cells release arginase, tumors with high arginase expression lead to local and systemic arginine deficiency. For example, patients with renal cell carcinoma and cervical cancer have a corresponding decrease in plasma arginine concentrations at diagnosis, which leads to a poorer prognosis (92, 93).

The increased uptake of arginine and high expression of arginase by tumor cells results in an immunosuppressive phenotype. As mentioned above, arginine deficiency leads to a series of inhibitory phenotypes such as decreased T-cell activation, impaired proliferation, and cycle arrest through multiple mechanisms. It was found that co-culture of Arg2-expressing cancer cells with T cells was sufficient to induce arginine depletion and lead to impaired T-cell proliferation, decreased IFN-γ release, and PD-1 upregulation (25). Moreover, T-cell and myeloid cell infiltration is reduced in head and neck squamous cell carcinomas with high arginase II expression (94). Likewise, in acute myeloid leukemia with high arginase II expression, the surrounding monocytes were more polarized to M2-like macrophages (91). Conversely, arginine replenishment (95) or the use of small-molecule inhibitors of arginase II (91) can alleviate arginine-deficient immunosuppression and reduce T-cell dysfunction (25).

In addition to tumor cells, immunosuppressive cells expressing arginase 1 form an inhibitory immune barrier. The accumulation of ARG1-expressing immunomodulatory cells, including M2-like tumor-associated macrophages, tolerogenic DCs, MDSCs, and Treg cells, in the tumor microenvironment (TME) may suppress antitumor immunity by degrading arginine, thus limiting the availability of this amino acid to T cells (96, 97). Mouse and human tumor cells can secrete soluble factors, such as GM-CSF and G-CSF, which lead to the recruitment and accumulation of MDSCs (98). In GBM patients, the number of circulating MDSCs with high Arg1 expression increases (99). Overexpression of Arg1 in MDSCs leads to downregulation of the CD3ζ chain, which adversely affects CD4+ and CD8+ T cells (100). Additionally, MDSCs exhibit functional similarities to M2-like macrophages (101), including IL-10, TGF-β, and IDO expression (102). This suggests that immunosuppressive cells are closely linked to arginine metabolism; however, this requires further investigations.

Depleting important nutrients such as arginine is a key strategy for cancer cells to evade immunity. Although many tumors are arginine auxotrophic (21), a large proportion can tolerate a low-arginine state (91, 103). This suggests that there must be a unique mechanism that allows these tumors to tolerate an arginine-deficient environment. These tumor cells can synthesize arginine from citrulline by upregulating ASS1. In the absence of arginine, ASS1 transcription is induced by the binding of ATF4 and CEBPb to the enhancer of ASS1. But in T cells, the situation is completely different. Arginine deficiency leads to chromatin compaction and inhibits histone methylation in T cells, which disrupts the binding of ATF4 and CEBPb to ASS1 enhancers and prevents the transcription of target genes (104). These findings help explain the differences in arginine metabolism between tumor cells and T cells and can aid in the development of more effective targeted therapies for the treatment of gliomas.

Arginine Deficiency Suppresses the Antitumor Function of T Lymphocytes

Tumor-infiltrating immune cells typically experience metabolic stress as a result of the dysregulated metabolic activity of tumor cells, leading to impaired antitumor immune responses. Activated T cells consume a large amount of arginine and rapidly convert it into downstream metabolites, resulting in a significant decrease in intracellular arginine levels. T cells are extremely sensitive to extracellular concentrations of arginine because of their low or absent expression of arginine synthase ASS1 and OTC (105, 106). Various studies have demonstrated that arginine deficiency leads to decreased T-cell activation, impaired proliferation, cycle arrest, decreased cytokine (IFN-γ) release, and increased expression of immunosuppressive molecules (PD-1) (**Figure 4**) (22–25). The low arginine levels in the TME also impairs the proliferation of chimeric antigen receptor T cells (CAR-T), limiting their therapeutic effects (107).

In contrast to the lack of arginine, high arginine levels can increase the antitumor activity of T cells, which may be due to a combination of phenotypic alterations, including increased T-cell viability, improved metabolic adaptability, and maintenance of a central memory-like phenotype (95). Therefore, sufficient extracellular arginine is critical for T-cell function. Researchers have exploited the beneficial effects of arginine on T-cell survival and antitumor function to improve adoptive T-cell therapy. For

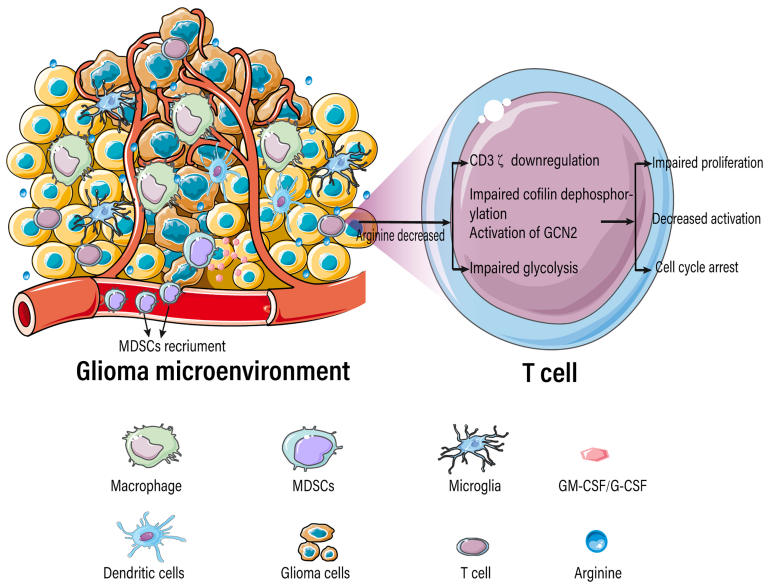


FIGURE 4 | Arginine-deficient glioma microenvironment suppresses T-cell function.

example, CAR-T cells have been reconstituted to express the enzymes ASS1 and OTC, which are required for arginine synthesis. This increases the arginine content in CAR-T cells, thus enhancing the activity of CAR-T cells *in vivo* against solid and hematological tumors (107).

Arginine deficiency-mediated suppression of T-cell function is caused due to a myriad of factors including downregulation of the CD3ζ subunit of the T-cell receptor complex (108–111), damage to cofilin dephosphorylation (112), blockade of protein translation by activation of general control nonderepressible 2 (GCN2) (113), blockage of glycolysis (114), decreased expression of early T-cell activation markers CD25 and CD69 (115), and aberrant expression of D-type cyclins (22, 116, 117).

It is important to note that most studies on the effects of arginine on T cells are based on interventions in extracellular arginine concentrations. For example, change in the concentration of arginine in the T-cell medium. However, extracellular and intracellular arginine pools are not freely interchangeable (118), which means that extracellular arginine supply may not be a reliable indicator of intracellular arginine availability. Recent studies on Arg2 in T cells further demonstrated that the intracellular metabolism of arginine profoundly alters T-cell function. Pharmacological inhibition of arginase increases activation and survival of human T cells *in vitro*. Since human T cells express ARG2, but not ARG1, this suggests that such effects are caused by Arg2 (48). Studies have also found that deletion of Arg2 germline and adoptive transfer of Arg2−/− CD8+ T cells significantly reduced tumor growth in preclinical cancer models by enhancing CD8+ T-cell activation, cytotoxic function, and persistence (48, 119). Importantly, these experiments were performed under arginine excess conditions and, therefore, did not depend on extracellular arginine availability. This indicates that the observed changes are mainly caused by a cellular autonomous mechanism, and that we should focus on the direct effects of intracellular arginine pools on T-cell functions.

ARGININE REPLENISHMENT THERAPY FOR GLIOMA

Researchers have attempted arginine replenishment therapy to treat tumors, by increasing the availability of arginine to improve antitumor immunity. One study found that oral administration of arginine and an anti-PD-L1 antibody restricted tumor growth and increased survival in mice, suggesting a synergistic effect of arginine and PD-L1 blockers. To achieve the desired antitumor effect, mice must be administered a relatively high dose of arginine (2 mg/g of body weight). In comparison, an adult patient weighing 75 kg would require 150g of arginine per day, which is unrealistic. Therefore, researchers have developed metabolically engineered bacteria called L-Arg bacteria, to be planted in the tumors, which will produce large amounts of arginine. L-Arg bacteria and PD-L1 blockers can synergistically inhibit tumor growth, increase the infiltration of CD4+ and CD8+ T cells, and reduce the infiltration of Treg cells in the tumor. Further studies also found that this combination reduced the percentage of PD-1+LAG-3+ T cells, indicating the persistence of effector T-cell function with the simultaneous increase in the formation of tumor-specific memory T cells (120). Similar studies have found that arginine increases radiosensitivity in patients with brain metastases. Additional oral administration of arginine before standard radiotherapy in 31 patients with brain metastases significantly improved the

therapeutic effect of radiotherapy. This therapeutic effect is due to NO-induced metabolic inhibition, which increases the susceptibility of NOS2-expressing cancer cells to DNA damage (121). NOS2 activity is required for tumor brain metastasis and it can decompose arginine to NO. Arginine increases radiosensitivity through an NO-mediated mechanism, and high intratumoral NO concentrations lead to a decrease in tumor glycolysis and thus a decrease in lactate levels. These metabolic changes ultimately impair the repair of radiation-induced DNA damage in cancer cells. In addition, the authors suggested that the enhanced overall antitumor effect may also be due to immune activation. In mouse tumor models, oral administration of arginine improved the metabolic adaptability of T cells, which is critical for antitumor responses (95). Administration of arginine prior to radiation therapy reversed radiation-induced T-cell and B-cell dependent immune dysfunction in mice (122). Although this mechanism has not been fully elucidated, it is speculated that arginine-induced reduction in lactate levels may also contribute to the enhanced antitumor activity of tumor-infiltrating lymphocytes (123).

ARGININE DEPRIVATION THERAPY FOR GLIOMA

Arginine deprivation therapy is a novel antimetabolic strategy that exploits the differential expression of key urea cycle enzymes to treat arginine auxotrophic tumors. Arginine deaminase (ADI), a metabolic enzyme extracted from Mycoplasma (124), catalyzes the conversion of arginine to citrulline. Owing to its instability, strong immunogenicity, and short half-life (5 h), ADI is combined with polyethylene glycol (ADI-PEG20) to reduce antigenicity and prolong half-life (125). Synthetic human arginase 1 (HuArgI) is another arginine deprivation agent used to treat arginine auxotrophic tumors. Its activity is also enhanced by adding polyethylene glycol and replacing Mn2+ with Co2+, resulting in HuArgI(CO)-PEG5000. HuArgI(CO)-PEG5000 lasts longer in serum, has better catalytic activity, and is less exposed to the immune system (126–128).

If the cells were not rescued by adding citrulline after arginine depletion, these cell lines were completely auxotrophic to arginine; however, when rescued after adding citrulline, the cell lines became partially auxotrophic. Pegylated recombinant human arginase I was used to target nine GBM cell lines and human fetal glial cells (SVG-p12), and was found to be cytotoxic to all GBM cell lines except SVG-p12 cells, which shows selective cytotoxicity induced by arginine deprivation. Subsequent addition of citrulline rescued these six GBM cell lines. The ability of citrulline to rescue cells was dependent on argininosuccinate synthase 1 expression, and cells that were not rescued were negative for ASS1 expression. Knockdown of ASS1 reversed the ability of citrulline to rescue GBM cells, further illustrating the dependence on ASS1 expression (129). Approximately 30% of GBMs lack ASS1 expression and can be targeted by arginase I, which has no cytotoxicity to normal glial cells. Likewise, depletion of arginine using pegylated arginine deaminase resulted in cell death *in vitro* and tumor regression in orthotopic xenograft models, whereas ASS1-expressing GBM cells were unaffected (84, 130). In addition, researchers also found that the use of arginine deprivation agents in combination with other treatments showed better therapeutic effects. Many studies have described the molecular mechanism of arginine deprivation in ASS1 deficient tumors, thereby uncovering additional vulnerabilities in these tumors. This has prompted the use of other drugs in combination with arginine deprivation therapy for more effective killing of tumor cells. For example, TRAIL is used for mesothelioma (131), cisplatin is used for various tumor types (132), and chloroquine is used for sarcoma (133). The combination of arginine deprivation and canavanine, a plant-derived arginine analog, is a novel approach to glioma treatment. This combination therapy profoundly affects cell viability, morphology, motility, and adhesion. It also disrupts the cytoskeleton and mitochondrial network, thereby inducing apoptosis. At the molecular level, canavanine inhibits pro-survival kinases such as FAK, AKT, and AMPK. Importantly, these effects are selective to GBM cells, as shown by their less pronounced effects on rat glial cells (134). Similarly, the combination of ADI and Palomid 529, an inhibitor of mTORC1 and mTORC2 complexes, showed a potent cytotoxic effect in glioma cell lines. In addition, ADI combined with chloroquine showed an enhanced antitumor effect. *In vivo*, ADI alone and the combination of ADI and SAHA, a protein deacetylase inhibitor, effectively inhibited the growth of xenograft tumors (135). A recent phase I clinical trial preliminarily verified the therapeutic effects of arginine deprivation therapy. Ten patients with severe ASS1-deficient recurrent high-grade gliomas were treated with ADI-PEG20 in combination with pemetrexed and cisplatin. The treatment was safe and well tolerated by the patients. The best overall response was stable disease in eight patients (80%). The results showed that the treatment was well tolerated and 80% of patients had stable overall efficacy, with plasma arginine significantly suppressed below baseline levels. However, the titers of anti-ADI-PEG20 antibodies in patients increased, indicating the production of neutralizing antibodies, which may affect the therapeutic effect of ADI-PEG20 (136). Additional clinical studies on arginine depletion in glioma treatment are presented in **Table 1**.

The above findings suggest that arginine deprivation therapy is only effective in ASS1-negative glioma and has little effect on ASS1-positive glioma (including adaptive transcriptional upregulation of ASS1 after treatment), which greatly limits the clinical applications of ADT. Therefore, researchers have attempted to combine ADT with other treatment modalities to improve the curative effect of the treatment for ASS1-positive gliomas. Animals bearing intracranial human GBM tumors of varying ASS status were treated with ADI-PEG20 alone or in combination with temozolomide and monitored for tumor growth and regression. ADI-PEG20 monotherapy significantly reduces intracranial growth of ASS1-negative GBM and extends survival of mice carrying ASS1 negative GBM without obvious toxicity. ADI-PEG20 combined with temozolomide shows enhanced antitumor effects in both ASS1-negative and ASS1-positive backgrounds. The mechanism underlying this effect is unclear, but these results suggest that ADI-PEG20 in combination with TMZ may be clinically useful in both ASS1-

TABLE 1 | Clinical studies on arginine depletion in glioma treatment.

Disease	Treatment	Clinical phase	No. of patients	Status	Clinical Trials.gov Identifier
Recurrent high-grade glioma (ASS1-deficient)	ADI-PEG20 with pemetrexed and cisplatin	Phase I	10	Terminated	NCT02029690
Glioblastoma multiforme	ADI-PEG 20 with Radiotherapy and Temozolomide	Phase I	32 (Estimated)	Recruiting	NCT04587830
Advanced solid cancers	ADI-PEG 20 with pembrolizumab	Phase I	33	Terminated	NCT03254732
High-grade gliomas and others	rhArg1peg5000	phase I/II	64 (Estimated; Children and Young Adults)	Unknown	NCT03455140
Advanced/Metastatic solid tumors	INCMGA00012 (PD-1 Inhibitor), INCB001158(Arginase Inhibitor), and the combination	Phase I	18	Completed	NCT03910530
Advanced/Metastatic solid tumors	INCB001158 with chemotherapy	phase I/II	149	Active, not recruiting	NCT03314935

negative and ASS1-positive settings (130). In addition, ADT combined with radiotherapy may be a new treatment strategy for patients with GBM. ADT caused significant radiosensitization, which was more pronounced in a GBM cell model with loss of function of p53 than in its p53- wildtype counterpart. This synergistic effect was independent of basic and induced ASS1 or ASL expression (137). ADI-PEG20 also significantly enhanced the efficacy of radiotherapy for ASS1-positive GBM *in vivo* (73). However, ADT combined with radiotherapy has not yet been studied in clinical trials.

The mechanism of glioma cell death in the absence of arginine has not yet been fully elucidated. GBM exhibits caspase-independent, non-apoptotic cell death upon arginine deprivation. The latter, a process known as autophagy, provides a temporary but limited supply of arginine through the destruction of intracellular organelles. Therefore, this process protects against cell death, but leads to non-apoptotic death in the long run. The autophagy inhibitor, chloroquine, was added to GBM cells treated with HuArgI(CO)-PEG5000. As expected, the inhibition of autophagy increased the sensitivity of cells to HuArgI(CO)-PEG5000 (129, 138). After emphasizing the effect of arginine deficiency on cell viability, it is important to observe the effect of arginine deficiency on cell motility and migration ability. Arginine deficiency affects tumor cell morphology and inhibits motility, invasiveness, and adhesion. Moreover, it has little effect on normal glial cells. This is because of specific changes in actin assembly caused by arginine deprivation in gliomas. Arginine deprivation reduces β-actin filament content and affects N-terminal arginylation. This suggests that arginine deprivation-based therapeutic strategies can inhibit the invasive process of highly malignant brain tumors (139).

The combined treatment with ADI-PEG20 significantly enhanced the efficacy of GBM radiotherapy in a non-arginine auxotrophic background. This combination results in a durable, complete radiological, and pathological response. It also prolonged disease-free survival in an *in situ* model of GBM with no apparent toxicity (73). Further studies found that the combination treatment resulted in downregulation of Arg1 and upregulation of inducible NOS. Under arginine-deficient conditions, inducible NOS has a higher affinity for arginine than for Arg1. Combination therapy increased the production of NO, which further formed cytotoxic peroxynitrite (140). This could enhance the sensitivity of ASS1-positive GBM to ionizing radiation (141). In addition, arginine deficiency greatly reduces vasogenic edema and neovascularization, which are typical features of GBM (142, 143). The antiangiogenic activity of ADI appears to be partly due to the twisting of actin filaments, which prevents blood vessels from sprouting, blooming, and growing. ADI-PEG20 also inhibits HIF, particularly HIF-1α (144). HIF-1α is associated with a decrease in the expression of vascular endothelial growth factor, which induces blood vessel growth. HIF has also been implicated in the pathogenesis of GBM (145). High HIF-1α levels also reduce glioma responsiveness to TMZ (146). Thus, ADI-PEG20 has antitumor effects, at least in part, due to its anti-HIF effects.

The antitumor properties of ADT have been extensively investigated. ADT inhibits the growth of auxotrophic cancers *in vitro* and *in vivo*. However, its impact on immune cells in the tumor microenvironment remains, largely, unknown. The removal of arginine can theoretically impair the immune function of T cells. Interestingly, ADI-PEG20 led to a marked increase in tumor-infiltrating CD4+ and CD8+ T cells in a syngeneic B16-F10-melanoma mouse model (147). Similarly, arginine deprivation combined with radiotherapy increased recruitment of microglia into tumors in a glioma model and enhanced their activity and phagocytic phenotype. Arginine deprivation switched the activation of tumor-associated macrophages/microglia from a tumor-supporting phenotype to a more phagocytosis-competent and, hence, tumor-inhibiting phenotype. Simultaneously, a significant increase in the number of CD4+ and CD8+ T cells and a corresponding decrease in FoxP3+ regulatory cells was observed in the glioma microenvironment (73). Despite the increased number of infiltrating T cells, it remains unknown whether T-cell function is affected. It is important to further explore changes in the tumor immune microenvironment after ADT treatment.

Studies on the potential mechanism of ADT resistance have found that the re-expression of ASS1, production of neutralizing antibodies to arginine deprivation agents, and autophagy are the main causes. ADT, by nutrient starvation or exposure to ADI-PEG20, induces adaptive transcriptional upregulation of ASS1 and

ASL in glioma cells *in vitro*, thereby conferring resistance to ADI-PEG20 treatment. The specific mechanism of the adaptive transcriptional upregulation of ASS1 and ASL is unknown, but studies in melanoma suggest that accumulated cMyc can induce ASS1 expression by interacting with the ASS1 promoter (148, 149). Although the modification of ADI by conjugation with polyethylene glycol can reduce its immunogenicity, the production of anti-ADI antibodies has also been observed in patients enrolled in clinical trials. This suggests that long-term treatment may lead to the development of resistance due to the production of neutralizing antibodies. This phenomenon may explain the negative correlation between plasma neutralizing antibody levels and duration of arginine depletion after ADI-PEG20 treatment (136). When arginase I was combined with the autophagy inhibitor, chloroquine, to treat GBM *in vitro*, the inhibition of autophagy increased cellular sensitivity to arginase I. This finding suggested that autophagy plays a supporting role in ADI resistance (129). Arginine deprivation agents for cancer treatment should have low toxicity, non-immunogenicity (to prevent antibody production and allergic reactions), rapid action (to delay the emergence of resistance), and long circulating half-lives (to achieve sustained arginine consumption) (150). It is worth investigating whether low arginine levels during arginine deprivation therapy can adversely affect antitumor immunity, since T-cell function is regulated by arginine. In addition, the reconstruction of adaptive immune function against the background of arginine-mediated tumor immune escape is a promising therapeutic strategy.

Metabolic reprogramming is often mediated by oncogenic signaling pathways. In particular, mTOR signaling is commonly activated in tumors and controls cancer cell metabolism by altering the expression and/or activity of several key metabolic enzymes (151). Conversely, metabolic alterations affect mTOR signaling. mTORC1 is one of the mechanisms that checks cellular amino acid levels and/or nutritional deprivation in cells. For example, arginine activates mTORC1 through the GATOR1/2-Rag pathway by directly binding to the arginine sensor CASTOR1 (Cellular Arginine Sensor for mTORC1) (151). Interestingly, ASS1 knockdown results in increased mTORC1 activity in osteosarcoma cells, potentially due to increased aspartate levels (86). Treatment with rhARG reduces mTORC1 activity and induces cytotoxicity and apoptosis in non-SCLC cells (152). However, resistance to arginine deprivation agents has been observed. ADI-PEG20-resistant tumor cells exhibited reduced mTOR signaling but enhanced AKT signaling, which led to the stabilization of MYC. MYC in turn induces ASS1 expression by competing with HIF1α for ASS1 promoter-binding sites (76). The molecular mechanism underlying the downregulation of mTOR signaling in ADI resistance remains unclear.

ADVANTAGES AND DISADVANTAGES OF TARGETING ARGININE THERAPIES FOR GLIOMA

We now summarize the advantages and disadvantages of Arginine deprivation therapy and Arginine replenishment therapy as follows:

Advantages of Arginine deprivation therapy (1): There are mainly five enzymatic agents catabolizing free arginine in theory (NOS, glycine amidinotransferase, arginine decarboxylase, arginase, and arginine deiminase) (153). This provides a variety of options for arginine deprivation therapy (2). Arginine deprivation therapy achieves its therapeutic effect by lowering the plasma arginine concentration, which is especially appropriate for intracranial tumors and is no longer hindered by the blood-brain barrier (3). Arginine deprivation therapy has completed different clinical trials in patients with metastatic melanoma and mesothelioma with promising results (154, 155) (4). Arginine deprivation therapies have different mechanisms in tumors, such as induction of autophagy, ROS overproduction, cell cycle arrest, and caspase-dependent/independent apoptosis in cells. Thus, AD therapy has the potential to treat tumors in combination with other treatments.

Disadvantages of Arginine deprivation therapy (1): The resistance of tumors to arginine deprivation agents is currently the biggest obstacle, mainly due to the re-expression of ASS1, production of neutralizing antibodies to arginine deprivation agents, and autophagy. We urgently need to elucidate the underlying mechanisms of drug resistance to increase their therapeutic efficacy against tumors (2). The therapeutic effect of arginine deprivation depends largely on whether the tumor is auxotrophic. In other words, it depends on the expression of ASS1 in tumor cells. This greatly limits the application of arginine deprivation agents. However, there are ongoing research studies to circumvent this problem. For example, a combination of arginine deprivation therapy with radiotherapy or TMZ has shown a good therapeutic effect on ASS1-positive gliomas.

Advantages of Arginine replenishment therapy: Arginine is an inexpensive, readily available amino acid that cancer patients only need to consume orally. This greatly increases the convenience of this treatment. Moreover, arginine is a nutrient needed by the body and does not produce toxic side effects like other chemotherapeutic drugs.

Disadvantages of Arginine replenishment therapy: Arginine replenishment therapy requires a high concentration of arginine in the tumor microenvironment in order to achieve a good therapeutic effect. Achieving the required concentration poses a challenge that needs to be addressed urgently.

DISCUSSION

The advantages of targeting arginine in the treatment of gliomas are evident. It kills tumor cells directly or indirectly by interfering with tumor cell metabolism, without affecting normal cell function. Concurrently, it can bypass the blood-brain barrier, which is especially suitable for intracranial diseases. Arginine deprivation therapy works directly on tumor cells. The combination of arginine deprivation with other treatments has shown great potential and application value, and requires further in-depth research. Arginine replenishment therapy is more likely to act on immune cells and affect tumor cells, which is an indirect

mechanism. Although the two treatments seem contradictory, differences in their mechanisms of action make us interested in finding ways to combine them. Currently, targeting arginine metabolism to treat glioma faces the dilemma of choosing arginine deprivation therapy or arginine replenishment therapy. The former achieves tumor inhibition by "starving" tumor cells, but its negative effects are often ignored by researchers. The arginine-deficient extracellular environment created by arginine deprivation agents undoubtedly exerts a strong inhibitory effect on antitumor T cells. Further studies are required to determine whether the suppressed T cells are responsible for the poor effects of arginine deprivation therapy. The latter increases the antitumor activity of T cell by fulfilling their arginine requirements. Likewise, the arginine replenishment therapy "feeds" tumor cells. The direct effect of excess arginine on glioma cells is unclear, but we do not want tumor cells to be "nutrient-rich." Another strategy to increase arginine levels in the body is to prevent its breakdown. In mouse tumor models, ARG1 inhibitors, which prevent the breakdown of arginine, increase CD8+ T-cell infiltration and stimulate the production of inflammatory cytokines in the TME (97, 156). Further studies are needed to determine the therapeutic effect of ARG1 inhibitors on glioma. Most existing studies describe arginine deprivation therapy as the chosen method to treat brain tumors; however, a few studies have also described arginine replenishment therapy to treat brain tumors. Here, we hope to adopt a suitable strategy to combine the two strategies, both "starving" tumor cells and enhancing antitumor immune response. CAR-T therapy combined with arginine deprivation therapy may be an effective strategy to circumvent this pitfall. CAR-T cells can recombinantly express ASS1 and OTC, increasing the arginine content in cells. This increases the persistence of CAR-T cells *in vivo* (107). However, we still need to conduct extensive preclinical studies to determine the effectiveness of this therapy. We hope that this will open new avenues for comprehensive treatment for glioma.

AUTHOR CONTRIBUTIONS

XH and BW contributed to the conceptual design, and helped in the writing and editing of the manuscript. SC, PZ, and DG revised the manuscript and commented on its previous versions. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China (Grant No. 82072797), the National Natural Science Foundation of China (Grant No. 81874086), and Hubei Provincial Natural Science Foundation of China (Grant No. 2020CFB671)

ACKNOWLEDGMENTS

I would like to thank smart servier (https://smart.servier.com/) for the vector material used for the drawing.

REFERENCES

1. Ostrom QT, Cioffi G, Gittleman H, Patil N, Waite K, Kruchko C, et al. Cbtrus Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2012-2016. *Neuro-oncology* (2019) 21(Suppl 5):v1–v100. doi: 10.1093/neuonc/noz150
2. Khasraw M, Fujita Y, Lee-Chang C, Balyasnikova IV, Najem H, Heimberger AB. New Approaches to Glioblastoma. *Annu Rev Med* (2022) 73:279–92. doi: 10.1146/annurev-med-042420-102102
3. Neftel C, Laffy J, Filbin MG, Hara T, Shore ME, Rahme GJ, et al. An Integrative Model of Cellular States, Plasticity, and Genetics for Glioblastoma. *Cell* (2019) 178(4):835–49.e21. doi: 10.1016/j.cell.2019.06.024
4. Louveau A, Smirnov I, Keyes TJ, Eccles JD, Rouhani SJ, Peske JD, et al. Structural and Functional Features of Central Nervous System Lymphatic Vessels. *Nature* (2015) 523(7560):337–41. doi: 10.1038/nature14432
5. Ransohoff RM, Kivisäkk P, Kidd G. Three or More Routes for Leukocyte Migration Into the Central Nervous System. *Nat Rev Immunol* (2003) 3 (7):569–81. doi: 10.1038/nri1130
6. Mrdjen D, Pavlovic A, Hartmann FJ, Schreiner B, Utz SG, Leung BP, et al. High-Dimensional Single-Cell Mapping of Central Nervous System Immune Cells Reveals Distinct Myeloid Subsets in Health, Aging, and Disease. *Immunity* (2018) 48(2):380–95.e6. doi: 10.1016/j.immuni.2018.01.011
7. Van Hove H, Martens L, Scheyltjens I, De Vlaminck K, Pombo Antunes AR, De Prijck S, et al. A Single-Cell Atlas of Mouse Brain Macrophages Reveals Unique Transcriptional Identities Shaped by Ontogeny and Tissue Environment. *Nat Neurosci* (2019) 22(6):1021–35. doi: 10.1038/s41593-019-0393-4
8. Radjavi A, Smirnov I, Derecki N, Kipnis J. Dynamics of the Meningeal Cd4 (+) T-Cell Repertoire Are Defined by the Cervical Lymph Nodes and Facilitate Cognitive Task Performance in Mice. *Mol Psychiatry* (2014) 19 (5):531–3. doi: 10.1038/mp.2013.79
9. Kida S, Pantazis A, Weller RO. Csf Drains Directly From the Subarachnoid Space Into Nasal Lymphatics in the Rat. Anatomy, Histology and Immunological Significance. *Neuropathol Appl Neurobiol* (1993) 19 (6):480–8. doi: 10.1111/j.1365-2990.1993.tb00476.x
10. Widner H, Möller G, Johansson BB. Immune Response in Deep Cervical Lymph Nodes and Spleen in the Mouse After Antigen Deposition in Different Intracerebral Sites. *Scandinavian J Immunol* (1988) 28(5):563–71. doi: 10.1111/j.1365-3083.1988.tb01488.x
11. Fischer HG, Reichmann G. Brain Dendritic Cells and Macrophages/ Microglia in Central Nervous System Inflammation. *J Immunol (Baltimore Md: 1950)* (2001) 166(4):2717–26. doi: 10.4049/jimmunol.166.4.2717
12. Yang I, Tihan T, Han SJ, Wrensch MR, Wiencke J, Sughrue ME, et al. Cd8+ T-Cell Infiltrate in Newly Diagnosed Glioblastoma Is Associated With Long-Term Survival. *J Clin neuroscience: Off J Neurosurg Soc Australasi* (2010) 17 (11):1381–5. doi: 10.1016/j.jocn.2010.03.031
13. Berghoff AS, Kiesel B, Widhalm G, Rajky O, Ricken G, Wöhrer A, et al. Programmed Death Ligand 1 Expression and Tumor-Infiltrating Lymphocytes in Glioblastoma. *Neuro-oncology* (2015) 17(8):1064–75. doi: 10.1093/neuonc/nou307
14. Quail DF, Joyce JA. The Microenvironmental Landscape of Brain Tumors. *Cancer Cell* (2017) 31(3):326–41. doi: 10.1016/j.ccell.2017.02.009
15. Rossi ML, Hughes JT, Esiri MM, Coakham HB, Brownell DB. Immunohistological Study of Mononuclear Cell Infiltrate in Malignant Gliomas. *Acta neuropatholo* (1987) 74(3):269–77. doi: 10.1007/ bf00688191
16. Cloughesy TF, Mochizuki AY, Orpilla JR, Hugo W, Lee AH, Davidson TB, et al. Neoadjuvant Anti-Pd-1 Immunotherapy Promotes a Survival Benefit With Intratumoral and Systemic Immune Responses in Recurrent Glioblastoma. *Nat Med* (2019) 25(3):477–86. doi: 10.1038/s41591-018-0337-7

17. Mohan AA, Tomaszewski WH, Haskell-Mendoza AP, Hotchkiss KM, Singh K, Reedy JL, et al. Targeting Immunometabolism in Glioblastoma. *Front Oncol* (2021) 11:696402. doi: 10.3389/fonc.2021.696402
18. Buck MD, Sowell RT, Kaech SM, Pearce EL. Metabolic Instruction of Immunity. *Cell* (2017) 169(4):570–86. doi: 10.1016/j.cell.2017.04.004
19. Kesarwani P, Kant S, Prabhu A, Chinnaiyan P. The Interplay Between Metabolic Remodeling and Immune Regulation in Glioblastoma. *Neuro-oncology* (2017) 19(10):1308–15. doi: 10.1093/neuonc/nox079
20. Biswas SK. Metabolic Reprogramming of Immune Cells in Cancer Progression. *Immunity* (2015) 43(3):435–49. doi: 10.1016/j.immuni.2015.09.001
21. Delage B, Fennell DA, Nicholson L, McNeish I, Lemoine NR, Crook T, et al. Arginine Deprivation and Argininosuccinate Synthetase Expression in the Treatment of Cancer. *Int J Cancer* (2010) 126(12):2762–72. doi: 10.1002/ijc.25202
22. Rodriguez PC, Quiceno DG, Ochoa AC. L-Arginine Availability Regulates T-Lymphocyte Cell-Cycle Progression. *Blood* (2007) 109(4):1568–73. doi: 10.1182/blood-2006-06-031856
23. Zea AH, Rodriguez PC, Culotta KS, Hernandez CP, DeSalvo J, Ochoa JB, et al. L-Arginine Modulates Cd3zeta Expression and T Cell Function in Activated Human T Lymphocytes. *Cell Immunol* (2004) 232(1-2):21–31. doi: 10.1016/j.cellimm.2005.01.004
24. Tate DJJr., Vonderhaar DJ, Caldas YA, Metoyer T, JRt P, DH A, et al. Effect of Arginase Ii on L-Arginine Depletion and Cell Growth in Murine Cell Lines of Renal Cell Carcinoma. *J Hematol Oncol* (2008) 1:14. doi: 10.1186/1756-8722-1-14
25. Mussai F, Wheat R, Sarrou E, Booth S, Stavrou V, Fultang L, et al. Targeting the Arginine Metabolic Brake Enhances Immunotherapy for Leukaemia. *Int J Cancer* (2019) 145(8):2201–8. doi: 10.1002/ijc.32028
26. Hanahan D, Weinberg RA. Hallmarks of Cancer: The Next Generation. *Cell* (2011) 144(5):646–74. doi: 10.1016/j.cell.2011.02.013
27. Al-Koussa H, El Mais N, Maalouf H, Abi-Habib R, El-Sibai M. Arginine Deprivation: A Potential Therapeutic for Cancer Cell Metastasis? A Review. *Cancer Cell Int* (2020) 20:150. doi: 10.1186/s12935-020-01232-9
28. Okada S, Hongo T, Yamada S, Watanabe C, Fujii Y, Ohzeki T, et al. *In Vitro* Efficacy of L-Asparaginase in Childhood Acute Myeloid Leukaemia. *Br J haematol* (2003) 123(5):802–9. doi: 10.1046/j.1365-2141.2003.04703.x
29. Kenny GE, Pollock ME. Mammalian Cell Cultures Contaminated With Pleuropneumonia-Like Organisms. I. Effect of Pleuropneumonia-Like Organisms on Growth of Established Cell Strains. *J Infect Dis* (1963) 112:7–16. doi: 10.1093/infdis/112.1.7
30. Kraemer PM, Defendi V, Hayflick L, Manson LA. Mycoplasma (Pplo) Strains With Lytic Activity for Murine Lymphoma Cells in Vitro. *Proc Soc Exp Biol Med Soc Exp Biol Med (New York NY)* (1963) 112:381–7. doi: 10.3181/00379727-112-28052
31. Schimke RT, Berlin CM, Sweeney EW, Carroll WR. The Generation of Energy by the Arginine Dihydrolase Pathway in Mycoplasma Hominis 07. *J Biol Chem* (1966) 241(10):2228–36. doi: 10.1016/s0021-9258(18)96610-x
32. Mörén L, Perryman R, Crook T, Langer JK, Oneill K, Syed N, et al. Metabolomic Profiling Identifies Distinct Phenotypes for Ass1 Positive and Negative Gbm. *BMC Cancer* (2018) 18(1):167. doi: 10.1186/s12885-018-4040-3
33. Wu G, Bazer FW, Davis TA, Kim SW, Li P, Marc Rhoads J, et al. Arginine Metabolism and Nutrition in Growth, Health and Disease. *Amino Acids* (2009) 37(1):153–68. doi: 10.1007/s00726-008-0210-y
34. Jahani M, Noroznezhad F, Mansouri K. Arginine: Challenges and Opportunities of This Two-Faced Molecule in Cancer Therapy. *Biomedicine pharmacother = Biomedecine pharmacother* (2018) 102:594–601. doi: 10.1016/j.biopha.2018.02.109
35. Yu YM, Ryan CM, Castillo L, Lu XM, Beaumier L, Tompkins RG, et al. Arginine and Ornithine Kinetics in Severely Burned Patients: Increased Rate of Arginine Disposal. *Am J Physiol Endocrinol Metab* (2001) 280(3):E509–17. doi: 10.1152/ajpendo.2001.280.3.E509
36. Wilmore D. Enteral and Parenteral Arginine Supplementation to Improve Medical Outcomes in Hospitalized Patients. *J Nutr* (2004) 134(10 Suppl):2863S–7S. doi: 10.1093/jn/134.10.2863S
37. Bode-Böger SM, Böger RH, Galland A, Tsikas D, Frölich JC. L-Arginine-Induced Vasodilation in Healthy Humans: Pharmacokinetic-Pharmacodynamic Relationship. *Br J Clin Pharmacol* (1998) 46(5):489–97. doi: 10.1046/j.1365-2125.1998.00803.x
38. Castillo L, Chapman TE, Sanchez M, Yu YM, Burke JF, Ajami AM, et al. Plasma Arginine and Citrulline Kinetics in Adults Given Adequate and Arginine-Free Diets. *Proc Natl Acad Sci United States America* (1993) 90 (16):7749–53. doi: 10.1073/pnas.90.16.7749
39. Featherston WR, Rogers QR, Freedland RA. Relative Importance of Kidney and Liver in Synthesis of Arginine by the Rat. *Am J Physiol* (1973) 224 (1):127–9. doi: 10.1152/ajplegacy.1973.224.1.127
40. Dhanakoti SN, Brosnan JT, Herzberg GR, Brosnan ME. Renal Arginine Synthesis: Studies in Vitro and in Vivo. *Am J Physiol* (1990) 259(3 Pt 1): E437–42. doi: 10.1152/ajpendo.1990.259.3.E437
41. Cheung CW, Cohen NS, Raijman L. Channeling of Urea Cycle Intermediates in Situ in Permeabilized Hepatocytes. *J Biol Chem* (1989) 264(7):4038–44. doi: 10.1042/bj3120717
42. Watford M. The Urea Cycle: A Two-Compartment System. *Essays Biochem* (1991) 26:49–58. doi: 10.1002/hep.510250429
43. Scrimshaw NS, Kitamura T, Ajami AM, Fukagawa NK Arginine Metabolism: Enzymology, Nutrition, and Clinical Significance. Proceedings of a Symposium Dedicated to the Memory of Vernon R. Young. April 5-6, 2004. Bermuda. *J Nutr* (2004) 134(10 Suppl):2741s–897s. doi: 10.1093/ajcn/83.2.496S
44. Wu GY, Brosnan JT. Macrophages Can Convert Citrulline Into Arginine. *Biochem J* (1992) 281(Pt 1):45–8. doi: 10.1042/bj2810045
45. Nussler AK, Billiar TR, Liu ZZ, Morris SMJr. Coinduction of Nitric Oxide Synthase and Argininosuccinate Synthetase in a Murine Macrophage Cell Line. *Implications Regul Nitric Oxide Production. J Biol Chem* (1994) 269 (2):1257–61. doi: 10.1016/s0014-4835(02)00274-9
46. Sørensen MD, Dahlrot RH, Boldt HB, Hansen S, Kristensen BW. Tumour-Associated Microglia/Macrophages Predict Poor Prognosis in High-Grade Gliomas and Correlate With an Aggressive Tumour Subtype. *Neuropathol Appl Neurobiol* (2018) 44(2):185–206. doi: 10.1111/nan.12428
47. Nieves CJr., Langkamp-Henken B. Arginine and Immunity: A Unique Perspective. *Biomedicine pharmacother = Biomedecine pharmacother* (2002) 56(10):471–82. doi: 10.1016/s0753-3322(02)00291-3
48. Peranzoni E, Marigo I, Dolcetti L, Ugel S, Sonda N, Taschin E, et al. Role of Arginine Metabolism in Immunity and Immunopathology. *Immunobiology* (2007) 212(9-10):795–812. doi: 10.1016/j.imbio.2007.09.008
49. Barile MF, Leventhal BG. Possible Mechanism for Mycoplasma Inhibition of Lymphocyte Transformation Induced by Phytohaemagglutinin. *Nature* (1968) 219(5155):750–2. doi: 10.1038/219751a0
50. Tong BC, Barbul A. Cellular and Physiological Effects of Arginine. *Mini Rev medicinal Chem* (2004) 4(8):823–32. doi: 10.2174/1389557043403305
51. Mandal A. Do Malnutrition and Nutritional Supplementation Have an Effect on the Wound Healing Process? *J Wound Care* (2006) 15(6):254–7. doi: 10.12968/jowc.2006.15.6.26923
52. Barbul A, Lazarou SA, Efron DT, Wasserkrug HL, Efron G. Arginine Enhances Wound Healing and Lymphocyte Immune Responses in Humans. *Surgery* (1990) 108(2):331–6.
53. Bansal V, Syres KM, Makarenkova V, Brannon R, Matta B, Harbrecht BG, et al. Interactions Between Fatty Acids and Arginine Metabolism: Implications for the Design of Immune-Enhancing Diets. *JPEN J parenteral enteral Nutr* (2005) 29(1 Suppl):S75–80. doi: 10.1177/01486071050290s1s75
54. Ochoa JB, Makarenkova V, Bansal V. A Rational Use of Immune Enhancing Diets: When Should We Use Dietary Arginine Supplementation? *Nutr Clin practice: Off Publ Am Soc Parenteral Enteral Nutr* (2004) 19(3):216–25. doi: 10.1177/0115426504019003216
55. Daly JM, Reynolds J, Thom A, Kinsley L, Dietrick-Gallagher M, Shou J, et al. Immune and Metabolic Effects of Arginine in the Surgical Patient. *Ann Surg* (1988) 208(4):512–23. doi: 10.1097/00000658-198810000-00013
56. Braga M, Vignali A, Gianotti L, Cestari A, Profili M, Carlo VD. Immune and Nutritional Effects of Early Enteral Nutrition After Major Abdominal Operations. *Eur J Surg = Acta chirurgica* (1996) 162(2):105–12.
57. Bower RH, Cerra FB, Bershadsky B, Licari JJ, Hoyt DB, Jensen GL, et al. Early Enteral Administration of a Formula (Impact) Supplemented With Arginine, Nucleotides, and Fish Oil in Intensive Care Unit Patients: Results of a Multicenter, Prospective, Randomized, Clinical Trial. *Crit Care Med* (1995) 23(3):436–49. doi: 10.1097/00003246-199503000-00006

58. Grimble RF. Immunonutrition. *Curr Opin Gastroenterol* (2005) 21(2):216–22. doi: 10.1097/01.mog.0000153360.90653.82
59. Wu G, Morris SMJr. Arginine Metabolism: Nitric Oxide and Beyond. *Biochem J* (1998) 336(Pt 1):1–17. doi: 10.1042/bj3360001
60. Morris SM Jr. Arginine Metabolism: Boundaries of Our Knowledge. *J Nutrition* (2007) 137(6 Suppl 2):1602s–9s. doi: 10.1093/jn/137.6.1602S
61. Li H, Meininger CJ, Hawker JRJr., Haynes TE, Kepka-Lenhart D, Mistry SK, et al. Regulatory Role of Arginase I and Ii in Nitric Oxide, Polyamine, and Proline Syntheses in Endothelial Cells. *Am J Physiol Endocrinol Metab* (2001) 280(1):E75–82. doi: 10.1152/ajpendo.2001.280.1.E75
62. Orlando GF, Wolf G, Engelmann M. Role of Neuronal Nitric Oxide Synthase in the Regulation of the Neuroendocrine Stress Response in Rodents: Insights From Mutant Mice. *Amino Acids* (2008) 35(1):17–27. doi: 10.1007/s00726-007-0630-0
63. Sullivan LB, Gui DY, Vander Heiden MG. Altered Metabolite Levels in Cancer: Implications for Tumour Biology and Cancer Therapy. *Nat Rev Cancer* (2016) 16(11):680–93. doi: 10.1038/nrc.2016.85
64. Morris SMJr. Arginine: Master and Commander in Innate Immune Responses. *Sci Signaling* (2010) 3(135):pe27. doi: 10.1126/scisignal.3135pe27
65. Wu G, Meininger CJ. Regulation of Nitric Oxide Synthesis by Dietary Factors. *Annu Rev Nutr* (2002) 22:61–86. doi: 10.1146/annurev.nutr.22.110901.145329
66. Lind DS. Arginine and Cancer. *J Nutr* (2004) 134(10 Suppl):2837S–41S. doi: 10.1093/jn/134.10.2837S
67. Pervin S, Singh R, Hernandez E, Wu G, Chaudhuri G. Nitric Oxide in Physiologic Concentrations Targets the Translational Machinery to Increase the Proliferation of Human Breast Cancer Cells: Involvement of Mammalian Target of Rapamycin/Eif4e Pathway. *Cancer Res* (2007) 67(1):289–99. doi: 10.1158/0008-5472.Can-05-4623
68. Bonavida B, Garban H. Nitric Oxide-Mediated Sensitization of Resistant Tumor Cells to Apoptosis by Chemo-Immunotherapeutics. *Redox Biol* (2015) 6:486–94. doi: 10.1016/j.redox.2015.08.013
69. Choi BS, Martinez-Falero IC, Corset C, Munder M, Modolell M, Müller I, et al. Differential Impact of L-Arginine Deprivation on the Activation and Effector Functions of T Cells and Macrophages. *J leukocyte Biol* (2009) 85 (2):268–77. doi: 10.1189/jlb.0508310
70. Dillon BJ, Prieto VG, Curley SA, Ensor CM, Holtsberg FW, Bomalaski JS, et al. Incidence and Distribution of Argininosuccinate Synthetase Deficiency in Human Cancers: A Method for Identifying Cancers Sensitive to Arginine Deprivation. *Cancer* (2004) 100(4):826–33. doi: 10.1002/cncr.20057
71. Wheatley DN, Kilfeather R, Stitt A, Campbell E. Integrity and Stability of the Citrulline-Arginine Pathway in Normal and Tumour Cell Lines. *Cancer Lett* (2005) 227(2):141–52. doi: 10.1016/j.canlet.2005.01.004
72. Feun LG, Kuo MT, Savaraj N. Arginine Deprivation in Cancer Therapy. *Curr Opin Clin Nutr Metab Care* (2015) 18(1):78–82. doi: 10.1097/mco.0000000000000122
73. Hajji N, Garcia-Revilla J, Soto MS, Perryman R, Symington J, Quarles CC, et al. Arginine Deprivation Alters Microglial Polarity and Synergizes With Radiation to Eradicate Non-Arginine-Auxotrophic Glioblastoma Tumors. *J Clin Invest* (2022) 132(6):e142137. doi: 10.1172/jci142137
74. Silberman A, Goldman O, Boukobza Assayag O, Jacob A, Rabinovich S, Adler L, et al. Acid-Induced Downregulation of Ass1 Contributes to the Maintenance of Intracellular Ph in Cancer. *Cancer Res* (2019) 79(3):518–33. doi: 10.1158/0008-5472.Can-18-1062
75. Miyamoto T, Lo PHY, Saichi N, Ueda K, Hirata M, Tanikawa C, et al. Argininosuccinate Synthase 1 Is an Intrinsic Akt Repressor Transactivated by P53. *Sci Adv* (2017) 3(5):e1603204. doi: 10.1126/sciadv.1603204
76. Long Y, Tsai WB, Wangpaichitr M, Tsukamoto T, Savaraj N, Feun LG, et al. Arginine Deiminase Resistance in Melanoma Cells Is Associated With Metabolic Reprogramming, Glucose Dependence, and Glutamine Addiction. *Mol Cancer Ther* (2013) 12(11):2581–90. doi: 10.1158/1535-7163.Mct-13-0302
77. Brashears CB, Barlin M, Ehrhardt WR, Rathore R, Schultze M, Tzeng SC, et al. Systems Level Profiling of Arginine Starvation Reveals Myc and Erk Adaptive Metabolic Reprogramming. *Cell Death Dis* (2020) 11(8):662. doi: 10.1038/s41419-020-02899-8
78. Kobayashi E, Masuda M, Nakayama R, Ichikawa H, Satow R, Shitashige M, et al. Reduced Argininosuccinate Synthetase Is a Predictive Biomarker for the Development of Pulmonary Metastasis in Patients With Osteosarcoma. *Mol Cancer Ther* (2010) 9(3):535–44. doi: 10.1158/1535-7163.Mct-09-0774
79. Allen MD, Luong P, Hudson C, Leyton J, Delage B, Ghazaly E, et al. Prognostic and Therapeutic Impact of Argininosuccinate Synthetase 1 Control in Bladder Cancer as Monitored Longitudinally by Pet Imaging. *Cancer Res* (2014) 74(3):896–907. doi: 10.1158/0008-5472.Can-13-1702
80. Huang HY, Wu WR, Wang YH, Wang JW, Fang FM, Tsai JW, et al. Ass1 as a Novel Tumor Suppressor Gene in Myxofibrosarcomas: Aberrant Loss *Via* Epigenetic DNA Methylation Confers Aggressive Phenotypes, Negative Prognostic Impact, and Therapeutic Relevance. *Clin Cancer research: an Off J Am Assoc Cancer Res* (2013) 19(11):2861–72. doi: 10.1158/1078-0432.Ccr-12-2641
81. Tsai CY, Chi HC, Chi LM, Yang HY, Tsai MM, Lee KF, et al. Argininosuccinate Synthetase 1 Contributes to Gastric Cancer Invasion and Progression by Modulating Autophagy. *FASEB journal: Off Publ Fed Am Societies Exp Biol* (2018) 32(5):2601–14. doi: 10.1096/fj.201700094R
82. Shan YS, Hsu HP, Lai MD, Yen MC, Luo YP, Chen YL. Increased Expression of Argininosuccinate Synthetase Protein Predicts Poor Prognosis in Human Gastric Cancer. *Oncol Rep* (2015) 33(1):49–57. doi: 10.3892/or.2014.3556
83. Huang CC, Tsai ST, Kuo CC, Chang JS, Jin YT, Chang JY, et al. Arginine Deprivation as a New Treatment Strategy for Head and Neck Cancer. *Oral Oncol* (2012) 48(12):1227–35. doi: 10.1016/j.oraloncology.2012.06.004
84. Syed N, Langer J, Janczar K, Singh P, Lo Nigro C, Lattanzio L, et al. Epigenetic Status of Argininosuccinate Synthetase and Argininosuccinate Lyase Modulates Autophagy and Cell Death in Glioblastoma. *Cell Death Dis* (2013) 4(1):e458. doi: 10.1038/cddis.2012.197
85. Li F, Simon MC. Cancer Cells Don't Live Alone: Metabolic Communication Within Tumor Microenvironments. *Dev Cell* (2020) 54(2):183–95. doi: 10.1016/j.devcel.2020.06.018
86. Rabinovich S, Adler L, Yizhak K, Sarver A, Silberman A, Agron S, et al. Diversion of Aspartate in Ass1-Deficient Tumours Fosters *De Novo* Pyrimidine Synthesis. *Nature* (2015) 527(7578):379–83. doi: 10.1038/nature15529
87. Keshet R, Szlosarek P, Carracedo A, Erez A. Rewiring Urea Cycle Metabolism In Cancer to Support Anabolism. *Nat Rev Cancer* (2018) 18 (10):634–45. doi: 10.1038/s41568-018-0054-z
88. Zea AH, Rodriguez PC, Atkins MB, Hernandez C, Signoretti S, Zabaleta J, et al. Arginase-Producing Myeloid Suppressor Cells in Renal Cell Carcinoma Patients: A Mechanism of Tumor Evasion. *Cancer Res* (2005) 65(8):3044–8. doi: 10.1158/0008-5472.Can-04-4505
89. Bronte V, Kasic T, Gri G, Gallana K, Borsellino G, Marigo I, et al. Boosting Antitumor Responses of T Lymphocytes Infiltrating Human Prostate Cancers. *J Exp Med* (2005) 201(8):1257–68. doi: 10.1084/jem.20042028
90. Mussai F, Egan S, Hunter S, Webber H, Fisher J, Wheat R, et al. Neuroblastoma Arginase Activity Creates an Immunosuppressive Microenvironment That Impairs Autologous and Engineered Immunity. *Cancer Res* (2015) 75(15):3043–53. doi: 10.1158/0008-5472.Can-14-3443
91. Mussai F, De Santo C, Abu-Dayyeh I, Booth S, Quek L, McEwen-Smith RM, et al. Acute Myeloid Leukemia Creates an Arginase-Dependent Immunosuppressive Microenvironment. *Blood* (2013) 122(5):749–58. doi: 10.1182/blood-2013-01-480129
92. Hasim A, Aili A, Maimaiti A, Mamtimin B, Abudula A, Upur H. Plasma-Free Amino Acid Profiling of Cervical Cancer and Cervical Intraepithelial Neoplasia Patients and Its Application for Early Detection. *Mol Biol Rep* (2013) 40(10):5853–9. doi: 10.1007/s11033-013-2691-3
93. Rodriguez PC, Ernstoff MS, Hernandez C, Atkins M, Zabaleta J, Sierra R, et al. Arginase I-Producing Myeloid-Derived Suppressor Cells in Renal Cell Carcinoma Are a Subpopulation of Activated Granulocytes. *Cancer Res* (2009) 69(4):1553–60. doi: 10.1158/0008-5472.Can-08-1921
94. Bron L, Jandus C, Andrejevic-Blant S, Speiser DE, Monnier P, Romero P, et al. Prognostic Value of Arginase-Ii Expression and Regulatory T-Cell Infiltration in Head and Neck Squamous Cell Carcinoma. *Int J Cancer* (2013) 132(3):E85–93. doi: 10.1002/ijc.27728
95. Geiger R, Rieckmann JC, Wolf T, Basso C, Feng Y, Fuhrer T, et al. L-Arginine Modulates T Cell Metabolism and Enhances Survival and Anti-

Tumor Activity. *Cell* (2016) 167(3):829–42.e13. doi: 10.1016/j.cell.2016.09.031
96. Mondanelli G, Ugel S, Grohmann U, Bronte V. The Immune Regulation in Cancer by the Amino Acid Metabolizing Enzymes Arg and Ido. *Curr Opin Pharmacol* (2017) 35:30–9. doi: 10.1016/j.coph.2017.05.002
97. Li X, Wenes M, Romero P, Huang SC, Fendt SM, Ho PC. Navigating Metabolic Pathways to Enhance Antitumour Immunity and Immunotherapy. *Nat Rev Clin Oncol* (2019) 16(7):425–41. doi: 10.1038/s41571-019-0203-7
98. Gabrilovich DI, Bronte V, Chen SH, Colombo MP, Ochoa A, Ostrand-Rosenberg S, et al. The Terminology Issue for Myeloid-Derived Suppressor Cells. *Cancer Res* (2007) 67(1):425. doi: 10.1158/0008-5472.Can-06-3037
99. Gielen PR, Schulte BM, Kers-Rebel ED, Verrijp K, Bossman SA, Ter Laan M, et al. Elevated Levels of Polymorphonuclear Myeloid-Derived Suppressor Cells in Patients With Glioblastoma Highly Express S100a8/9 and Arginase and Suppress T Cell Function. *Neuro-oncology* (2016) 18(9):1253–64. doi: 10.1093/neuonc/now034
100. Sinha P, Clements VK, Ostrand-Rosenberg S. Interleukin-13-Regulated M2 Macrophages in Combination With Myeloid Suppressor Cells Block Immune Surveillance Against Metastasis. *Cancer Res* (2005) 65(24):11743–51. doi: 10.1158/0008-5472.Can-05-0045
101. Montero AJ, Diaz-Montero CM, Kyriakopoulos CE, Bronte V, Mandruzzato S. Myeloid-Derived Suppressor Cells in Cancer Patients: A Clinical Perspective. *J immunother (Hagerstown Md: 1997)* (2012) 35(2):107–15. doi: 10.1097/CJI.0b013e318242169f
102. Smith C, Chang MY, Parker KH, Beury DW, DuHadaway JB, Flick HE, et al. Ido Is a Nodal Pathogenic Driver of Lung Cancer and Metastasis Development. *Cancer Discov* (2012) 2(8):722–35. doi: 10.1158/2159-8290.Cd-12-0014
103. Szlosarek PW, Grimshaw MJ, Wilbanks GD, Hagemann T, Wilson JL, Burke F, et al. Aberrant Regulation of Argininosuccinate Synthetase by Tnf-Alpha in Human Epithelial Ovarian Cancer. *Int J Cancer* (2007) 121(1):6–11. doi: 10.1002/ijc.22666
104. Crump NT, Hadjinicolaou AV, Xia M, Walsby-Tickle J, Gileadi U, Chen JL, et al. Chromatin Accessibility Governs the Differential Response of Cancer and T Cells to Arginine Starvation. *Cell Rep* (2021) 35(6):109101. doi: 10.1016/j.celrep.2021.109101
105. Werner A, Koschke M, Leuchtner N, Luckner-Minden C, Habermeier A, Rupp J, et al. Reconstitution of T Cell Proliferation Under Arginine Limitation: Activated Human T Cells Take Up Citrulline *Via* L-Type Amino Acid Transporter 1 and Use It to Regenerate Arginine After Induction of Argininosuccinate Synthase Expression. *Front Immunol* (2017) 8:864. doi: 10.3389/fimmu.2017.00864
106. Sugimura K, Kimura T, Arakawa H, Ohno T, Wada Y, Kimura Y, et al. Elevated Argininosuccinate Synthetase Activity in Adult T Leukemia Cell Lines. *Leukemia Res* (1990) 14(10):931–4. doi: 10.1016/0145-2126(90)90184-b
107. Fultang L, Booth S, Yogev O, Martins da Costa B, Tubb V, Panetti S, et al. Metabolic Engineering Against the Arginine Microenvironment Enhances Car-T Cell Proliferation and Therapeutic Activity. *Blood* (2020) 136(10):1155–60. doi: 10.1182/blood.2019004500
108. Weissman AM, Ross P, Luong ET, Garcia-Morales P, Jelachich ML, Biddison WE, et al. Tyrosine Phosphorylation of the Human T Cell Antigen Receptor Zeta-Chain: Activation *Via* Cd3 But Not Cd2. *J Immunol (Baltimore Md: 1950)* (1988) 141(10):3532–6.
109. Rodriguez PC, Zea AH, Culotta KS, Zabaleta J, Ochoa JB, Ochoa AC. Regulation of T Cell Receptor Cd3zeta Chain Expression by L-Arginine. *J Biol Chem* (2002) 277(24):21123–9. doi: 10.1074/jbc.M110675200
110. Minami Y, Weissman AM, Samelson LE, Klausner RD. Building a Multichain Receptor: Synthesis, Degradation, and Assembly of the T-Cell Antigen Receptor. *Proc Natl Acad Sci United States America* (1987) 84(9):2688–92. doi: 10.1073/pnas.84.9.2688
111. Munder M, Schneider H, Luckner C, Giese T, Langhans CD, Fuentes JM, et al. Suppression of T-Cell Functions by Human Granulocyte Arginase. *Blood* (2006) 108(5):1627–34. doi: 10.1182/blood-2006-11-010389
112. Feldmeyer N, Wabnitz G, Leicht S, Luckner-Minden C, Schiller M, Franz T, et al. Arginine Deficiency Leads to Impaired Cofilin Dephosphorylation in Activated Human T Lymphocytes. *Int Immunol* (2012) 24(5):303–13. doi: 10.1093/intimm/dxs004
113. Holcik M, Sonenberg N. Translational Control in Stress and Apoptosis. *Nat Rev Mol Cell Biol* (2005) 6(4):318–27. doi: 10.1038/nrm1618
114. Fletcher M, Ramirez ME, Sierra RA, Raber P, Thevenot P, Al-Khami AA, et al. L-Arginine Depletion Blunts Antitumor T-Cell Responses by Inducing Myeloid-Derived Suppressor Cells. *Cancer Res* (2015) 75(2):275–83. doi: 10.1158/0008-5472.Can-14-1491
115. Ochoa JB, Strange J, Kearney P, Gellin G, Endean E, Fitzpatrick E. Effects of L-Arginine on the Proliferation of T Lymphocyte Subpopulations. *JPEN J parenteral enteral Nutr* (2001) 25(1):23–9. doi: 10.1177/014860710102500123
116. Kato JY. Control of G1 Progression by D-Type Cyclins: Key Event for Cell Proliferation. *Leukemia* (1997) 11 Suppl 3:347–51.
117. Rodriguez PC, Hernandez CP, Morrow K, Sierra R, Zabaleta J, Wyczechowska DD, et al. L-Arginine Deprivation Regulates Cyclin D3 Mrna Stability in Human T Cells by Controlling Hur Expression. *J Immunol (Baltimore Md: 1950)* (2010) 185(9):5198–204. doi: 10.4049/jimmunol.1001224
118. Topal G, Brunet A, Walch L, Boucher JL, David-Dufilho M. Mitochondrial Arginase Ii Modulates Nitric-Oxide Synthesis Through Nonfreely Exchangeable L-Arginine Pools in Human Endothelial Cells. *J Pharmacol Exp Ther* (2006) 318(3):1368–74. doi: 10.1124/jpet.106.103747
119. Martí i Líndez AA, Dunand-Sauthier I, Conti M, Gobet F, Núñez N, Hannich JT, et al. Mitochondrial Arginase-2 Is a Cell–Autonomous Regulator of Cd8+ T Cell Function and Antitumor Efficacy. *JCI Insight* (2019) 4(24):e132975. doi: 10.1172/jci.insight.132975
120. Marullo R, Castro M, Yomtoubian S, Calvo-Vidal MN, Revuelta MV, Krumsiek J, et al. The Metabolic Adaptation Evoked by Arginine Enhances the Effect of Radiation in Brain Metastases. *Sci Adv* (2021) 7(45):eabg1964. doi: 10.1126/sciadv.abg1964
121. Canale FP, Basso C, Antonini G, Perotti M, Li N, Sokolovska A, et al. Metabolic Modulation of Tumours With Engineered Bacteria for Immunotherapy. *Nature* (2021) 598(7882):662–6. doi: 10.1038/s41586-021-04003-2
122. Shukla J, Chatterjee S, Thakur VS, Premachandran S, Checker R, Poduval TB. L-Arginine Reverses Radiation-Induced Immune Dysfunction: The Need for Optimum Treatment Window. *Radiat Res* (2009) 171(2):180–7. doi: 10.1667/rr1241.1
123. Harmon C, O'Farrelly C, Robinson MW. The Immune Consequences of Lactate in the Tumor Microenvironment. *Adv Exp Med Biol* (2020) 1259:113–24. doi: 10.1007/978-3-030-43093-1_7
124. Takaku H, Matsumoto M, Misawa S, Miyazaki K. Anti-Tumor Activity of Arginine Deiminase From Mycoplasma Argini and Its Growth-Inhibitory Mechanism. *Japanese J Cancer research: Gann* (1995) 86(9):840–6. doi: 10.1111/j.1349-7006.1995.tb03094.x
125. Holtsberg FW, Ensor CM, Steiner MR, Bomalaski JS, Clark MA. Poly (Ethylene Glycol) (Peg) Conjugated Arginine Deiminase: Effects of Peg Formulations on Its Pharmacological Properties. *J Controlled release: Off J Controlled Release Soc* (2002) 80(1-3):259–71. doi: 10.1016/s0168-3659(02)00042-1
126. Nasreddine G, El-Sibai M, Abi-Habib RJ. Cytotoxicity of [Huargi (Co)-Peg5000]-Induced Arginine Deprivation to Ovarian Cancer Cells Is Autophagy Dependent. *Investigational New Drugs* (2020) 38(1):10–9. doi: 10.1007/s10637-019-00756-w
127. Harris JM, Chess RB. Effect of Pegylation on Pharmaceuticals. *Nat Rev Drug Discov* (2003) 2(3):214–21. doi: 10.1038/nrd1033
128. Glazer ES, Stone EM, Zhu C, Massey KL, Hamir AN, Curley SA. Bioengineered Human Arginase I With Enhanced Activity and Stability Controls Hepatocellular and Pancreatic Carcinoma Xenografts. *Trans Oncol* (2011) 4(3):138–46. doi: 10.1593/tlo.10265
129. Khoury O, Ghazale N, Stone E, El-Sibai M, Frankel AE, Abi-Habib RJ. Human Recombinant Arginase I (Co)-Peg5000 [Huargi (Co)-Peg5000]-Induced Arginine Depletion Is Selectively Cytotoxic to Human Glioblastoma Cells. *J neuro-oncol* (2015) 122(1):75–85. doi: 10.1007/s11060-014-1698-5
130. Przystal JM, Hajji N, Khozoie C, Renziehausen A, Zeng Q, Abaitua F, et al. Efficacy of Arginine Depletion by Adi-Peg20 in an Intracranial Model of Gbm. *Cell Death Dis* (2018) 9(12):1192. doi: 10.1038/s41419-018-1195-4
131. Wangpaichitr M, Wu C, Bigford G, Theodoropoulos G, You M, Li YY, et al. Combination of Arginine Deprivation With Trail Treatment as a Targeted-Therapy for Mesothelioma. *Anticancer Res* (2014) 34(12):6991–9.

132. Phillips MM, Sheaff MT, Szlosarek PW. Targeting Arginine-Dependent Cancers With Arginine-Degrading Enzymes: Opportunities and Challenges. *Cancer Res Treat* (2013) 45(4):251–62. doi: 10.4143/crt.2013.45.4.251
133. Bean GR, Kremer JC, Prudner BC, Schenone AD, Yao JC, Schultze MB, et al. A Metabolic Synthetic Lethal Strategy With Arginine Deprivation and Chloroquine Leads to Cell Death in Ass1-Deficient Sarcomas. *Cell Death Dis* (2016) 7(10):e2406. doi: 10.1038/cddis.2016.232
134. Karatsai O, Shliaha P, Jensen ON, Stasyk O, Rędowicz MJ. Combinatory Treatment of Canavanine and Arginine Deprivation Efficiently Targets Human Glioblastoma Cells *Via* Pleiotropic Mechanisms. *Cells* (2020) 9 (10):2217. doi: 10.3390/cells9102217
135. Fiedler T, Strauss M, Hering S, Redanz U, William D, Rosche Y, et al. Arginine Deprivation by Arginine Deiminase of Streptococcus Pyogenes Controls Primary Glioblastoma Growth in Vitro and in Vivo. *Cancer Biol Ther* (2015) 16(7):1047–55. doi: 10.1080/15384047.2015.1026478
136. Hall PE, Lewis R, Syed N, Shaffer R, Evanson J, Ellis S, et al. A Phase I Study of Pegylated Arginine Deiminase (Pegargiminase), Cisplatin, and Pemetrexed in Argininosuccinate Synthetase 1-Deficient Recurrent High-Grade Glioma. *Clin Cancer research: an Off J Am Assoc Cancer Res* (2019) 25 (9):2708–16. doi: 10.1158/1078-0432.Ccr-18-3729
137. Hinrichs CN, Ingargiola M, Käubler T, Löck S, Temme A, Köhn-Luque A, et al. Arginine Deprivation Therapy: Putative Strategy to Eradicate Glioblastoma Cells by Radiosensitization. *Mol Cancer Ther* (2018) 17 (2):393–406. doi: 10.1158/1535-7163.Mct-16-0807
138. Kim RH, Bold RJ, Kung HJ. Adi, Autophagy and Apoptosis: Metabolic Stress as a Therapeutic Option for Prostate Cancer. *Autophagy* (2009) 5(4):567–8. doi: 10.4161/auto.5.4.8252
139. Pavlyk I, Rzhepetskyy Y, Jagielski AK, Drozak J, Wasik A, Pereverzieva G, et al. Arginine Deprivation Affects Glioblastoma Cell Adhesion, Invasiveness and Actin Cytoskeleton Organization by Impairment of B-Actin Arginylation. *Amino Acids* (2015) 47(1):199–212. doi: 10.1007/s00726-014-1857-1
140. Roe ND, Ren J. Nitric Oxide Synthase Uncoupling: A Therapeutic Target in Cardiovascular Diseases. *Vasc Pharmacol* (2012) 57(5-6):168–72. doi: 10.1016/j.vph.2012.02.004
141. Szabó C, Ischiropoulos H, Radi R. Peroxynitrite: Biochemistry, Pathophysiology and Development of Therapeutics. *Nat Rev Drug Discov* (2007) 6(8):662–80. doi: 10.1038/nrd2222
142. Villanueva-Meyer JE, Barajas RFJr., Mabray MC, Chen W, Shankaranarayanan A, Koon P, et al. Differentiation of Brain Tumor-Related Edema Based on 3d T1rho Imaging. *Eur J Radiol* (2017) 91:88–92. doi: 10.1016/j.ejrad.2017.03.022
143. Ahir BK, Engelhard HH, Lakka SS. Tumor Development and Angiogenesis in Adult Brain Tumor: Glioblastoma. *Mol Neurobiol* (2020) 57(5):2461–78. doi: 10.1007/s12035-020-01892-8
144. Burrows N, Cane G, Robson M, Gaude E, Howat WJ, Szlosarek PW, et al. Hypoxia-Induced Nitric Oxide Production and Tumour Perfusion Is Inhibited by Pegylated Arginine Deiminase (Adi-Peg20). *Sci Rep* (2016) 6:22950. doi: 10.1038/srep22950
145. Wang G, Wang JJ, Fu XL, Guang R, To ST. Advances in the Targeting of Hif-1α and Future Therapeutic Strategies for Glioblastoma Multiforme (Review). *Oncol Rep* (2017) 37(2):657–70. doi: 10.3892/or.2016.5309
146. Lo Dico A, Martelli C, Diceglie C, Lucignani G, Ottobrini L. Hypoxia-Inducible Factor-1α Activity as a Switch for Glioblastoma Responsiveness to Temozolomide. *Front Oncol* (2018) 8:249. doi: 10.3389/fonc.2018.00249
147. Brin E, Wu K, Lu HT, He Y, Dai Z, He W. Pegylated Arginine Deiminase Can Modulate Tumor Immune Microenvironment by Affecting Immune Checkpoint Expression, Decreasing Regulatory T Cell Accumulation and Inducing Tumor T Cell Infiltration. *Oncotarget* (2017) 8(35):58948–63. doi: 10.18632/oncotarget.19564
148. Tsai WB, Aiba I, Lee SY, Feun L, Savaraj N, Kuo MT. Resistance to Arginine Deiminase Treatment in Melanoma Cells Is Associated With Induced Argininosuccinate Synthetase Expression Involving C-Myc/Hif-1alpha/Sp4. *Mol Cancer Ther* (2009) 8(12):3223–33. doi: 10.1158/1535-7163.Mct-09-0794
149. Tsai WB, Aiba I, Long Y, Lin HK, Feun L, Savaraj N, et al. Activation of Ras/Pi3k/Erk Pathway Induces C-Myc Stabilization to Upregulate Argininosuccinate Synthetase, Leading to Arginine Deiminase Resistance in Melanoma Cells. *Cancer Res* (2012) 72(10):2622–33. doi: 10.1158/0008-5472.Can-11-3605
150. Fung MKL, Chan GC. Drug-Induced Amino Acid Deprivation as Strategy for Cancer Therapy. *J Hematol Oncol* (2017) 10(1):144. doi: 10.1186/s13045-017-0509-9
151. Saxton RA, Sabatini DM. Mtor Signaling in Growth, Metabolism, and Disease. *Cell* (2017) 168(6):960–76. doi: 10.1016/j.cell.2017.02.004
152. Shen W, Zhang X, Fu X, Fan J, Luan J, Cao Z, et al. A Novel and Promising Therapeutic Approach for Nsclc: Recombinant Human Arginase Alone or Combined With Autophagy Inhibitor. *Cell Death Dis* (2017) 8(3):e2720. doi: 10.1038/cddis.2017.137
153. Fultang L, Vardon A, De Santo C, Mussai F. Molecular Basis and Current Strategies of Therapeutic Arginine Depletion for Cancer. *Int J Cancer* (2016) 139(3):501–9. doi: 10.1002/ijc.30051
154. Szlosarek PW, Steele JP, Nolan L, Gilligan D, Taylor P, Spicer J, et al. Arginine Deprivation With Pegylated Arginine Deiminase in Patients With Argininosuccinate Synthetase 1-Deficient Malignant Pleural Mesothelioma: A Randomized Clinical Trial. *JAMA Oncol* (2017) 3(1):58–66. doi: 10.1001/jamaoncol.2016.3049
155. Ascierto PA, Scala S, Castello G, Daponte A, Simeone E, Ottaiano A, et al. Pegylated Arginine Deiminase Treatment of Patients With Metastatic Melanoma: Results From Phase I and Ii Studies. *J Clin oncology: Off J Am Soc Clin Oncol* (2005) 23(30):7660–8. doi: 10.1200/jco.2005.02.0933
156. Steggerda SM, Bennett MK, Chen J, Emberley E, Huang T, Janes JR, et al. Inhibition of Arginase by Cb-1158 Blocks Myeloid Cell-Mediated Immune Suppression in the Tumor Microenvironment. *J immunother Cancer* (2017) 5(1):101. doi: 10.1186/s40425-017-0308-4

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

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 08 August 2022
DOI 10.3389/fonc.2022.913803

OPEN ACCESS

EDITED BY
Yapeng Su,
Fred Hutchinson Cancer Research Center, United States

REVIEWED BY
Bishal Paudel,
University of Virginia, United States
Da Zhou,
Xiamen University, China

*CORRESPONDENCE
Mohit Kumar Jolly
mkjolly@iisc.ac.in

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Cancer Metabolism, a section of the journal Frontiers in Oncology

RECEIVED 06 April 2022
ACCEPTED 11 July 2022
PUBLISHED 08 August 2022

CITATION
Pillai M, Rajaram G, Thakur P, Agarwal N, Muralidharan S, Ray A, Barbhaya D, Somarelli JA and Jolly MK (2022) Mapping phenotypic heterogeneity in melanoma onto the epithelial-hybrid-mesenchymal axis. *Front. Oncol.* 12:913803. doi: 10.3389/fonc.2022.913803

Mapping phenotypic heterogeneity in melanoma onto the epithelial-hybrid-mesenchymal axis

Maalavika Pillai[1,2†], Gouri Rajaram[3†], Pradipti Thakur[3], Nilay Agarwal[1,2], Srinath Muralidharan[1], Ankita Ray[3], Dev Barbhaya[4], Jason A. Somarelli[5] and Mohit Kumar Jolly[1*]

[1]Centre for BioSystems Science and Engineering, Indian Institute of Science, Bangalore, India, [2]Undergraduate Programme, Indian Institute of Science, Bangalore, India, [3]Department of Biotechnology, Indian Institute of Technology, Kharagpur, India, [4]Department of Biological Sciences and Bioengineering, Indian Institute of Technology, Kanpur, India, [5]Department of Medicine, Duke University, Durham, NC, United States

Epithelial to mesenchymal transition (EMT) is a well-studied hallmark of epithelial-like cancers that is characterized by loss of epithelial markers and gain of mesenchymal markers. Melanoma, which is derived from melanocytes of the skin, also undergo phenotypic plasticity toward mesenchymal-like phenotypes under the influence of various micro-environmental cues. Our study connects EMT to the phenomenon of de-differentiation (i.e., transition from proliferative to more invasive phenotypes) observed in melanoma cells during drug treatment. By analyzing 78 publicly available transcriptomic melanoma datasets, we found that de-differentiation in melanoma is accompanied by upregulation of mesenchymal genes, but not necessarily a concomitant loss of an epithelial program, suggesting a more "one-dimensional" EMT that leads to a hybrid epithelial/mesenchymal phenotype. Samples lying in the hybrid epithelial/mesenchymal phenotype also correspond to the intermediate phenotypes in melanoma along the proliferative-invasive axis - neural crest and transitory ones. As melanoma cells progress along the invasive axis, the mesenchymal signature does not increase monotonically. Instead, we observe a peak in mesenchymal scores followed by a decline, as cells further de-differentiate. This biphasic response recapitulates the dynamics of melanocyte development, suggesting close interactions among genes controlling differentiation and mesenchymal programs in melanocytes. Similar trends were noted for metabolic changes often associated with EMT in carcinomas in which progression along mesenchymal axis correlates with the downregulation of oxidative phosphorylation, while largely maintaining glycolytic capacity. Overall, these results provide an explanation for how EMT and de-differentiation axes overlap with respect to their transcriptional and metabolic programs in melanoma.

KEYWORDS

phenotypic plasticity, EMT, melanoma, metabolic reprogramming, dedifferentiation, phenotypic heterogeneity

Introduction

Epithelial to mesenchymal transition (EMT) is a well-characterized phenomenon involved in multiple axes of cancer progression, such as metastasis and drug resistance. EMT is commonly associated with morphological changes, functional changes (increased migration, invasion, and immune invasion) (1–3) and molecular changes, including upregulation of EMT markers and transcription factors (TFs), such as *VIM, ZEB1, SNAI1* and *TWIST1*. While the phenomenon of EMT has largely been characterized for epithelial cancers (such as breast cancer and lung adenocarcinoma), similar molecular, functional and morphological changes have also been observed in non-epithelial cancers, such as sarcomas (4, 5), glioblastoma (6), myeloma (7), lymphoma (8, 9), leukemia (10, 11) and melanoma (12) in preclinical and clinical settings.

Treatment of melanoma tumors harboring BRAFV600E mutation often involves targeted therapy strategies that inhibit BRAF or MEK signaling. While these targeted agents provide clinical benefit to melanoma patients, resistance to these therapies is common. Therapy-resistant melanomas often undergo de-differentiation, which is characterized by loss of melanocytic markers such as *MLANA, TRPM1* and *TYR* and gain of invasive molecular markers such as *c-JUN, NGFR* and *ZEB1* (13–16). The de-differentiation trajectory of melanoma cells is characterized by a transition along the proliferation-invasion axis, from a melanocytic phenotype to an undifferentiated phenotype while passing through the intermediate transitory and neural crest stem cell-like (NCSC) phenotypes (Figure 1A). This trajectory is the reverse of the differentiation that occurs during melanocyte development, where undifferentiated tissue in the embryonic neural plate give rise to highly migratory and mesenchymal neural crest cells, some of which differentiate into melanocytes upon reaching the epidermis (17). Therapy resistant melanoma is also commonly associated with a mesenchymal-like phenotype with more invasive and aggressive features (13, 16, 18–20). These relationships between de-differentiation, invasion, and EMT pathways in response to therapy suggest EMT and de-differentiation programs in melanoma may be linked.

The similarity between EMT and de-differentiation programs extends beyond cell-intrinsic alterations and impacts cell-extrinsic changes as well. EMT often leads to varied extracellular matrix (ECM) stiffness and density (21–23) and altered cell-matrix and cell-cell interactions (24, 25). In melanoma, acquisition of de-differentiated and invasive phenotypes is often accompanied with changes in composition and physical properties of ECM, and modified cell-matrix interactions and cell morphology (26–28). Increased expression of matrix metalloproteases (MMPs), immune evasion (characterized by both signaling-mediated immune suppression (e.g. by TGF-ß release) and prevention of immune cell entry into tumors by dense collagen matrix/low α-SMA expression), increased inflammatory markers (such as TNF-α, NF-kB and AP-1) and cytoskeleton remodeling have been closely linked to the acquisition of an invasive phenotype and loss of melanocytic differentiation regulator MITF (29–34). All of these changes are reported with EMT progression as well in multiple epithelial cancers (35–37). Such extensive similarity between EMT and de-differentiation programs in cancer-microenvironment cross-talk and niche construction underscore the potential of common regulatory pathways involved in both EMT and de-differentiation.

Another common feature that links EMT in epithelial cancers to de-differentiation in melanoma is the presence of intermediate or hybrid phenotypes. Hybrid epithelial/mesenchymal (E/M) cells express molecular and functional characteristics of both epithelial (high proliferation and cell-cell adhesion, low invasion) and mesenchymal (low proliferation and cell-cell adhesion, high invasion) cells (38). On the other hand, melanoma intermediate phenotypes, which comprise transitory and neural crest-like stem cell-like (NCSC) phenotypes, exhibit combined features of proliferative and invasive phenotypes (39, 40) (Figure 1A). Gene regulatory networks for EMT and melanoma provide a mechanistic basis for explaining the existence of these hybrid/intermediate states (41, 42). An overlap in key regulators and stabilizers for hybrid E/M phenotypes and melanoma phenotypes (such as *ZEB1, NFATC2, CDH1, SNAI2, NRF2*) suggest common regulatory links (13, 43–49). For instance, *SNAI2*, a stabilizer of the hybrid E/M phenotype, is a key regulator of the NCSC phenotype and metastasis in melanoma, suggesting its involvement in regulating the intermediate phenotypes in melanoma as well (45, 49). However, certain regulators show opposite trends in melanoma and EMT. For instance, *ZEB2* is considered an inducer of EMT in epithelial cancers, but in the context of melanoma, it inhibits the mesenchymal phenotype (19, 50). Other molecules that show opposite effects include *KLF4* (51, 52) and *TFAP2A* (53, 54). Thus, understanding the mechanistic underpinning of how the de-differentiation and EMT programs are linked can help decipher reasons for the similarities and differences between these pathways across cancers.

In this study, we map the de-differentiation axis in melanoma (also called proliferative-invasive/P-I axis) to the EMT axis using previously defined scoring metrics (3, 55–57). We compare the extent to which a gain in a mesenchymal signature corresponds to a loss in the epithelial signature during de-differentiation of melanoma. By deciphering the interdependencies between de-differentiation and mesenchymal programs, the differences in molecular regulation between EMT and de-differentiation can be better understood. We have identified that the mesenchymal program, but not the epithelial program, is closely linked to de-differentiation. Although the mesenchymal signature enrichment shows a strong negative correlation with a differentiated/

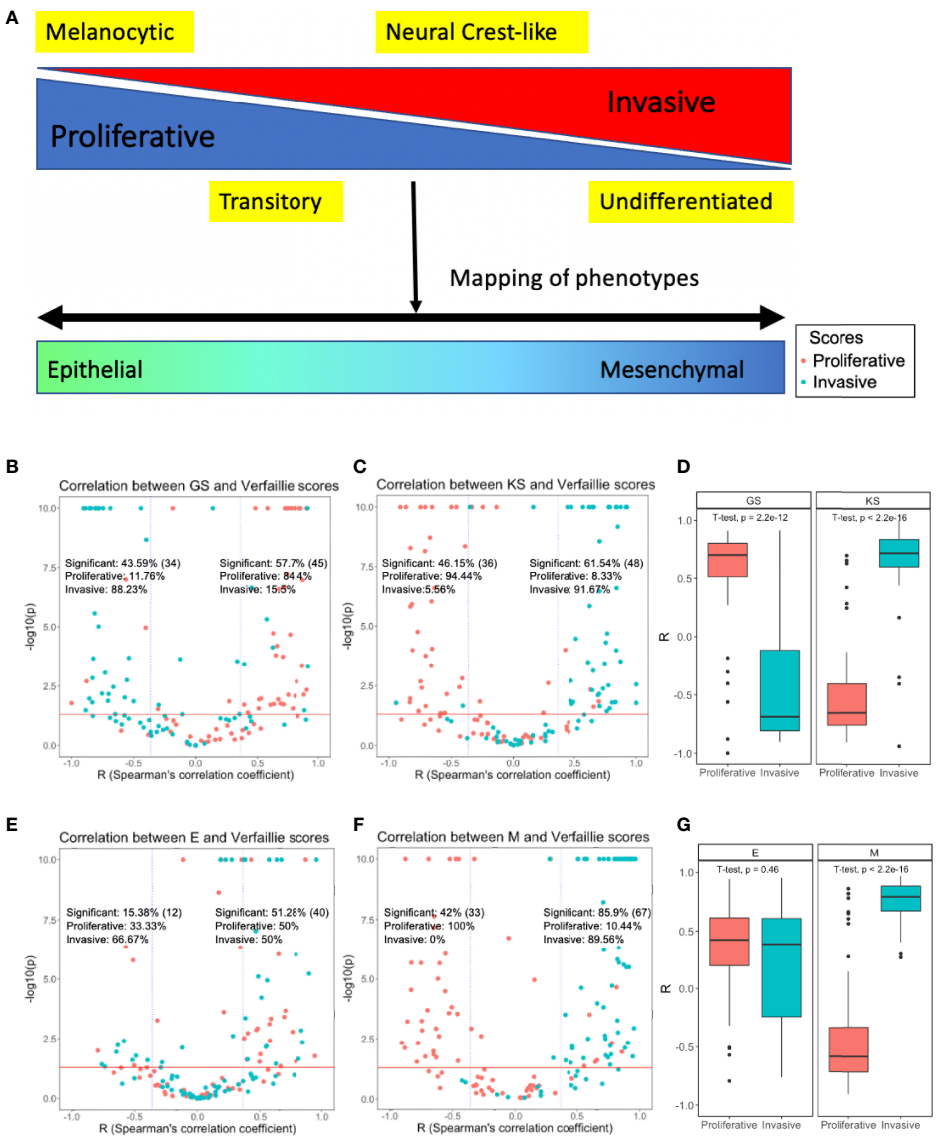


FIGURE 1
Mapping phenotypic heterogeneity in melanoma onto the EMT axis. **(A)** A schematic representation. Volcano plots depicting Spearman's correlation coefficients and -$\log_{10}$(p-value) of 78 datasets for the Verfaillie proliferative and invasive gene set with **(B)** 76GS EMT scoring metric, and with **(C)** KS EMT scoring metric **(D)** Boxplot depicting range of correlation coefficients for KS and 76GS with Verfaillie invasive and proliferative gene sets. Volcano plots depicting the Spearman's correlation coefficient and -$\log_{10}$(p-value) of 78 datasets for Verfaillie proliferative and invasive gene set with **(E)** Epithelial gene set (E scores) and **(F)** Mesenchymal gene set (M scores). **(G)** Boxplot depicting range of correlation coefficients for E and M scores with Verfaillie invasive and proliferative gene sets. Inset labelled "Significant" is calculated as the fraction of datasets (out of 78) which show a significant correlation trend ($r < -0.36$ or $r > 0.36$, $p < 0.05$). The absolute number of significant points (datasets) for the specified cut-off is indicated in brackets. "Proliferative" and "Invasive" labels represent the percentage of significant correlations that are between the EMT score and proliferative score or invasive score, respectively.

melanocytic transcriptional program, it does not increase monotonically during de-differentiation. This non-monotonic trend is also captured by metabolic programs associated with EMT, such as glycolysis and HIF1α, but not with metabolic programs associated with differentiation/melanocytic genes, such as the MITF-regulated OXPHOS pathway. Our results indicate that phenotypic heterogeneity in melanoma occurs along a proliferative-invasive axis that correlates with a "one-dimensional EMT" in which cells transition along a mesenchymal axis without an alteration in epithelial phenotype. Deciphering such inter-connections among multiple axes of plasticity in a cancer cell population may guide potent combinatorial therapeutic strategies aimed at controlling transitions to a more hybrid cell type with combined features of both proliferation and invasion.

Materials and methods

Software and datasets

Publicly available datasets from Gene Expression Omnibus (GEO), The Cancer Genome Atlas (TCGA), Cancer Cell Line Encyclopedia (CCLE- Broad Institute) (58), and National Cancer Institute-60 (NCI-60) databases were analyzed. Microarray data were downloaded from GEO using GEOquery Bioconductor R package. All analyses done on R version 4.1.0. *ggplot2*, and *ggpubr* R packages were used to create and customize plots.

Pre-processing of datasets

Microarray datasets, with un-mapped probe IDs, were pre-processed by mapping the probe IDs onto their gene symbols using the relevant platform annotation table. In the case of multiple probes mapping to the same gene, the mean expression of all the probes was considered for that gene. For non-normalized RNA-Seq datasets TPM normalization followed by *log2* transformation with an offset value of 1 was used.

ssGSEA

Single-sample Gene Set Enrichment Analysis, an extension of Gene Set Enrichment Analysis (GSEA) (56, 59), calculates separate enrichment scores for each sample and a gene set. Each score represents the degree to which genes in a gene set are up or down-regulated in a sample. We calculated ssGSEA scores for the Verfaillie proliferative and Verfaillie invasive gene sets (60), Hoek proliferative and Hoek invasive gene sets (39), the epithelial (E) and mesenchymal (M) gene sets of the EM tumor gene signature genes and cell lines gene signatures in the KS scoring metric (57), and the

TABLE 1 List of scores used for quantifying various axes of heterogeneity.

Score	Description	Significance	Reference
76 Gene Signature (76GS)	Metric for how epithelial a sample is. Calculated by using a weighted sum of gene expression for 76 genes.	Shows weak correlation with de-differentiation scores	55
Kolmogorov -Smirnov test (KS)	Metric for how mesenchymal a sample is. Ranges from -1 to +1. Calculated by subtracting expression-level based scores for epithelial genes from that for mesenchymal genes,	Shows weak correlation with de-differentiation scores	57
E score	ssGSEA score for only epithelial genes used in KS score. No mesenchymal genes are used for quantification.	Shows no correlation with de-differentiation scores	3
M Score	ssGSEA score for mesenchymal genes used in KS score. Na epithelial genes are used for quantification.	Shows strong overall correlation with de-differentiation scores, non-monotonic	3
OXPHOS	ssGSEA score for oxidative phosphorylation geneset	Shows strong overall correlation with de-differentiation scores, monotonic	61
Glyco	ssGSEA score for glycolysis geneset	Shows strong overall correlation with de-differentiation scores, non-monotonic	61
HIF-1	Singscore calculation for 59 downstream targets of HIF-1	Shows strong overall correlation with de-differentiation scores, non-monotonic	62
FAO	Average Z-scores for 14 FAO enzyme genes	Shows strong overall correlation with de-differentiation scores, monotonic	63
Verfaillie proliferative score	ssGSEA score for proliferative geneset	NA	60
Verfaillie invasive score	ssGSEA score for invasive geneset	NA	60
Hook proliferative score	ssGSEA score for proliferative geneset	NA	39
Hook invasive score	ssGSEA score for invasive geneset	NA	39
Tsoi Melanocytic score	ssGSEA score for melanocytic geneset	NA	40
Tsoi Transitory score	ssGSEA score for Transitory geneset	NA	40
Tsoi NCSC score	ssGSEA score for NCSC geneset	NA	40
Tsoi Undifferentiated score	ssGSEA score for undifferentiated geneset	NA	40

Not Applicable.

Tsoi melanocytic, transitory, NCSC, and undifferentiated gene set (40) (Table 1).

Calculation of EMT scores

We calculated EMT scores of datasets using four metrics- 76 Gene Signature (76GS), Kolmogorov -Smirnov test (KS), E score and M score (Table 1). 76GS and KS were calculated as defined earlier (1, 55, 57). 76GS score is a metric for how epithelial a sample is, i.e., higher scores reflect greater association with an epithelial phenotype. The KS score is a metric for how mesenchymal a sample is. The higher the KS score of a sample, the greater is its association with a mesenchymal phenotype. While 76GS scores do not have a pre-defined range of scores, KS scores lie within a +1 to -1 range. The E and M scores are ssGSEA scores for epithelial and mesenchymal gene lists, respectively, for the KS scoring metric (3). For calculating KS, E and M scores, datasets were classified based on whether the samples were derived from cell-lines or tumors and the appropriate gene sets were used.

Correlations

All correlation values were calculated using Spearman's correlation coefficient, unless mentioned otherwise. Spearman's correlation coefficient method generates a coefficient ranging between −1 to +1, where +1 indicates a strong positive correlation, and −1 indicates a strong negative correlation between two variables. It determines the correlation between any monotonically related variables- linear or non-linear. Correlations with R >0.36 and p<0.05 are considered significant.

Moving window average

A moving window average is used to quantify the gradient for a variable along a given axis. A window covering 60% of the entire range of the axis is created and the average value of the variable for all samples in the window is calculated. Then the window is then shifted by 1% and the average is re-calculated. This is iteratively repeated to cover the entire range. The slope of the averages determines the magnitude and direction of the gradient.

Conditional probabilities

Once the cell lines were sorted into their respective phenotypes and the conditional probabilities were obtained, the statistical significance and p-values for the conditional probabilities were calculated using the one-proportion Z test.

The z-score was calculated using the equation

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

where $\hat{p}$ is the observed proportion, p_0 is the null probability, and n is the sample size. The obtained value of z was then converted into the corresponding p-value using the standard normal distribution. If the obtained p-value < 0.05, it was considered significant.

Assigning phenotypes to samples

In order to identify samples belonging to the 4 phenotypes (melanocytic, transitory, NCSC and undifferentiated), we calculated ssGSEA scores based on gene sets for each of these phenotypes (40). Samples lying in the top 10% scores were assigned that particular phenotype. Taking a cut-off value of less than 10% would can enable only one point being selected for each phenotypes in datasets having less than 20 samples (e.g. int in Figure 4D, GSE101434) while increasing this threshold might lead to non-specific phenotype cells being selected in larger datasets. Thus, 10% was chosen as an optimal cut-off.

Metabolic scores

The oxidative phosphorylation (OXPHOS) and glycolysis (Glyco) scores in our study were calculated using ssGSEA carried out with the corresponding hallmark gene sets for these pathways [obtained from Molecular Signature Database (MSigDB) (61)]. The HIF-1 signature - which is a surrogate for glycolysis - was quantified based on a method previously reported (64). This method uses expression levels of their downstream target genes to capture the respective enzyme activities. A total of 59 downstream genes for HIF-1 were used and the scores were obtained using the Singscore method performed on these gene sets (62, 65). The fatty acid oxidation (FAO) scores were calculated based on the equation previously reported (63) which uses expression levels of 14 FAO enzyme genes.

Results

Enrichment of mesenchymal genes can capture the extent of de-differentiation in melanoma

To test whether EMT and de-differentiation in melanoma programs are correlated with one another, we used previously-defined EMT scores – KS and 76GS (55, 57) – and ssGSEA scores for Verfaillie proliferative and invasive (60) and Hoek

proliferative and invasive melanoma gene sets (39) and investigated their correlation coefficients across 78 datasets. Additionally, to dissect the contributions of epithelial and mesenchymal gene set separately, we calculated the ssGSEA scores (56, 59) for corresponding gene sets individually too (57), referred here as E and M scores, respectively (3). A sample with a higher 76GS or E score is more epithelial while a higher KS or M score refers to more mesenchymal phenotype. Thus, given the overlap between mesenchymal and invasive programs, we expected invasive scores to correlate positively with KS and M scores and negatively with 76GS and E scores. We also expected opposite trends for proliferation scores: negative correlations with KS and M scores and positive correlation with 76GS and E scores. We visualized the relationships between these pathways as volcano plots in which each dot corresponds to a dataset analysed. For positively-correlated metrics, we expect the majority of data sets to lie in the top right rectangle, while those displaying a significant negative correlation are expected to lie in the top left rectangle.

A total of 34 out of 78 datasets (43.59%) showed a significant negative correlation ($r < -0.36$, $p < 0.05$) between 76GS and one of the two Verfaillie (proliferative, invasive) scores (66). In 30 out of those 34 datasets (88.23%), 76GS scores correlated negatively with invasive scores, while in remaining 4 datasets (11.76%), 76GS scores correlated negatively with proliferative scores (Figure 1B, left). Similarly, among 45 datasets that showed a positive correlation ($r > 0.36$, $p < 0.05$) between 76GS scores and one of Verfaillie scores, 38 (84.4%) cases had a positive correlation between 76GS and proliferative scores, and in the remaining seven datasets, 76GS scores correlated positively with invasive scores (Figure 1B, right). Overall, both the scoring metrics (76GS and KS) displayed correlations with Verfaillie and Hoek proliferative and invasive scores across the 78 datasets to support a relationship between E/M plasticity and the proliferative/invasive axis (Figures 1B–D, S1A–C).

Because gain of mesenchymal features is reported more commonly in melanoma as compared to loss of epithelial features, we decoupled the epithelial and mesenchymal components of the scoring metrics (E and M scores, respectively). The KS method provides separate information on genes that are associated with an epithelial phenotype and those with a mesenchymal state. Using the genes from the KS scoring method we segregated the genes and calculated individual ssGSEA scores for epithelial and mesenchymal gene lists and re-evaluated their correlation with proliferative and invasive scores in melanoma. While epithelial genes continued to show random distributions of samples throughout the plot, mesenchymal genes showed clear segregation of proliferative and invasive scores based on Spearman's correlation coefficients, i.e., invasive scores were positively correlated with M score while proliferative scores were negatively correlated with the M scores (Figures 1E-G, S1D–F). This observation suggests that mesenchymal genes, but not epithelial genes, can capture the phenotypic heterogeneity displayed by melanoma along the proliferative-invasive axis.

To provide further support for these observations, we focused only on Verfaillie gene sets, since they have levels of overlap with gene sets for the intermediate phenotypes that were previously identified (40) (Figure S1G). Thus, a continuous scoring metric defined for the Verfaillie gene set is expected to be more sensitive for capturing intermediate phenotypes as compared to the Hoek gene set.

Because correlation coefficients only provide an overall trend in data, we wished to determine how proliferative and invasive scores vary along the E and the M axis. For this purpose, we generated two dimensional EMT plots of the data sets in which E and M scores are represented along each of the two axes. These plots display the relative position of a sample along an epithelial or mesenchymal axis (3, 56, 59). We then overlay information on the proliferative and invasive scores for each sample. As expected, across various datasets, proliferative and invasive scores for samples had a stronger visible gradient along the M axis as compared to the E axis (Figures 2A–B). To quantify this gradient, we used a rolling window to estimate the increase of average proliferative and invasive scores across the E and M axis. For this, we start with a rolling window covering 60% of the entire range along a given axis and calculate the average proliferative (P) or invasive (I) score within that window. Then the window is shifted by 1% and the average is re-calculated. This process is repeated until the entire range is covered, and the change in averages is plotted. For an axis that strongly correlates with the change in scores, we expect a steeper slope. The nature of a slope (positive or negative) is determined by the correlation between the axis and the score. Both axes trend in the expected direction, with a positive slope for invasive scores and negative slope for proliferative scores along the M axis and vice versa for the E axis (Figure 2C). This analysis also reveals that the M axis has a steeper curve than the E axis for both P and I scores. These results suggest that proliferative-invasive heterogeneity in melanoma can be considered as a "one-dimensional form" of EMT where the mesenchymal program enrichment increases as cells become more invasive, but the epithelial program need not be suppressed concomitantly (Figures 2, S2), as often tacitly assumed for the case of EMT. Other non-epithelial cancers, such as glioblastoma (GBM) and sarcoma, also display larger variation along the M-score axis as compared to the E-score axis, suggesting that "one-dimensional EMT" might not be specific to melanoma alone (Figures S3A–B). Moreover, we also observe that more de-differentiated phenotypes in sarcoma display higher M scores, while in GBM a switch from proneural to mesenchymal phenotypes is clearly visualised along the M-score axis. Thus, phenotypic plasticity along a mesenchymal axis in non-epithelial cancers can take various trajectories.

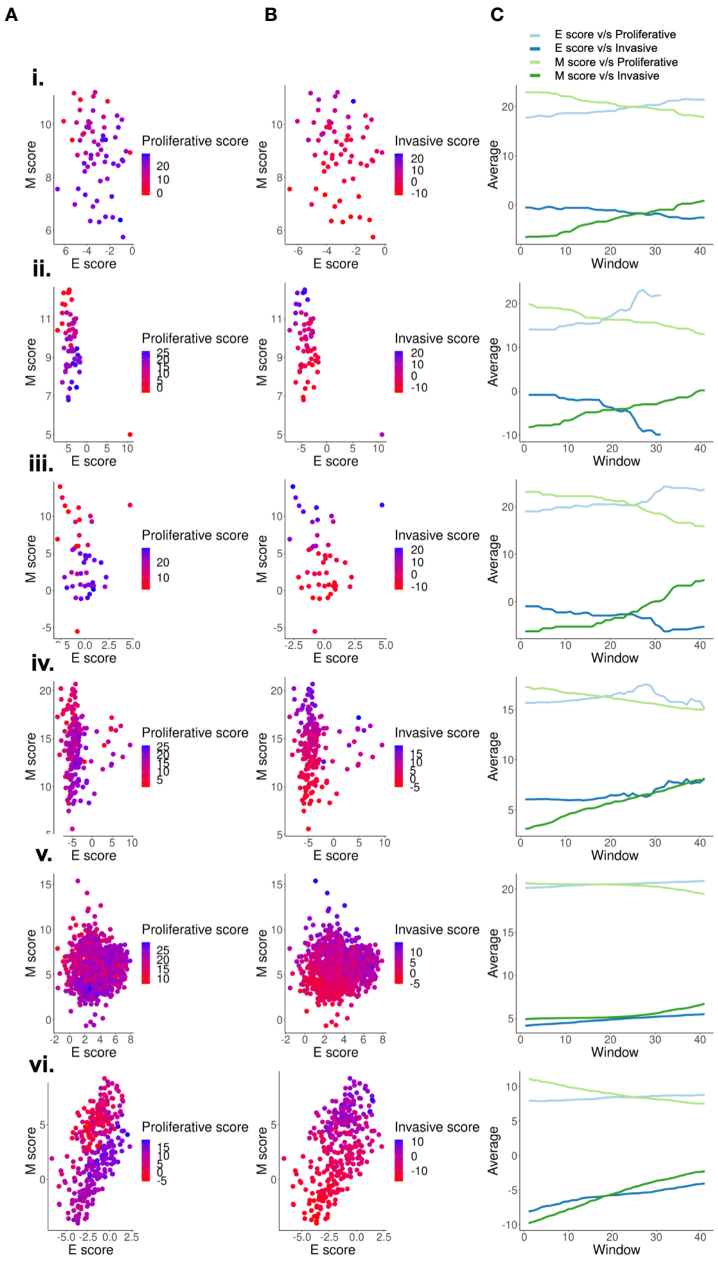


FIGURE 2
Scoring metrics based on mesenchymal genes capture de-differentiation better than metrics based on epithelial genes. Two dimensional EMT plots of different types of datasets- (i) GSE7127 (63 melanoma cell lines - microarray), ii. CCLE (59 cell lines - microarray), iii.GSE4843 (45 tumor samples - microarray), iv.GSE65904 (214 tumor samples - microarray),v. GSE72056 (1257 single-cell tumor samples), vi.GSE81383 (307 single-cell tumor sample) depicting the variation of **(A)** Proliferative scores along the E and M score axes. **(B)** Invasive scores along the E and M score axes. **(C)** Quantifying the proliferative and invasive score gradient along the E-M axes using a rolling window.

The mesenchymal axis follows a non-monotonic relationship with de-differentiation

Because the M score axis was able to capture the phenomenon of de-differentiation quantified by continuous scoring metrics, such as the proliferative and invasive scores, we next tested if the discretized phenotypes also arrange themselves in order of appearance along the two dimensional EMT plane. The classification of samples into four categories - melanocytic, transitory, NCSC and undifferentiated (in order of increasing de-differentiation) - for GSE80829, GSE10916,

GSE4843, GSE7127 and GSE116237 was done as previously defined (15, 42). Along the proliferative-invasive plane, samples displayed a strong negative relationship between the two scores, i.e., proliferative scores of samples decreased as their invasive score increased. The four phenotypes also appeared in the expected order (18, 40), with the melanocytic samples having the highest proliferative scores and lowest invasive scores, and the undifferentiated samples displaying the lowest invasive scores and highest proliferative scores (Figure 3A). However, the two dimensional EMT plane failed to resolve the four phenotypes in terms of these four phenotypes showing non-overlapping scores. Since the E score axis performed poorly previously (Figures 1E, G) in these metrics, we quantified the ability of M score axis alone to resolve the four phenotypes by quantifying the conditional probability of a sample to belong to the intermediate phenotypes (transitory and NCSC), given that they lie in an intermediate M score range. Interestingly, samples with intermediate M scores were significantly likely to belong to

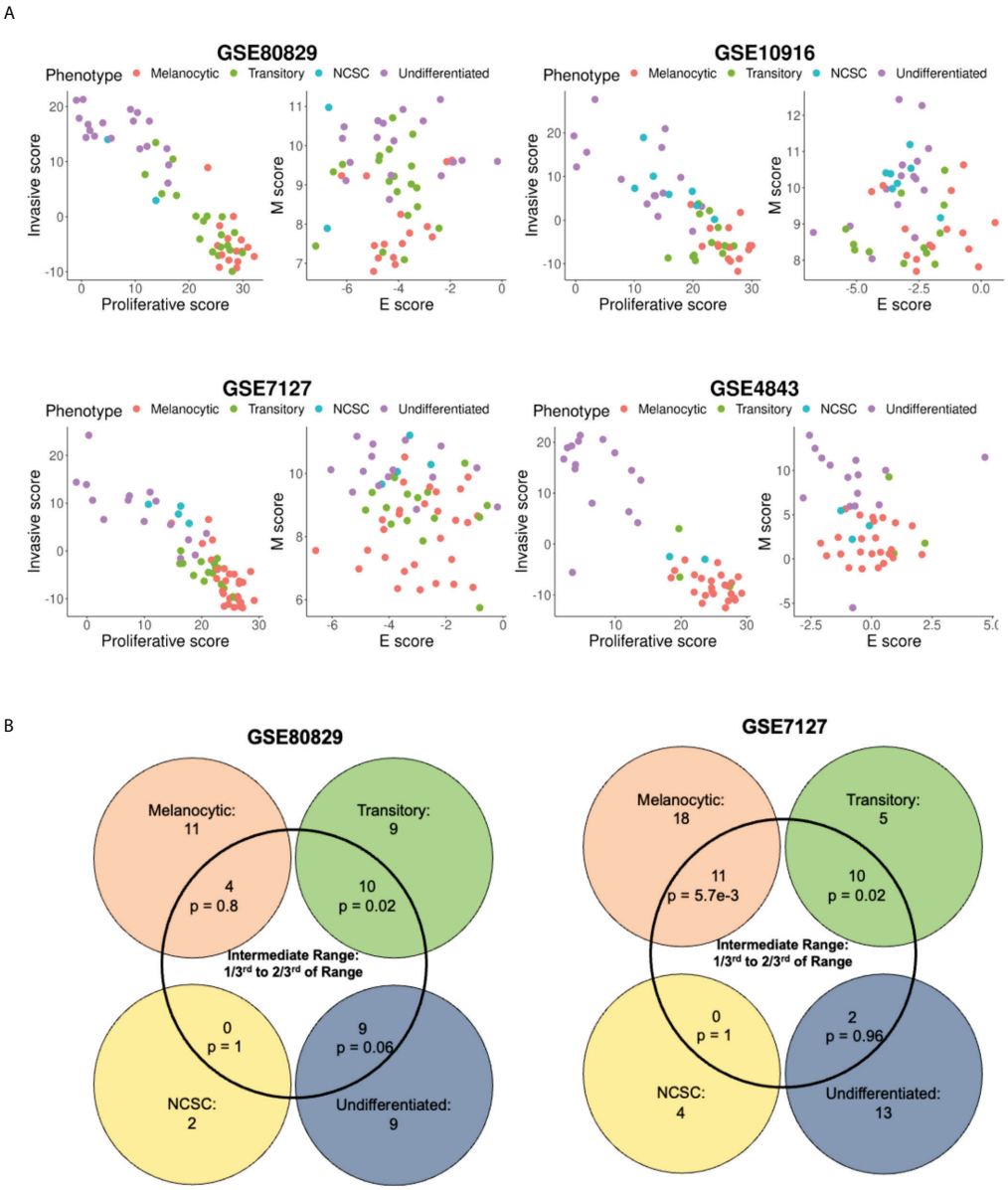


FIGURE 3
Variation of the four molecular phenotype scores along the epithelial, mesenchymal, proliferative, and invasive axes. **(A)** Plotting samples classified into four phenotypes onto the E-M, proliferative-invasive score axes. **(B)** Venn diagram depicting the intersection of the four phenotype scores of samples and intermediate M scores. p represents p-value for the conditional probability that a sample belongs to the phenotype given that they lie in the intermediate M score range.

10.3389/fonc.2022.913803

the transitory phenotype (Figures 3B, S3C, Table 2). However, the probability of these samples to belong to the NCSC phenotype was negligible. In some datasets (GSE7127, GSE116237), the melanocytic phenotype was also significantly enriched in the intermediate M score populations. However, the melanocytic phenotype cells were enriched in the bottom M score population as well, and were not uniquely present in the intermediate score range like the transitory phenotype cells (Figures S3D–F).

To further dissect the relationship between the four phenotypes and the M score axis, we quantified the change in M score with respect to the invasive scores for the four phenotypes. To identify the four phenotypes, we used ssGSEA scores for gene sets defined for each of the four phenotypes (40). The top 10% of samples that had the highest scores for a particular gene set, were assigned the label of that particular phenotype. We observed that in these samples there was a non-monotonic increase in M scores as invasive score/de-differentiation increased. As samples progressed from NCSC to undifferentiated, M scores either decreased (Figures 4C–E) or remained the same (Figures 4A–B, F). In the context of melanocyte development, neural crest cells are precursors for melanocytes with high migratory potential and high levels of EMT markers (17, 67, 68). Thus, the non-monotonic increase in the mesenchymal program seen here is reminiscent of the differentiation of melanocytes.

Metabolic reprogramming along the proliferative-invasive axis in melanoma

The EMT status of epithelial cancer cells is often associated with distinct metabolic programs. Generally speaking, EMT is negatively correlated with the enrichment of oxidative phosphorylation (OXPHOS) and fatty acid oxidation (FAO), but positively correlated with glycolysis (62). In melanoma, the proliferative state is associated with high levels of OXPHOS and the invasive phenotype is associated with high levels of glycolysis (69–72), reinforcing the commonalities between these two different instances of phenotypic plasticity. Computational analysis has suggested the existence of four metabolic sub-populations (63): 1) OXPHOS-high/glycolysis-low, 2) OXPHOS-low/glycolysis-high, 3) OXPHOS- low/glycolysis-low, and 4) OXPHOS high/glycolysis-high.

To assess whether the OXPHOS-glycolysis metabolism axis can be mapped onto the proliferation-invasion axis, we calculated Spearman's correlation coefficients between the metabolic scores (OXPHOS and glycolysis) and the de-differentiation scores (proliferative and invasive scores) (Figures 5A–C) across the 78 datasets. In 38 out of 78 datasets where the OXPHOS scores correlate significantly with proliferative scores, 34 datasets show a positive correlation. Similarly, among 43 datasets showing a significant correlation of OXPHOS scores with invasive scores, all of them showed negative correlation. Thus, overall, OXHOS scores corelated positively with proliferative scores and negatively with invasive scores (Figure 5A). Glycolysis scores, on the other hand, did not show a clear relationship with EMT status, with a subset of datasets showing trends in both the directions (positive and negative correlation) both for proliferative and invasive scores (Figure 5B). This difference is reminiscent of prior observations for the association of EMT with OXPHOS and glycolysis in which glycolysis is only moderately correlated with EMT status, but OXPHOS is consistently negatively correlated with EMT (62). This trend is substantiated by observations that in cases where OXPHOS is positively correlated with proliferative scores or negatively correlated with invasive scores, glycolysis scores do not show any particular direction of enrichment with either proliferative or invasive axes (Figure 5C).

We next sought to dissect whether intermediate melanoma phenotypes might be enriched for a specific metabolic profile. To investigate this trend, we calculated the Spearman's correlation coefficients for metabolic scores and ssGSEA scores for gene signatures corresponding to each of the four molecular phenotypes of melanoma (Figure 5D–F). OXPHOS showed a clear shift from datasets with a significant positive correlation with a melanocytic phenotype to a significant negative correlation for the undifferentiated phenotype (Figure 5D). On the contrary, glycolysis scores do not show a clear shift from negative to positive correlations with de-differentiation (Figure 5E). Similar to the non-monotonic trend observed for M-scores, the glycolysis scores show the strongest positive correlation trends for the NCSC phenotype. Undifferentiated phenotype scores have a mixture of positively correlated and negatively correlated datasets with respect to glycolysis scores. Put together, these observations suggest that the regulatory modules controlling the switch to glycolysis are likely linked to the mesenchymal program rather than the de-differentiation one. On the other hand, regulatory modules for OXPHOS are likely to be closely linked to the melanocytic

TABLE 2 Conditional probabilities for a sample belonging to a particular phenotype given that it lies in the intermediate M score range.

Dataset	P (Melanocytic| Intermediate M Score)	p-value	P (Transitory| Intermediate M Score)	p-value	P (NCSC| Intermediate M Score)	p-value	P (Undifferentiated| Intermediate M Score)	p-value
GSE80829	0.17	0.8	0.43	0.02	0	1	0.39	0.06
GSEE7127	0.48	0.01	0.43	0.02	0	1	0.09	0.96
GSE116237	0.36	0	0.49	0	0.09	1	0.06	1

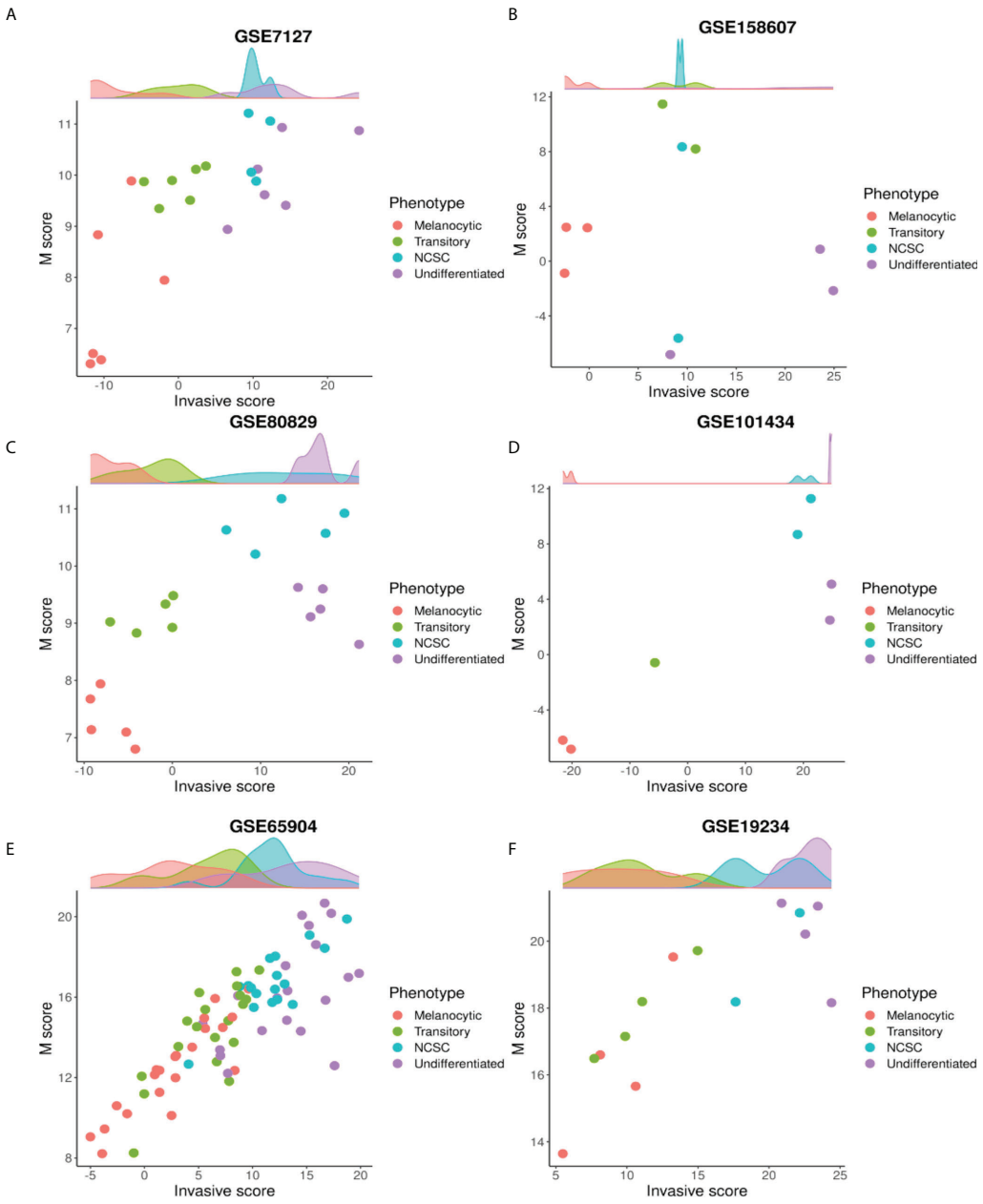


FIGURE 4
The mesenchymal axis follows a non-monotonic relationship with de-differentiation. Plotting M scores against invasive scores for different phenotypes along the P-I axis in many datasets: **(A)** GSE7127 **(B)** GSE158607 **(C)** GSE80829 **(D)** GSE101434 **(E)** GSE65904 **(F)** GSE19234.

differentiation program. This trend is in accordance with experimental evidence that suggests that OXPHOS in melanoma cells is regulated by PGC1α, a downstream target of MITF, a key regulator of melanocyte differentiation (73, 74). Interestingly, fatty acid oxidation, which is also directly controlled by MITF *via* SCD (75), also displays trends similar to OXPHOS (Figure S4A) while a HIF1α signature, that is commonly associated with the invasive phenotype follows a non-linear trend similar to glycolysis (Figure S4B), suggesting that it is linked to the mesenchymal program rather than the de-differentiation program.

Discussion

De-differentiation in melanoma occurs in response to targeted therapy. This process may be mediated by transitions

10.3389/fonc.2022.913803

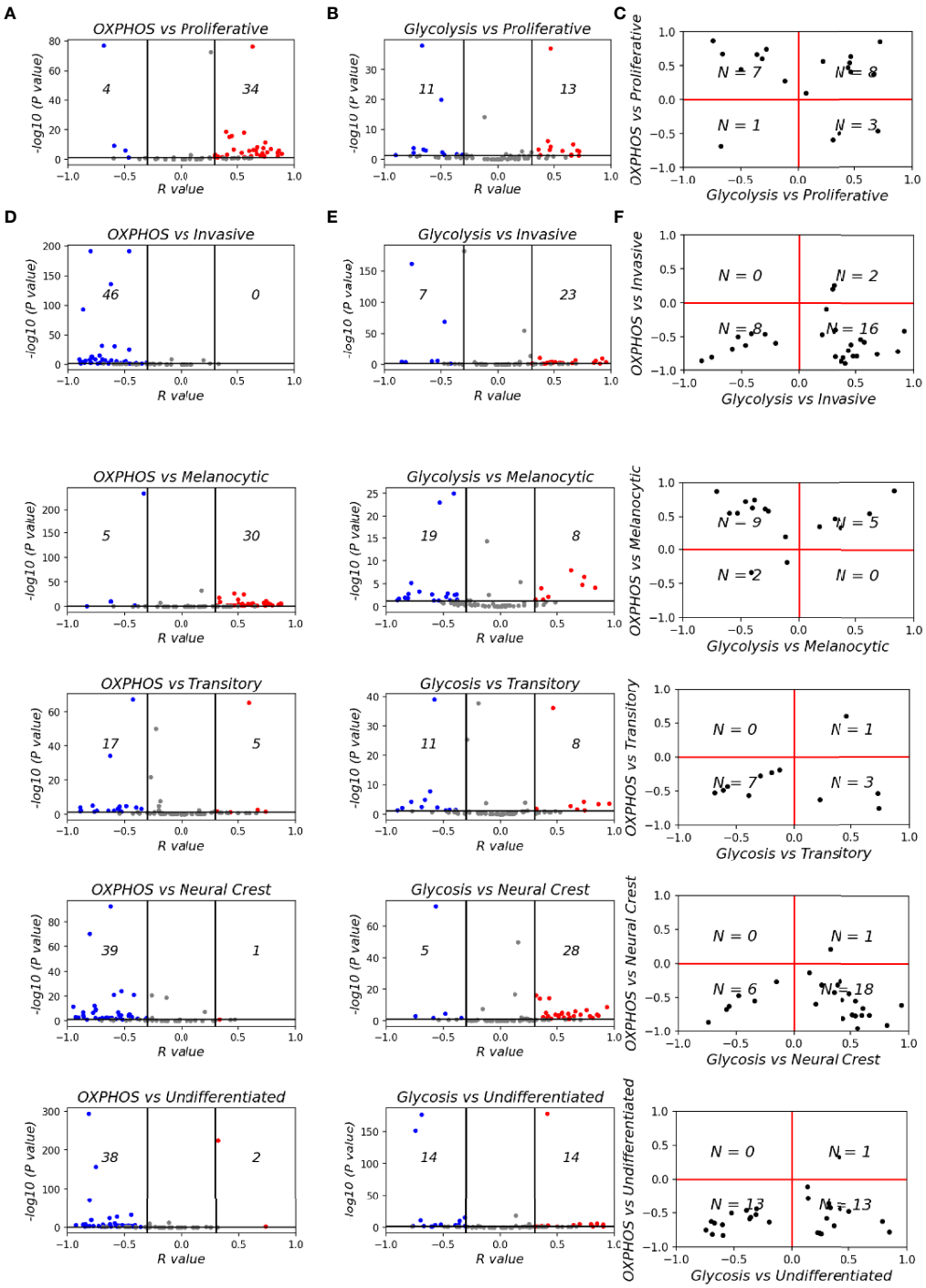


FIGURE 5
Mapping metabolic programs associated with EMT onto the de-differentiation program axes. Volcano plots depicting Spearman's correlation coefficient and -$\log_{10}$(p-value) of 78 datasets for **(A)** Hallmark OXPHOS and Verfaillie gene set. **(B)** Hallmark glycolysis and Verfaillie gene set. **(C)** Spearman's correlation coefficient between OXPHOS and Glycolysis and Verfaillie scores. **(D)** Hallmark OXPHOS and Tsoi gene set. **(E)** Hallmark glycolysis and Tsoi gene set. **(F)**. Spearman's correlation coefficient between OXPHOS and Glycolysis and Tsoi scores. N represents number of samples present in a given quadrant.

across a spectrum of phenotypes in which melanocytic cells treated with BRAF/MEK inhibitors pass through a transitory phenotype, followed by the NCSC phenotype, before becoming completely un-differentiated (15, 18, 40). This trajectory is accompanied by loss of a proliferative signature and gain of invasive characteristics. Here, we decipher the relationship between de-differentiation and EMT in melanoma. These processes are often considered to co-occur during drug treatment (14, 16, 34); however, comparison of EMT and de-differentiation scores reveal that the two processes may be more closely related to the mesenchymal program rather than the loss of an epithelial-like state or an EMT program *per se*. This observation is reminiscent of previous results in breast cancer and melanoma in which regulatory genes for the mesenchymal and de-differentiated phenotypes overlapped while those corresponding to epithelial and differentiated (melanocytic) phenotypes did not overlap and were tissue-specific (76). Previous pan-cancer studies have also highlighted that downregulation of epithelial components and upregulation of mesenchymal features need not always be as strongly coupled as often assumed (77, 78). Moreover, differences along these two axes need not be necessarily reflected at a transcriptional level (79). Together, these observations highlight the need to analyze epithelial and mesenchymal axes independently, rather than as a conventional single metric for EMT.

Our results also indicate that metabolic programs can be linked either with the de-differentiation program or the mesenchymal program. OXPHOS and fatty acid oxidation are both indirectly regulated by MITF. In the case of OXPHOS, MITF regulates PGC-1α (74); in the fatty acid oxidation pathway, MITF regulates SCD (75). MITF, which controls both metabolic pathways, decreases with increasing de-differentiation. This trend is explained by the decline in MITF associated with de-differentiation, in accordance with the MITF rheostat model (80). On the other hand, glycolysis and HIF-1α signatures seem to be co-regulated with the mesenchymal program. Previous studies in epithelial cancers have shown how well-established EMT transcription factors (EMT-TFs) regulate the metabolic profile of a cell and cause a switch to glycolysis (also called Warburg effect) (81). Consistently, neural crest cells also display decay of glycolytic capabilities as they differentiate into melanocytes (82). Our analysis suggests that the metabolic state of a cell is closely linked to the transcriptional program governing it at a given time point. Thus, de-differentiation captures the transcriptional and metabolic states observed during melanocyte development.

Although our study focuses on melanoma, EMT-like phenotypic switching is also characteristic of other non-epithelial cancers and de-differentiation of melanocytes independent of malignant transformation. De-differentiation of melanocytes into pluripotent stem cells demonstrated a reduction in expression levels of E-Cadherin, an epithelial marker, and similarities to mesenchymal stem cells (83). Molecular subtypes of glioblastoma multiforme (GBM), a non-epithelial cancer, include the pro-neural, classical, and mesenchymal phenotypes, which exist along a spectrum of worsening prognosis (84). Single-cell analysis reveals that these molecular subtypes recapitulate neurodevelopmental trajectories, with proneural cells forming a major composition of proliferative glial progenitor-like cells (85, 86). A proneural-to-mesenchymal transition (PMT) is characterized by an increase in mesenchymal markers, such as SNAI1 and ZEB1. Similarly, glioma stem cells (GSCs) exist as proneural GSCs and mesenchymal GSCs, which can give rise to the complete spectrum of intra-tumor heterogeneity, including the classical phenotype (87), reminiscent of epithelial and mesenchymal CSCs reported in breast cancer (88). Moreover, samples belonging to the classical subtype are depleted of pro-neural GSCs and enriched for mesenchymal GSCs, possibly suggesting that mesenchymal GSCs are more likely to give rise to the classical subtype. This trend strengthens the semi-independent nature of EMT and stemness as seen in epithelial cancers (78). Another study in GBM cell lines reports that loss of N-cadherin (a well-established mesenchymal marker) increases invasiveness (89), reinforcing the trends that increased migration and invasion is not always an inevitable consequence of carcinoma-associated EMT (90). These scenarios of non-overlapping behaviors in terms of invasiveness, stemness and EMT, seen both for epithelial and non-epithelial cancers, advocate for improving existing therapeutic strategies by targeting multiple axes of cellular plasticity simultaneously rather than individually.

Our study focuses on the overlap between the de-differentiation and the EMT axis during drug treatment in melanoma samples. However, de-differentiation is not the only trajectory altered by drug treatment. Cells can follow multiple paths to therapy resistance, one of which is by attaining a hyper-pigmented phenotype (15, 91, 92). The mapping of these trajectories and states to the E-M axis remains to be studied. In addition, another axis of cellular plasticity commonly associated with EMT is immune suppression and immune evasion. Previous studies have shown that the expression levels of programmed death-ligand 1 transmembrane protein (PD-L1) – a driver of immune evasion - does not increase monotonically with EMT (3). Consistently, in melanoma, the expected trend of worse response to anti-PD-1 therapy with increasing de-differentiation is not observed; rather, results from the CheckMate 038 clinical trial indicate that the NCSC phenotype is associated with a better outcome to immune checkpoint blockade therapy as compared to the melanocytic phenotype (93). The extent of overlap between the axes of EMT, immune evasion, and de-differentiation require further study to design temporally-sequenced effective combination therapies that can shift the differentiation and EMT status of melanoma toward a less invasive and more immune activated state. Recent *in vitro* investigations in melanoma have shown proof-of-principle evidence of phenotypic plasticity driven drug resensitization as a

mechanism underlying the beneficial impact of intermittent therapy (94).

Despite providing the abovementioned insights, our study suffers from various limitations. First, no mechanism-based models have been developed to gain insights into the emergent dynamics for the observed trends. A better understanding of the dynamics can help identify more effective therapeutic strategies by fine-tuning the interval, sequence, and dosage for combinatorial and/or sequential therapeutic strategies (95). Second, our analysis only characterizes phenotypes at a transcriptomic level, although preliminary investigation supports consistent trends at a proteomic level too (Figure S5). Third, due to limited availability of longitudinal transcriptomic data for varying treatment durations, our analysis is not restricted to time-resolved data exclusively. Preclinical data shows that short duration of drug treatment can induce a NCSC phenotype that is highly mesenchymal (14, 16), while prolonged treatment (8-12 weeks) can drive an undifferentiated phenotype. Our study indicates that a prolonged treatment can induce further de-differentiation but not always a concomitant increase in mesenchymal status, a prediction that needs detailed experimental validation. However, this observation of the NCSC phenotype being the most mesenchymal is in accordance with melanocyte development. Neural crest cells are progenitors of melanocytes that undergo EMT during development to delaminate and migrate from the neural tube to the epidermis, where they lose their EMT signature and differentiate into melanocytes (17, 67, 68). Thus, the non-monotonic variation in EMT during development (the initial increase during migration followed by decrease during differentiation) can be possibly recapitulated during treatment-induced de-differentiation. We propose that the often-presumed overlap between the mesenchymal and invasive axes may arise from the lack of information for longer time scales (since most *in vitro* drug treatment studies are performed in under three weeks), and often held assumptions about linearly increasing trends. However, increasing evidence suggests that maximum stemness is associated with hybrid E/M phenotypes rather than 'extreme' mesenchymal or epithelial phenotypes, suggesting that many such associations among axes of plasticity can be non-monotonic in nature (96–98).

Overall, our transcriptomic data-based analysis highlights the partially overlapping nature of EMT with molecular attributes of de-differentiation and metabolism during drug treatment in melanoma. We provide a framework for studying multiple intertwined axes of plasticity and heterogeneity (EMT, metabolic reprogramming, proliferative-invasive status) and identifying the degree to which these axes overlap.

Code Availability

All codes and scores generated for this paper can be found at: https://github.com/csbBSSE/EMT_Melanoma
https://github.com/priyanka8993/EMT_score_calculation

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/ Supplementary Material.

Author contributions

MP and GR performed research, analyzed data and wrote the first draft of the manuscript. PT, NA, SM, AR and DB performed research. JS and MKJ conceived research and worked on manuscript revisions. MKJ supervised research. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by Ramanujan Fellowship awarded to MKJ by Science and Engineering Research Board (SERB), Department of Science and Technology (DST), Government of India (SB/S2/RJN-049/2018) and by Infosys Young Investigator award to MJ supported by Infosys Foundation, Bangalore. MP is supported by KVPY fellowship (DST). JS is supported by NIH 1R01CA233585-03, Department of Defense W81XWH-18-1-0189, and the Triangle Center for Evolutionary Medicine.

Conflict of interest

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

Publisher's Note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.913803/full#supplementary-material

References

1. Chakraborty P, George JT, Tripathi S, Levine H, Jolly MK. Comparative study of transcriptomics-based scoring metrics for the epithelial-Hybrid-Mesenchymal spectrum. *Front Bioengineer Biotechnol* (2020) 8:220. doi: 10.3389/fbioe.2020.00220/bibtex

2. Hanahan D, Weinberg RA. Hallmarks of cancer: The next generation. *Cell* (2011) 144(5):646–74.

3. Sahoo S, Nayak SP, Hari K, Purkait P, Mandal S, Kishore A, et al. Immunosuppressive traits of the hybrid Epithelial/Mesenchymal phenotype. *Front Immunol* (2021) 12:797261/bibtex. doi: 10.3389/fimmu.2021.797261/bibtex

4. Jolly MK, Ware KE, Xu S, Gilja S, Shetler S, Yang Y, et al. E-Cadherin Represses Anchorage-Independent Growth in Sarcomas through Both Signaling and Mechanical Mechanisms E-Cadherin Represses Anchorage-Independent Growth. *Molecular Cancer Research* (2019) 17(6):1391–402.

5. Somarelli JA, Shetler S, Jolly MK, Wang X, Bartholf Dewitt S, Hish AJ, et al. Mesenchymal-epithelial transition in sarcomas is controlled by the combinatorial expression of microRNA 200s and GRHL2. *Molecular and cellular biology* (2016) 36(19):2503–13.

6. Siebzehnrubl FA, Silver DJ, Tugertimur B, Deleyrolle LP, Siebzehnrubl D, Sarkisian MR, et al. The ZEB1 pathway links glioblastoma initiation, invasion and chemoresistance. *EMBO Mol Med* (2013) 5(8):1196–212. doi: 10.1002/emmm.201302827

7. Roccaro AM, Mishima Y, Sacco A, Moschetta M, Tai YT, Shi J, et al. CXCR4 regulates extra-medullary myeloma through epithelial-Mesenchymal-Transition-like transcriptional activation. *Cell Rep* (2015) 12(4):622–35. doi: 10.1016/j.celrep.2015.06.059

8. Lemma S, Karihtala P, Haapasaari KM, Jantunen E, Soini Y, Bloigu R, et al. Biological roles and prognostic values of the epithelial–mesenchymal transition-mediating transcription factors twist, ZEB1 and slug in diffuse large b-cell lymphoma. *Histopathology* (2013) 62(2):326–33. doi: 10.1111/his.12000

9. Sánchez-Tilló E, Fanlo L, Siles L, Montes-Moreno S, Moros A, Chiva-Blanch G, et al. The EMT activator ZEB1 promotes tumor growth and determines differential response to chemotherapy in mantle cell lymphoma. *Cell Death Different* (2013) 21(2):247–57. doi: 10.1038/cdd.2013.123

10. Stavropoulou V, Kaspar S, Brault L, Sanders MA, Juge S, Morettini S, et al. MLL-AF9 expression in hematopoietic stem cells drives a highly invasive AML expressing EMT-related genes linked to poor outcome. *Cancer Cell* (2016) 30 (1):43–58. doi: 10.1016/j.ccell.2016.05.011

11. Wu S, Du Y, Beckford J, Alachkar H. Upregulation of the EMT marker vimentin is associated with poor clinical outcome in acute myeloid leukemia. *J Trans Med* (2018) 16(1):1–9. doi: 10.1186/s12967-018-1539-y/figures/7

12. Kahlert UD, Joseph J, Kruyt FAE. EMT- and MET-related processes in nonepithelial tumors: importance for disease progression, prognosis, and therapeutic opportunities. *Mol Oncol* (2017) 11(7):860–77. doi: 10.1002/1878-0261.12085

13. Denecker G, Vandamme N, Akay Ö., Koludrovic D, Taminau J, Lemeire K, et al. Identification of a ZEB2-MITF-ZEB1 transcriptional network that controls melanogenesis and melanoma progression. *Cell Death Different* (2014) 21(8):1250–61. doi: 10.1038/cdd.2014.44

14. Fallahi-Sichani M, Becker V, Izar B, Baker GJ, Lin J-R, Boswell SA, et al. Adaptive resistance of melanoma cells to RAF inhibition *via* reversible induction of a slowly dividing de-differentiated state. *Mol Syst Biol* (2017) 13:905. doi: 10.15252/msb

15. Rambow F, Rogiers A, Marin-Bejar O, Aibar S, Femel J, Dewaele M, et al. Toward minimal residual disease-directed therapy in melanoma. *Cell* (2018) 174 (4):843–55. doi: 10.1016/j.cell.2018.06.025

16. Ramsdale R, Jorissen RN, Li FZ, Al-Obaidi S, Ward T, Sheppard KE, et al. The transcription cofactor c-JUN mediates phenotype switching and BRAF inhibitor resistance in melanoma. *Sci Signaling* (2015) 8(390):ra82. doi: 10.1126/scisignal.aab1111/suppl_file/8_ra82_sm.pdf

17. Dupin E, le Douarin NM. Development of melanocyte precursors from the vertebrate neural crest. *Oncogene 2003 22:20* (2003) 22(20):3016–23. doi: 10.1038/sj.onc.1206460

18. Su Y, Wei W, Robert L, Xue M, Tsoi J, Garcia-Diaz A, et al. Single-cell analysis resolves the cell state transition and signaling dynamics associated with melanoma drug-induced resistance. *Proc Natl Acad Sci USA* (2017) 114(52):13679–84. doi: 10.1073/pnas.1712064115

19. Vandamme N, Denecker G, Bruneel K, Blancke G, Akay O, Taminau J, et al. The EMT transcription factor ZEB2 promotes proliferation of primary and metastatic melanoma while suppressing an invasive, mesenchymal-like phenotype. *Cancer Res* (2020) 80(14):2983–95. doi: 10.1158/0008-5472.can-19-2373

20. Wouters J, Kalender-Atak Z, Minnoye L, Spanier KI, de Waegeneer M, Bravo González-Blas C, et al. Robust gene expression programs underlie recurrent cell states and phenotype switching in melanoma. *Nat Cell Biol* (2020) 22(8):986–98. doi: 10.1038/s41556-020-0547-3

21. Deng Y, Chakraborty P, Jolly MK, Levine H. A theoretical approach to coupling the epithelial-mesenchymal transition (EMT) to extracellular matrix (ECM) stiffness *via* LOXL2. *Cancers 2021* (2021) 13(7):1609. doi: 10.3390/cancers13071609

22. Fattet L, Jung HY, Matsumoto MW, Aubol BE, Kumar A, Adams JA, et al. Matrix rigidity controls epithelial-mesenchymal plasticity and tumor metastasis *via* a mechanoresponsive EPHA2/LYN complex. *Dev Cell* (2020) 54(3):302–16.e7. doi: 10.1016/j.devcel.2020.05.031

23. Kumar S, Das A, Sen S. Extracellular matrix density promotes EMT by weakening cell-cell adhesions. *Mol Biosyst* (2014) 10(4):838–50. doi: 10.1039/c3mb70431a

24. Bianchi A, Gervasi ME, Bakin A. Role of β5-integrin in epithelial-mesenchymal transition in response to TGF-β. *Cell cycle (Georgetown, Tex.)* (2010) 9(8):1647–59. doi: 10.4161/cc.9.8.11517

25. Kilinc AN, Han S, Barrett LA, Anandasivam N, Nelson CM. Integrin-linked kinase tunes cell-cell and cell-matrix adhesions to regulate the switch between apoptosis and EMT downstream of TGFβ1. *Mol Biol Cell* (2021) 32(5):402–12. doi: 10.1091/mbc.E20-02-0092

26. Kaur A, Ecker BL, Douglass SM, Kugel CH, Webster MR, Almeida F, et al. Remodeling of the collagen matrix in aging skin promotes melanoma metastasis and affects immune cell motility. *Cancer Discov* (2019) 9(1):64–81. doi: 10.1158/2159-8290.CD-18-0193

27. Long JE, Wongchenko MJ, Nickles D, Chung W-J, Wang B, Riegler J, et al. Therapeutic resistance and susceptibility is shaped by cooperative multi-compartment tumor adaptation. *Cell Death Different* (2019) 26(11):2416–29. doi: 10.1038/s41418-019-0310-0

28. Spoerri L, Tonnessen-Murray CA, Gunasingh G, Hill DS, Beaumont KA, Jurek RJ, et al. Phenotypic melanoma heterogeneity is regulated through cell-matrix interaction-dependent changes in tumor microarchitecture. *BioRxiv* (2021) 06(09)2020:141747. doi: 10.1101/2020.06.09.141747

29. Dilshat R, Fock V, Kenny C, Gerritsen I, Lasseur RMJ, Travnickova J, et al. Mitf reprograms the extracellular matrix and focal adhesion in melanoma. *ELife* (2021) 10:e63093. doi: 10.7554/elife.63093

30. Jensen C, Madsen DH, Hansen M, Schmidt H, Svane IM, Karsdal MA, et al. Non-invasive biomarkers derived from the extracellular matrix associate with response to immune checkpoint blockade (anti-CTLA-4) in metastatic melanoma patients. *J Immunother Cancer* (2018) 6(1):1–10. doi: 10.1186/s40425-018-0474-z/figures/4

31. Kim MH, Kim J, Hong H, Lee S-H, Lee J-K, Jung E, et al. Actin remodeling confers BRAF inhibitor resistance to melanoma cells through YAP/TAZ activation. *EMBO J* (2016) 35(5):462–78. doi: 10.15252/embj.201592081

32. Lal G, Contreras PG, Kulak M, Woodfield G, Bair T, Domann FE, et al. Human melanoma cells over-express extracellular matrix 1 (ECM1) which is regulated by TFAP2C. *PLoS One* (2013) 8(9):e73953. doi: 10.1371/journal.pone.0073953

33. Miskolczi Z, Smith MP, Rowling EJ, Ferguson J, Barriuso J, Wellbrock C. Collagen abundance controls melanoma phenotypes through lineage-specific microenvironment sensing. *Oncogene* (2018) 37(23):3166. doi: 10.1038/S41388-018-0209-0

34. Riesenberg S, Groetchen A, Siddaway R, Bald T, Reinhardt J, Smorra D, et al. MITF and c-jun antagonism interconnects melanoma dedifferentiation with pro-inflammatory cytokine responsiveness and myeloid cell recruitment. *Nat Commun* (2015) 6:8755. doi: 10.1007/s10911-010-9177-x

35. Radisky ES, Radisky DC. Matrix metalloproteinase-induced epithelial-mesenchymal transition in breast cancer. *J Mammary Gland Biol Neoplas* (2010) 15(2):201–12. doi: 10.1007/s10911-010-9177-x/figures/3

36. Suarez-Carmona M, Lesage J, Cataldo D, Gilles C. EMT and inflammation: inseparable actors of cancer progression. *Mol Oncol* (2017) 11(7):805–23. doi: 10.1002/1878-0261.12095

37. Tripathi SC, Peters HL, Taguchi A, Katayama H, Wang H, Momin A, et al. Immunoproteasome deficiency is a feature of non-small cell lung cancer with a mesenchymal phenotype and is associated with a poor outcome. *Proc Natl Acad Sci USA* (2016) 113(11):E1555–64. doi: 10.1073/pnas.1521812113/suppl_file/pnas.1521812113.sd07.xlsx

38. Jolly MK, Murphy RJ, Bhatia S, Whitfield HJ, Redfern A, Davis MJ, et al. Measuring and modelling the epithelial- mesenchymal hybrid state in cancer: Clinical implications. *Cells Tiss Organ* (2022) 211(2):110–33. doi: 10.1159/000515289

39. Hoek KS, Eichhoff OM, Schlegel NC, Döbbeling U, Kobert N, Schaerer L, et al. *In vivo* switching of human melanoma cells between proliferative and invasive states. *Cancer Res* (2008) 68(3):650–6. doi: 10.1158/0008-5472.can-07-2491

40. Tsoi J, Robert L, Paraiso K, Galvan C, Sheu KM, Lay J, et al. Multi-stage differentiation defines melanoma subtypes with differential vulnerability to drug-induced iron-dependent oxidative stress. *Cancer Cell* (2018) 33(5):890–904. doi: 10.1016/j.ccell.2018.03.017

41. Jolly MK, Tripathi SC, Jia D, Mooney SM, Celiktas M, Hanash SM, et al. Stability of the hybrid epithelial/mesenchymal phenotype. *Oncotarget* (2016) 7 (19):27067–84. doi: 10.18632/oncotarget.8166

42. Pillai M, Jolly MK. Systems-level network modeling deciphers the master regulators of phenotypic plasticity and heterogeneity in melanoma. *IScience* (2021) 24(10):103111. doi: 10.1016/j.isci.2021.103111

43. Bocci F, Tripathi SC, Vilchez Mercedes SA, George JT, Casabar JP, Wong PK, et al. NRF2 activates a partial epithelial-mesenchymal transition and is maximally present in a hybrid epithelial/mesenchymal phenotype. *Integr Biol* (2019) 11(6):251–63. doi: 10.1093/intbio/zyz021

44. Caramel J, Papadogeorgakis E, Hill L, Browne GJ, Richard G, Wierinckx A, et al. A switch in the expression of embryonic EMT-inducers drives the development of malignant melanoma. *Cancer Cell* (2013) 24(4):466–80. doi: 10.1016/j.ccr.2013.08.018

45. Gupta PB, Kuperwasser C, Brunet JP, Ramaswamy S, Kuo WL, Gray JW, et al. The melanocyte differentiation program predisposes to metastasis after neoplastic transformation. *Nat Genet* (2005) 37(10):1047–54. doi: 10.1038/ng1634

46. Jessen C, Kreß JKC, Baluapuri A, Hufnagel A, Schmitz W, Kneitz S, et al. The transcription factor NRF2 enhances melanoma malignancy by blocking differentiation and inducing COX2 expression. *Oncogene* (2020) 39(44):6841. doi: 10.1038/S41388-020-01477-8

47. Perotti V, Baldassari P, Molla A, Nicolini G, Bersani I, Grazia G, et al. An actionable axis linking NFATc2 to EZH2 controls the EMT-like program of melanoma cells. *Oncogene 2019 38:22* (2019) 38(22):4384–96. doi: 10.1038/s41388-019-0729-2

48. Subbalakshmi AR, Kundnani D, Biswas K, Ghosh A, Hanash SM, Tripathi SC, et al. NFATc acts as a non-canonical phenotypic stability factor for a hybrid Epithelial/Mesenchymal phenotype. *Front Oncol* (2020) 10:553342. doi: 10.3389/fonc.2020.553342

49. Subbalakshmi AR, Sahoo S, Biswas K, Jolly MK. A computational systems biology approach identifies SLUG as a mediator of partial epithelial-mesenchymal transition (EMT). *Cells Tiss Organ* (2022) 211(6):105–18. doi: 10.1159/000512520

50. Vandewalle C, Comijn J, de Craene B, Vermassen P, Bruyneel E, Andersen H, et al. SIP1/ZEB2 induces EMT by repressing genes of different epithelial cell–cell junctions. *Nucleic Acids Res* (2005) 33(20):6566–78. doi: 10.1093/nar/gki965

51. Subbalakshmi AR, Sahoo S, McMullen I, Saxena AN, Venugopal SK, Somarelli JA, et al. KLF4 induces mesenchymal–epithelial transition (MET) by suppressing multiple EMT-inducing transcription factors. *Cancers* (2021) 13 (20):5135. doi: 10.3390/cancers13205135

52. Zhang D, Lin J, Chao Y, Zhang L, Jin L, Li N, et al. Regulation of the adaptation to ER stress by KLF4 facilitates melanoma cell metastasis *via* upregulating NUCB2 expression. *J Exp Clin Cancer Res* (2018) 37(1):1–14. doi: 10.1186/s13046-018-0842-z

53. Campbell NR, Rao A, Hunter M, Sznurkowska MK, Briker L, Zhang M, et al. Cooperation between melanoma cell states promotes metastasis through heterotypic cluster formation. *Dev Cell* (2021) 56(20):2808–25.e10. doi: 10.1016/j.devcel.2021.08.018

54. Dimitrova Y, Gruber AJ, Mittal N, Ghosh S, Dimitriades B, Mathow D, et al. TFAP2A is a component of the ZEB1/2 network that regulates TGFB1-induced epithelial to mesenchymal transition. *Biol Direct* (2017) 12(1):8. doi: 10.1186/s13062-017-0180-7/figures/5

55. Byers LA, Diao L, Wang J, Saintigny P, Girard L, Peyton M, et al. An epithelial-mesenchymal transition gene signature predicts resistance to EGFR and PI3K inhibitors and identifies axl as a therapeutic target for overcoming EGFR inhibitor resistance. *Clin Cancer Res* (2013) 19(1):279–90. doi: 10.1158/1078-0432.ccr-12-1558

56. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* (2005) 102(43):15545–550. doi: 10.1073/pnas.0506580102

57. Tan TZ, Miow QH, Miki Y, Noda T, Mori S, Huang RY, et al. Epithelial-mesenchymal transition spectrum quantification and its efficacy in deciphering survival and drug responses of cancer patients. *EMBO Mol Med* (2014) 6(10):1279–93. doi: 10.15252/emmm.201404208

58. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin A, Kim S, et al. The cancer cell line encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* (2012) 483(7391):603–7. doi: 10.1038/nature11003

59. Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn IF, et al. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nat* (2009) 462(7269):108–12. doi: 10.1038/nature08460

60. Verfaillie A, Imrichova H, Atak ZK, Dewaele M, Rambow F, Hulselmans G, et al. Decoding the regulatory landscape of melanoma reveals TEADS as regulators of the invasive cell state. *Nat Commun* (2015) 6:6683. doi: 10.1038/ncomms7683

61. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The molecular signatures database hallmark gene set collection. *Cell Syst* (2015) 1 (6):417–25. doi: 10.1016/j.cels.2015.12.004

62. Muralidharan S, Sahoo S, Saha A, Chandran S, Majumdar SS, Mandal S, et al. Quantifying the patterns of metabolic plasticity and heterogeneity along the epithelial–Hybrid–Mesenchymal spectrum in cancer. *Biomolecules* (2022) 12 (2):297. doi: 10.3390/biom12020297

63. Jia D, Paudel BB, Hayford CE, Hardeman KN, Levine H, Onuchic JN, et al. Drug-tolerant idling melanoma cells exhibit theory-predicted metabolic low-low phenotype. *Front Oncol* (2020) 10:1426. doi: 10.3389/fonc.2020.01426

64. Yu L, Lu M, Jia D, Ma J, Ben-Jacob E, Levine H, et al. Modeling the genetic regulation of cancer metabolism: Interplay between glycolysis and oxidative phosphorylation. *Cancer Res* (2017) 77(7):1564–74. doi: 10.1158/0008-5472.can-16-2074/652665/am/modeling-the-genetic-regulation-of-cancer

65. Foroutan M, Bhuva DD, Lyu R, Horan K, Cursons J, Davis MJ. Single sample scoring of molecular phenotypes. *BMC Bioinf* (2018) 19(1):1–10. doi: 10.1186/s12859-018-2435-4/figures/2

66. Taylor R. Interpretation of the correlation coefficient: A basic review. *J Diagn Med Sonogr* (1990) 6(1):35–9. doi: 10.1177/875647939000600106

67. Tang Y, Durand S, Dalle S, Caramel J. EMT-inducing transcription factors, drivers of melanoma phenotype switching, and resistance to treatment. *Cancers* (2020) 12(8):2154. doi: 10.3390/cancers12082154

68. Wessely A, Steeb T, Berking C, Heppt MV. How neural crest transcription factors contribute to melanoma heterogeneity, cellular plasticity, and treatment resistance. *Int J Mol Sci* (2021) 22(11):5761. doi: 10.3390/ijms22115761

69. Abildgaard C, Guldberg P. Molecular drivers of cellular metabolic reprogramming in melanoma. *Trends Mol Med* (2015) 21(3):164–71. doi: 10.1016/j.molmed.2014.12.007

70. Bettum IJ, Gorad SS, Barkovskaya A, Pettersen S, Moestue SA, Vasiliauskaite K, et al. Metabolic reprogramming supports the invasive phenotype in malignant melanoma. *Cancer Lett* (2015) 366(1):71–83. doi: 10.1016/j.canlet.2015.06.006

71. Gelato KA, Schöckel L, Klingbeil O, Rückert T, Lesche R, Toedling J, et al. Super-enhancers define a proliferative PGC-1α-expressing melanoma subgroup sensitive to BET inhibition. *Oncogene 2018 37:4* (2017) 37(4):512–21. doi: 10.1038/onc.2017.325

72. Laurenzana A, Chillà A, Luciani C, Peppicelli S, Biagioni A, Bianchini F, et al. uPA/uPAR system activation drives a glycolytic phenotype in melanoma cells. *Int J Cancer* (2017) 141(6):1190–200. doi: 10.1002/ijc.30817

73. Haq R, Shoag J, Andreu-Perez P, Yokoyama S, Edelman H, Rowe GC, et al. Oncogenic BRAF regulates oxidative metabolism *via* PGC1α and MITF. *Cancer Cell* (2013) 23(3):302–15. doi: 10.1016/j.ccr.2013.02.003

74. Vazquez F, Lim JH, Chim H, Bhalla K, Girnun G, Pierce K, et al. PGC1α expression defines a subset of human melanoma tumors with increased mitochondrial capacity and resistance to oxidative stress. *Cancer Cell* (2013) 23 (3):287–301. doi: 10.1016/j.ccr.2012.11.020

75. Vivas-García Y, Falletta P, Liebing J, Louphrasitthiphol P, Feng Y, Chauhan J, et al. Lineage-restricted regulation of SCD and fatty acid saturation by MITF controls melanoma phenotypic plasticity. *Mol Cell* (2020) 77(1):120–137.e9. doi: 10.1016/j.molcel.2019.10.014

76. Klinke DJ, Torang A. An unsupervised strategy for identifying epithelial-mesenchymal transition state metrics in breast cancer and melanoma. *IScience* (2020) 23(5):101080. doi: 10.1016/j.isci.2020.101080

77. Cook DP, Vanderhyden BC. Context specificity of the EMT transcriptional response. *Nat Commun* (2020) 11(1):2142. doi: 10.1038/s41467-020-16066-2

78. Sahoo S, Ashraf B, Duddu AS, Biddle A, Jolly MK. Interconnected high-dimensional landscapes of epithelial–mesenchymal plasticity and stemness in cancer. *Clin Exp Metast* (2022) 1:1–12. doi: 10.1007/S10585-021-10139-2

79. Norgard RJ, Pitarresi JR, Maddipati R, Aiello-Couzo NM, Balli D, Li J, et al. Calcium signaling induces a partial EMT. *EMBO Rep* (2021) 22(9):e51872. doi: 10.15252/embr.202051872

80. Rambow F, Marine JC, Goding CR. Melanoma plasticity and phenotypic diversity: Therapeutic barriers and opportunities. *Genes Dev* (2019) 33(19–20):1295–1318. doi: 10.1101/gad.329771.119

81. Youssef KK, Nieto MA. Glucose metabolism takes center stage in epithelial-mesenchymal plasticity. *Dev Cell* (2020) 53(2):133–5. doi: 10.1016/J.Devcel.2020.03.021

82. Zheng X, Boyer L, Jin M, Mertens J, Kim Y, Ma L, et al. Metabolic reprogramming during neuronal differentiation from aerobic glycolysis to neuronal oxidative phosphorylation. *ELife* (2016) 5(JUN2016):e13374. doi: 10.7554/elife.13374

83. Vidács DL, Veréb Z, Bozó R, Flink LB, Polyánka H, Németh IB, et al. Phenotypic plasticity of melanocytes derived from human adult skin. *Pigment Cell & Melanoma Res* (2022) 35(1):38–51. doi: 10.1111/pcmr.13012

84. Fedele M, Cerchia L, Pegoraro S, Sgarra R, Manfioletti G. Proneural-mesenchymal transition: Phenotypic plasticity to acquire multitherapy resistance in glioblastoma. *Int J Mol Sci* (2019) 20(11):2746. doi: 10.3390/ijms20112746

85. Couturier CP, Ayyadhury S, Le PU, Nadaf J, Monlong J, Riva G, et al. Single-cell RNA-seq reveals that glioblastoma recapitulates a normal neurodevelopmental hierarchy. *Nat Commun* (2020) 11(1):3406. doi: 10.1038/s41467-020-17186-5

86. Phillips HS, Kharbanda S, Chen R, Forrest WF, Soriano RH, Wu TD, et al. Molecular subclasses of high-grade glioma predict prognosis, delineate a pattern of disease progression, and resemble stages in neurogenesis. *Cancer Cell* (2006) 9 (3):157–73. doi: 10.1016/j.ccr.2006.02.019

87. Wang L, Babikir H, Müller S, Yagnik G, Shamardani K, Catalan F, et al. The phenotypes of proliferating glioblastoma cells reside on a single axis of variation. *Cancer Discov* (2019) 9(12):1708–19. doi: 10.1158/2159-8290.CD-19-0329

88. Liu S, Cong Y, Wang D, Sun Y, Deng L, Liu Y, et al. Breast cancer stem cells transition between epithelial and mesenchymal states reflective of their normal counterparts. *Stem Cell Rep* (2013) 2(1):78–91. doi: 10.1016/j.stemcr.2013.11.009

89. Camand E, Peglion F, Osmani N, Sanson M, Etienne-Manneville S. N-cadherin expression level modulates integrin-mediated polarity and strongly impacts on the speed and directionality of glial cell migration. *J Cell Sci* (2012) 125(4):844–57. doi: 10.1242/jcs.087668

90. Schaeffer D, Somarelli JA, Hanna G, Palmer GM, Garcia-Blanco MA. Cellular migration and invasion uncoupled: Increased migration is not an inexorable consequence of epithelial-to-Mesenchymal transition. *Mol Cell Biol* (2014) 34(18):3486–99. doi: 10.1128/mcb.00694-14/suppl_file/zmb999100574so2.pdf

91. Goyal Y, Dardani IP, Busch GT, Emert B, Fingerman D, Kaur A, et al. Pre-determined diversity in resistant fates emerges from homogenous cells after anti-cancer drug treatment. *BioRxiv* (2021) 2021:471833. doi: 10.1101/2021.12.08.471833

92. Su Y, Ko ME, Cheng H, Zhu R, Xue M, Wang J, et al. Multi-omic single-cell snapshots reveal multiple independent trajectories to drug tolerance in a melanoma cell line. *Nat Commun* (2020) 11(1):2345. doi: 10.1038/s41467-020-15956-9

93. Kim YJ, Sheu KM, Tsoi J, Abril-Rodriguez G, Medina E, Grasso CS, et al. Melanoma dedifferentiation induced by IFN-γ epigenetic remodeling in response to anti–PD-1 therapy. *J Clin Invest* (2021) 131(12). doi: 10.1172/jci145859

94. Kavran AJ, Stuart SA, Hayashi KR, Basken JM, Brandhuber BJ, Ahn NG. Intermittent treatment of BRAF V600E melanoma cells delays resistance by adaptive resensitization to drug rechallenge. *Proc Natl Acad Sci* (2022) 119(12): e2113535119. doi: 10.1073/PNAS.2113535119

95. Goldman A, Majumder B, Dhawan A, Ravi S, Goldman D, Kohandel M, et al. Temporally sequenced anticancer drugs overcome adaptive resistance by targeting a vulnerable chemotherapy-induced phenotypic transition. *Nat Commun* (2015) 6(1):2015. doi: 10.1038/ncomms7139

96. Grosse-Wilde A, D'Hérouël AF, McIntosh E, Ertaylan G, Skupin A, Kuestner RE, et al. Stemness of the hybrid Epithelial/Mesenchymal state in breast cancer and its association with poor survival. *PLoS One* (2015) 10(5): e0126522. doi: 10.1371/journal.pone.0126522

97. Kröger C, Afeyan A, Mraz J, Eaton EN, Reinhardt F, Khodor YL, et al. Acquisition of a hybrid E/M state is essential for tumorigenicity of basal breast cancer cells. *Proc Natl Acad Sci USA* (2019) 116(15):7353–62. doi: 10.1073/pnas.1812876116

98. Pasani S, Sahoo S, Jolly MK. Hybrid E/M phenotype(s) and stemness: A mechanistic connection embedded in network topology. *J Clin Med* (2021) 10 (1):60. doi: 10.3390/jcm10010060

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 30 September 2022
DOI 10.3389/fonc.2022.954512

OPEN ACCESS

EDITED BY
Dongya Jia,
National Cancer Institute, (NIH), United States

REVIEWED BY
Valentina Audrito,
Università del Piemonte Orientale, Italy
Xuefei Li,
Institute of Synthetic Biology, (CAS), China
Aleš Dvořák,
Charles University, Czechia

*CORRESPONDENCE
Nagasuma Chandra
nchandra@iisc.ac.in

†PRESENT ADDRESS
Madhulika Mishra,
European Molecular Biology Laboratory, European Bioinformatics Institutes, The Wellcome Trust Genome Campus, Cambridge SD, United Kingdom
Omkar Kulkarni,
Department of Biological Sciences, National University of Singapore, Science Drive 4, Singapore, Singapore

SPECIALTY SECTION
This article was submitted to Cancer Metabolism, a section of the journal Frontiers in Oncology

RECEIVED 27 May 2022
ACCEPTED 15 August 2022
PUBLISHED 30 September 2022

CITATION
Chedere A, Mishra M, Kulkarni O, Sriraman S and Chandra N (2022) Personalized quantitative models of NAD metabolism in hepatocellular carcinoma identify a subgroup with poor prognosis.
Front. Oncol. 12:954512.
doi: 10.3389/fonc.2022.954512

Personalized quantitative models of NAD metabolism in hepatocellular carcinoma identify a subgroup with poor prognosis

Adithya Chedere[1], Madhulika Mishra[1†], Omkar Kulkarni[1†], Shrisruti Sriraman[2] and Nagasuma Chandra[1,2*]

[1]Department of Biochemistry, Biological Science Division, Indian Institute of Science, Bengaluru, Karnataka, India, [2]IISc Mathematics Initiative, Indian Institute of Science, Bengaluru, Karnataka, India

Cancer cells are known to undergo metabolic adaptation to cater to their enhanced energy demand. Nicotinamide adenine dinucleotide (NAD) is an essential metabolite regulating many cellular processes within the cell. The enzymes required for NAD synthesis, starting from the base precursor - tryptophan, are expressed in the liver and the kidney, while all other tissues convert NAD from intermediate precursors. The liver, being an active metabolic organ, is a primary contributor to NAD biosynthesis. Inhibition of key enzymes in the NAD biosynthetic pathways is proposed as a strategy for designing anti-cancer drugs. On the other hand, NAD supplementation has also been reported to be beneficial in cancer in some cases. As metabolic adaptation that occurs in cancer cells can lead to perturbations to the pathways, it is important to understand the exact nature of the perturbation in each individual patient. To investigate this, we use a mathematical modelling approach integrated with transcriptomes of patient samples from the TCGA-LIHC cohort. Quantitative profiling of the NAD biosynthesis pathway helps us understand the NAD biosynthetic status and changes in the controlling steps of the pathway. Our results indicate that NAD biosynthesis is heterogeneous among liver cancer patients, and that Nicotinate phosphoribosyl transferase (NAPRT) levels are indicative of the NAD biosynthetic status. Further, we find that reduced NAPRT levels combined with reduced Nicotinamide phosphoribosyl transferase (NAMPT) levels contribute to poor prognosis. Identification of the precise subgroup who may benefit from NAD supplementation in subgroup with low levels of NAPRT and NAMPT could be explored to improve patient outcome.

KEYWORDS

NAD metabolism, pathway model, NAPRT, NAMPT, liver cancer, precision medicine, patient subtyping

Introduction

Nicotinamide adenine dinucleotide (NAD) is an essential cofactor for the cell. It mediates various biological processes such as energy metabolism, DNA repair, signalling, and gene-expression regulation. NAD regulates energy metabolism pathways, including glycolysis, fatty acid oxidation (β-oxidation), the tricarboxylic acid (TCA) cycle, and oxidative phosphorylation (1). NAD exists in both oxidised (NAD^+) as well as in reduced (NADH) forms; the total concentration of NAD^+ and NADH is considered as the NAD pool in the cell (2). The NAD^+/NADH ratio maintains the redox potential of the cell and thereby acts as a metabolic regulator of various NAD-dependent reactions (3). This includes more than 600 metabolic reactions as well as some involved in the signalling. The utilisation of NAD at the global level in the cell makes it an indispensable currency metabolite for the cell (1). Three routes that lead to NAD biosynthesis are well characterised, the first route from tryptophan as a precursor (kynurenine pathway) (Figure 1A and Table 1: reactions J1-J10), the second from nicotinic acid (Preiss-Handler pathway) (Figure 1A and Table 1: reaction J22), and the third, a salvage pathway that utilizes several alternative precursors (4–6) (Figure 1A and Table 1: reactions J8, J17-J20). All the known genes involved in the NAD biosynthesis are expressed in the liver (7–9). In particular, hepatocytes can utilise all precursors from vitamin B3 and from tryptophan to NAD^+, indicating that the precursors and the synthesis of NAD are high in the liver. The liver also serves as a reservoir of the NAD pool by converting NAD precursors from nutrient sources to nicotinamide (Nam) that can be released into the bloodstream when required (8, 9). Thus, it can be said that the liver regulates the overall physiological requirement of this essential energy currency. NAD does not get degraded in metabolic processes but only interconverts between oxidised NAD^+ form to the reduced NADH form. On the other hand, processes related to DNA repair, MAPK signalling, Ca^{+2} signalling, and gene expression utilise NAD and degrade it to Nam, which can be later converted back to NAD (7, 10). Hence, any imbalance in the NAD pool will lead to global perturbations in the cell and are known to be associated with various disease conditions such as ageing, inflammation, and cancer.

In cancer, three routes of NAD utilisation are known to be perturbed and are associated with carcinogenesis (7). The first route of NAD utilisation is the NAD mediated central carbon metabolism which is highly altered in cancer (Warburg effect). Reduced values of the NAD^+/NADH ratio lead to activation of HIF1α through oxidative stress, which in turn activates transcriptional expression of glucose metabolism. Interestingly, not only the central carbon metabolism but other NAD utilising metabolic processes such as fatty acid oxidation, bile acid synthesis, glycerophospholipid metabolism, amino sugar metabolism, etc., are also known to be altered in cancer, especially in hepatocellular carcinoma (HCC) (1, 11). The second utilisation route of the NAD pool is the phosphorylation of NAD^+ to $NADP^+$ by the NAD kinase enzyme (Figure 1A and Table 1: reaction J11). $NADP^+$ also has a variety of cellular functions which are associated with carcinogenesis, such as the reactive oxygen species (ROS) defence, detoxification, and oxidative burst in an immune response. The third and the most important utilisation route of NAD is the NAD degrading ADP-ribose (ADPR) transfer reactions (Figure 1A and Table 1: reaction J13-J15). NAD acts as a co-substrate for three families of proteins namely Sirtuins, PARPs and cADPR synthases (5). These reactions are critical for CD38 signalling, P53, FOXO, MAPK-dependent growth signalling, mono-ADP-ribosylation reactions, and many more (7). Most of the above-mentioned processes are not only altered in cancer but also have a carcinogenic role in tumour progression. The preferred route of NAD synthesis and utilization is tissue-dependent, and is epigenetically encoded in the cells (5, 10–12). Extracellular NAD pools, partially produced by extracellular NAMPT and NAPRT, aid in inflammation and immune suppression further enhancing the tumour progression (13). The enzymes and metabolites involved in *de novo* pathway are known to be associated with inflammation and immune response (14, 15). Gut bacteria also help in maintaining the NAD pool in the body by producing NAD pathway intermediates, like NaAD, which can be directly converted into NAD (refer to Figure 1A and Table 1: reactions J9, J10), especially in the liver as well as other organs (16). Therefore, synthesis and NAD utilising reactions, in particular the NAMPT, Sirtuins, and PARPs, have been explored as potential drug targets in the last two decades (17–21). However, many drugs targeting NAD and associated pathways have failed in clinical trials for various reasons such as drug toxicity, patient heterogeneity, and alternate routes of signalling (7, 22).

On the other hand, reports from various epidemiologic studies suggest an association between low NAD precursor diets with an increased rate of cancer incidence (23). NAD

Abbreviations: ADPR, ADP-ribose; cADPR, Cyclic ADP Ribose; COPASI, COmplex PAthway Simulator; CoRC, COPASI R Connector; DNA, Deoxyribonucleic acid; HCC, hepatocellular carcinoma; HR, Hazards ratio value; log_2FC, log_2 of Fold Change value; Na, Nicotinic acid; NaAD, Nicotinic acid adenine dinucleotide; NAD, Nicotinamide adenine dinucleotide; NAD_{net}, quantitative kinetic model of the NAD biosynthetic pathway; NADP, Nicotinamide adenine dinucleotide phosphate; Nam, Nicotinamide; NaMN, Nicotinic acid adenine mononucleotide; NAMPT, Nicotinamide phosphoribosyl transferase; NAPRT, Nicotinate phosphoribosyl transferase; NAR, Nicotinic acid riboside; NL, normal liver; NMN, Nicotinamide mononucleotide; NMNAT, NMN adenylyltransferase; NR, Nicotinamide riboside; ODE, Ordinary differential equations; PRPP, Phosphoribosyl diphosphate; PARP, Poly (ADP-ribose) polymerase; PaxDb, Protein Abundance Database; RNASeq, RNA sequencing; ROS, reactive oxygen species; TCA, tricarboxylic acid cycle; TCGA-LIHC, The Cancer Genome Atlas-Liver Hepatocellular Carcinoma.

10.3389/fonc.2022.954512

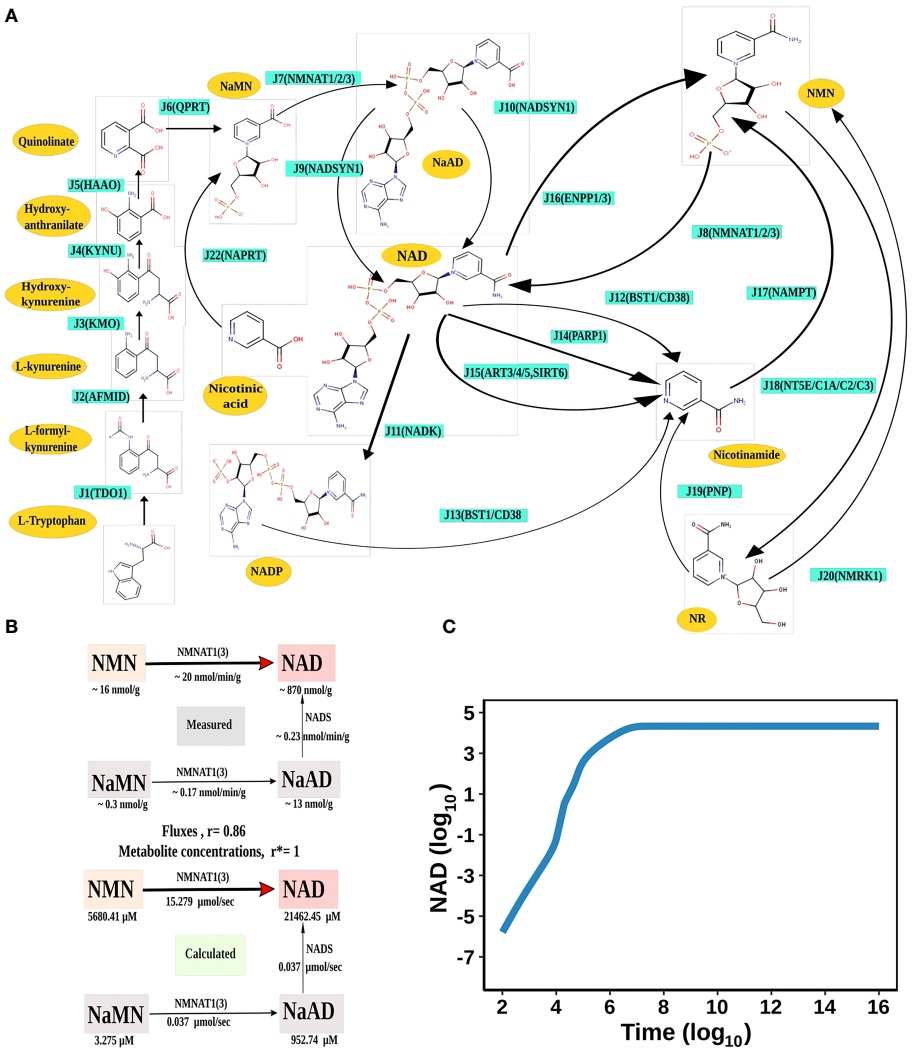


FIGURE 1
NAD Biosynthesis is perturbed in HCC. **(A)** Diagrammatic representation of the liver NAD_{net}. Metabolites are represented as yellow nodes; enzymes are in cyan colour. NAD^+ can be synthesised by the three different routes - (A) Route I: *de novo* biosynthesis pathway starting with the precursor tryptophan (B) Route II: Preiss–Handler pathway from nicotinic acid and (C) Route III: NAD salvage pathway utilising nicotinamide and nicotinamide riboside for NAD^+ biosynthesis. Note: NAD represents total NAD in the system (both oxidised and reduced form). **(B)** Validation of the liver NAD_{net}. The above panel (Measured) is adopted from the report of Mori *et al.* and represents the metabolic reconstruction of NAD biosynthesis in mouse liver tissue and reflects the main route of NAD generation is *via* NMN. The below panel (Calculated) is the reconstruction of NAD biosynthesis from steady-state concentrations and fluxes obtained from model simulation. r is the Spearman correlation between measured and calculated fluxes. r* is the Spearman correlation between measured and calculated metabolite concentrations. **(C)** Time course simulation of NAD levels for the base model. The blue curve shows the changes in NAD concentration (µM) (log_{10} scale) with respect to time(s) (log_{10} scale).

levels decrease with ageing, thus forming an association with diseases related to ageing, such as neurodegenerative diseases, cardiovascular diseases, bone dysfunctions, and cancer (24–26). Studies using mice models of cancer and ageing also exhibit low NAD levels and therefore are more prone to oxidative stress. Current knowledge about this suggests that a low NAD level can lead to oxidative stress-induced DNA damage and thereby promote mutagenesis and tumour initiation (10, 27). Thus, a high level of NAD can have a preventive role in tumorigenesis. Another recent study Tummala et al. reported that an increased expression of unconventional prefoldin RPB5 interactor (URI) leads to AhR- and/or ER-mediated reduction of the NAD pool and thereby promotes HCC tumorigenesis due to increased DNA damage (28). This seemed to suggest that NAD supplementation can have a protective role against HCC development and progression in cirrhotic patients (29, 30).

TABLE 1 NAD_{net} pathway reactions to gene, enzyme, and Factor mappings information.

No.	Reaction ID	Gene	Enzyme name	EC	Reaction	Factor
1	J1	TDO2, IDO2, IDO1	Tryptophan 2,3-dioxygenase	1.13.11.11	L-Tryptophan + O_2 = N-Formyl-L-kynurenine	F1
2	J2	AFMID	Formylkynurenine formamidase	3.5.1.9	N-Formyl-L-kynurenine + H_2O = Formate + L-Kynurenine	F2
3	J3	KMO	Kynurenine 3-hydroxylase	1.14.13.9	L-Kynurenine + NADPH + H^+ + O_2 = 3-Hydroxy-L-kynurenine + $NADP^+$ + H_2O	F3
4	J4	KYNU	Kynureninase	3.7.1.3	3-Hydroxy-L-kynurenine + H_2O = 3-Hydroxyanthranilate + L-Alanine	F4
5	J5	HAAO	3-Hydroxyanthranilate 3,4-dioxygenase	1.13.11.6	3-Hydroxyanthranilate + O_2 = 2-Amino-3-carboxymuconate semialdehyde → Quinolinate	F5
6	J6	QPRT	Quinolinate phosphoribosyltransferase	2.4.2.19	Quinolinate + PRPP = NMN + PPi + CO_2	F6
7	J7	NMNAT1, NMNAT2, NMNAT3	NMN adenylyltransferase	2.7.7.1	NaMN + ATP = $NaAD^+$ + PPi	F7
8	J8	NMNAT1, NMNAT2, NMNAT3	NMN adenylyltransferase	2.7.7.1	NMN + ATP = NAD^+ + PPi	F8
9	J9	NADSYN1	NAD^+ synthetase (glutamine-hydrolyzing)	6.3.5.1	$NaAD^+$ + ATP + L-Gln + H_2O = NAD^+ + AMP + PPi + L-Glu	F9
10	J10	NADSYN1	NAD^+ synthetase (ammonia-dependent)	6.3.1.5	$NaAD^+$ + ATP + NH_3 = NAD^+ + AMP + PPi	F10
11	J11	NADK	NAD^+ kinase	2.7.1.23	NAD^+ + ATP = $NADP^+$ + ADP	F11
12	J12	BST1, CD38	NAD^+ glycohydrolase	3.2.2.5	NAD^+ + H2O = Nicotinamide + ADP-ribose	F12
13	J13	BST1, CD38	$NAD(P)^+$ nucleosidase	3.2.2.6	$NAD(P)^+$+ H2O = Nicotinamide + ADP-ribose(2¢-phosphate)	F13
14	J14	PARP1	Poly (ADP-ribose) polymerase	2.4.2.30	NAD^+ + $(ADP\text{-}ribose)_n$ = Nicotinamide + $(ADP\text{-}ribose)_{n+1}$	F14
15	J15	ART3, ART4, ART5, SIRT6	Mono ADPribosyltransferase	2.4.2.31	NAD^++ L-Arg = Nicotinamide + N(2)-(ADP-ribosyl)-L-Arg	F15
16	J16	ENPP1, ENPP3	NAD^+ pyrophosphatase	3.6.1.22	NAD^+ + H_2O = AMP + NMN	F16
17	J17	NAMPT	Nicotinamide phosphoribosyltransferase	2.4.2.12	Nicotinamide + PRPP = NMN + PPi	F17
18	J18	NT5E, NT5C1A, NT5C2	5'-Nucleotidase	3.1.3.5	NMN + H_2O = Nicotinamide riboside + P	F18
19	J19	PNP	Nicotinamide nucleoside phosphorylase	2.4.2.1	Nicotinamide riboside + P = Nicotinamide + R-1-P	F19
20	J20	NMRK1	Ribosylnicotinamide kinase	2.7.1.22	ATP + Nicotinamide riboside = ADP + NMN	F20
21	J22	NAPRT	Nicotinate phosphoribosyltransferase	2.4.2.11	Nicotinate + PRPP = NaMN + PPi	F22

However, while on the one hand, an increase in NAD related activities are linked to metabolic and signalling alterations in cancers, leading to the hypothesis that the pathway is an attractive drug target for tumour killing. On the other hand, NAD protects cells from DNA damage and is found to be downregulated in many cancers including HCC, hence suggesting that supplementation of NAD can stop carcinogenesis. These seemingly opposite findings of NAD imbalance in HCC have led to a controversy about the role of NAD in tumorigenesis. Therefore, it is important to first address the role of NAD in tumorigenesis in individual HCC patients and understand whether it acts as a tumour suppressor or a tumour promoter.

The main objective of this study is to decipher the role of NAD in HCC and to understand whether NAD biosynthesis inhibition or alternately NAD supplementation will be beneficial in treating HCC. To address this objective, it is important to understand the NAD pathway profile in HCC and whether there is any heterogeneity among HCC patients. Pathway modelling offers a useful method to profile the individual enzymes in the pathway and decide whether it is altered in a given individual as compared to the normal liver (NL) and, if so, in which direction. Knowledge of the NAD pathway profile in individual HCC patients will enable precision targeting and ultimately aid the clinician in decision-making for the management of HCC. To answer these questions, in this work, a quantitative kinetic model of the NAD biosynthetic pathway (NAD_{net}) is constructed and simulated for each individual patient by integrating patient-specific transcriptomics data available through TCGA-LIHC.

Materials and methods

Model building of NAD_{net}

The base structure of the network

The liver NAD_{net}, a liver-specific NAD biosynthesis network, was reconstructed in the laboratory using the previously published NAD model on glioma from our laboratory (31). The model comprehensively captures known reactions in the NAD biosynthesis, covering the *de novo* pyridine ring formation *via* the kynurenine pathway, the utilisation of dietary precursor nicotinate through the Preiss–Handler pathway, and the utilisation of nicotinamide and nicotinamide riboside through the NAD salvage pathway for NAD^+ synthesis in cancer. A base model was first constructed by considering reactions that can occur in any human tissue, which was subsequently curated to check for the feasibility of each reaction in the human liver. It was observed that out of 24 enzymatic reactions from the base model, only 21 enzymatic reactions were feasible in the liver and were therefore retained. Other than the enzymatic reactions, four non-enzymatic sink reactions were added to the model for the model stability. A list of enzymatic reactions in the model is provided in Table 1. Kinetic parameters for each of the enzymes in the model were manually curated from the primary literature. K_{cat}, K_m, and K_i for inhibitory interactions were also obtained from the same primary sources, and wherever possible, the parameters for the liver tissue were used. A full list of parameters is provided in Supplementary Table 1. A list of fixed metabolites for simulation purposes is provided in Supplementary Table 2.

Estimation of F_{Kcat} and F_{conc}

One of the most challenging tasks in kinetic modelling is to deal with various types of inconsistencies in units of reported parameters. In order to get all parameters in a comparable framework, a factor (F_{Kcat}) was calculated so as to represent K_{cat} in 1/sec units for all enzymes. The detailed calculation of (F_{Kcat}) factor for each type of specific activity is provided in Supplementary Table 3. For each enzyme in the model, enzyme concentration was estimated from the PaxDB database (32). Then, using these estimated K_{cat} values and enzyme concentrations, V_{max} for each reaction was calculated. The estimated V_{max} values are listed in Supplementary Table 3.

Transcriptome data: TCGA dataset

The Cancer Genome Atlas (TCGA) Liver Hepatocellular Carcinoma (LIHC) RNASeq HT-Seq gene expression (counts data) and phenotype data were collected through UCSC Xena (http://xena.ucsc.edu) (TCGA-LIHC cohort) (33, 34). The dataset contains 374 HCC tissue biopsy samples, out of which three are samples from recurrent HCC samples, and the rest are from primary HCC. The dataset included 50 normal liver (NL) tissue biopsy samples as well. For this analysis, we have considered only primary HCC and Normal liver samples. RNASeq counts data was normalised using edgeR package (V 3.34.1) (35). The mean of normalised gene count was calculated for all normal samples and was used as a control to calculate the fold change of each gene for each tumour sample. All ensembl ids were mapped to gene symbols using org.Hs.eg.db package in R Language (V 3.15.0) (36).

Integration of gene expression data into the model

Fold change values of gene expression of each enzyme were integrated into the model as described previously (37). The correlation between RNA and protein is ~ 0.5, indicating that the transcript levels of genes and the corresponding proteins follow the same trend in their concentrations, justifying the use of RNA levels an indicator for the protein levels (38). For reactions catalysed by multiple genes, the cumulative sum of fold change values in the expression of all associated genes was used. The F-factor for each reaction across all patients was calculated using the mean expression profile of Normal Liver (39). The F-factors differ between the NL model and any patient model. For the NL model, F-factors are all equal to one. For patient models, F-factors are substituted as the cumulative sum of FC values of genes involved in the corresponding reaction. The changes in F-factor values influence the reaction rates. For a given reaction, if the F-factor value is greater than one, the reaction rate is increased by the F-factor value times as compared to the NL model; similarly, if the F-factor value is less than one, the reaction rate is decreased by the F-factor value times as compared to the NL model.

Mutation analysis

Pre-processed mutation data for each cancer type was obtained from the cBioPortal resource (40). From this, mutation frequencies of genes from the NAD_{net} were retrieved.

ODE simulation

Ordinary differential equations (ODE) of the reconstructed models were solved to obtain steady-state values using the getSS function, with the modified option of resolution to 1E-03 and the maximum duration for forward integration to 1E+20, in the CoRC (V 0.11.0) package (41) in R (V 4.1.3) (42). Steady-state metabolite concentrations and reaction flux values were analysed.

Sensitivity analysis

Sensitivity analysis provides a measure of how much a selected model variable (the effect) changes when a selected parameter (the cause) is changed. Sensitivity was calculated for the perturbation effect of individual parameters on the steady-state concentration of NAD. Therefore, it can identify parameters having an effect on NAD concentration. For models which gave results for steady-state analysis, parameter sensitivity analysis was performed. In the current model, there are 114 parameters, so each parameter for a given simulation was only varied by 1 % from the original value, thus resulting in a total of 229 simulations for each model, i.e., one unaltered parameter simulation and 2*114 single parameters altered either by +1 % or -1 % of the original parameter value. All the altered parameter simulations were scaled by taking the percentage change compared to the unaltered parameter simulation. Results were summarised separately for concentrations of metabolites and flux of reactions in the form of a 2D plot with a colour scale representing the percentage change in the simulation value, using the corrplot (V 0.92) package in R (43). The red colour represents an increase in the concentration of the metabolite as compared to the unaltered parameter simulation, whereas the blue colour represents a decrease in concentration of the metabolite. The extent of colour filled in the squares represents the extent of percentage change in the metabolite due to the change in the parameter value.

For a summary of all patient model changes, the percentage change in the altered parameter simulation is calculated and represented as a pie chart of the percentage of models with alterations among all the steady-state models.

Correlation analysis

The correlation analysis was performed between NAD genes and NAD metabolites using Pearson correlation (log_2FC values - for numeric variables) in cor.test function from the stats package in R. Corrplot function from the corrplot (V 0.92) package was used to represent the correlogram (43).

Clustering and heatmap

Hierarchical clustering of gene expression data and metabolite steady-state values with the patient profiles was carried out using the Heatmap function in ComplexHeatmap (V 2.8.0) package in R (44). Gene expression data and calculated metabolite fold changes were categorized into three groups; Up ($log_2FC \geq +1$), No_change ($-1 < log_2FC < +1$) and Down ($log_2FC \leq -1$); and substituted with an integer value before Hierarchical clustering; Up (+1), No_change (0) and Down (-1).

Survival analysis

Survival analysis and univariate Cox regression analysis were performed using the survival (V 3.3-1) package in R (45, 46). Genes and metabolites for each patient were classified as upregulated ($log_2FC \geq +1$) or downregulated ($log_2FC \leq -1$). Hazards ratio values (HR) were obtained using the coxph function from the survival package. HR in survival analysis is the hazard ratio which essentially is the ratio of the hazard rates corresponding to the conditions described by two levels of gene expression. If the gene has a value HR > +1, the given gene is a poor prognostic marker (over-expression of the gene is associated with high mortality of the patients) and vice-versa. The survdiff function from the survival package was used to identify the significant genes/metabolites (p-value < 0.05) associated with patient survival. Survival plots were generated using the ggsurvplot function from the survminer (V 0.4.9) package (47). The survival analysis was performed among the distinct groups.

Results

NAD biosynthesis is perturbed in HCC

NAD biosynthesis network in the human liver

The first objective was to reconstruct a NAD biosynthesis network that captures the physiological processes in the human liver tissue. The liver NAD_{net} model consisting of 26 reactions, 29 genes, 31 metabolites, and 138 parameters (Figure 1A) was reconstructed using the model published in Padiadpu et al. (31). The liver NAD_{net} model retains all three routes of NAD^+ biosynthesis - (a) Route I - production of NAD^+ from tryptophan through the kynurenine pathway, which is known to be active in the liver (b) Route II - utilisation of Nicotinic acid (Na) as a substrate for NAD^+ generation through the Preiss–Handler pathway and (c) Route III - the salvage pathway of synthesising NAD from extracellular precursors provided by the diet (for, e.g., Nicotinamide (Nam) and nucleosides (Nicotinamide riboside (NR) and Nicotinic acid riboside (NAR)) (4–6, 48). Nam, Na, NR, and NAR are collectively referred to as Vitamin B3. Detailed information about enzymatic reactions is given in Table 1 and Supplementary Table 1. A steady-state analysis of the NAD_{net} was performed using CoRC (V 0.11.0). Steady-state values of metabolites and fluxes of the corresponding reactions are given in Table 2. Kinetic stability analysis of the model revealed that it was asymptotically stable.

Validation of NAD_{net}

The liver NAD_{net} was first inspected for validity by (a) Steady-state metabolite concentrations from the simulations

10.3389/fonc.2022.954512

TABLE 2 Steady-state concentration of metabolite and fluxes of reaction.

A) Steady-state metabolite concentrations		B) Steady-state fluxes of reactions	
Metabolite	**Concentration (μM)**	**Reaction ID**	**Flux (μM/s)**
L-Tryptophan	1.60E + 01	J1	1.90E-03
L-Formyl-kynurenine	6.12E + 01	J2	1.90E - 03
O_2	1.00E + 03	J3	1.90E - 03
Hydroxy-L-kynurenine	1.34E - 01	J4	1.90E - 03
NADPH	4.00E + 02	J5	1.90E - 03
Hydroxyanthranilate	6.06E - 07	J6	1.90E - 03
Quinolinate	6.70E - 01	J7	3.74E - 02
NaMN	3.28E + 00	J8	1.53E + 01
PRPP	1.00E + 03	J9	3.63E - 02
PPi	1.54E + 04	J10	1.11E - 03
NaAD	9.53E + 02	J11	1.59E - 02
NAD	2.15E + 04	J12	1.70E - 08
NMN	5.68E + 03	J13	1.98E - 11
ATP	1.00E + 03	J14	9.88E + 00
NH_3	1.00E+02	J15	5.03E + 00
Glutamine	6.00E + 02	J16	3.72E - 01
NADP	1.59E + 01	J17	1.49E + 01
Nam	4.98E + 04	J18	4.81E - 12
ADPribose	1.70E - 05	J19	2.75E - 24
ADPriboseP	1.98E - 08	J20	4.81E - 12
L-Kynurenine	9.21E - 02	J22	3.55E - 02
Arginine_protein	1.00E + 03	re23	2.15E - 02
NR	4.12E - 10	re25	1.59E - 02
P	1.00E + 03	re26	1.54E + 01
R1P	5.00E + 01	re27	1.98E - 11
Na	1.00E + 01	re28	1.70E - 08

were compared with the experimentally determined values reported in the literature (8); (b) Steady-state fluxes of enzymes of the enzymatic reactions were compared with the experimentally determined values. The available experimental data about metabolite concentrations were from diverse sources, including liver tissue, blood, and cerebral fluid. Moreover, some were from humans, and some from mice and other model organisms. To add to this difficulty, some were reported as nmol/gm while some others were given as μmol/gm of the liver tissue, making direct comparisons difficult. To overcome this problem, a rank-based correlation, using the Spearman correlation metric, was calculated for both experimental and simulated data. Experimentally constructed NAD biosynthesis rate for mouse liver reported by Mori et al. (8) was compared with the model-predicted metabolite level and flux rate for human hepatocytes (Figure 1B). A relative ranking of the metabolites (NAD, NaAD, NaMN and, NMN) and separately of the fluxes obtained from the experimental and from the simulation profile were used to compare the correlation between experimental and computational predictions. For the metabolites, the correlation was found to be +1, and for enzymatic reactions, it was found to be +0.86, suggesting that the model is consistent with experimental data. Utilisation of NMN through NMNAT1 (3) is seen to be the main route of NAD biosynthesis in the liver (Figure 1B). Further time-course analysis was performed on the NAD_{net}, and the NAD time course was plotted (Figure 1C) and compared against experimentally determined time course after labelled NAD supplementation in a mouse model (9). The time course profiles of NAD were in excellent agreement with that reported in the Liu et al. model.

HCC patients exhibit heterogeneity in their NAD profiles

Construction of personalised NAD_{net} models

Our next goal was to build personalised NAD_{net} models for each HCC patient by integrating the transcriptomics data available in the TCGA-LIHC cohort. First, we studied the transcriptomic variation in the enzymes of the NAD_{net} in 365 patients in the dataset. The fold change of each gene for a patient

sample was calculated by dividing the gene expression values of tumour tissue by the mean gene expression value of the normal liver tissues in the TCGA-LIHC cohort. Most of the NAD_{net} gene expressions in tumours were observed to be significantly different from the normal tissue (Supplementary Tables 4, 5). The distribution of log_2FC gene expression values among the NAD_{net} genes was analysed, as shown in Supplementary Figure 1A, and heterogeneity among patients is shown in Figure 2A. To build personalised models, the fold change in expression value of each gene was converted as an expression factor (F1 - F22), which was further integrated into their corresponding reaction of the pathway (by utilising gene-protein-reaction association). As J1, J7, J8, J12, J13, J15, J16, and J18 reactions are associated with multiple genes, a cumulative sum of the fold change of genes associated with each reaction was considered as expression factors. For example, for the J1 reaction, the cumulative sum of *TDO2, IDO1*, and *IDO2* gene expression values was taken as the expression factor (F1).

A distribution of expression factors, shown in Figure 2B, clearly indicates high heterogeneity across the TCGA-LIHC cohort for this model. The variation was seen to be the highest for F7 and F8 (NMN adenylyltransferase), F15 (Mono ADPribosyltransferase), F1 (Tryptophan 2,3- dioxygenase), F18 (5'-Nucleotidase), F16 (NAD^+ pyrophosphatase), F12 (NAD^+ glycohydrolase), F13 ($NAD(P)^+$ nucleosidase), and F22 (Nicotinate phosphoribosyltransferase) reactions (refer to Supplementary Table 6). The observed gene expression variations also suggest that there is likely to be variation in the reaction flux and the metabolite levels across different patients in the cohort (Supplementary Figure S1A and Supplementary Table 7). The mutation frequencies of the genes related to NAD_{net} were also obtained and analysed using the cBioPortal. The most frequent of them, which was in the *PARP1* gene, was seen to occur only in ~1% of the patients, while the rest of them were mutated in less than 1% of the patients (Supplementary Figure S2), clearly indicating that alterations in the NAD biosynthesis network are because of alteration in gene expression values, and not because of mutations.

Personalised NAD_{net} models indicate high patient heterogeneity in the dataset

The previous analysis (Supplementary Figure 1) reflected that the alterations in NAD_{net} profile in HCC could be attributed to variations in gene expression of the associated enzymes and also that there was no indication of any significant alteration in enzyme kinetics (K_m, K_{cat}). To construct personalised models for each HCC patient, the corresponding gene expression data was integrated into the base liver NAD_{net} as a surrogate measure of the enzyme abundance. Kinetic simulations and steady-state analysis of each personalised model were performed. Steady-state analysis was performed for all the models using different resolution thresholds. With the default COPASI resolution of 1E-09, we obtained 39 models with stable states. When the resolution threshold was lowered to 1E-03, we got 168 models obtaining stable states. With 1E-01 resolution threshold, we obtained 326 models. Considering the accuracy of defining a steady state and the number of models obtaining steady states, we used 1E-03 as the final resolution threshold. After steady-state analysis with the resolution threshold of 1E-03, 168 models out of 365 models were found to reach a stable state. As F-factors were the only difference among all the personalised models, the distribution of F-factors was compared between both the models, stable (models which obtained a stable state) and unstable (models which did not obtain a stable state) (refer to Supplementary Figure S1B and Supplementary Table 7). The unstable models had a significant difference in the values of F1, F2, F3, F12, F13, F14, F17, and F22 as compared to the stable models. Also, most of the F-factors distributions of the unstable models had a higher mean compared to the stable models. Notably, rate limiting reactions of the three routes of NAD synthesis, i.e., F1, F17, and F22, are significantly different and have higher values in the unstable models compared to stable models. Further, the NAD-consuming reactions F13 and F14 are also significantly higher in unstable models. Only the stable models are included for further analysis. Further, fold change values of each metabolite and reaction flux were calculated by dividing their respective steady-state values by the corresponding values in the NL model for the stable models. In the distribution of log_2FC steady-state metabolites among stable models, shown in Figure 2C, all metabolites, except NADP, show high variance, clearly indicating high heterogeneity at the metabolite level as well (Supplementary Table 8).

A comparison of the steady-state concentrations of metabolites and the reaction fluxes in the pathway of individual HCC patients with that of NL revealed that the patients could be classified into three groups: (a) the pathway, on the whole, is downregulated, and the NAD pool is low (NAD_low), (b) the pathway, on the whole, is upregulated and the NAD level is high (NAD_high) and (c) the pathway does not show any significant change with respect to NL (NAD_No_change) (Figures 2D, E). The analysis clearly indicated that (a) the kynurenine pathway (Route I) was observed to be significantly downregulated or unchanged [J1, J2, J3, J4, J5, and J6] in all except six patients. (b) biosynthesis of NAD from NA (Route II) - [J22, J9, and J7] was found to be upregulated in one subset and downregulated in another subset of patients, while it is unchanged in all others (c) the salvage route of NAD biosynthesis (Route III) - [J18, J19, and J20] was also found to be upregulated in a subset of patients and downregulated in the rest (Figure 2D). These changes put together result in an accumulation of Nam in most patients. (d) Interestingly Route II and Route III are not upregulated together in any given patient (in one sub-subset of patients, Route II is upregulated, while in another Route III is upregulated), suggesting that upregulation of NAD biosynthesis occurs through different routes. Hierarchical clustering analysis of the fluxes and metabolites led to the identification of 4 clusters

10.3389/fonc.2022.954512

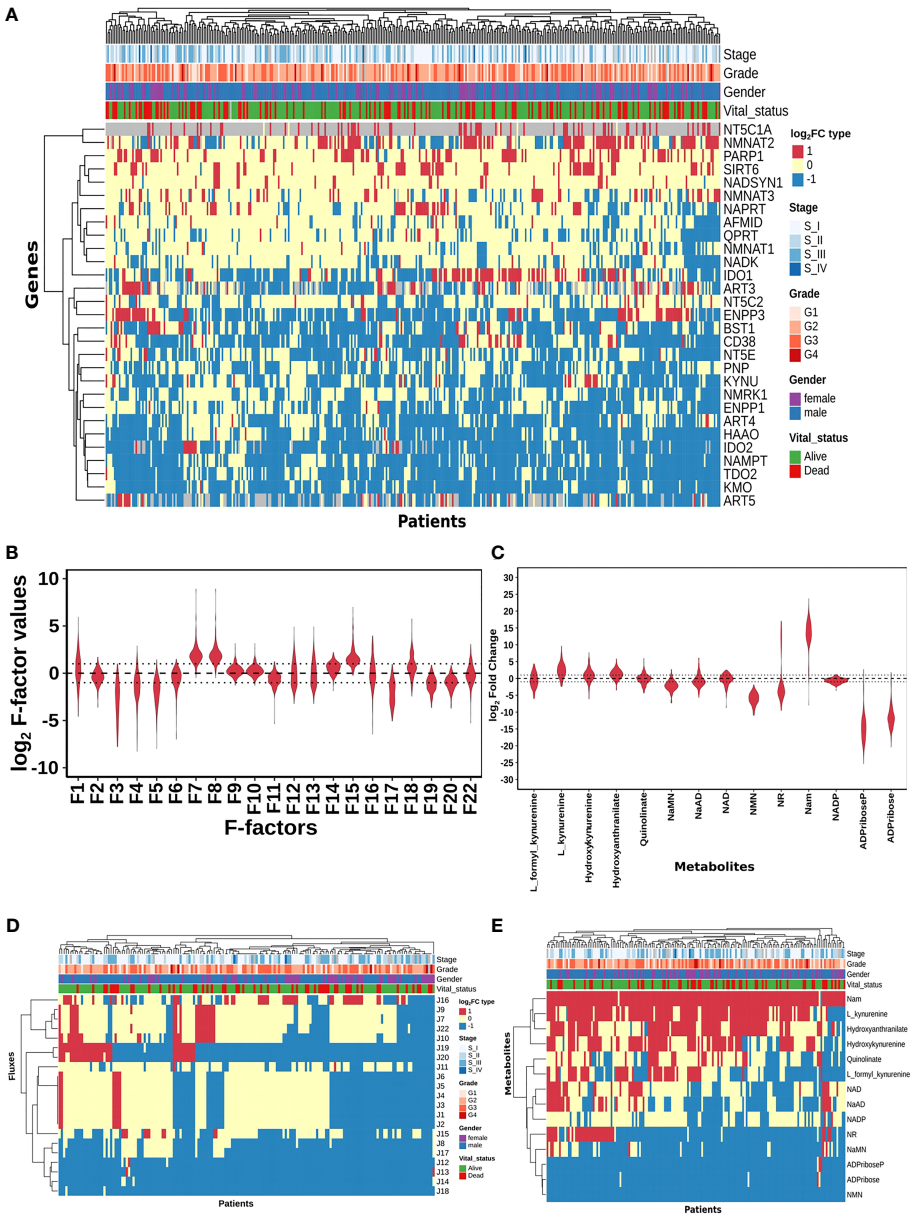


FIGURE 2
HCC patients exhibit heterogeneity in their NAD profiles. **(A)** Heatmap of NAD_{net} genes log_2FC values in tumour tissue calculated with respect to the normal tissue, rows correspond to genes (n = 29) and columns correspond to patients (n = 371). The red colour represents the upregulation of gene expression in tumour tissue compared to the normal tissue ($log_2FC \geq +1$), the blue colour represents a downregulation of gene expression in tumour tissue compared to the normal tissue ($log_2FC \leq -1$), yellow colour represents no change of gene expression in tumour tissue compared to the normal tissue ($-1 < log_2FC < +1$). Rows and columns are arranged based on the complete hierarchical clustering method. Annotations on the top of the heatmap are Stage, Grade, Gender, and Vital status (refer to the key in the image for more details). **(B)** Violin plot of log_2 F-factor values of NAD_{net} model. The X-axis shows the F-factors and the Y-axis shows the log_2 (fold change) values. **(C)** Violin plot of log_2FC values of NAD_{net} metabolites. The X-axis shows the Genes, and the Y-axis shows the log_2 (fold change) values. Metabolites are arranged according to the routes mentioned in Figure 1A. **(D)** Heatmap of reaction Fluxes obtained after steady-state analysis, rows correspond to reaction flux (n = 21), and columns correspond to patients (n = 168). The red colour represents an increase in flux compared to the base model, the blue colour represents a decrease in flux compared to the base model, and the yellow colour represents no change in flux compared to the base model. Rows and columns are arranged based on the complete hierarchical clustering method. Annotations on the top of the heatmap are Stage, Grade, Gender, and Vital status (refer to the key in the image for more details). **(E)** Heatmap of Metabolites obtained after steady-state analysis, rows correspond to metabolites and columns correspond to patients (n=168). The red colour represents an increase in the concentration of metabolite compared to the base model, the blue colour represents a decrease in the concentration of metabolite compared to the base model, yellow colour represents no change in concentration of metabolite compared to the base model. Rows and columns are arranged based on the complete hierarchical clustering method. Annotations on the top of the heatmap are Stage, Grade, Gender, and Vital status (refer to the key in the image for more details).

among the genes, largely corresponding to the route of NAD synthesis and utilisation, the fluxes, and metabolites among each route correlating positively within the same route (Supplementary Figures S1C, S1D).

NAPRT levels are suggestive of NAD biosynthetic status

A steady-state concentration of a metabolite depends not only on the enzyme concentration but also on various other parameters such as the concentration of the input metabolite for the given reaction, K_m, and K_{cat} of the enzymatic step, as well as on any feedback or feedforward loops for the given reaction. Therefore, metabolite abundance depends on the gene expression of the enzyme, metabolite inputs into the system, product metabolites, and the kinetics of the enzyme. The extent of correlation between the gene expression of all NAD_{net} genes and metabolites obtained after steady-state analysis was estimated for all patients in the TCGA-LIHC cohort using the Pearson correlation method (Figure 3A). Further, NAPRT was seen to have the highest correlation, which was statistically significant. The correlation values for all genes and metabolites, along with the statistical significance values [p-value and rho(r^2)], are given in Supplementary Table 9. This study clearly demonstrated that NAD steady-state levels in a cell are correlated to the NAPRT gene expression (r^2 = 0.92) (Figure 3B), and therefore, NAPRT gene expression can be used as a readout for NAD biosynthesis in the cell. As NAPRT levels are indicative of NAD levels, all the patients can be grouped based on NAPRT levels into three groups a) NAPRT_Up group, where NAD levels are high as compared to normal liver, b) NAPRT_Down group, where NAD levels are low as compared to normal liver, and c) NAPRT_No_change group, where NAD levels are comparable with the normal liver (Figures 3C, D). The analysis also clearly shows that Route II is the critical determinant of NAD status in HCC patients.

NAPRT is a control point in NAD_{net}

Our next goal was to identify reactions that wielded the highest control on the NAD_{net}, so as (a) to understand how the pathway dynamics are controlled and (b) to explore possible intervention points to manipulate the pathway. Further, it was of interest to investigate if the pathway control points varied significantly in different individuals in the cohort. Although the overall topology of the network remains the same, the weights associated with nodes (metabolites) and edges (reactions) change based on the gene expression patterns in different individuals, leading to the possibility of altering the control structures. To address this, the individual patient-wise kinetic models were used, and a parameter sensitivity analysis was performed on each of them using CoRC (refer to methods section parameter sensitivity analysis) and those reactions (and their corresponding genes) that had the highest influence on NAD levels were identified (Figure 4 and Supplementary Figure S3). Each model parameter sensitivity was calculated as a percentage change from the unaltered model, and models showing greater than one percent are concerned as altered models. If any parameter had greater than +1 percent change, it was taken that it positively influences the metabolite concentrations, whereas parameters with less than -1 percent have a negative influence on the metabolite concentrations. F22 (NAPRT, (Nicotinate phosphoribosyltransferase)) was observed to have a positive influence on NaAD, NAD, and Nam metabolites, and F22 had a negative influence on ADPriboseP across all patients in the cohort. Therefore, an increase in the gene expression of the NAPRT gene will lead to enhanced levels of the NaAD, NAD, and Nam metabolites. F22 has a positive influence on NaMN, NMN, and NR in only a subset of patients. F11 (NAD^+ Kinase) showed a negative influence on NAD in only a subset of patients but was not a control point in other patients. F1 (Tryptophan 2,3 dioxygenase) has a positive influence on the *de novo* pathway metabolites across all the patients in the cohort (Figure 4 and Supplementary Figure S3). F1 is known to be the rate-limiting step of the *de novo* pathway (15), and the same was observed in our analysis. Further, even in the base model, F1, F17 and F22 were identified as the key factors controlling metabolite concentrations of their respective routes (Supplementary Figure S3).

Survival analysis suggests potential benefits of NAD supplementation in NAPRT down subgroup

With the previous analysis, we identified NAPRT to be an indicator of NAD levels. We were interested in testing if there was any variation in survival in the groups based on NAPRT levels. For this, we performed a univariate cox-regression analysis using the predicted NAPRT level of individuals in the TCGA-LIHC cohort and calculated the extent of association of NAPRT level with HCC progression. Patients were divided into three groups based on NAPRT levels, NAPRT_Down ($\log_2FC \leq -1$), NAPRT_No_change ($-1 < \log_2FC < +1$) and NAPRT_Up ($\log_2FC \geq +1$) (Figure 3D). A Kaplan-Meier analysis was performed, and a log-rank test was used to determine significant differences in the overall disease progression in all group pairs. We first tested if NAPRT levels by themselves have any prognostic value, but the correlation with the risk of patient mortality was non-significant when compared between NAPRT UP and Down subgroups (p-value = 0.407 and HR (Up) = 0.783) (Supplementary Table 10), but NAPRT Down subgroup was correlated significantly with patient mortality when compared with NAPRT No_change subgroup (p-value = 0.0158 and HR

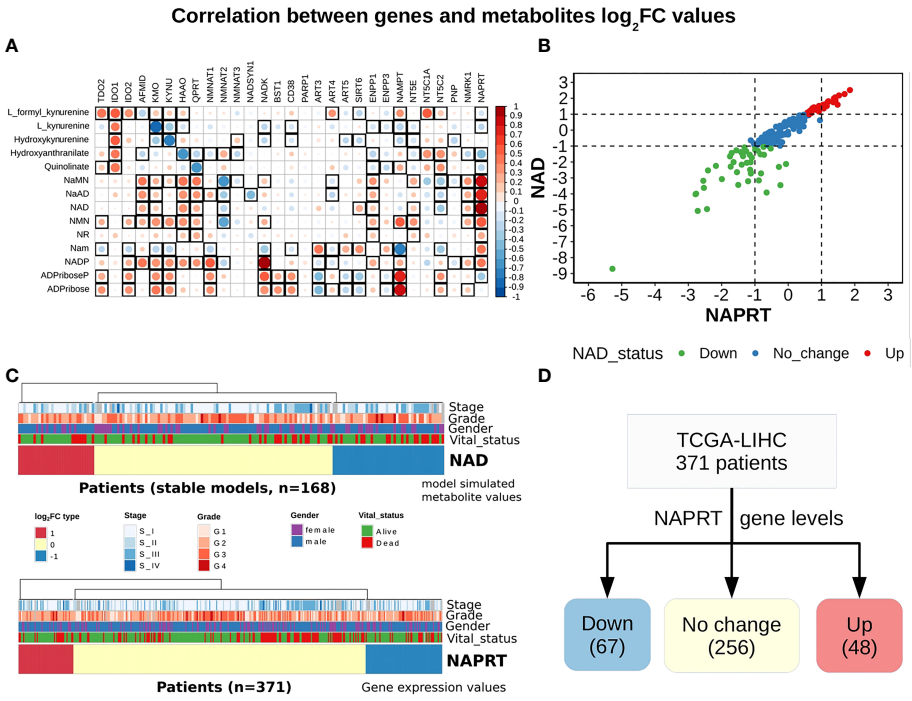


FIGURE 3
NAPRT alone is sufficient to indicate the NAD status in patients. **(A)** Correlogram between gene log_2FC values and metabolite log_2FC values. Rows represent metabolites and columns represent genes. The red colour corresponds to positive correlation, the blue colour corresponds to negative correlation, the area covered in the square corresponds to the absolute value of the correlation, and the black squares correspond to significant correlations (p-value < 0.05). Rows and columns are arranged based on the routes mentioned in Figure 1A. **(B)** Correlation plot showing the NAPRT and NAD log_2FC values. The X-axis represents NAPRT log_2FC values and the Y-axis represents NAD log_2FC values. The points are coloured based on the NAD status of the samples. Red points indicate NAD up samples, blue points indicate NAD no change samples, and green points indicate NAD down samples. (n = 168) **(C)** Above panel shows the heatmap of NAD log_2FC values obtained after steady-state analysis, rows correspond to NAD groups and columns correspond to patients (n = 168). The Red colour represents up NAD levels ($log_2FC \geq +1$), blue colour represents NAPRT down levels ($log_2FC \leq -1$), yellow colour represents NAD no change levels ($-1 < log_2FC < +1$). The below panel shows the heatmap of NAPRT log_2FC values in tumour tissue calculated with respect to the normal tissue, rows correspond to NAPRT groups and columns correspond to patients (n = 371). The Red colour represents up NAPRT levels ($log_2FC \geq +1$), the blue colour represents NAPRT down levels ($log_2FC \leq -1$), yellow colour represents NAPRT no change levels ($-1 < log_2FC < +1$). Columns are arranged based on the complete hierarchical clustering method. Annotations on the top of the heatmap are Stage, Grade, Gender, and Vital status (refer to the key in the image for more details). **(D)** Schematic showing the division of patients into three groups based on NAPRT gene levels.

(No_change) = 0.596) (Figure 5A). NAMPT is known to be a poor prognosis marker and a known drug target for the NAD pathway in many cancers, and we tested if the levels of this gene had any prognostic value (49). Here too, we found the correlation with the risk of mortality to be non-significant (p-value = 0.451, HR (No_change) = 0.854) (Figure 5B). Clearly, neither NAPRT nor NAMPT did not have any survival prognosis by itself. We then tested if pairs of groups with different NAPRT and NAMPT statuses exhibited any survival difference. In total, six groups were tested (Figure 5). NAMPT did not have any upregulated patients in the TCGA-LIHC cohort; also, Route III was downregulated in most of the patients. Upon Kaplan-Meier analysis, we found that the NAPRT_NAMPT Down_Down group has a significantly poorer prognosis than other groups (Figures 5C, D; Supplementary Table 10). This suggests that, in the NAPRT_NAMPT Down_down group, the prognosis could be improved by NAD supplementation to improve survival.

Discussion

Nicotinamide adenine dinucleotide (NAD), being an important cofactor in various biochemical reactions, plays a pivotal role in enabling and governing essential cellular activities. The levels of NAD are used by the cell as sensors for deciding what metabolic state it attains. A systems' understanding of the pathways involved in NAD biosynthesis that provide quantitative insights is therefore important. The enzymes in the pathway have been well studied individually, and a wealth of biochemical information is available on each of them, enabling the reconstruction of a systems model of a NAD

10.3389/fonc.2022.954512

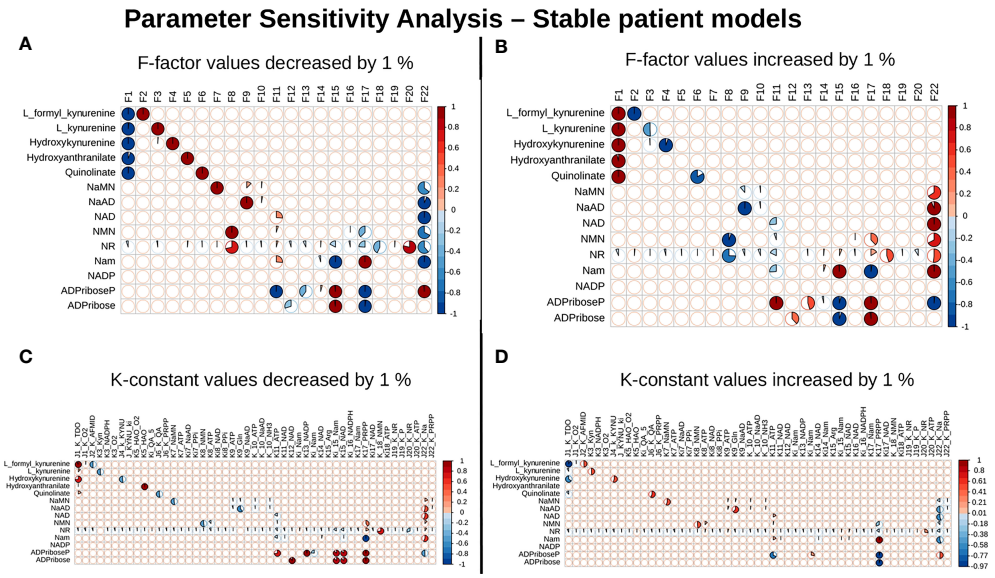


FIGURE 4
NAPRT is the control point in the NAD biosynthesis pathway in HCC patients. Correlogram between Metabolites and F-factors summarising the extent of patients affected with changes in the parameter values by -1 % in F-factors **(A)** by +1 % in F-factors **(B)** -1 % in K-constants **(C)**, and +1 % in K-constants **(D)**. The X-axis represents the parameters, and the Y-axis represents the metabolites. The red colour represents an increase in concentration and the blue colour represents a decrease in concentration. The area occupied by the coloured pie shows percentage of stable models, with greater than 1 % change in the concentration due to the change in the parameter value, out of 168 stable models.

biosynthetic network (NAD_{net}). We then use a kinetic modelling approach to study if there is variation in the NAD levels in HCC patients. Using NAD_{net} as a base model, we then construct personalised models for each patient by integrating with patient-specific gene expression values for all the enzymes in the network. While normal liver cells are known to use *de novo* NAD biosynthetic routes to maintain intracellular NAD levels, our model suggests that cancer cells are primarily dependent on the Preiss-Handler pathway (Route II in NAD_{net}). NAPRT being a rate-limiting step of this route, is clearly seen to have altered gene expression in several HCC patients. While most studies provide population or cohort-level insights, our modelling approach of constructing personalised models has a unique advantage of providing insights at the individual patient level.

The effect of NAD on disease progression presents a complex picture. The analysis carried out here by studying perturbations at a patient level provides insights leading to sub-grouping. This, in turn, serves as a framework to resolve some of the inconsistencies evident in the literature. A subgroup of HCC patients with high NAD biosynthetic status responds differently to the subgroup that has low NAD biosynthetic status. While the first subgroup can be envisaged to benefit from an inhibitor of NAD biosynthesis, the latter subgroup will benefit from supplementation. Enhanced levels of NAD have been shown to support tumour proliferation. Inhibition of the pathway, specifically with NAMPT and NAPRT as drug targets, has been suggested as a strategy for reducing NAD levels (26). Both NAPRT and NAMPT are critical enzymes of two different routes of NAD synthesis. A combination of both gives a better representation of NAD levels. A group with NAMPT and NAPRT (Down_Down) group, has low levels of NAD, and they require supplementation; on the other hand, a group with NAMPT Down and other combinations of NAPRT (except NAPRT Down) can still maintain NAD levels. Our study shows that identifying the precise subgroup is essential for determining whether NAD inhibition or NAD supplementation would be beneficial.

Supplementation using readily available vitamin B3 supplements is an easy intervention to achieve if the subgroup is correctly identified. There are multiple lines of evidence in support of the supplementation. First, inhibition of NAD production has been associated with higher levels of DNA damage and triggering of hepatocarcinogenesis (28). Boosting NAD^+ levels with supplements has been shown to have prophylactic effects in a genetically engineered mouse model of unconventional prefoldin RPB5 interactor (URI) used to study the mechanism of HCC development (28, 50). Second, in some other cancers, such as colorectal cancer, NAMPT and NAPRT high expression are seen to be associated with poor prognosis for the patient (51). Third, NAD levels were reported to be declining with age as well as implicated in a few liver diseases, including NAFLD (30). Due to this, several studies have proposed supplementation with NAD, and many NAD precursors are tested as supplements to increase NAD levels (52, 53). Among the precursors, Na was reported to be one of the

10.3389/fonc.2022.954512

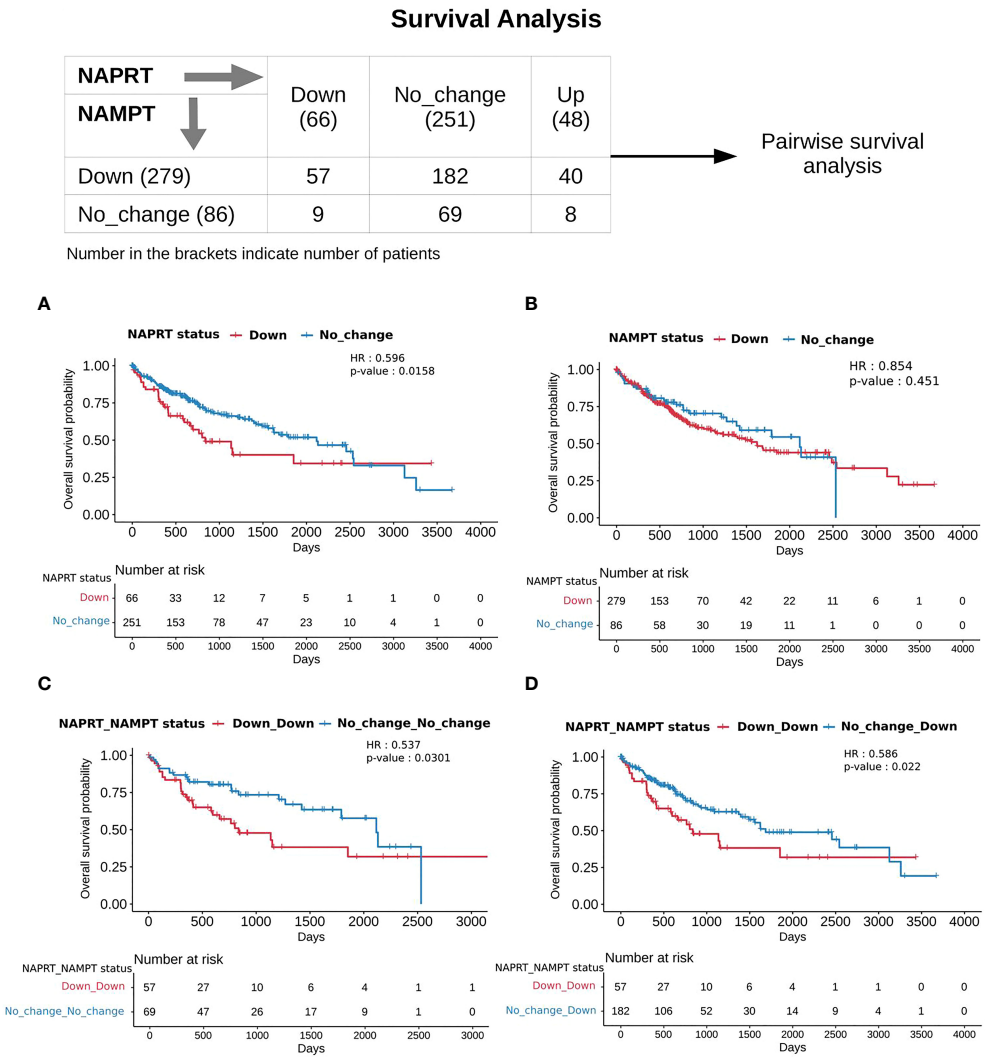


FIGURE 5
NAPRT_NAMPT Down_Down status corresponds to poorer survival. The panel on the top shows the distribution of patients into groups based on the NAPRT and NAMPT gene expression status. Kaplan-Meier Overall survival curve for HCC patients classified based on NAPRT Down and No_change groups **(A)** NAMPT Down and No_change groups **(B)**, NAPRT_NAMPT Down_Down versus No_change_No_change groups **(C)**, and NAPRT_NAMPT Down_Down versus Down_No_change groups **(D)**. HR and p-values reported in the figure panels are for the group represented in blue colour.

best precursors with the least side effects and greater potential of getting converted into NAD (54).

The reconstructed model has the following three major limitations- (a) As the model has only biosynthesis reactions but not all the utilisation reactions of NAD, the model fails to capture the quantitative level of NAD in the cell, and (b) even though NAD metabolism is known to have subcellular compartmentalization of NAD pools both at the metabolite as well as the enzyme level; the reconstructed model considers the total NAD pool only, and there is no subcellular compartment in the model and therefore it cannot capture intracellular compartmental dynamics of NAD biosynthesis and (c) As metabolism is one of the most tightly regulated processes in the cell, regulatory interactions i.e., transcription factors that may govern the gene expression of enzymes of the NAD biosynthesis pathway are not included here, and therefore the effect of perturbation at transcription regulation cannot be modelled directly here.

Nevertheless, the model is useful for understanding the extent of variation in NAD biosynthesis at an individual patient level. From the correlation analysis, it is evident that the changes in gene expression are captured at the metabolite level. NAPRT levels are found to indicate the NAD biosynthetic status in the sample. Furthermore, NAPRT levels are regulated

by MYC and TP53 transcription factors which are involved in cell growth and proliferation. NAPRT is also involved in immune and inflammation signalling (55).

In conclusion, we find high levels of heterogeneity in the NAD levels in HCC patients, and NAPRT gene expression levels are sufficient to indicate the NAD levels. Based on the NAPRT status, HCC patients can be subtyped into three categories corresponding to upregulation, no change, and downregulation of NAPRT with respect to a healthy liver. The NAPRT_Down group, when combined with NAMPT_Down, is seen to show poorer survival as compared to a group of HCC patients where the levels of these two enzymes are unaltered. Lower NAD levels correlate with lower levels of NAPRT and suggest that supplementation of NAD may be beneficial in this group of patients. Our study provides a rationale, and a means to explore subgrouping in HCC patients, paving the way for precision diagnosis and intervention.

Data availability statement

The model generated in this study is deposited in EBI-BioModels and assigned the identifier MODEL2205250001: "https://www.ebi.ac.uk/biomodels/MODEL2205250001.1"; Scripts used in the study are made available on Github: "https://github.com/Adithya-C/NAD_LIHC_project"; Publicly available TCGA-LIHC dataset was analyzed in this study, which was retrieved from the Xena Browser, cohort: GDC TCGA Liver Cancer (LIHC): "https://xenabrowser.net/datapages/?cohort=GDC%20TCGA%20Liver%20Cancer%20(LIHC)&removeHub=https%3A%2F%2Fxena.treehouse.gi.ucsc.edu%3A443".

Author contributions

NC conceptualized and supervised the study. MM and OK constructed the initial computational model and AC refined the model. AC, MM, and SS performed the simulations. AC, MM, and NC analysed the results. AC, MM, and NC wrote the first draft of the manuscript. All authors have read and approved the final manuscript. All authors contributed to the article and approved the submitted version.

Acknowledgments

We thank Department of Biotechnology (DBT), Government of India (Bioinformatics Centre Grant, BT/PR40187/BTIS/137/3/2021) for general support. AC acknowledges Council for Scientific Research (CSIR), Government of India for Senior Research Fellowship (SRF).

Conflict of interest

NC is a co-founder of the companies qBiome Research Pvt. Ltd. and HealthSeq Precision Medicine Pvt. Ltd. They had no role in this manuscript.

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.954512/full#supplementary-material

References

1. Cantó C, Menzies KJ, Auwerx J. NAD+ metabolism and the control of energy homeostasis: A balancing act between mitochondria and the nucleus. *Cell Metab* (2015) 22:31–53. doi: 10.1016/j.cmet.2015.05.023

2. Magni G, Amici A, Emanuelli M, Orsomando G, Raffaelli N, Ruggieri S. Enzymology of NAD+ homeostasis in man. *Cell Mol Life Sci* (2004) 61: 19–34. doi: 10.1007/s00018-003-3161-1

3. Nikiforov A, Kulikova V, Ziegler M. The human NAD metabolome: Functions, metabolism and compartmentalization. *Crit Rev Biochem Mol Biol* (2015) 50:284–97. doi: 10.3109/10409238.2015.1028612

4. Belenky P, Bogan KL, Brenner C. NAD+ metabolism in health and disease. *Trends Biochem Sci* (2007) 32:12–9. doi: 10.1016/j.tibs.2006.11.006

5. Katsyuba E, Auwerx J. Modulating NAD + metabolism, from bench to bedside. *EMBO J* (2017) 36:2670–83. doi: 10.15252/embj.201797135

6. Zapata-Pérez R, Wanders RJA, van Karnebeek CDM, Houtkooper RH. NAD + homeostasis in human health and disease. *EMBO Mol Med* (2021) 13:e13943. doi: 10.15252/emmm.202113943

7. Chiarugi A, Dölle C, Felici R, Ziegler M. The NAD metabolome — a key determinant of cancer cell biology. *Nat Rev Cancer* (2012) 12:741–52. doi: 10.1038/nrc3340

8. Mori V, Amici A, Mazzola F, di Stefano M, Conforti L, Magni G, et al. Metabolic profiling of alternative NAD biosynthetic routes in mouse tissues. *PLoS One* (2014) 9:e113939. doi: 10.1371/journal.pone.0113939

9. Liu L, Su X, Quinn WJ, Hui S, Krukenberg K, Frederick DW, et al. Quantitative analysis of NAD synthesis-breakdown fluxes. *Cell Metab* (2018) 27:1067–80.e5. doi: 10.1016/j.cmet.2018.03.018

10. Imai S-I, Guarente L. It takes two to tango: NAD + and sirtuins in aging/longevity control. *NPJ Aging Mech Dis* (2016) 2:16017. doi: 10.1038/npjamd.2016.17

11. Nwosu ZC, Megger DA, Hammad S, Sitek B, Roessler S, Ebert MP, et al. Identification of the consistently altered metabolic targets in human hepatocellular carcinoma. *Cell Mol Gastroenterol Hepatol* (2017) 4:303–323.e1. doi: 10.1016/j.jcmgh.2017.05.004

12. Chowdhry S, et al. NAD metabolic dependency in cancer is shaped by gene amplification and enhancer remodelling. *Nature* (2019) 569(7757):570–575. doi: 10.1038/s41586-019-1150-2

13. Audrito V, Messana VG, Deaglio S. NAMPT and NAPRT: Two metabolic enzymes with key roles in inflammation. *Front Oncol* (2020) 10:358. doi: 10.3389/fonc.2020.00358

14. Rodriguez Cetina Biefer H, Vasudevan A, Elkhal A. Aspects of tryptophan and nicotinamide adenine dinucleotide in immunity: A new twist in an old tale. *Int J Tryptophan Res* (2017) 10:1178646917713491. doi: 10.1177/1178646917713491

15. Badawy AA-B. Kynurenine pathway of tryptophan metabolism: Regulatory and functional aspects. *Int J Tryptophan Res* (2017) 10:117864691769193. doi: 10.1177/1178646917691938

16. Shats I, Williams JG, Liu J, Makarov MV, Wu X, Lih FB, et al. Bacteria boost mammalian host NAD metabolism by engaging the deamidated biosynthesis pathway. *Cell Metab* (2020) 31:564–79.e7. doi: 10.1016/j.cmet.2020.02.001

17. Lucas S, Soave C, Nabil G, Othman Ahmed ZS, Chen G, El-Banna HA, et al. Pharmacological inhibitors of NAD biosynthesis as potential an ticancer agents. *Recent Patents Anti-Cancer Drug Discovery* (2017) 12:190–207. doi: 10.2174/1574892812666170619125503

18. Piacente F, Caffa I, Ravera S, Sociali G, Passalacqua M, Vellone VG, et al. Nicotinic acid phosphoribosyltransferase regulates cancer cell metabolism, susceptibility to NAMPT inhibitors, and DNA repair. *Cancer Res* (2017) 77:3857–69. doi: 10.1158/0008-5472.CAN-16-3079

19. Tangutoori S, Baldwin P, Sridhar S. PARP inhibitors: A new era of targeted therapy. *Maturitas* (2015) 81:5–9. doi: 10.1016/j.maturitas.2015.01.015

20. Tan B, Young DA, Lu Z-H, Wang T, Meier TI, Shepard RL, et al. Pharmacological inhibition of nicotinamide phosphoribosyltransferase (NAMPT), an enzyme essential for NAD+ biosynthesis, in human cancer cells. *J Biol Chem* (2013) 288:3500–11. doi: 10.1074/jbc.M112.394510

21. Watson M, Roulston A, Bélec L, Billot X, Marcellus R, Bédard D, et al. The small molecule GMX1778 is a potent inhibitor of NAD+ biosynthesis: strategy for enhanced therapy in nicotinic acid phosphoribosyltransferase 1-deficient tumors. *Mol Cell Biol* (2009) 29:5872–88. doi: 10.1128/MCB.00112-09

22. Ghanem MS, Monacelli F, Nencioni A. Advances in NAD-lowering agents for cancer treatment. *Nutrients* (2021) 13:1665. doi: 10.3390/nu13051665

23. Surjana D, Halliday GM, Damian DL. Role of nicotinamide in DNA damage, mutagenesis, and DNA repair. *J Nucleic Acids* (2010) 2010:1–13. doi: 10.4061/2010/157591

24. Garrido A, Djouder N. NAD + deficits in age-related diseases and cancer. *Trends Cancer* (2017) 3:593–610. doi: 10.1016/j.trecan.2017.06.001

25. Srivastava S. Emerging therapeutic roles for NAD + metabolism in mitochondrial and age-related disorders. *Clin Trans Med* (2016) 5:25. doi: 10.1186/s40169-016-0104-7

26. Yaku K, Okabe K, Nakagawa T. NAD metabolism: Implications in aging and longevity. *Ageing Res Rev* (2018) 47:1–17. doi: 10.1016/j.arr.2018.05.006

27. Mederacke I, Schwabe RF. NAD(+) supplementation as a novel approach to cURIng HCC? *Cancer Cell* (2014) 26:777–778. doi: 10.1016/j.ccell.2014.11.011

28. Tummala KS, Gomes AL, Yilmaz M, Graña O, Bakiri L, Ruppen I, et al. Inhibition of *de novo* NAD(+) synthesis by oncogenic URI causes liver tumorigenesis through DNA damage. *Cancer Cell* (2014) 26:826–839. doi: 10.1016/j.ccell.2014.10.002

29. Martens CR, Denman BA, Mazzo MR, Armstrong ML, Reisdorph N, McQueen MB, et al. Chronic nicotinamide riboside supplementation is well-tolerated and elevates NAD+ in healthy middle-aged and older adults. *Nat Commun* (2018) 9:1286. doi: 10.1038/s41467-018-03421-7

30. Zhou C-C, Yang X, Hua X, Liu J, Fan M-B, Li G-Q, et al. Hepatic NAD(+) deficiency as a therapeutic target for non-alcoholic fatty liver disease in ageing. *Br J Pharmacol* (2016) 173:2352–68. doi: 10.1111/bph.13513

31. Padiadpu J, Mishra M, Sharma E, Mala U, Somasundaram K, Chandra N. Probing the druggability limits for enzymes of the NAD biosynthetic network in glioma. *J Chem Inf Modeling* (2016) 56:843–53. doi: 10.1021/acs.jcim.5b00733

32. Wang M, Herrmann CJ, Simonovic M, Szklarczyk D, Mering C. Version 4.0 of PaxDb: Protein abundance data, integrated across model organisms, tissues, and cell-lines. *PROTEOMICS* (2015) 15:3163–8. doi: 10.1002/pmic.201400441

33. Ally A, Balasundaram M, Carlsen R, Chuah E, Clarke A, Dhalla N, et al. Comprehensive and integrative genomic characterization of hepatocellular carcinoma. *Cell* (2017) 169:1327–41.e23. doi: 10.1016/j.cell.2017.05.046

34. Goldman MJ, Craft B, Hastie M, Repečka K, McDade F, Kamath A, et al. Visualizing and interpreting cancer genomics data *via* the xena platform. *Nat Biotechnol* (2020) 38:675–8. doi: 10.1038/s41587-020-0546-8

35. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* (2010) 26:139–40. doi: 10.1093/bioinformatics/btp616

36. Carlson M, et al. _org.Hs.eg.db: Genome wide annotation for Human_. R package version 3.15.0. (2022)

37. Mishra M, Jayal P, Karande AA, Chandra N. Identification of a co-target for enhancing efficacy of sorafenib in HCC through a quantitative modeling approach. *FEBS J* (2018) 285:3977–3992. doi: 10.1111/febs.14641

38. Nusinow DP, Szpyt J, Ghandi M, Rose CM, McDonald ER, Kalocsay M, et al. Quantitative proteomics of the cancer cell line encyclopedia. *Cell* (2020) 180:387–402.e16. doi: 10.1016/j.cell.2019.12.023

39. Stavrum A-K, Heiland I, Schuster S, Puntervoll P, Ziegler M. Model of tryptophan metabolism, readily scalable using tissue-specific gene expression data. *J Biol Chem* (2013) 288:34555–66. doi: 10.1074/jbc.M113.474908

40. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discovery* (2012) 2:401–4. doi: 10.1158/2159-8290.CD-12-0095

41. Förster J, Bergmann FT, Pahle J. CoRC: the COPASI r connector. *Bioinformatics* (2021) 37:2778–9. doi: 10.1093/bioinformatics/btab033

42. R Core Team. *R: A language and environment for statistical computing*. (R Foundation for Statistical Computing:Vienna, Austria) (2022).

43. Wei T, Simko V. *GitHub - taiyun/corrplot: A visual exploratory tool on correlation matrix* (2021). Available at: https://github.com/taiyun/corrplot.

44. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* (2016) 32:2847–9. doi: 10.1093/bioinformatics/btw313

45. Therneau T. _A Package for Survival Analysis in R_. R package version 3.3-1, (2022). Available at: https://CRAN.R-project.org/package=survival.

46. Therneau TM, Grambsch PM. Modeling Survival Data: Extending the Cox Model. Springer, New York, (2000).

47. Kassambara A, Kosinski M, Biecek P. *GitHub - kassambara/survminer: Survival analysis and visualization* (2021). Available at: https://rpkgs.datanovia.com/survminer/index.html.

48. Katsyuba E, Romani M, Hofer D, Auwerx J. NAD+ homeostasis in health and disease. *Nat Metab* (2020) 2:9–31. doi: 10.1038/s42255-019-0161-5

49. Galli U, Colombo G, Travelli C, Tron GC, Genazzani AA, Grolla AA. Recent advances in NAMPT inhibitors: A novel immunotherapic strategy. *Front Pharmacol* (2020) 11:656. doi: 10.3389/fphar.2020.00656

50. Djouder N. Boosting NAD+ for the prevention and treatment of liver cancer. *Mol {\&} Cell Oncol* (2015) 2:e1001199. doi: 10.1080/23723556.2014.1001199

51. Li X, Lei J, Mao L, Wang Q, Xu F, Ran T, et al. NAMPT and NAPRT, Key Enzymes in NAD Salvage Synthesis Pathway, Are of Negative Prognostic Value in Colorectal Cancer. *Front Oncol* (2019) 9:736. doi: 10.3389/fonc.2019.00736

52. She J, Sheng R, Qin Z-H. Pharmacology and potential implications of nicotinamide adenine dinucleotide precursors. *Aging Dis* (2021) 12:1879–1897. doi: 10.14336/AD.2021.0523

53. Rajman L, Chwalek K, Sinclair DA. Therapeutic potential of NAD-boosting molecules: The *In vivo* evidence. *Cell Metab* (2018) 27:529–547. doi: 10.1016/j.cmet.2018.02.011

54. Palmer RD, Elnashar MM, Vaccarezza M. Precursor comparisons for the upregulation of nicotinamide adenine dinucleotide. novel approaches for better aging. *Aging Med* (2021) 4:214–220. doi: 10.1002/agm2.12170

55. Duarte-Pereira S, Fajarda O, Matos S, Luís Oliveira J, Silva RM. NAPRT expression regulation mechanisms: Novel functions predicted by a bioinformatics approach. *Genes (Basel)* (2021) 12:2022. doi: 10.3390/genes12122022

Frontiers in Oncology

Advances knowledge of carcinogenesis and tumor progression for better treatment and management

The third most-cited oncology journal, which highlights research in carcinogenesis and tumor progression, bridging the gap between basic research and applications to imrpove diagnosis, therapeutics and management strategies.

Discover the latest Research Topics

See more →

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact