Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Computational integration of single-cell transcriptomics and functional dependency data reveals metabolic-regulatory vulnerabilities of drug-tolerant persister cells in EGFR-mutant non-small-cell lung cancer

Linghan Meng
Guang'anmen Hospital, China Academy of Chinese Medical Sciences
Jingna Tao
Dongzhimen Hospital Affiliated to Beijing University of Chinese Medicine
Yue Li
Guang'anmen Hospital, China Academy of Chinese Medical Sciences
Yue Luo
Guang'anmen Hospital, China Academy of Chinese Medical Sciences
Zehang Lei
Guang'anmen Hospital, China Academy of Chinese Medical Sciences
Di Zhang
Guang'anmen Hospital, China Academy of Chinese Medical Sciences
Xin Chen
Guang'anmen Hospital, China Academy of Chinese Medical Sciences
Haixiao Liu
Guang'anmen Hospital, China Academy of Chinese Medical Sciences
Yanju Bao
byjtxyx@163.com
Guang'anmen Hospital, China Academy of Chinese Medical Sciences
Baojin Hua
Guang'anmen Hospital, China Academy of Chinese Medical Sciences

Research Article

Keywords: drug-tolerant persister cells, metabolic reprogramming, transcription factor regula tory network, ferroptosis, GPX4, thioredoxin reductase, single-cell RNA sequencing, EGFR-mutant NSCLC, cancer drug resistance, DepMap, computational biology

Posted Date: September 11th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9489377/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Computational integration of single-cell transcriptomics and functional dependency data reveals metabolic–regulatory vulnerabilities of drug-tolerant persister cells in EGFR-mutant non-small-cell lung cancer

Linghan Meng[1,3,†], Jingna Tao[2,3,†], Yue Li[1], Yue Luo[1], Zehang Lei[1], Di Zhang[1], Xin Chen[1], Haixiao Liu[1], Yanju Bao[1,*], and Baojin Hua[1,*]

[1]Department of Oncology, Guang'anmen Hospital, China Academy of Chinese Medical Sciences, Beijing 100053, China
[2]Dongzhimen Hospital, Beijing University of Chinese Medicine, Beijing 100700, China
[3]Beijing University of Chinese Medicine, Beijing 100029, China
[†]These authors contributed equally to this work.
[*]Corresponding authors: Yanju Bao, byjtxyx@163.com; Baojin Hua, huabaojinxs@126.com.

Abstract

Drug-tolerant persister (DTP) cells survive targeted therapy through reversible, non-genetic adaptations and can seed eventual relapse, yet how their metabolic reprogramming unfolds over time—and what regulatory logic governs it—remains poorly understood. We addressed this gap computationally, analyzing time-resolved single-cell RNA-seq data from PC9 EGFR-mutant lung cancer cells treated with erlotinib across six time points (3,046 cells after quality control). Rather than a single metabolic switch, we found a staged transition: an acute shift toward oxidative phosphorylation within the first 24 hours, a subsequent collapse of lipogenic programs linked to inferred SREBF1/2 transcription factor repression, and a late consolidation of aldehyde-detoxification and ferroptosis-defense pathways. These changes persist after cell cycle regression, ruling out proliferative arrest as the sole driver. Upstream, a regulatory network anchored by NFE2L2 (NRF2) connects 12 hub transcription factors to the metabolic effector genes. When we cross-referenced these candidates against DepMap CRISPR screens stratified by a DTP-like expression signature, GPX4 emerged as preferentially essential in DTP-signature-high cell lines ($P = 4.7 \times 10^{-4}$), whereas the thioredoxin reductase TXNRD1 was broadly essential regardless of DTP status. An independent PC9 dataset treated with gefitinib and osimertinib reproduced the core metabolic changes. Because this is a purely computational study, every target and therapeutic hypothesis—including the prioritization of auranofin for combination with EGFR inhibitors—requires experimental validation. The data and framework are offered as a resource for designing such experiments.

Keywords: drug-tolerant persister cells, metabolic reprogramming, transcription factor regulatory network, ferroptosis, GPX4, thioredoxin reductase, single-cell RNA sequencing, EGFR-mutant NSCLC, cancer drug resistance, DepMap, computational biology

1 Introduction

The emergence of drug resistance remains the principal barrier to durable therapeutic responses and long-term survival in cancer. While decades of research have focused on acquired genetic mutations—such as the EGFR T790M gatekeeper mutation in non-small-cell lung cancer (NSCLC) [1] or BRAF splice variants in melanoma—as the primary drivers of therapeutic failure, a growing body of evidence now implicates a fundamentally distinct, non-genetic mechanism mediated by drug-tolerant persister (DTP) cells [2, 3]. First described in the landmark 2010 study by Sharma et al., DTP cells constitute a small subpopulation—typically 0.3–1% of the initial tumor cell population—that survives intensive drug exposure not through pre-existing or acquired genetic alterations, but through rapid, reversible phenotypic adaptations enacted at the epigenetic, transcriptional, and metabolic levels [4]. Upon treatment withdrawal, these cells can re-enter the cell cycle within days to weeks, re-establishing a drug-sensitive population that is capable of seeding eventual disease relapse [5]. This reversibility distinguishes DTPs from classically resistant clones and has profound therapeutic implications: if the persister reservoir can be eliminated during the window of drug tolerance—before irreversible genetic resistance mutations arise—durable remission may be achievable.

DTP cells have now been documented across a broad range of cancer types and therapeutic modalities. In EGFR-mutant NSCLC, DTP cells survive both first-generation (erlotinib, gefitinib) and third-generation (osimertinib) EGFR tyrosine kinase inhibitors (TKIs) through non-genetic mechanisms involving chromatin remodeling, signaling pathway rewiring, and metabolic adaptation [6–8]. In BRAF-mutant melanoma, persister cells emerge within days of vemurafenib treatment and adopt a neural crest stem cell-like state driven by ALDH1A3-dependent acetaldehyde metabolism [9]. In triple-negative breast cancer (TNBC), DTP cells surviving chemotherapy exhibit a shared persistence program characterized by AP-1/NF-κB stress response activation and basal keratin expression that is conserved across treatments and patients [10, 11]. The universality of the persister phenomenon across cancer lineages and drug classes suggests the existence of a core adaptive program that may be therapeutically targetable.

Two interconnected biological programs underlie the DTP state: epigenetic remodeling and metabolic reprogramming. At the epigenetic level, DTPs engage chromatin-modifying enzymes—among them the histone H3K4 demethylase KDM5A—to establish a transcriptionally quiescent, slow-cycling state characterized by global reduction of H3K4me3 at promoters of proliferation-associated genes [2]. This epigenetic remodeling is reinforced by additional mechanisms: genome-wide CRISPR screens have identified the Hippo/YAP/TEAD pathway as a non-genetic driver of persister cell survival in EGFR-mutant lung cancer [6], and integrin-linked kinase (ILK) signaling through SFK–YAP promotes EMT-mediated drug tolerance in lung adenocarcinoma [7]. Recent work has further revealed that APOBEC3 mutagenesis is induced in DTP cells, potentially accelerating the transition from reversible tolerance to irreversible genetic resistance [8], underscoring the urgency of eliminating DTPs before this evolutionary trajectory is completed.

At the metabolic level, DTP cells undergo a fundamental rewiring of energy and biosynthetic metabolism. The most consistently reported change is a shift from aerobic glycolysis (the Warburg phenotype characteristic of proliferating cancer cells) toward mitochondrial oxidative phosphorylation (OXPHOS) and fatty acid β-oxidation (FAO) [4, 12]. This metabolic switch is accompanied by upregulation of antioxidant defense systems, particularly the glutathione peroxidase GPX4, which suppresses ferroptosis by reducing phospholipid hydroperoxides [13–15], and ALDH family enzymes, which detoxify reactive aldehydes generated by lipid peroxidation [12]. Functional validation of these dependencies has been demonstrated in several systems: combined inhibition of glutathione synthesis (using buthionine sulfoximine) and induction of reactive oxygen species

(using arsenic trioxide) completely prevented relapse in pancreatic cancer xenograft models [16], while BET bromodomain inhibition selectively killed cycling persister cells by suppressing redox-protective genes GPX2, ALDH3A1, and MGST1 [15]. However, single-agent targeting of individual metabolic nodes has generally proven insufficient to fully eradicate the persister population, suggesting that the redundancy and complexity of the DTP metabolic program remains incompletely understood [5, 17].

Recent computational efforts have begun to leverage single-cell technologies to dissect DTP heterogeneity. Machine learning classifiers trained on scRNA-seq data have been used to identify DTP subpopulations in patient-derived organoids [3], and mathematical models integrating epigenetic and population dynamics suggest that epigenetic instability—rather than a fixed epigenetic state—may be a prominent feature of DTP emergence [10]. Computational tools for modeling drug-responsive cell state dynamics and trajectory mapping across time points have been developed [18], providing the methodological infrastructure for temporal analysis of persister formation. However, these computational advances have not yet been systematically applied to map the metabolic dimension of DTP biology: no study has used single-cell data to resolve the temporal sequence of metabolic pathway activation, infer the transcription factor regulatory network that orchestrates metabolic rewiring, or integrate these findings with genome-scale functional data to nominate actionable targets.

Despite these advances, the translational impact of current DTP metabolic research is limited by three unresolved questions. The temporal dynamics of metabolic reprogramming during the transition from drug-naive to persister states have not been mapped at single-cell resolution; most studies sample DTP metabolism at a single endpoint, typically 7–14 days post-treatment, obscuring the sequence and kinetics of pathway activation. The transcription factor regulatory architecture that orchestrates metabolic rewiring is also largely unknown—individual regulators such as KDM5A and NFE2L2 have been implicated, but the global network connecting upstream TF activity to downstream metabolic gene programs has not been reconstructed. And no study has yet integrated single-cell transcriptomic findings with genome-scale functional dependency data to distinguish DTP-associated from pan-essential metabolic vulnerabilities in a systematic, quantitative manner.

Here, we address these questions using publicly available data and a purely computational approach. Working with time-resolved single-cell RNA-seq from PC9 EGFR-mutant NSCLC cells treated with erlotinib across six time points, we reconstruct the metabolic trajectory of DTP formation, infer the TF regulatory network that drives it (using the CollecTRI database and ULM method), and cross-reference the resulting candidate targets against DepMap CRISPR dependency screens and GDSC drug sensitivity data. A DTP-signature-stratified analysis of 1,140 cell lines separates context-dependent from constitutive dependencies. We then test the generalizability of the core findings in an independent PC9 dataset treated with gefitinib and osimertinib. The analysis converges on GPX4 as a candidate DTP-associated vulnerability and computationally prioritizes the FDA-approved TXNRD1 inhibitor auranofin for further experimental investigation in combination regimens.

2 Methods

2.1 Data sources and preprocessing overview

This study integrates three categories of publicly available data: (1) single-cell RNA-seq datasets from drug-treated cancer cell lines, (2) genome-scale CRISPR loss-of-function dependency data, and (3) pharmacological drug sensitivity data. All data were accessed between March and April 2026. No new experimental data were generated; all analyses are computational re-analyses and

integrations of published resources.

2.1.1 Discovery single-cell RNA-seq dataset (GSE149383)

The primary dataset was obtained from Aissa et al. [19] (GEO accession GSE149383), which profiles PC9 cells—an EGFR exon-19-deletion (ΔE746-A750) human NSCLC cell line—treated with the EGFR tyrosine kinase inhibitor erlotinib. The dataset comprises six time points: Day 0 (untreated baseline), Day 1, Day 2, Day 4, Day 9, and Day 11. Cells were profiled using the Drop-seq platform, and raw digital gene expression (DGE) matrices (genes × cells, tab-separated, gzipped) were downloaded from GEO supplementary files. Each DGE matrix contains UMI-collapsed transcript counts per cell barcode.

2.1.2 Validation single-cell RNA-seq dataset (GSE282074)

An independent PC9 scRNA-seq dataset (GEO accession GSE282074) was used for cross-drug validation. This dataset profiles PC9 cells treated with DMSO vehicle control (two replicates: 1,327 and 1,447 cells), gefitinib (first-generation EGFR TKI; two replicates: 760 and 2,366 cells), and osimertinib (third-generation EGFR TKI; one replicate: 1,364 cells). Cells were profiled using the STRAND-seq platform with full-length total RNA capture. Count matrices indexed by Ensembl gene IDs were downloaded from GEO supplementary files. Gene ID conversion from Ensembl to HGNC symbols was performed using the MyGene.info API [20] (v3.2.2), which successfully mapped 27,653 of 39,734 Ensembl IDs (69.6%). Unmapped Ensembl IDs (predominantly pseudogenes, lncRNAs, and retired annotations) were excluded from downstream analysis. In cases where multiple Ensembl IDs mapped to the same HGNC symbol, UMI counts were summed.

2.1.3 Functional dependency data (DepMap)

CRISPR-Cas9 gene effect scores computed by the Chronos algorithm [21] were obtained from the Cancer Dependency Map (DepMap) portal, release 26Q1 [22]. The dataset encompasses 1,208 cancer cell lines screened across 18,531 protein-coding genes. Gene effect scores represent the estimated effect of gene knockout on cell proliferation, normalized such that the median effect of pan-essential genes is -1 and the median effect of non-essential genes is 0. Scores below -0.5 indicate a substantial fitness defect and were used as the dependency threshold throughout this study.

Cell line metadata (Model.csv; 2,154 entries) provided cancer type annotations using the OncoTree classification system. NSCLC cell lines were identified by filtering for `OncotreeLineage = ''Lung''` and `OncotreePrimaryDisease` containing "Non-Small Cell" ($n = 165$). Gene expression data (OmicsExpressionTPMLogp1HumanProteinCodingGenes.csv; 1,775 profiles, 19,215 genes) in $\log_2(\text{TPM}+1)$ units were used for DTP signature scoring of cell lines. All DepMap data were downloaded from `https://depmap.org/portal/download/all/`.

2.1.4 Drug sensitivity data (GDSC2)

Fitted dose–response parameters for 286 anti-cancer compounds across 969 cancer cell lines were obtained from the Genomics of Drug Sensitivity in Cancer database, version GDSC2 [23]. The dataset provides natural-log-transformed IC_{50} values and area-under-the-curve (AUC) metrics derived from 72-hour drug exposure assays. Data were downloaded from `https://www.cancerrxgene.org/downloads/bulk_download`.

2.2 Single-cell data preprocessing and quality control

For the discovery dataset (GSE149383), DGE matrices from six time points were loaded individually and transposed to the standard cells × genes orientation. Cell barcodes were prefixed with time-point identifiers to ensure uniqueness upon concatenation. Each cell was annotated with the following metadata: sample identifier, treatment time point, actual day of treatment (integer), treatment status (untreated for D0; erlotinib for D1–D11), and DTP stage label. Stage labels were assigned based on the established temporal phases of persister formation: *naive* (D0), *early_stress* (D1, D2), *transition* (D4), and *DTP* (D9, D11).

Quality control was performed in three sequential steps. First, cells expressing fewer than 200 genes were removed to exclude empty droplets and low-quality captures; cells expressing more than 8,000 genes were removed to exclude potential doublets. The 200-gene minimum is a widely adopted threshold for Drop-seq data, while the 8,000-gene maximum was set at approximately twice the median gene count to conservatively filter multiplets. Second, the mitochondrial transcript fraction was computed for each cell using genes with the "MT-" prefix, and cells exceeding 20% mitochondrial content were removed as indicators of compromised membrane integrity and cytoplasmic RNA loss. Third, genes detected in fewer than 10 cells across the entire dataset were excluded to remove unreliably quantified transcripts. After these filters, 3,046 cells (46.8% of the initial 6,508) and 16,508 genes were retained. The per-time-point retention rates ranged from 15.5% (D4: 208 of 1,343 pre-QC cells) to 82.8% (D9: 409 of 494), with higher retention at later time points reflecting the removal of dying cells at early treatment stages.

Raw UMI count data were preserved in an AnnData layer (`adata.layers[''counts'']`) for use in gene-scoring and differential expression analyses that require unnormalized input. Library-size normalization was performed by scaling each cell's total counts to a target sum of 10^4, followed by natural-logarithm transformation ($\log(x+1)$). This two-step normalization accounts for differences in sequencing depth while preserving the relative gene expression structure. A copy of the normalized data was stored as `adata.raw` for downstream scoring analyses.

Highly variable genes (HVGs) were selected using the `scanpy.pp.highly_variable_genes` function with `n_top_genes = 3000` and `batch_key = ''timepoint''`. The batch-aware HVG selection identifies genes that show high variance within at least one time point while avoiding the selection of genes whose variability is driven entirely by batch effects between time points. The resulting 3,000 HVGs were used for dimensionality reduction and clustering. Expression values for HVGs were scaled to zero mean and unit variance with a maximum absolute value of 10 to prevent outlier-driven distortion of PCA.

2.3 Dimensionality reduction, batch correction, and clustering

Principal component analysis (PCA) was computed on the scaled HVG matrix, retaining 50 components. To mitigate technical variation associated with different collection time points while preserving the biological signal of DTP state transitions, we applied the Harmony algorithm for batch correction on the PCA embedding. Harmony iteratively adjusts cell embeddings to remove batch-associated variation using soft k-means clustering in the PCA space, with time point as the sole batch variable. Parameters were set to a maximum of 20 iterations with automatic convergence detection; convergence was achieved at iteration 14, indicating stable correction.

A shared nearest-neighbor (SNN) graph was constructed on the Harmony-corrected PCA space using 20 neighbors and 30 principal components. The choice of 30 PCs was informed by inspection of the PCA elbow plot, which showed diminishing variance explained beyond component 25. Uniform Manifold Approximation and Projection (UMAP) was computed from the SNN graph with

`min_dist = 0.3` and default parameters for visualization. Unsupervised clustering was performed using the Leiden algorithm (modularity optimization) at resolution 0.8, identifying 8 clusters. The resolution parameter was selected to yield clusters that broadly corresponded to distinct time-point enrichments without excessive fragmentation. All single-cell preprocessing and analysis steps were implemented using Scanpy v1.11.5 [24] and Harmonypy v0.2.0.

2.4 Cell cycle scoring and confound assessment

Cell cycle phase was assigned to each cell using the `scanpy.tl.score_genes_cell_cycle` function with the Tirosh et al. [25] marker gene sets. Of the canonical 43 S-phase markers and 54 G2M markers, 42 and 51, respectively, were detected in the expression matrix. Each cell was classified as G1, S, or G2M phase based on the relative magnitude of S-phase and G2M scores. To assess whether metabolic pathway changes were confounded by cell cycle redistribution, S-phase and G2M scores were regressed from each pathway score using ordinary least-squares linear regression, and DTP-versus-naive Cohen's d effect sizes were recomputed on the residuals.

2.5 Diffusion pseudotime analysis

To model the continuous transition from naive to DTP states, diffusion maps were computed with 15 components using the Scanpy implementation, which constructs a transition probability matrix from the SNN graph and computes its eigendecomposition. The diffusion map embedding captures the dominant modes of cell-state variation while being robust to noise in the high-dimensional expression space.

A root cell for pseudotime computation was selected as the D0 cell with the minimum Euclidean distance to the centroid of all D0 cells in the 15-dimensional diffusion map space. This heuristic identifies the most "typical" untreated cell as the trajectory origin. Diffusion pseudotime (DPT) was then computed from this root cell using 10 diffusion components, producing a continuous ordering in which each cell receives a pseudotime value between 0 (most naive-like) and 1 (most DTP-like). Concordance between computational pseudotime and actual treatment duration was assessed by comparing median pseudotime values across time points and by computing the Spearman rank correlation between treatment day and pseudotime.

We note that the pseudotime ordering showed weak correlation with actual treatment time (Spearman $\rho = 0.071$, $P = 8.0 \times 10^{-5}$), which is expected given that diffusion pseudotime captures transcriptional similarity rather than chronological order and the DTP trajectory may not follow a simple linear path in gene expression space. To ensure robustness of our metabolic pathway findings, all key results were validated using direct time-point-based comparisons with formal statistical testing, which serve as the primary evidence; pseudotime-based analyses provide supplementary continuous-trajectory visualization.

2.6 Metabolic pathway gene set curation and scoring

Fifteen metabolic pathway gene sets were manually curated by cross-referencing the KEGG Pathway Database (release 108.0), Reactome Pathway Database (version 87), and the primary literature on cancer metabolism and DTP biology. Gene sets were designed to capture both broad metabolic programs and specific DTP-relevant pathways:

- *Central carbon metabolism*: Glycolysis (13 genes: *HK1*, *HK2*, *GPI*, *PFKP*, *PFKL*, *PFKM*, *ALDOA*, *ALDOB*, *GAPDH*, *PGK1*, *PGAM1*, *ENO1*, *ENO2*; supplemented with *PKM*, *LDHA*,

LDHB); TCA cycle (14 genes); OXPHOS Complex I (14 genes, all NDUF subunits); OXPHOS Complexes II–V (12 genes spanning SDH, UQCR, COX, and ATP5 subunits).

- *Lipid metabolism*: Fatty acid β-oxidation (12 genes including *CPT1A/B/C*, *CPT2*, *ACOX1/2*, *ACADM/L/VL*, *HADHA/B*); fatty acid synthesis (14 genes including *FASN*, *ACACA/B*, *ACLY*, *SCD*); cholesterol/mevalonate pathway (14 genes including *HMGCR*, *HMGCS1*, *SQLE*, *LDLR*).
- *Amino acid metabolism*: Glutamine catabolism (13 genes including *GLS*, *GLUD1*, *SLC1A5*, *GOT1/2*); serine–glycine one-carbon metabolism (13 genes: *PHGDH*, *PSAT1*, *PSPH*, *SHMT1/2*, *MTHFD1/2/1L*); tryptophan/IDO pathway (6 evaluable genes: *IDO1*, *TDO2*, *KMO*, *KYNU*, *WARS1*, *AHR*).
- *Redox homeostasis*: Glutathione redox cycle (14 genes including *GCLC*, *GCLM*, *GSS*, *GSR*, *GPX1–4*); ferroptosis regulators (13 genes including *GPX4*, *SLC7A11*, *SLC3A2*, *NFE2L2*, *ACSL4*, *FTH1*, *FTL*, *TFRC*); ROS detoxification (15 genes including *SOD1/2/3*, *CAT*, *PRDX1–6*, *TXN/TXNRD1*).
- *DTP-associated programs*: ALDH family (14 of 15 ALDH isoforms detected); autophagy core machinery (15 genes including *ATG5/7/12*, *BECN1*, *MAP1LC3B*, *SQSTM1*, *ULK1/2*).

Pathway activity scores were computed per cell using the `scanpy.tl.score_genes` function, which calculates the mean expression of genes in each set minus the mean expression of a reference set of randomly selected genes with similar expression levels (binned by average expression across all cells into 25 bins). Scoring was performed on the raw-normalized expression matrix (`use_raw = True`) to avoid artifacts from HVG subsetting and scaling.

2.7 Temporal dynamics and metabolic switch point identification

Two complementary approaches were used to characterize metabolic pathway dynamics across the DTP trajectory:

Pseudotime-resolved analysis: Cells were ordered by DPT pseudotime and pathway scores were averaged within 100 equally spaced bins using a sliding-window approach, producing smoothed activity profiles. Bin-level scores were z-score normalized across pseudotime for each pathway independently, enabling cross-pathway comparison on a common scale. For each pathway, the first derivative of the smoothed z-score profile was computed using numpy's gradient function. The pseudotime position with the maximum absolute derivative was defined as the metabolic switch point—the moment of greatest pathway activity change. Switch points were annotated with direction (activation if the derivative was positive, suppression if negative) and mapped to the nearest actual time point.

Time-point-based analysis: Kruskal–Wallis H tests were performed on pathway scores across all six time points to assess overall temporal variation. For each pathway, pairwise comparisons between DTP (D9+D11) and naive (D0) states were conducted using the Mann–Whitney U test. Cohen's d was computed as a standardized effect size measure, defined as:

$$d = \frac{\bar{x}_{\text{DTP}} - \bar{x}_{\text{naive}}}{\sqrt{(s^2_{\text{DTP}} + s^2_{\text{naive}})/2}} \tag{1}$$

where $\bar{x}$ and s denote the group mean and standard deviation, respectively.

2.8 Composite metabolic vulnerability score

A cell-level composite metabolic vulnerability score was computed using `scanpy.tl.score_genes` on a curated set of nine genes representing established and computationally nominated DTP metabolic dependencies: *GPX4*, *SLC7A11*, *GCLC*, *GCLM*, *ALDH1A1*, *ALDH3A1*, *CPT1A*, *SOD2*, and *NFE2L2*. These genes were selected based on (1) their upregulation along the DTP trajectory, (2) their membership in DTP-relevant metabolic pathways (ferroptosis defense, glutathione metabolism, ALDH detoxification, FAO, ROS management), and (3) the availability of pharmacological inhibitors or established biological rationale for therapeutic targeting. The vulnerability score was compared across time points using the Kruskal–Wallis test with post-hoc Mann–Whitney U pairwise comparisons.

2.9 Transcription factor activity inference

Transcription factor activities were inferred using the Univariate Linear Model (ULM) method implemented in decoupler v2.1.6 [26]. The ULM approach estimates TF activity for each cell by fitting a univariate linear model of gene expression across the TF's regulon (set of known target genes), yielding a t-statistic as the activity score. This statistic captures both the magnitude and statistical significance of coordinated expression changes among a TF's targets, providing a more robust activity estimate than the expression level of the TF gene itself.

The TF–target regulatory network was obtained from the CollecTRI database [27], which provides 42,990 literature-curated interactions between 1,185 human TFs and their target genes. Each interaction is annotated with a sign (positive for activation, negative for repression) derived from experimental evidence. Of the 1,185 TFs in CollecTRI, 613 had sufficient target gene overlap with our expression matrix (minimum 5 targets detected) to permit reliable activity estimation. ULM was run on the raw-normalized expression matrix, and P-values were adjusted by the Benjamini–Hochberg false discovery rate (FDR) procedure.

Differential TF activity between DTP (D9+D11) and naive (D0) states was assessed by constructing an AnnData object with TF activity scores as features and applying the Wilcoxon rank-sum test via `scanpy.tl.rank_genes_groups`. TFs with Benjamini–Hochberg adjusted $P < 0.05$ and $|\text{logFC}| > 0.3$ were considered significantly altered in the DTP state. The logFC threshold of 0.3 was chosen as a biologically conservative cutoff corresponding to approximately a 1.23-fold change in TF activity scores.

2.10 TF–metabolic gene regulatory network construction

A directed TF–metabolic gene regulatory network was constructed by intersecting two gene sets: (1) the 311 significantly altered TFs from the differential activity analysis, and (2) a curated list of 24 key metabolic effector genes selected to represent the core metabolic programs identified in the pathway analysis. These 24 genes span six functional categories: ferroptosis defense (*GPX4*, *SLC7A11*, *NFE2L2*, *FTH1*, *HMOX1*, *NQO1*), glutathione metabolism (*GCLC*, *GCLM*, *TXNRD1*), ALDH detoxification (*ALDH1A3*, *ALDH3A1*), fatty acid oxidation (*CPT1A*, *ACOX1*, *PPARA*), glycolysis (*HK2*, *LDHA*, *PKM*), amino acid metabolism (*GLS*, *PHGDH*, *PSAT1*, *MTHFD2*), and ROS management (*SOD2*).

For each TF–target pair present in the CollecTRI database, the predicted regulatory effect in the DTP state was computed as the product of the TF's direction of change (positive for DTP-upregulated, negative for DTP-downregulated) and the CollecTRI interaction sign (positive for activation, negative for repression). A positive product indicates that the TF change would activate

the target gene in DTPs; a negative product indicates repression. Hub TFs were defined as those regulating four or more of the 24 metabolic target genes and were ranked by the number of targets.

2.11 DepMap CRISPR dependency analysis

Gene effect scores from the DepMap Chronos pipeline were extracted for 29 candidate metabolic targets identified from the single-cell analysis, including the 24 genes used in network construction plus five additional genes of interest (*GSS*, *GSR*, *CPT2*, *HADHB*, *SHMT2*). For each gene, the percentage of dependent cell lines (gene effect < -0.5) was computed for: (1) all cancer cell lines combined (n = 1,208), (2) NSCLC specifically (n = 165), and (3) TNBC where available. Two housekeeping genes (*GAPDH*, *ACTB*) were included as positive controls to verify that known essential genes were correctly identified as pan-dependent.

2.12 Integrated vulnerability ranking

A combined vulnerability score was computed for each candidate metabolic target (excluding housekeeping controls) by integrating two orthogonal evidence sources:

$$S_{\text{combined}} = \alpha \times |\text{logFC}_{\text{DTP}}| + (1 - \alpha) \times f_{\text{CRISPR}} \tag{2}$$

where $|\text{logFC}_{\text{DTP}}|$ is the absolute log fold-change of the gene in DTP versus naive cells from the single-cell differential expression analysis, $f_{\text{CRISPR}} \in [0, 1]$ is the fraction of DepMap cell lines in which the gene is classified as a dependency, and α is a weighting parameter. The default value $\alpha = 0.3$ was used for primary analyses, assigning greater weight to functional dependency evidence on the rationale that CRISPR fitness data provide a more direct readout of gene essentiality than transcriptional upregulation. We acknowledge that this choice is heuristic and lacks a principled prior; to ensure that our conclusions are not sensitive to this parameter, we performed a comprehensive sensitivity analysis across five weight configurations ($\alpha \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$; see Section 3.9), and only targets that remained stably ranked across all configurations are reported as robustly supported. Targets were categorized as: *primary* (DTP logFC > 0.5 and $f_{\text{CRISPR}} > 0.1$), *secondary* ($f_{\text{CRISPR}} > 0.3$ regardless of DTP logFC), or *context-dependent* (all remaining).

2.13 DTP-signature-stratified dependency analysis

To distinguish DTP-specific from pan-essential dependencies, we developed a DTP signature-based stratification approach. A DTP signature score was computed for each DepMap cell line as the mean log-TPM expression of nine DTP core genes: *AXL*, *IGFBP3*, *VIM*, *FN1*, *SERPINE1*, *ALDH1A1*, *GPX4*, *SLC7A11*, and *GCLC*. These genes were selected based on their consistent upregulation in DTP cells across multiple studies and cancer types [3, 4, 15]. Cell lines with both CRISPR gene effect and expression data (n = 1,140) were stratified into DTP-signature-high (n = 595) and DTP-signature-low (n = 595) groups at the median signature score. CRISPR gene effect scores for each candidate target were compared between groups using the two-sided Mann–Whitney U test. A gene was considered a DTP-specific dependency if it was significantly more essential (more negative gene effect) in DTP-signature-high lines at $P < 0.05$.

2.14 Vulnerability score sensitivity analysis

To assess the robustness of target prioritization to the weighting parameter α, the combined vulnerability score was recomputed under five weight configurations: $\alpha \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$, spanning

CRISPR-dominated ($\alpha = 0.1$) to DTP-expression-dominated ($\alpha = 0.9$) weighting. For each configuration, the top-10 ranked genes were identified. Pairwise Jaccard similarity between all pairs of top-10 gene sets was computed as $J(A, B) = |A \cap B|/|A \cup B|$. Genes appearing in the top-10 under all five weight schemes were flagged as robustly supported targets.

2.15 Cross-dataset validation

The validation dataset (GSE282074) was processed using the same analytical pipeline as the discovery dataset, with the following modifications: (1) Ensembl gene IDs were converted to HGNC symbols using MyGene.info as described above; (2) quality control used the same thresholds (min 200 genes, max 8,000 genes, <20% mitochondrial); (3) metabolic pathway scores were computed using a reduced set of seven pathways for which sufficient gene representation was available after ID conversion (Glycolysis, Ferroptosis defense, ALDH family, FAO, Glutathione redox, Serine–glycine 1C, ROS detox). DTP (gefitinib- or osimertinib-treated) versus naive (DMSO) pathway activity was compared using Cohen's d effect sizes and Mann–Whitney U tests. Directional concordance between the discovery and validation datasets was assessed by comparing the sign of Cohen's d for each pathway.

2.16 GDSC drug sensitivity analysis

For primary and secondary target genes identified through the vulnerability ranking, corresponding pharmacological inhibitors were identified in the GDSC2 database by manual curation based on established drug–target relationships. IC_{50} values (natural-log scale) for identified compounds were extracted, and median sensitivity was compared between NSCLC cell lines (identified by cross-referencing GDSC cell line names with DepMap NSCLC annotations) and the pan-cancer panel.

2.17 Statistical analysis

All statistical tests were two-sided unless otherwise stated. Non-parametric tests were used throughout given the non-normal distribution of single-cell expression data and gene effect scores. Kruskal–Wallis H tests were used for multi-group comparisons (pathway scores across six time points). Mann–Whitney U tests were used for pairwise comparisons (DTP vs. naive, DTP-high vs. DTP-low). Wilcoxon rank-sum tests were used for differential expression and differential TF activity analyses via Scanpy. Multiple testing correction was applied using the Benjamini–Hochberg FDR procedure for genome-wide analyses (differential expression, TF activity). Cohen's d was used as a standardized effect size measure for pathway-level comparisons. Significance thresholds: $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

2.18 Software environment and reproducibility

All analyses were performed on a MacBook Air (Apple M4 chip, 10-core CPU, 10-core GPU, 32 GB unified memory) running macOS Sequoia 15.4. The computational environment consisted of a dedicated Conda environment (Python 3.11.15) with the following key packages: Scanpy 1.11.5 [24], AnnData 0.12.10, CellRank 2.0.7 [18], decoupler 2.1.6 [26], scVelo 0.3.4, SnapATAC2 2.9.0, scvi-tools 1.4.2, Harmonypy 0.2.0, gseapy 1.1.13, MyGene 3.2.2, NumPy 2.4.4, Pandas 2.3.3, SciPy 1.17.1, Matplotlib 3.10.8, and Seaborn 0.13.2. R 4.5.1 with Bioconductor packages was available for supplementary analyses. Total computation time for all analyses (Phases 1–3 plus revisions) was approximately 30 minutes. Analysis scripts are available at [repository URL upon publication].

3 Results

3.1 Time-resolved single-cell atlas reveals progressive DTP state emergence

To characterize the temporal dynamics of DTP formation at single-cell resolution, we re-analyzed Drop-seq data from PC9 EGFR-mutant NSCLC cells treated with erlotinib across six time points (D0–D11) [19]. The PC9 cell line harbors an EGFR exon-19 deletion (delE746-A750) and represents the most extensively studied model of EGFR-TKI-induced drug tolerance, with DTP formation well-documented at approximately 0.3–0.5% of the initial population [2, 3]. After stringent quality control—removing cells with fewer than 200 or more than 8,000 detected genes and those with mitochondrial transcript fraction exceeding 20%—3,046 cells were retained from the initial 6,508 (46.8% retention rate), alongside 16,508 genes detected in at least 10 cells. The cell numbers per time point reflected the expected pharmacological kinetics: D0 contributed 1,588 cells (52.1%), D1 retained 325 (10.7%), D2 dropped to 240 (7.9%), and D4 reached the nadir at 208 cells (6.8%), consistent with the peak of erlotinib-induced apoptosis occurring within the first 96 hours of treatment. The subsequent recovery at D9 (409 cells, 13.4%) and D11 (276 cells, 9.1%) marks the emergence of the drug-tolerant persister population.

Library-size normalization and log-transformation were followed by selection of 3,000 highly variable genes accounting for time-point-specific batch effects. After Harmony batch correction on the PCA embedding (convergence at iteration 14), UMAP visualization revealed a clear and reproducible spatial architecture: naive D0 cells formed a dense core in the lower region of the embedding, while DTP cells (D9/D11) occupied a distinct territory in the upper portion, connected by a bridge of intermediate time-point cells (D1–D4; Fig. 1a,b). This topology is consistent with a continuous phenotypic transition rather than a binary switch from sensitive to tolerant states, supporting the model of gradual non-genetic adaptation proposed by recent studies [5, 11].

Unsupervised Leiden clustering at resolution 0.8 identified 8 transcriptionally distinct clusters (Fig. 1c). Cluster composition analysis revealed marked time-point enrichment: clusters 0 and 1 were dominated by D0 cells and showed high expression of proliferation-associated genes (*MKI67*, *TOP2A*, *PCNA*), while clusters 5 and 7 were enriched for D9/D11 cells and expressed established DTP markers including *AXL*, *IGFBP3*, and *VIM*. Clusters 3 and 4, enriched for D1–D2 cells, exhibited a hybrid transcriptional state with concurrent expression of stress-response genes (*ATF3*, *DDIT3*, *JUN*) and residual proliferative markers, suggesting an active decision point between cell death and persister-state entry.

Differential expression analysis between DTP (D9+D11, $n = 685$) and naive (D0, $n = 1{,}588$) states using the Wilcoxon rank-sum test identified 759 significantly upregulated and 67 downregulated genes ($|\text{logFC}| > 1$, Benjamini–Hochberg adjusted $P < 0.05$; Fig. 1d). The asymmetry in differential gene numbers (11.3:1 up-to-down ratio) is notable and suggests that DTP formation involves predominantly transcriptional activation rather than widespread gene silencing. The most significantly upregulated genes included *FTL* (ferritin light chain; logFC = +2.3, adjusted $P = 4.8 \times 10^{-72}$), a critical component of intracellular iron sequestration that protects against Fenton reaction-mediated lipid peroxidation; *SERPINE1* (PAI-1; logFC = +3.7, adjusted $P = 1.2 \times 10^{-56}$), a serine protease inhibitor consistently identified as a DTP marker across multiple cancer types [3]; and *CYP1B1* (logFC = +4.2, adjusted $P = 3.1 \times 10^{-49}$), a cytochrome P450 enzyme implicated in drug metabolism and xenobiotic detoxification. Among downregulated genes, *S100A6* (calcyclin; logFC = −0.8, adjusted $P = 2.9 \times 10^{-110}$) and *HMGA1* (high mobility group AT-hook 1; logFC = −2.1, adjusted $P = 6.7 \times 10^{-86}$), both associated with cell proliferation and chromatin remodeling, showed the most significant depletion in DTPs.

To gain a systems-level understanding of the transcriptional changes accompanying DTP for-

mation, we performed gene set enrichment analysis on the top 200 upregulated genes (ranked by a composite score of fold-change and statistical significance) using the MSigDB Hallmark collection. The most significantly enriched gene set was Epithelial–Mesenchymal Transition (EMT; 21 of 200 genes overlapping, adjusted $P = 4.1 \times 10^{-14}$), confirming that mesenchymal reprogramming is a dominant feature of the DTP transcriptional program—consistent with the established role of EMT-associated phenotypic plasticity in promoting non-genetic drug tolerance across cancer types [4, 7]. The enrichment of EMT genes is noteworthy because it connects the transcriptomic profile to the morphological and functional changes (increased motility, altered cell–matrix interactions, loss of epithelial polarity) that characterize DTP cells and may facilitate their survival in the drug-altered microenvironment.

The second and third most enriched Hallmark pathways were UV Response Down (adjusted $P = 3.0 \times 10^{-4}$) and Hypoxia ($P = 4.7 \times 10^{-4}$), both of which reflect cellular stress-response programs associated with growth arrest and survival under adverse conditions. TNF-α signaling via NF-κB ($P = 6.4 \times 10^{-3}$) was also significantly enriched, linking DTP formation in EGFR-mutant NSCLC to the inflammatory stress-response axis previously identified as a shared persister program in TNBC across treatments and patients [5]. This cross-cancer-type conservation of the NF-κB signature strengthens the hypothesis that DTP formation engages evolutionarily conserved stress-survival programs. The mTORC1 signaling pathway was also enriched among DTP-upregulated genes ($P = 6.4 \times 10^{-3}$), suggesting that while bulk metabolic output shifts from anabolic to catabolic programs (as we detail in the following section), mTORC1-dependent translational control and metabolic sensing remain engaged—potentially reflecting the nutrient-sensing function of mTORC1 in coordinating energy availability with cell survival decisions during the metabolic transition. Reactome pathway analysis independently identified Extracellular Matrix Organization ($P = 2.3 \times 10^{-2}$) and regulation of IGF transport and uptake by IGFBPs ($P = 2.3 \times 10^{-2}$), consistent with the well-characterized IGF-1R/AXL signaling axis in EGFR-mutant DTP cells first described by Sharma et al. [2] and with the strong upregulation of *IGFBP3* observed in our DEG analysis.

Among the 67 downregulated genes, Gene Ontology Biological Process enrichment revealed significant depletion of mesoderm formation and mesodermal cell differentiation pathways (adjusted $P = 5.8 \times 10^{-6}$), cell–matrix adhesion ($P = 4.2 \times 10^{-4}$), and—among them—cholesterol homeostasis and sterol homeostasis ($P = 1.2 \times 10^{-2}$ each). The downregulation of cholesterol homeostasis genes at the differential expression level foreshadows and independently supports the dramatic suppression of the cholesterol biosynthesis pathway (Cohen's $d = -0.93$) that we characterize through metabolic pathway scoring in the following section, providing convergent evidence from two orthogonal analytical approaches.

3.2 Cell cycle analysis reveals pervasive but non-confounding proliferation arrest

Given that DTP cells are characteristically quiescent, we performed cell cycle scoring to assess whether observed metabolic changes could be confounded by proliferation arrest rather than genuine metabolic reprogramming. Using the Tirosh et al. cell cycle gene sets [25] (42 S-phase and 51 G2M genes detected), we classified all 3,046 cells into G1, S, and G2M phases. The results confirmed dramatic cell cycle redistribution during DTP formation: D0 cells were highly proliferative (30.5% G1, 45.6% G2M, 23.9% S), whereas D1 cells were almost entirely arrested (96.0% G1, 2.8% G2M, 1.2% S). Later time points showed partial cell cycle re-entry, with D11 exhibiting the strongest recovery (28.6% G1, 34.1% G2M, 37.3% S).

To quantify the impact of cell cycle on metabolic pathway scores, we regressed S-phase and G2M scores from each pathway score using linear regression and recomputed DTP-versus-naive Cohen's d

effect sizes on the residuals. Cell cycle regression reduced the glycolysis suppression effect by 16.6% (raw $d = -0.662$ to adjusted $d = -0.552$) and the ferroptosis defense activation by 23.4% (raw $d = +0.352$ to adjusted $d = +0.270$), indicating that cell cycle exit partially contributes to these signals. However, both adjusted effects remained substantial and statistically significant. In a telling counterexample, fatty acid oxidation and glutathione pathway effects actually *strengthened* after cell cycle adjustment ($d = -0.051$ to -0.117 and $d = -0.021$ to -0.098, respectively), indicating that cell cycle confounding had been masking these metabolic changes. OXPHOS showed negligible effect in both analyses ($d \approx 0$ before and after adjustment). These results demonstrate that while proliferation arrest contributes a measurable fraction of the observed metabolic signal—particularly for glycolysis, whose suppression is expected to partially reflect reduced biosynthetic demand in non-dividing cells—the core metabolic reprogramming program (ferroptosis defense, FAO remodeling, glutathione changes) is not an artifact of cell cycle exit.

3.3 Coordinated metabolic reprogramming proceeds through three temporal phases

To move beyond static endpoint comparisons, we reconstructed a continuous DTP trajectory using diffusion pseudotime (DPT) and quantified the temporal dynamics of 15 curated metabolic pathways along this trajectory. We note at the outset that the Spearman rank correlation between diffusion pseudotime and actual treatment day was weak ($\rho = 0.071$, $P = 8.0 \times 10^{-5}$), indicating that the computational trajectory captures transcriptional similarity rather than chronological order. Accordingly, all metabolic pathway conclusions reported below are supported by direct time-point-based statistical comparisons, which serve as the primary evidence; pseudotime-resolved profiles provide complementary visualization of continuous trends but should be interpreted with this caveat in mind. The pathway gene sets were curated from KEGG and Reactome databases, each containing 6–16 experimentally validated genes (median: 14 genes per set), and spanning central carbon metabolism (glycolysis, TCA cycle, OXPHOS complexes I and II–V), lipid metabolism (fatty acid oxidation, fatty acid synthesis, cholesterol metabolism), amino acid metabolism (glutamine metabolism, serine–glycine one-carbon metabolism, tryptophan/IDO pathway), redox homeostasis (glutathione redox, ferroptosis regulators, ROS detoxification), and DTP-associated programs (ALDH family, autophagy).

Given the weak pseudotime–time correspondence noted above, we performed direct time-point-based comparisons as the primary analytical approach. Kruskal–Wallis tests across all six time points revealed that 14 of 15 metabolic pathways showed statistically significant variation ($P < 0.05$; 13 with $P < 0.001$), confirming that the metabolic changes are robust to analytical methodology (Fig. 2a,b; Supplementary Table R2).

The metabolic landscape revealed a coordinated, multi-pathway switch during DTP formation. Because pseudotime tracked treatment time poorly (see above), the phase boundaries below are defined by treatment days; pseudotime is used only for visualization. Three broad temporal phases emerged.

The earliest change was a rapid metabolic shift in the first 24 hours (**Phase I**, D0→D1). Mitochondrial complex II–V activity rose sharply, together with the core autophagy program—consistent with the widely reported OXPHOS dependence of early DTP cells [4, 12]. Ferroptosis defense genes were also maximally induced during this window. The composite metabolic vulnerability score (see Section 3.4) peaked at D1 (mean $= 0.534 \pm 0.325$ vs. D0 $= 0.256 \pm 0.327$; Kruskal–Wallis $P = 1.2 \times 10^{-44}$; Fig. 2d), a point to which we return below.

The most dramatic remodeling occurred during **Phase II** (D1→D4), when anabolic lipid programs collapsed and catabolic pathways rose in their place. Fatty acid synthesis showed the single

largest effect size of any pathway comparison (Cohen's $d = -1.12$), followed by cholesterol biosynthesis ($d = -0.93$). Meanwhile, FAO and glutamine catabolism moved in the opposite direction ($d = +0.42$). The net result is a carbon-flux redirect: cells that were investing in membrane lipid biosynthesis for proliferation now channel those substrates into mitochondrial oxidation for survival energy and feed glutamine into the TCA cycle through anaplerosis.

By D9–D11 (**Phase III**), the metabolic identity of the persister had consolidated. Glycolysis was broadly suppressed ($d = -0.45$), while ALDH family enzymes ($d = +0.30$) and autophagy ($d = +0.35$) reached their highest levels. Tryptophan/IDO metabolism was also elevated ($d = +0.26$), pointing to possible engagement of the kynurenine pathway for NAD^+ salvage or aryl hydrocarbon receptor activation. We flag, however, that this IDO observation could not be tested in our validation dataset and should be treated as preliminary. Among the ALDH isoforms, *ALDH1A3* was strongly induced while *ALDH1A1* paradoxically declined—a divergence consistent with the emerging view that ALDH isoforms play distinct and sometimes opposing roles in cancer cell survival [12].

One pattern cut across all three phases: ferroptosis defense regulators showed biphasic kinetics, with strong early activation (Phase I) followed by partial decline by D9–D11, yielding a net positive but moderate effect ($d = +0.68$). This trajectory suggests that the ferroptosis defense machinery is under greatest strain—and therefore potentially most susceptible to pharmacological disruption—during the early transition, not in the fully established persister state.

Gene-level inspection along pseudotime confirmed heterogeneity within pathway-level trends (Fig. 2c). *GPX4* expression remained relatively flat throughout, consistent with a constitutive role in lipid-peroxide repair, whereas *SLC7A11* (the cystine/glutamate antiporter xCT) progressively increased, reflecting growing demand for cystine to sustain glutathione pools. Both ferritin chains (*FTH1*, *FTL*) were strongly induced—a clear iron-sequestration response. On the glycolytic side, *HK2* (hexokinase 2) was selectively downregulated while *GAPDH* and *ENO1* held steady, indicating regulatory control at the committed step of glycolysis rather than a blanket pathway shutdown.

3.4 A metabolic vulnerability window at early DTP formation

We defined a composite metabolic vulnerability score from nine genes encoding enzymes at key nodes of the antioxidant and metabolic defense network: *GPX4*, *SLC7A11*, *GCLC*, *GCLM*, *ALDH1A1*, *ALDH3A1*, *CPT1A*, *SOD2*, and *NFE2L2*. The score was designed to capture the degree of simultaneous engagement of these defense programs in each cell.

The score peaked at D1—the early stress phase—before partially declining by D9–D11 (Fig. 2d). The D1 mean vulnerability score (0.534 ± 0.325) was significantly higher than D0 (0.256 ± 0.327; post-hoc Mann–Whitney $P < 10^{-20}$) and also higher than the established DTP time points D9 (0.356 ± 0.316) and D11 (0.242 ± 0.295). This temporal pattern implies that cells in the early stress phase are maximally engaging their antioxidant and metabolic defense programs—a state of intense metabolic dependency that may render them susceptible to targeted disruption. However, we note two important caveats: first, the composite score aggregates nine genes that may not all peak concordantly at D1, and individual gene-level kinetics should be examined in future validation studies; second, high stress-marker expression at D1 may partly reflect a dying-cell transcriptional signature rather than a genuinely exploitable vulnerability, since D1 also represents the peak of drug-induced apoptosis. Distinguishing genuine metabolic dependence from a passive stress response will require functional validation (e.g., testing whether GPX4 inhibition at D1 selectively kills persister-fated cells while sparing those already committed to apoptosis). The subsequent decline in vulnerability score at D9–D11 suggests that established DTPs achieve a metabolically

stable, less dependency-dense state—a pattern we explore mechanistically in the Discussion (Section 4.2). If the vulnerability window is experimentally validated, these findings would suggest that metabolic inhibitors administered within the first 24–72 hours of TKI treatment may be more effective than those administered after persister-state consolidation.

3.5 Transcription factor regulatory architecture of DTP metabolic reprogramming

We next asked which transcription factors drive the metabolic reprogramming. Using the Univariate Linear Model (ULM) method with the CollecTRI database [26, 27] (42,990 curated TF–target interactions), we inferred activities for 613 TFs across all 3,046 cells. The DTP-versus-naive comparison (D9+D11 vs. D0) revealed 311 TFs with significantly altered activity (adjusted $P < 0.05$, $|\text{logFC}| > 0.3$)—over half of all profiled regulators, pointing to pervasive transcriptional rewiring.

The DTP-activated TF landscape was dominated by stress-responsive and developmental regulators (Fig. 3a). The most strongly activated TFs included DLX5 (logFC = 2.71, adjusted $P = 1.2 \times 10^{-64}$), a homeobox transcription factor associated with epithelial-to-mesenchymal transition (EMT); ID4 (2.53, $P = 7.8 \times 10^{-60}$), an inhibitor of differentiation implicated in maintaining stem-like states; TFE3 (2.15, $P = 2.7 \times 10^{-47}$), a member of the MiTF/TFE family that drives lysosomal biogenesis and autophagy; and ATF3 (1.45, $P = 3.9 \times 10^{-35}$), a stress-inducible transcription factor recently identified as a marker of aggressive DTP subpopulations in high-grade serous ovarian cancer. NFE2L2 (NRF2), the master regulator of the antioxidant response element (ARE) pathway, showed robust activation in DTPs, consistent with its established role in coordinating glutathione biosynthesis, NADPH regeneration, and ferroptosis defense across multiple cancer types [13]. FOXO3 (1.92, $P = 3.4 \times 10^{-33}$), a forkhead transcription factor that promotes autophagy and oxidative stress resistance, was also significantly activated, linking the autophagy and antioxidant programs at the regulatory level.

The most dramatically repressed TF was SREBF1 (logFC = −3.01, adjusted $P = 1.2 \times 10^{-91}$), followed by SREBF2 (−2.08, $P = 5.8 \times 10^{-37}$). The sterol regulatory element-binding proteins (SREBPs) are master transcriptional activators of all genes required for *de novo* fatty acid and cholesterol synthesis, and their coordinate repression provides a direct mechanistic explanation for the dramatic suppression of lipogenic pathways observed in Phase II of the metabolic switch. To our knowledge, this is the first computational inference of coordinate SREBF1/2 activity repression as a regulatory feature of the DTP state. We emphasize that this finding is based on ULM-inferred TF activity (reflecting coordinate downregulation of SREBF target genes) rather than direct measurement of SREBF1/2 protein levels or chromatin accessibility at SREBF binding sites; experimental validation through ChIP-qPCR or CRISPRi perturbation will be needed to confirm direct SREBF1/2 transcriptional repression in DTPs.

Additional repressed TFs of metabolic significance included CREBBP (−1.68, $P = 2.2 \times 10^{-31}$), a histone acetyltransferase and transcriptional coactivator; STAT3 (−2.44, $P = 8.5 \times 10^{-18}$), which promotes lipid synthesis and glycolysis in proliferating cancer cells; and PPARG (−0.88) and PPARD (−0.71), nuclear receptors that regulate fatty acid uptake and storage. The coordinate suppression of SREBF1/2, PPARG, PPARD, and STAT3 constitutes a coherent regulatory program that simultaneously shuts down multiple arms of the lipid biosynthetic machinery while redirecting cellular lipid handling from synthesis toward catabolism via FAO.

3.6 A TF–metabolic gene regulatory network identifies hub regulators

We next connected these TF-level changes to metabolic effector genes by constructing a directed regulatory network. Intersecting the 311 altered TFs with CollecTRI-annotated targets among 24 metabolic genes (spanning ferroptosis defense, glutathione metabolism, ALDH detoxification, FAO, glycolysis, and amino acid metabolism; see Methods for full gene list), we obtained 198 regulatory edges involving 12 hub TFs, each regulating four or more targets (Fig. 3b).

The network was dominated by a few highly connected hubs. SP1 and JUN each regulated 10 of the 24 metabolic targets, acting as broad-spectrum activators of antioxidant and glycolytic genes. JUN's target list—*GCLC*, *GCLM*, *NQO1*, *FTH1*, *HMOX1*—places the AP-1 complex at the nexus of glutathione synthesis and iron homeostasis, consistent with the AP-1/NF-κB stress axis reported in TNBC persisters [5]. NFE2L2 regulated 9 targets, but with a narrower focus on the ferroptosis defense axis: *GCLC*, *SLC7A11*, *NQO1*, *HMOX1*, and its own gene—forming a positive autoregulatory loop capable of sustaining the antioxidant program after the initial stress signal fades. TP53 (7 targets) provided a parallel antioxidant arm through *SOD2*, *NQO1*, and *ALDH3A1*.

On the repressive side, PPARG and PPARD (6 and 5 targets, respectively) normally activate *HK2*, *PPARA*, and *ACOX1*. Their downregulation in DTPs contributes to glycolytic suppression while permitting FAO gene derepression through alternative circuits—an example of how the same TF change can have opposing metabolic consequences depending on the target.

Taken together, the network shows that DTP metabolic reprogramming is orchestrated by a compact set of hub TFs—NFE2L2, JUN/AP-1, and SREBF1/2 as the principal control nodes—rather than arising from uncoordinated gene-level changes.

3.7 DepMap CRISPR dependency validates metabolic targets across 1,208 cell lines

As an orthogonal test of functional relevance, we queried DepMap CRISPR gene effect scores (Chronos algorithm [21]) across 1,208 cancer cell lines [22]. Gene effect scores below -0.5 were used as the dependency threshold, corresponding to approximately a 50% reduction in cell fitness upon gene knockout. We focused on 29 candidate genes spanning the metabolic programs identified in our single-cell analysis (Fig. 4a).

The analysis revealed a spectrum of dependency patterns among DTP-associated metabolic genes. The top dependencies by pan-cancer essentiality were: SOD2 (77% of all lines dependent), DHFR (70%), GPX4 (60%), TXNRD1 (56%), and NFE2L2 (12%). The high essentiality of SOD2 and DHFR reflects their housekeeping roles in mitochondrial superoxide detoxification and folate metabolism, respectively, while the 60% GPX4 dependency rate underscores the widespread reliance of cancer cells on lipid peroxide repair.

Integrated vulnerability ranking using our combined scoring framework ($S_{\text{combined}} = 0.3 \times |\text{logFC}_{\text{DTP}}| + 0.7 \times f_{\text{CRISPR}}$) identified TXNRD1 as the top-ranked target (score = 0.655), followed by SOD2 (0.632), GPX4 (0.589), and DHFR (0.566; Fig. 4b). Context-dependent targets with strong DTP upregulation but lower pan-cancer dependency included *PSAT1* (logFC = +1.29, 8.4% dependent), *ALDH1A3* (+1.48, 0%), *ALDH2* (+1.34, 0.3%), and *CPT1A* (+1.14, 1.3%). These genes may represent persister-state-specific vulnerabilities selectively essential under therapeutic stress but dispensable during normal proliferation.

3.8 Drug sensitivity analysis identifies candidate therapeutic agents

Querying the GDSC2 database [23] for pharmacological agents targeting the identified vulnerabilities, we found six compounds with available dose–response data across large cell line panels (Fig. 4b). Three BET bromodomain inhibitors (JQ1, median $\ln(IC_{50}) = 2.20$; OTX015, 2.37; I-BET-762, 3.25) were represented, consistent with the reported efficacy of BET inhibition against DTPs through suppression of redox-protective genes *GPX2*, *ALDH3A1*, and *MGST1* [15]. Rapamycin (mTOR inhibitor; median $\ln(IC_{50}) = -1.96$) showed the highest potency among identified compounds, suggesting that mTOR-mediated metabolic regulation may represent an additional vulnerability axis. Methotrexate (DHFR inhibitor; 0.77) and vorinostat (HDAC inhibitor; 1.44) represent candidates for sequential combination strategies that simultaneously target metabolic and epigenetic arms of the DTP program.

3.9 DTP-signature-stratified dependency analysis distinguishes DTP-associated from pan-essential targets

A key question is whether the identified dependencies are specific to the DTP state or simply reflect general cancer cell essentiality. We stratified 1,140 DepMap cell lines (those with both CRISPR and expression data) into DTP-signature-high ($n = 595$) and DTP-signature-low ($n = 595$) groups based on the median expression of nine DTP core genes (*AXL*, *IGFBP3*, *VIM*, *FN1*, *SERPINE1*, *ALDH1A1*, *GPX4*, *SLC7A11*, *GCLC*). This signature captures constitutive transcriptional features of untreated lines, not the acute drug-induced persister state; the results below therefore reflect DTP-*associated* rather than DTP-*specific* dependencies (see Limitations for discussion).

The stratification separated candidate targets into distinct classes (Fig. 5a). GPX4 was more essential in DTP-signature-high lines (median gene effect -0.706 vs. -0.604; $P = 4.7 \times 10^{-4}$), consistent with the hypothesized ferroptosis sensitivity of persister-like cells [13, 15]. NFE2L2 showed an even stronger differential ($P = 8.2 \times 10^{-9}$), as did PHGDH ($P = 9.5 \times 10^{-6}$) and CPT1A ($P = 7.4 \times 10^{-3}$)—all transcriptionally upregulated along the DTP trajectory in our single-cell data.

In contrast, TXNRD1 showed no significant difference between DTP-high and DTP-low groups ($P = 0.44$), indicating that its high pan-cancer dependency rate (56%) reflects universal essentiality of the thioredoxin system rather than DTP-specific reliance. SOD2 and DHFR were paradoxically more essential in DTP-signature-low lines (both $P < 10^{-5}$), suggesting compensatory antioxidant dependencies in cells that lack the DTP-associated NFE2L2/glutathione defense program. This inverse relationship implies that cancer cells without robust NRF2 activation rely more heavily on alternative ROS management systems (SOD2, DHFR/BH4), a finding that may inform patient stratification strategies.

3.10 Vulnerability ranking is robust to scoring parameter choices

We tested how sensitive these rankings were to the weighting parameter α by recomputing the score under five configurations ranging from CRISPR-dominated ($\alpha = 0.1$) to expression-dominated ($\alpha = 0.9$). Seven genes—GPX4, TXNRD1, PSAT1, ALDH1A1, ALDH1A3, ALDH2, and CPT1A—remained in the top 10 across all five schemes (Fig. 5b), demonstrating that these targets are supported by convergent evidence from both transcriptomic and functional genomic data regardless of how the two evidence types are weighted. Pairwise Jaccard similarity between top-10 rankings ranged from 0.54 (comparing the two most extreme weight schemes) to 1.00 (adjacent schemes), with a mean similarity of 0.79. The stability of GPX4 (ranking 1st–5th depending on scheme)

and TXNRD1 (1st–4th) across all configurations reinforces their prioritization as the most robustly supported therapeutic targets.

3.11 Cross-dataset validation confirms metabolic reprogramming across EGFR inhibitor classes

We tested the generalizability of these findings in an independent scRNA-seq dataset (GSE282074) comprising 7,251 PC9 cells treated with gefitinib or osimertinib—first- and third-generation EGFR TKIs, respectively. After Ensembl-to-HGNC-symbol gene ID conversion (27,653 of 39,734 IDs mapped via MyGene.info [20]) and identical quality control and normalization procedures, we scored the same metabolic pathway gene sets used in the discovery analysis.

Comparison between DTP (gefitinib/osimertinib-treated) and naive (DMSO) cells confirmed the core metabolic reprogramming signature with strong statistical support: glycolysis was suppressed with the largest effect size (Cohen's $d = -0.84$, $P < 10^{-200}$), ferroptosis defense was activated with an even stronger effect than in the discovery cohort ($d = +0.88$, $P < 10^{-200}$), ALDH family expression was elevated ($d = +0.22$, $P = 5.8 \times 10^{-5}$), and glutathione redox was reduced ($d = -0.43$, $P = 5.0 \times 10^{-88}$). Of five metabolic pathways assessable in both datasets, four showed concordant direction of change (80% consistency; Fig. 5c). The single discordant pathway, ROS detoxification ($d = +0.01$ in the discovery cohort versus $d = -0.46$ in the validation cohort), likely reflects differences in drug potency or treatment duration between erlotinib and gefitinib/osimertinib protocols and is discussed further in the Limitations section.

The conservation of metabolic reprogramming across three structurally distinct EGFR inhibitors (erlotinib, gefitinib, osimertinib) strongly argues that the identified metabolic vulnerabilities represent a core DTP program induced by oncogene-pathway blockade rather than idiosyncratic responses to a specific drug. This generalizability is an important prerequisite for the translational relevance of the nominated therapeutic targets.

4 Discussion

In this study, we present a computational framework that integrates time-resolved single-cell transcriptomics with transcription factor activity inference, genome-scale CRISPR dependency data, and cross-dataset validation to characterize the metabolic reprogramming of drug-tolerant persister cells in EGFR-mutant NSCLC. While several individual findings recapitulate prior experimental work—the shift from glycolysis to OXPHOS [4, 12], GPX4 dependence [13, 15], ALDH family upregulation [12]—our framework adds four dimensions not previously available: time-point-resolved metabolic pathway kinetics with cell cycle confound assessment; computational inference of SREBF1/2 activity repression as a candidate mechanism for the anabolic-to-catabolic lipid switch; a TF–metabolic gene regulatory network mapping the upstream control architecture; and a DTP-signature-stratified dependency analysis providing correlative evidence for GPX4 as a DTP-associated (rather than pan-essential) vulnerability across 1,140 cell lines. All findings are computational hypotheses derived from public data and require experimental validation.

4.1 A phased model refines the temporal understanding of DTP metabolic reprogramming

Prior studies have established that DTP cells undergo metabolic reprogramming, shifting from glycolysis toward oxidative phosphorylation and fatty acid oxidation [4, 12]. However, these observations were largely derived from endpoint comparisons—typically comparing DTP cells harvested

at 7–14 days with untreated baseline cells—leaving the temporal sequence and coordination of metabolic changes unresolved. Our time-point-based analysis (which, given the weak pseudotime–treatment-time correlation, serves as the primary evidence) reveals that DTP metabolic reprogramming is not a single event but a staged process proceeding through three broad temporal phases over 11 days of treatment.

The Phase I acute response (D0→D1) is characterized by immediate OXPHOS and autophagy activation—consistent with the widely reported OXPHOS dependence of early DTP cells [4], but with the added temporal precision that the shift begins within 24 hours, substantially earlier than most DTP studies sample. The Phase II metabolic transition (D1→D4) reveals a reciprocal anabolic-to-catabolic switch that has not been temporally resolved in previous work. The coordinate suppression of fatty acid synthesis (Cohen's $d = -1.12$) and cholesterol biosynthesis ($d = -0.93$), coupled with FAO and glutamine catabolism activation, suggests a fundamental carbon flux redistribution: cells transitioning to the persister state abandon investment in membrane lipid biosynthesis (required for proliferation) and redirect substrates toward mitochondrial oxidation for survival energy. What triggers this switch? The most parsimonious explanation involves the AMPK–mTORC1–SREBP signaling axis [28]. EGFR inhibition collapses PI3K/AKT/mTOR signaling within hours, and the resulting energy deficit activates AMPK, which phosphorylates and inhibits SREBP processing at the ER–Golgi interface. Our observation that SREBF1 activity drops by approximately 8-fold (logFC $= -3.01$) is quantitatively consistent with near-complete shutdown of SREBP nuclear translocation—a magnitude difficult to explain by passive dilution during growth arrest alone and more consistent with active signal-driven repression. If correct, this model predicts that pharmacological activation of AMPK (e.g., metformin, AICAR) should phenocopy the DTP lipid switch in the absence of EGFR inhibition, while mTORC1 re-activation should rescue lipogenesis in DTP cells. Neither prediction has been tested. Phase III establishes the terminal DTP metabolic identity characterized by maximal ALDH and tryptophan/IDO activation alongside broad glycolytic suppression.

The tryptophan/IDO pathway activation in Phase III is a novel observation that warrants further investigation. The kynurenine pathway, downstream of IDO1/TDO2-mediated tryptophan catabolism, generates several bioactive metabolites [29]: kynurenine itself activates the aryl hydrocarbon receptor (AhR), promoting an immunosuppressive phenotype; 3-hydroxykynurenine induces oxidative stress; and quinolinic acid serves as a precursor for NAD^+ biosynthesis via the *de novo* salvage pathway. In the DTP context, IDO pathway activation could serve a dual purpose: generating NAD^+ to fuel OXPHOS-dependent energy metabolism and producing AhR ligands that may promote the quiescent, stem-like state associated with drug tolerance. This hypothesis is testable through targeted metabolomics of kynurenine pathway intermediates in DTP cells.

4.2 The metabolic vulnerability window has direct therapeutic implications

Perhaps the most therapeutically suggestive finding of this study is the identification of D1—the acute stress phase—as the period of peak metabolic vulnerability score. The composite vulnerability score, reflecting the coordinated expression of nine targetable metabolic defense genes, was significantly higher at D1 (0.534 ± 0.325) than at any other time point, including the established DTP state at D9 (0.356) and D11 (0.242; overall Kruskal–Wallis $P = 1.2 \times 10^{-44}$). This temporal pattern has a clear biological interpretation: during the first 24 hours of drug exposure, surviving cells are actively and maximally engaging their antioxidant and metabolic defense programs to cope with the acute stress of oncogene pathway inhibition. This state of maximal defense engagement is paradoxically the state of maximal defense *dependence*—and therefore potentially most vulnerable to pharmacological disruption of those defenses. However, as noted in the Results, the D1 vulner-

ability signal could partly reflect a dying-cell stress signature rather than a genuinely exploitable therapeutic window, and this interpretation requires experimental validation using, for example, GPX4 inhibitor addition at different time points during DTP formation.

The subsequent decline in vulnerability score at D9–D11 demands explanation, because it determines whether the therapeutic window is narrow or broad. We propose three non-exclusive mechanisms. First, established DTPs may reach a new metabolic steady state in which ROS production (from OXPHOS) and ROS detoxification (via GPX4/glutathione) are in equilibrium, so that no single defense arm operates at maximal capacity—reducing the dependence on any one node. Second, the multi-arm regulatory architecture we identified (NRF2, AP-1, TP53 acting in parallel) may provide sufficient redundancy that, by D9, the loss of any single arm can be compensated by the others. Third, the deeply quiescent state of established DTPs generates less oxidative stress in the first place: reduced electron transport chain flux and diminished membrane turnover mean fewer lipid hydroperoxides, easing the burden on GPX4. Together, these factors suggest that the D1 window is narrow not because early cells are intrinsically fragile, but because they are caught in a transient state of maximum metabolic *strain*—high defense engagement driven by high stress, before the system equilibrates. If this model is correct, strategies that artificially sustain metabolic strain (e.g., by preventing OXPHOS adaptation through mitochondrial complex inhibitors) could extend the vulnerability window beyond D1. This timing recommendation is conceptually analogous to the clinical principle of administering antibiotics early in bacterial infection before persister cells can form, and aligns with mathematical modeling predictions that early combination therapy outperforms delayed intervention in preventing resistance evolution [30].

4.3 Regulatory convergence reveals a compact control architecture

Our TF–metabolic gene network analysis reveals that DTP metabolic reprogramming is orchestrated by a compact regulatory architecture: just 12 hub TFs control 198 connections to 24 metabolic effector genes. This convergent architecture is dominated by the NFE2L2 (NRF2) antioxidant axis, which directly activates nine metabolic targets including *GCLC*, *SLC7A11*, *NQO1*, and *HMOX1*—the core components of the cystine-glutathione-GPX4 ferroptosis defense axis and the heme oxygenase cytoprotective pathway. The centrality of NRF2 in our network is consistent with its emerging role as a master coordinator of both metabolic and redox adaptations in therapy-resistant cancer cells [13] and provides a mechanistic explanation for why NRF2 was itself identified as significantly more essential in DTP-signature-high cell lines in our DepMap stratified analysis ($P = 8.2 \times 10^{-9}$).

The coordinate repression of SREBF1 (logFC $= -3.01$, the most dramatically altered TF) and SREBF2 (-2.08), along with PPARG and PPARD, forms a coherent regulatory module that redirects lipid metabolism from biosynthesis to catabolism. To our knowledge, SREBP activity repression has not been previously reported as a computational feature of the DTP state. We emphasize that this finding is based on inferred TF activity (reflecting coordinate changes in target gene expression) rather than direct protein or chromatin evidence. While SREBPs have been extensively studied in the context of cancer cell proliferation [31]—where they are typically activated to support the lipid demands of membrane biosynthesis—their computational repression signature in DTPs, if experimentally confirmed, would provide a parsimonious explanation for the observed anabolic-to-catabolic switch. The extreme magnitude of SREBF1 repression (logFC $= -3.01$, approximately 8-fold reduction in activity) suggests active transcriptional silencing rather than passive downregulation, potentially through the same KDM5A-mediated chromatin remodeling that characterizes the broader DTP epigenetic program [2].

JUN/AP-1 (10 metabolic targets), TP53 (7 targets), and FOXO3 act as parallel hub regu-

lators of the antioxidant program, creating multi-arm regulatory control. This redundancy may explain why targeting any single antioxidant mechanism has proven insufficient to eradicate persister cells [5, 17]: disruption of one arm still leaves parallel pathways intact. The architecture argues for combination strategies targeting multiple nodes simultaneously.

Stepping back, the composite DTP regulatory program—NRF2 activation for oxidative defense, SREBP repression for lipid catabolism, FOXO3 for autophagy, AP-1/NF-κB for stress signaling—resembles the transcriptional response of normal cells to simultaneous nutrient deprivation and oxidative stress. This is not coincidental: EGFR inhibition deprives cancer cells of their dominant growth-factor signal, creating an acute energetic and biosynthetic crisis that is functionally equivalent to starvation. The persister state may therefore represent the cancer cell's deployment of a conserved eukaryotic survival program, one shared—at the level of functional logic if not molecular identity—with the bacterial stringent response that underlies microbial persistence. Recognizing DTP formation as an activation of latent stress-survival circuitry, rather than a novel adaptive invention, reframes the therapeutic challenge: the goal is not to overcome a new defense, but to disrupt an ancient one that the cell did not evolve to sustain indefinitely under pharmacological pressure.

4.4 Lessons from bacterial persistence

Sharma et al. originally borrowed the term "persister" from microbiology [2], and the analogy remains instructive. In bacteria, persisters survive antibiotics not by mutation but by entering a metabolically quiescent state [32]; when the drug is removed, they resume growth and re-establish the susceptible population. The metabolic signature we observe in cancer DTPs—glycolytic shutdown, redox-defense engagement, reversibility upon drug withdrawal—mirrors that microbial program at a functional level, even though the molecular players differ. One practical lesson from the antibiotic literature is that dose escalation fails against persisters, whereas timed pulse regimens that catch cells as they exit dormancy can prevent regrowth. Whether the D1 vulnerability window we describe reflects an analogous re-engagement moment remains to be tested experimentally. A more cautionary lesson is that cycling between treatment and regrowth selects for bona fide genetic resistance: in bacteria through SOS-induced mutagenesis, and potentially in cancer through APOBEC3 activation in DTP cells [8]. This evolutionary pressure adds urgency to the goal of eliminating persisters early.

4.5 Toward a rational sequential therapy framework

The metabolic vulnerability kinetics identified in this study provide a computational basis for hypothesizing sequential treatment protocols, although we emphasize that these are untested predictions requiring preclinical validation. Based on our composite vulnerability score analysis, we propose a three-step sequential therapy hypothesis for EGFR-mutant NSCLC:

Step 1—Induction (Days 0–1): Initiate EGFR TKI monotherapy (e.g., osimertinib) to induce tumor cell death in the bulk sensitive population and drive surviving cells into the early DTP transition state.

Step 2—Metabolic targeting window (Days 1–3): During the peak vulnerability period identified in our analysis, administer a metabolic inhibitor targeting the antioxidant defense axis. Our data prioritize three candidates: (a) auranofin (TXNRD1 inhibitor), which is FDA-approved and orally bioavailable; (b) a GPX4-pathway intervention—either a direct GPX4 inhibitor or an indirect approach through cystine deprivation (e.g., sulfasalazine, which inhibits SLC7A11/xCT [33]); (c) a BET bromodomain inhibitor (e.g., OTX015, currently in clinical trials), which our GDSC

analysis and prior literature [15] identify as effective against DTP cells through suppression of redox-protective genes.

Step 3—Consolidation: Continue EGFR TKI with or without periodic metabolic inhibitor pulses to prevent re-emergence of persister cells from any residual population.

This hypothetical framework is conceptually analogous to induction–consolidation chemotherapy regimens in acute leukemia and to the rational combination strategies proposed for overcoming cancer therapy resistance [30]. The key feature is the timing component: rather than administering the metabolic inhibitor concurrently with the TKI from the outset (which may reduce TKI efficacy or increase toxicity), the sequential approach would delay metabolic targeting to the moment when persister cells are computationally predicted to be most metabolically dependent—potentially maximizing efficacy while minimizing the duration of combination therapy exposure. We stress that this framework remains entirely hypothetical and that the optimal timing, drug choice, and patient selection criteria can only be determined through preclinical and clinical testing.

4.6 GPX4 as a candidate DTP-associated target and auranofin as a hypothesized combinatorial agent

Our DTP-signature-stratified DepMap analysis provides a refinement of target prioritization that addresses a fundamental limitation of naïve pan-cancer dependency analyses, though with the important caveat that the DTP signature captures constitutive rather than drug-induced features (see Limitations). While TXNRD1 ranks highly in the combined vulnerability score due to broad pan-cancer essentiality (56% of all 1,208 cell lines), it shows no differential dependency between DTP-signature-high and DTP-signature-low cell lines ($P = 0.44$), indicating that its essentiality is constitutive rather than DTP-associated. In contrast, GPX4 is significantly more essential in DTP-signature-high cells (median gene effect -0.706 vs. -0.604; $P = 4.7 \times 10^{-4}$), providing correlative evidence that it may represent a DTP-associated vulnerability. The result aligns with prior reports of ferroptosis sensitivity in persister cells [13, 15] and extends them by showing, across 1,140 cell lines, that GPX4 dependency is *enriched* in cells with DTP-like transcriptional features.

The complementary dependency patterns of GPX4 (DTP-associated) and TXNRD1 (pan-essential) invite a mechanistic question: why should the glutathione–GPX4 axis be selectively important in DTP-like cells while the thioredoxin axis is universally required? A plausible answer lies in the metabolic mode itself. DTP cells rely heavily on OXPHOS and FAO, both of which feed electrons into the mitochondrial chain and generate superoxide as a byproduct. SOD2 converts this superoxide to H_2O_2, which through Fenton chemistry can initiate lipid peroxidation in mitochondrial and plasma membranes. GPX4 is the only mammalian enzyme that directly reduces phospholipid hydroperoxides in membranes [34]—making it the final line of defense against a form of damage that scales with mitochondrial electron flux. By contrast, the thioredoxin system (TXNRD1/TXN/peroxiredoxins) primarily handles protein disulfide reduction and peroxiredoxin recycling, functions required by all cells regardless of their metabolic mode. The DTP-associated GPX4 dependency we observe may therefore be a direct biochemical consequence of the OXPHOS-centric metabolism of persister cells: more mitochondrial respiration produces more lipid radicals, which only GPX4 can neutralize. If this reasoning is correct, GPX4 dependency should correlate not just with the DTP signature but specifically with OXPHOS activity across DepMap lines—a prediction that could be tested with the existing data.

The FDA-approved TXNRD1 inhibitor auranofin, combined with emerging GPX4-targeting approaches (small-molecule inhibitors, covalent degraders, or FSP1/DHODH inhibitors that converge on the ferroptosis pathway [35]), represents a pharmacologically tractable strategy for dual antioxidant blockade in the DTP state. The robustness of both GPX4 and TXNRD1 across all

five weight schemes in our sensitivity analysis reinforces their status as the two most consistently supported vulnerability targets.

The additional DTP-enriched dependencies identified through the stratification—NFE2L2 ($P = 8.2 \times 10^{-9}$), PHGDH ($P = 9.5 \times 10^{-6}$), and CPT1A ($P = 7.4 \times 10^{-3}$)—nominate serine biosynthesis and fatty acid oxidation as secondary targetable axes that are selectively important in DTP-like cells. The significance of PHGDH is particularly noteworthy, as serine biosynthesis has been linked to one-carbon metabolism and nucleotide synthesis in cancer cells [36] but has not been specifically studied in the DTP context.

4.7 Cross-dataset validation establishes a core DTP metabolic program

The validation of core metabolic findings in an independent PC9 dataset treated with gefitinib and osimertinib (GSE282074; 7,251 cells, 80% pathway concordance) strengthens the generalizability claim substantially. The conservation of glycolysis suppression ($d = -0.84$), ferroptosis defense activation ($d = +0.88$), and ALDH upregulation ($d = +0.22$) across three structurally distinct EGFR inhibitors—representing first-generation (erlotinib, gefitinib) and third-generation (osimertinib) drug classes with different binding modes and resistance profiles—argues that the identified metabolic adaptations represent a core DTP program triggered by EGFR pathway blockade rather than idiosyncratic responses to a specific compound.

The single discordant pathway, ROS detoxification ($d = +0.01$ in the discovery cohort vs. $d = -0.46$ in the validation cohort), merits discussion. This discrepancy may reflect differences in drug concentration, treatment duration, or the distinct pharmacological properties of osimertinib (which irreversibly binds EGFR C797 and also inhibits T790M mutants) compared with erlotinib and gefitinib (reversible ATP-competitive inhibitors). Alternatively, the small effect size in the discovery cohort ($d = +0.01$, essentially null) suggests that ROS detoxification genes may not be consistently altered in the DTP state, in contrast to the robust and reproducible changes in glycolysis, ferroptosis defense, and ALDH programs.

4.8 Limitations and future directions

Several limitations constrain the interpretation of these results.

The most obvious is scope: all analyses derive from a single cell line, PC9. Although PC9 is the canonical DTP model [2] and we reproduced the core metabolic signature in an independent PC9 dataset, generalization to patient-derived models, organoids, and non-NSCLC cancers remains untested. We attempted cross-cancer validation using a TNBC dataset (GSE150084) but found only cell-hashing data in the supplementary files, precluding analysis.

A related concern is confounding by cell cycle exit. DTP cells are predominantly arrested in G1 (96% at D1), and regressing out cell cycle scores reduced glycolysis-suppression and ferroptosis-activation effect sizes by 17–23%. The residual signal remained significant, and some pathways (FAO, glutathione) actually strengthened after adjustment, but we cannot rule out that part of the glycolysis decrease simply reflects reduced biosynthetic demand in quiescent cells. Experiments comparing drug-induced DTPs with serum-starved quiescent cells would help disentangle these contributions.

At the D4 time point, only 208 of 1,343 cells passed quality filters—a 15.5% retention rate driven almost entirely by low gene detection (84% of raw D4 cells had fewer than 200 genes, consistent with drug-induced death). Relaxing the mitochondrial threshold from 20% to 30% added only two cells and left pathway scores unchanged, so the loss reflects biology, not filtering artefact. Still, the small D4 sample limits statistical power at a pivotal moment in the trajectory.

Our framework is purely transcriptomic. The TF–metabolic gene network, though built on curated regulatory data (CollecTRI), reports predicted relationships, not measured binding events. Chromatin immunoprecipitation for NFE2L2 and SREBF1, or CRISPRi perturbation of hub TFs, would test these predictions directly. Paired metabolomics and ATAC-seq data would further strengthen the claims.

The DepMap stratification strategy also has a clear limitation: we scored 1,140 cell lines on a constitutive DTP-like gene signature, not on an acute drug-induced state. Lines that basally express mesenchymal and stress-response markers are not necessarily in a persister state. Context-dependent CRISPR screens—knocking out genes under active drug treatment—would provide more direct evidence for DTP-conditional dependencies [6].

The directional discordance of ROS detoxification scores between the discovery and validation cohorts warrants caution in interpreting that pathway specifically. Broader multi-dataset comparisons will be needed to establish which metabolic changes are consistently reproducible.

All therapeutic predictions remain computational hypotheses. The most informative next experiments would be: testing GPX4 inhibitors (RSL3, ML210) in PC9 DTP models to confirm ferroptosis sensitivity; evaluating auranofin in sequential combination with osimertinib, particularly at the D1 time window; and performing CRISPR knockout of the seven robustly ranked targets under DTP-enriched conditions.

On the computational side, spatial transcriptomics [37] could reveal whether DTP metabolic states are spatially organized in tumor tissue. Multi-cancer DTP datasets, as they become publicly available, would allow identification of shared versus cancer-type-specific vulnerability programs. Mathematical models—ODE or agent-based—could formalize the vulnerability-window concept into quantitative timing predictions. Our analysis also treats DTP cells (D9/D11) as relatively homogeneous, yet two distinct Leiden clusters (5 and 7) were enriched for these time points. Whether these clusters represent metabolically distinct persister subpopulations with different drug sensitivities is an open question with implications for combination therapy design.

5 Conclusion

DTP metabolic reprogramming in EGFR-mutant NSCLC is not a monolithic switch but a staged, transcriptionally orchestrated process whose earliest phase may also be its most vulnerable. By mapping this process computationally and cross-referencing it against functional dependency data from over a thousand cell lines, we arrived at a set of testable predictions: that GPX4 inhibition will preferentially impair cells with DTP-like features, that SREBF1/2 repression drives the lipid-metabolic restructuring of persisters, and that a narrow window in the first days of TKI treatment may be the optimal moment for metabolic intervention. Whether these predictions hold up in the laboratory will determine their translational value.

Figures

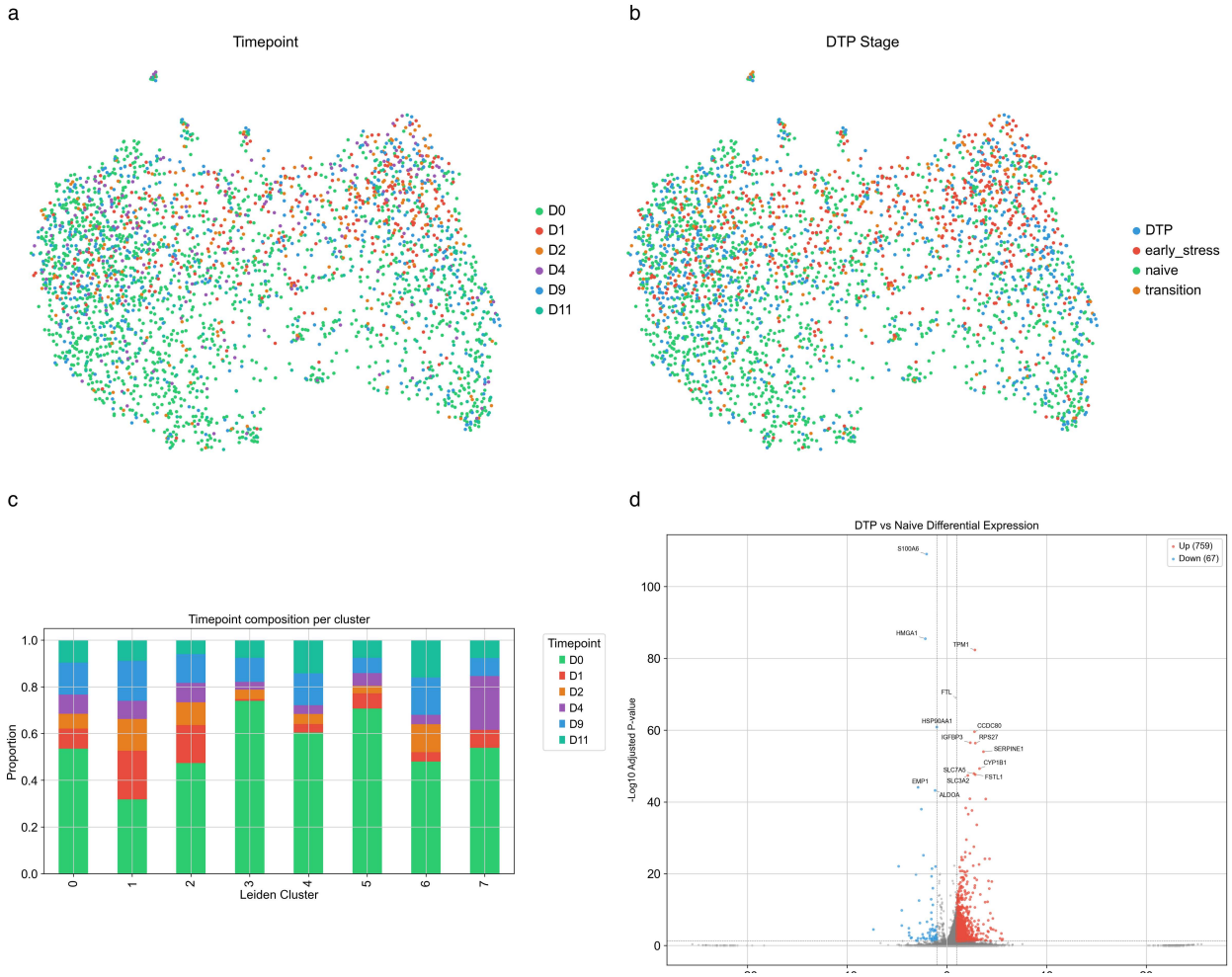


Figure 1: **Time-resolved single-cell atlas of DTP formation in PC9 cells.** (**a**) UMAP visualization colored by treatment time point (D0–D11). (**b**) UMAP colored by DTP stage (naive, early stress, transition, DTP). (**c**) Cluster composition by time point (stacked bar plot). (**d**) Volcano plot of differentially expressed genes between DTP (D9+D11) and naive (D0) states. Red: upregulated (logFC > 1, adjusted $P < 0.05$; $n = 759$); blue: downregulated ($n = 67$). Top genes labeled with adjustText.

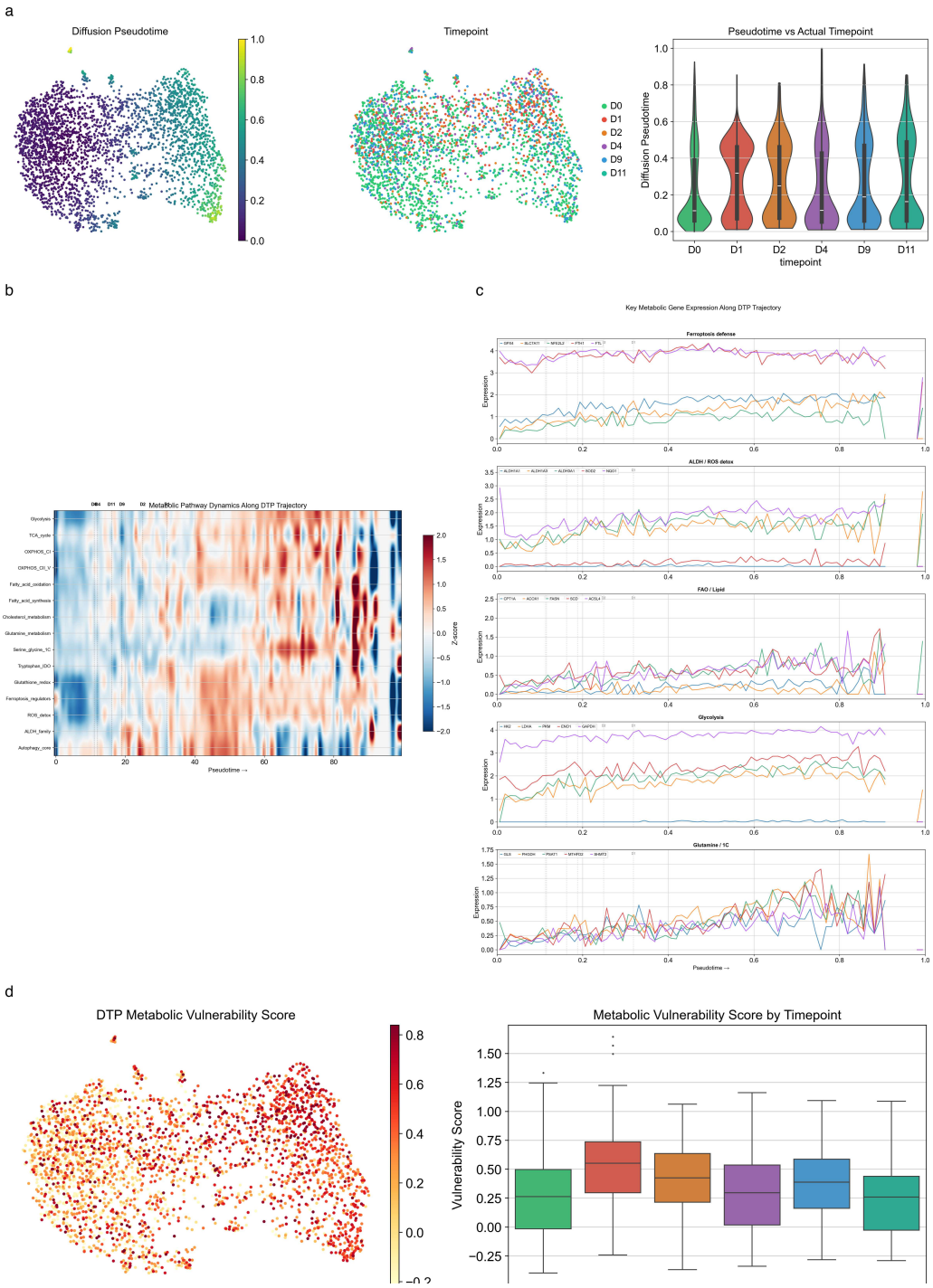


Figure 2: **Metabolic pathway dynamics along the DTP trajectory.** (**a**) Diffusion pseudotime validation: UMAP colored by pseudotime (left), by time point (center), and violin plot of pseudotime distribution per time point (right). (**b**) Metabolic pathway activity heatmap along pseudotime (z-scored, 15 pathways × 100 bins). Dashed lines indicate median pseudotime of each actual time point. (**c**) Key metabolic gene expression along pseudotime, grouped by functional category (ferroptosis defense, ALDH/ROS, FAO/lipid, glycolysis, glutamine/1C). (**d**) Metabolic vulnerability score: UMAP visualization (left) and boxplot by time point (right). Kruskal–Wallis $P = 1.2 \times 10^{-44}$; D1 shows highest vulnerability.

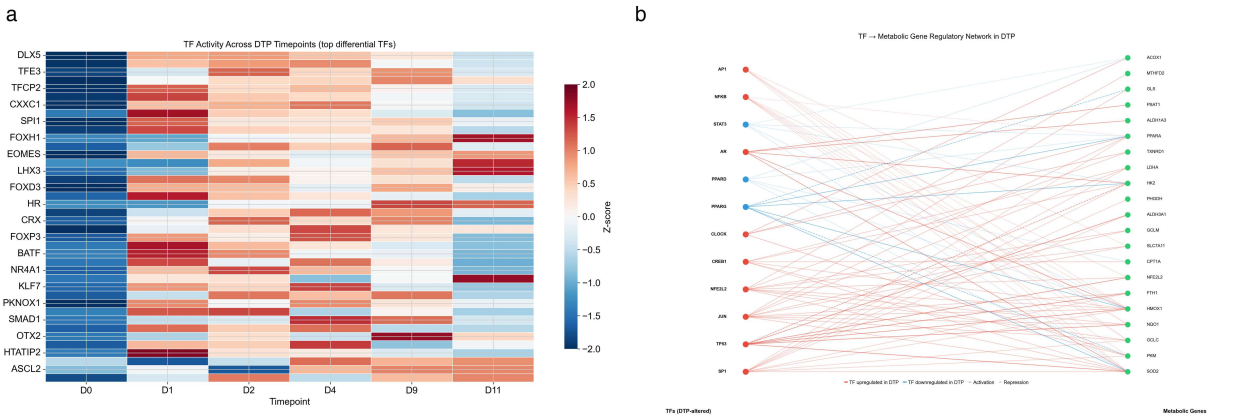


Figure 3: **Transcription factor regulatory architecture of DTP metabolic reprogramming.** (**a**) TF activity heatmap across time points (top differential TFs, z-scored by TF). (**b**) TF–metabolic gene regulatory network. Left: hub TFs (red = DTP-activated, blue = DTP-repressed); right: metabolic effector genes (green). Solid lines: activation; dashed lines: repression. 198 total edges derived from CollecTRI.

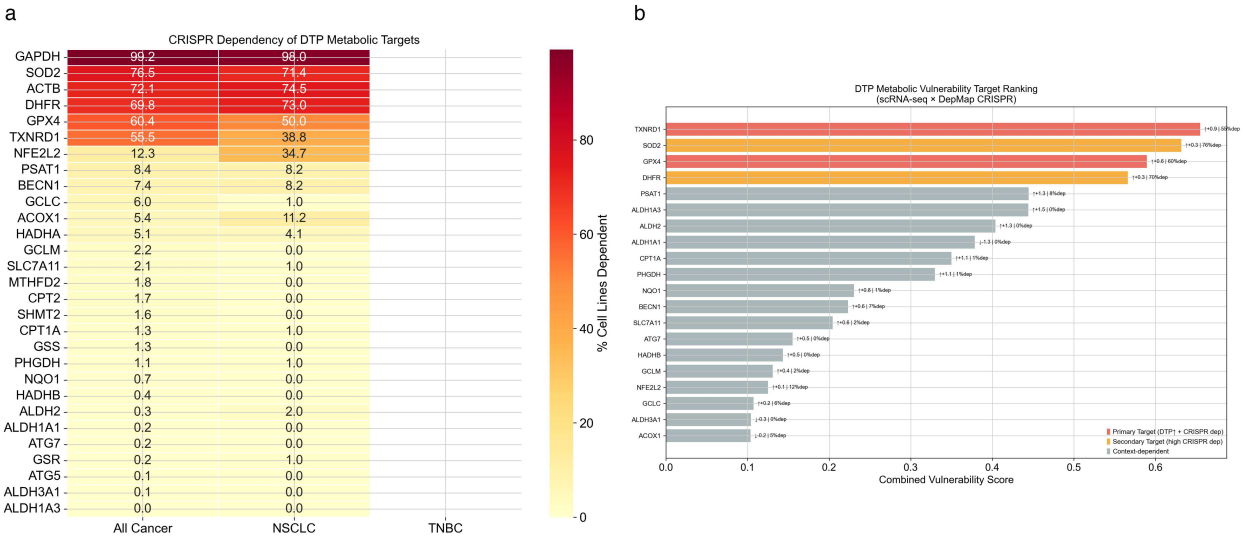


Figure 4: **DepMap and GDSC vulnerability validation.** (**a**) CRISPR dependency heatmap of 29 candidate metabolic targets across all cancer and NSCLC cell lines (% lines with gene effect < -0.5). (**b**) Integrated vulnerability target ranking (combined score $= 0.3 \times |\text{logFC}|_{\text{DTP}} + 0.7 \times f_{\text{CRISPR}}$). Red: primary targets; orange: secondary; grey: context-dependent.

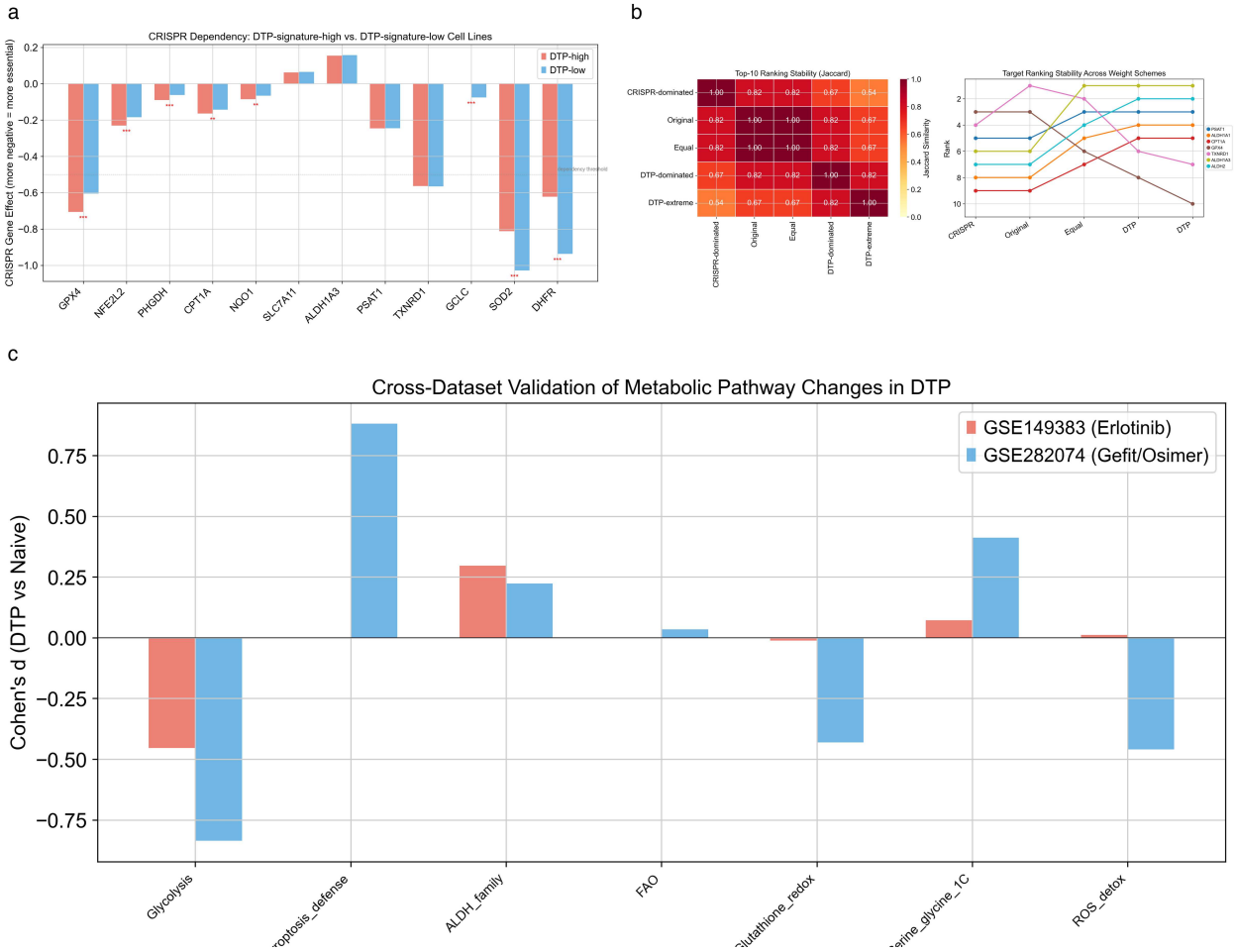


Figure 5: **Cross-validation, stratified dependency, and robustness analyses.** (**a**) DTP-signature-stratified CRISPR dependency: comparison of gene effect scores between DTP-signature-high ($n = 595$) and DTP-signature-low ($n = 595$) cell lines. Asterisks denote Mann–Whitney significance ($^{***}P < 0.001$). GPX4 is more essential in DTP-high lines ($P = 4.7 \times 10^{-4}$). (**b**) Vulnerability ranking sensitivity analysis. Left: Jaccard similarity heatmap of top-10 rankings across five weight schemes. Right: bump chart showing rank stability of seven genes present in all top-10 lists. (**c**) Cross-dataset validation: Cohen's d effect sizes for metabolic pathway changes in the discovery (GSE149383, erlotinib) and validation (GSE282074, gefitinib/osimertinib) cohorts. Four of five assessable pathways show concordant direction (80%).

Supplementary Tables

Table 1: **Table S1. Metabolic pathway statistics across DTP trajectory.** Kruskal–Wallis H statistic and P-value across all six time points; Cohen's d effect size for DTP (D9+D11) vs. naive (D0) comparison.

Pathway	KW H	KW P	Cohen's d	Direction
Ferroptosis regulators	393.6	7.1×10^{-83}	+0.68	UP
Glutamine metabolism	113.2	8.5×10^{-23}	+0.42	UP
Autophagy core	232.4	3.3×10^{-48}	+0.35	UP
ALDH family	283.3	2.8×10^{-74}	+0.30	UP
Tryptophan/IDO	114.9	1.5×10^{-23}	+0.26	UP
Serine–glycine 1C	47.5	4.5×10^{-9}	+0.05	UP
ROS detox	27.1	5.5×10^{-5}	+0.01	–
TCA cycle	19.5	1.5×10^{-3}	−0.03	–
OXPHOS CI	76.1	1.1×10^{-14}	−0.02	–
OXPHOS CII–V	7.5	0.19	−0.15	–
FAO	51.8	5.7×10^{-10}	−0.05	–
Glycolysis	301.2	7.7×10^{-78}	−0.45	DOWN
Glutathione redox	208.2	4.2×10^{-42}	−0.42	DOWN
Cholesterol metabolism	486.9	3.1×10^{-118}	−0.93	DOWN
Fatty acid synthesis	563.4	3.4×10^{-166}	−1.12	DOWN

Table 2: **Table S2. DepMap CRISPR dependency of candidate metabolic targets.** Percentage of cell lines with gene effect < -0.5 (Chronos). Top 15 genes shown; full table in supplementary data files.

Gene	All cancer (%)	NSCLC (%)	Combined score
TXNRD1	55.5	38.8	0.655
SOD2	76.5	71.4	0.632
GPX4	60.4	50.0	0.589
DHFR	69.8	73.0	0.567
PSAT1	8.4	8.2	0.446
ALDH1A3	0.2	0.0	0.443
ALDH2	0.3	2.0	0.407
CPT1A	1.3	1.0	0.350
ALDH1A1	0.2	0.0	0.345
NFE2L2	12.3	34.7	0.315
PHGDH	1.1	1.0	0.313
NQO1	0.7	0.0	0.247
BECN1	7.4	8.2	0.229
SLC7A11	2.1	1.0	0.224
GCLC	6.0	1.0	0.206

Declarations

Ethics approval and consent to participate. Not applicable.

Consent for publication. Not applicable.

Table 3: **Table S3. DTP-signature-stratified CRISPR dependency analysis.** Comparison of gene effect scores between DTP-signature-high ($n = 595$) and DTP-signature-low ($n = 595$) cell lines.

Gene	DTP-high median	DTP-low median	P (MWU)	More essential in
GPX4	−0.706	−0.604	4.7×10^{-4}	DTP-high
NFE2L2	−0.230	−0.185	8.2×10^{-9}	DTP-high
PHGDH	−0.090	−0.062	9.5×10^{-6}	DTP-high
CPT1A	−0.039	−0.022	7.4×10^{-3}	DTP-high
NQO1	−0.016	−0.025	1.4×10^{-2}	DTP-low
SLC7A11	−0.039	−0.050	7.1×10^{-3}	DTP-low
TXNRD1	−0.569	−0.573	0.44	n.s.
SOD2	−0.845	−0.959	7.5×10^{-6}	DTP-low
DHFR	−0.694	−0.815	1.2×10^{-6}	DTP-low

Table 4: **Table S4. GDSC2 drug sensitivity of candidate therapeutic agents.** Median $\ln(\mathrm{IC}_{50})$ values for compounds targeting DTP-associated metabolic vulnerabilities.

Drug	Target	NSCLC median	Pan-cancer median	n NSCLC	n other
Rapamycin	mTOR	−1.83	−1.99	103	847
JQ1	BRD4	2.55	2.15	101	848
I-BET-762	BRD4	3.70	3.16	96	797
OTX015	BRD4	2.78	2.31	99	804
Methotrexate	DHFR	0.36	0.82	99	814
Vorinostat	HDAC	1.80	1.40	101	834

Table 5: **Table S5. Cross-dataset validation of metabolic pathway changes.** Cohen's d effect sizes and Mann–Whitney P-values for DTP vs. naive comparison in the validation cohort (GSE282074).

Pathway	Cohen's d	P-value	Concordant with discovery
Glycolysis	−0.84	$< 10^{-200}$	Yes
Ferroptosis defense	+0.88	$< 10^{-200}$	Yes
ALDH family	+0.22	5.8×10^{-5}	Yes
Glutathione redox	−0.43	5.0×10^{-88}	Yes
ROS detox	−0.46	—	No (discovery $d = +0.01$)

Competing interests. The authors declare no competing interests.

Funding. This work was supported by the National Natural Science Foundation of China (grant no. 82174465); the Central High-Level Traditional Chinese Medicine Hospital Clinical Research and Achievement Transformation Capacity Enhancement Project (grant nos. HLCMHPP2023085 and HLCMHPP2023078); and the Qihuang Scholar Support Project of the National Administration of Traditional Chinese Medicine. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions. L.M. and J.T. conceived the study, designed the analytical framework, performed all computational analyses, and wrote the manuscript. Y.Li, Y.Luo, Z.L., D.Z., and X.C. contributed to data curation, literature review, and figure preparation. H.L. contributed to data interpretation and manuscript editing. Y.B. and B.H. supervised the study, provided critical intellectual input, and revised the manuscript. All authors read and approved the final manuscript.

Table 6: **Table S6. MSigDB Hallmark pathway enrichment of DTP-upregulated genes.** Top 10 enriched gene sets among the 200 most upregulated genes (DTP vs. naive).

Hallmark gene set	Adjusted P	Overlap
Epithelial–Mesenchymal Transition	4.1×10^{-14}	21/200
UV Response Down	3.0×10^{-4}	9/144
Hypoxia	4.7×10^{-4}	10/200
TNF-α Signaling via NF-κB	6.4×10^{-3}	10/200
mTORC1 Signaling	6.4×10^{-3}	9/200
Coagulation	6.4×10^{-3}	6/138
Apoptosis	9.5×10^{-3}	8/161
Myogenesis	1.0×10^{-2}	8/200
Complement	3.6×10^{-2}	7/200
Xenobiotic Metabolism	4.9×10^{-2}	6/200

Table 7: **Table S7. Hub transcription factors in the TF–metabolic gene regulatory network.** TFs regulating ≥ 4 of 24 metabolic target genes. logFC: DTP vs. naive TF activity fold-change.

TF	Targets (n)	Activity logFC (DTP vs. naive)
SP1	10	+0.59
JUN	10	+0.69
NFE2L2	9	+0.69
TP53	7	+0.43
PPARG	6	−0.88
PPARD	5	−0.71
AR	5	+0.26
CLOCK	5	+0.69
CREB1	5	−1.68
NFKB1	4	+0.57
STAT3	4	−2.44
SREBF1	4	−3.01

Acknowledgements. We thank the authors of the original datasets (Aissa et al., GSE149383; GSE282074) for making their single-cell RNA-seq data publicly available. We acknowledge the Broad Institute Cancer Dependency Map (DepMap) and the Genomics of Drug Sensitivity in Cancer (GDSC) projects for providing open-access functional genomics and pharmacological data. Computational analyses were performed using open-source bioinformatics tools including Scanpy, decoupler, and gseapy.

Data Availability

The single-cell RNA-seq data analyzed in this study are publicly available from the Gene Expression Omnibus: GSE149383 (discovery cohort) and GSE282074 (validation cohort). DepMap CRISPR gene effect data and CCLE expression data were obtained from the DepMap portal (`https://depmap.org`). GDSC2 drug sensitivity data were obtained from `https://www.cancerrxgene.org`. All analysis scripts and processed data are available at [repository URL upon publication].

Code Availability

All analysis scripts (Python) used in this study are available at [repository URL upon publication]. The computational environment and package versions are detailed in the Methods section.

References

[1] Susumu Kobayashi, Titus J Boggon, Tajhal Dayaram, Pasi A Janne, Otmar Kocher, Matthew Meyerson, et al. EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. *New England Journal of Medicine*, 352(8):786–792, 2005. doi: 10.1056/NEJMoa044238.

[2] Sreenath V Sharma, Diana Y Lee, Bihua Li, Margaret P Quinlan, Fumiyuki Takahashi, Shyamala Maheswaran, Ultan McDermott, Nina Azizian, Lee Zou, Michael A Fischbach, et al. A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. *Cell*, 141(1): 69–80, 2010. doi: 10.1016/j.cell.2010.02.027.

[3] Jean-Christophe Marine, Sarah-Jane Dawson, and Mark A Dawson. Cancer drug-tolerant persister cells: from biological questions to clinical opportunities. *Nature Reviews Cancer*, 24: 694–717, 2024. doi: 10.1038/s41568-024-00714-4.

[4] Qingqing Wang et al. Drug tolerant persister cell plasticity in cancer: a revolutionary strategy for more effective anticancer therapies. *Signal Transduction and Targeted Therapy*, 9:191, 2024. doi: 10.1038/s41392-024-01891-4.

[5] Xinyu Li, Miaomiao Xu, Jie Cheng, Zhengwei Yu, and Wen Tang. Drug-tolerant persister cell in cancer: reversibility, microenvironmental interplay, and therapeutic strategies. *Frontiers in Pharmacology*, 16:1612089, 2025. doi: 10.3389/fphar.2025.1612089.

[6] Kari J Kurppa et al. Genome-wide CRISPR screens identify the YAP/TEAD axis as a driver of persister cells in EGFR mutant lung cancer. *Communications Biology*, 7:6190, 2024. doi: 10.1038/s42003-024-06190-w.

[7] Hyun Woo Park et al. Drug tolerance and persistence to EGFR inhibitor treatment are mediated by an ILK-SFK-YAP signaling axis in lung adenocarcinoma. *Oncogene*, 44:3461, 2025. doi: 10.1038/s41388-025-03461-6.

[8] Aviad Weiss et al. APOBEC3 activity promotes the survival and evolution of drug-tolerant persister cells during EGFR inhibitor resistance in lung cancer. *Cancer Research Communications*, 5(5):825, 2025.

[9] Florian Rambow, Aljosja Rogiers, Oskar Marin-Bejar, Sara Aibar, Julia Femel, Michael Dewaele, et al. Toward minimal residual disease-directed therapy in melanoma. *Cell*, 174(4): 843–855, 2018. doi: 10.1016/j.cell.2018.06.025.

[10] Shensi Shen et al. Drug-tolerant persister cells in cancer: bridging the gaps between bench and bedside. *Nature Communications*, 16:66376, 2025. doi: 10.1038/s41467-025-66376-6.

[11] Yi Zhang et al. Drug-induced tolerant persisters in tumor: mechanism, vulnerability and perspective implication for clinical treatment. *Molecular Cancer*, 24:323, 2025. doi: 10.1186/s12943-025-02323-9.

[12] Wei Chen et al. Metabolic orchestration of drug-tolerant persister cells in cancer. *Life Medicine*, 3(6):lnae040, 2024. doi: 10.1093/lifemedi/lnae040.

[13] Xuejun Jiang, Brent R Stockwell, and Marcus Conrad. Decoding ferroptosis for cancer therapy. *Nature Reviews Cancer*, 25:864, 2025. doi: 10.1038/s41568-025-00864-1.

[14] Scott J Dixon et al. Ferroptosis as an approach to leverage cancer metabolism. *Trends in Cell Biology*, 36(3), 2026. doi: 10.1016/j.tcb.2026.00039.

[15] Yaara Oren et al. Targeting of vulnerabilities of drug-tolerant persisters identified through functional genetics delays tumor relapse. *Cell Reports Medicine*, 5(4):100117, 2024. doi: 10.1016/j.xcrm.2024.00117.

[16] Javier Garcia-Bermudez et al. Adaptation of redox metabolism in drug-tolerant persister cells is a vulnerability to prevent relapse in pancreatic cancer. *Oncogenesis*, 14:591, 2025. doi: 10.1038/s41389-025-00591-0.

[17] Yaara Oren et al. Targeting metabolism to combat anticancer and antibacterial resistance. *Trends in Pharmacological Sciences*, 47(2), 2026. doi: 10.1016/j.tips.2026.00017.

[18] Marius Lange et al. CellRank 2: unified fate mapping in multiview single-cell data. *Nature Methods*, 21:1196–1205, 2024. doi: 10.1038/s41592-024-02303-9.

[19] Amjad F Aissa, Abul B M M K Islam, Maha M Ariss, Anand K Rishi, et al. Single-cell transcriptional changes associated with drug tolerance and response to combination therapies in cancer. *Nature Communications*, 12:1628, 2021. doi: 10.1038/s41467-021-21884-z.

[20] Jiwen Xin, Adam Mark, Cyrus Afrasiabi, Ginger Tsueng, Moritz Juchler, Kristina Gober, Nils Schuth, et al. High-performance web services for querying gene and variant annotation. *Genome Biology*, 17:91, 2016. doi: 10.1186/s13059-016-0953-9.

[21] Joshua M Dempster, Isabella Boyle, Francisca Vazquez, David E Root, Jesse S Boehm, William C Hahn, Aviad Tsherniak, and James M McFarland. Chronos: a cell population dynamics model of CRISPR experiments that improves inference of gene fitness effects. *Genome Biology*, 22:343, 2021. doi: 10.1186/s13059-021-02540-7.

[22] Clare Pacini et al. A comprehensive clinically informed map of dependencies in cancer cells and framework for target prioritization. *Cancer Cell*, 42:301–316, 2024. doi: 10.1016/j.ccell.2023.12.009.

[23] Wanjuan Yang, Jorge Soares, Patricia Greninger, Elena J Edelman, Howard Lightfoot, Simon Forbes, et al. Genomics of drug sensitivity in cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Research*, 41:D955–D961, 2013. doi: 10.1093/nar/gks1111.

[24] F Alexander Wolf, Philipp Angerer, and Fabian J Theis. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biology*, 19:15, 2018. doi: 10.1186/s13059-017-1382-0.

[25] Itay Tirosh, Benjamin Izar, Sanjay M Prakadan, Marc H Wadsworth, Daniel Treacy, John J Trombetta, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science*, 352(6282):189–196, 2016. doi: 10.1126/science.aad0501.

[26] Pau Badia-i Mompel, Jessé Velez Santiago, Jana Braunger, Celina Geiss, Ricardo O. Ramirez Flores, and Julio Saez-Rodriguez. decoupler: ensemble of computational methods to infer biological activities from omics data. *Bioinformatics Advances*, 2:vbac016, 2022. doi: 10.1093/bioadv/vbac016.

[27] Sophia Muller-Dott et al. CollecTRI: a database of transcription factor–target gene interactions curated from the literature. *Nucleic Acids Research*, 52:D1321–D1329, 2024. doi: 10.1093/nar/gkad1044.

[28] Yingjiang Li, Shuangxi Xu, Maria M Mihaylova, Bin Zheng, Xiuyun Hou, Bingbing Jiang, et al. AMPK phosphorylates and inhibits SREBP activity to attenuate hepatic steatosis and atherosclerosis in diet-induced insulin-resistant mice. *Cell Metabolism*, 13(4):376–388, 2011. doi: 10.1016/j.cmet.2011.03.009.

[29] Michael Platten, Ellen A A Nollen, Ute F Rohrig, Francesca Fallarino, and Christiane A Opitz. Tryptophan metabolism as a common therapeutic target in cancer, neurodegeneration and beyond. *Nature Reviews Drug Discovery*, 18:379–401, 2019. doi: 10.1038/s41573-019-0016-5.

[30] Santosh Kumar et al. Therapeutic resistance and combination therapy for cancer. *Scientific Reports*, 15:12725, 2025. doi: 10.1038/s41598-025-12725-w.

[31] Jay D Horton, Joseph L Goldstein, and Michael S Brown. SREBPs: activators of the complete program of cholesterol and fatty acid synthesis in the liver. *Journal of Clinical Investigation*, 109(9):1125–1131, 2002. doi: 10.1172/JCI15593.

[32] Nathalie Q Balaban, Jack Merrin, Remy Chait, Lukasz Kowalik, and Stanislas Leibler. Bacterial persistence as a phenotypic switch. *Science*, 305(5690):1622–1625, 2004. doi: 10.1126/science.1099390.

[33] Peter W Gout, Anthony R Buckley, Cameron R Simms, and Nicholas Bruchovsky. Sulfasalazine, a potent suppressor of lymphoma growth by inhibition of the x(c)- cystine transporter: a new action for an old drug. *Leukemia*, 15:1633–1640, 2001. doi: 10.1038/sj.leu.2402238.

[34] Fulvio Ursini, Valentina Bosello Travain, Giorgio Cozza, Giorgia Miotto, Antonella Roveri, Stefano Toppo, and Matilde Maiorino. A white paper on phospholipid hydroperoxide glutathione peroxidase (GPx4) forty years later. *Free Radical Biology and Medicine*, 188:117–133, 2022. doi: 10.1016/j.freeradbiomed.2022.06.227.

[35] Kirill Bersuker, Joseph M Hendricks, Zhipeng Li, Leslie Magtanong, Brent Ford, Peter H Tang, et al. The CoQ oxidoreductase FSP1 acts parallel to GPX4 to inhibit ferroptosis. *Nature*, 575:688–692, 2019. doi: 10.1038/s41586-019-1705-2.

[36] Jason W Locasale, Alexandra R Grassian, Tamar Melman, Costas A Lyssiotis, Katie R Mattaini, Adam J Bass, et al. Phosphoglycerate dehydrogenase diverts glycolytic flux and contributes to oncogenesis. *Nature Genetics*, 43(9):869–874, 2011. doi: 10.1038/ng.890.

[37] Yuxiang Liu et al. Spatial omics at the forefront: emerging technologies, analytical innovations, and clinical applications. *Cancer Cell*, 44(1):24–49, 2026. doi: 10.1016/j.ccell.2025.12.009.