ALDH1A3-acetaldehyde re-wires neural crest stem cell and high metabolism states to potentiate melanoma heterogeneity

Yuting Lu[1, 2], Jana Travnickova[1, 2], Mihaly Badonyi[1], Florian Rambow[3, 4], Andrea Coates[1, 2], Zaid Khan[2], Jair Marques[2], Laura Murphy[1], Pablo Garcia-Martinez[5, 6], Richard Marais[5, 7], Pakavarin Louphrasitthiphol[8], Alex Von Kriegsheim[2], Joseph A Marsh[1], Valeria Pavet[5, 9], Owen J Sansom[9,10], Robert Illingworth[11], E. Elizabeth Patton[1, 2*]

1. MRC Human Genetics Unit, Institute of Genetics and Cancer, The University of Edinburgh, EH4 2XU, United Kingdom.
2. Edinburgh Cancer Research, CRUK Scotland Centre, Institute of Genetics and Cancer, The University of Edinburgh, EH4 2XR, United Kingdom.
3. Department of Applied Computational Cancer Research, Institute for AI in Medicine (IKIM), University Hospital Essen, 45131 Essen, Germany
4. University of Duisburg-Essen, 45141, Essen, Germany
5. Cancer Research UK Manchester Institute, The University of Manchester, Alderley Park, SK10 4TG, United Kingdom.
6. The Dermatology Centre, Northern Care Alliance NHS Foundation Trust, Salford, M6 8HD, United Kingdom
7. Oncodrug Ltd, Alderley Park, Macclesfield, SK10 4TG, United Kingdom.
8. Ludwig Institute for Cancer Research, Nuffield Department of Clinical Medicine, University of Oxford, Headington, Oxford OX3 7DQ, United Kingdom
9. Cancer Research UK Beatson Institute**,** CRUK Scotland Centre, Garscube Estate, Switchback Road, Bearsden Glasgow, G61 1BD, United Kingdom
10. School of Cancer Sciences, University of Glasgow, Glasgow, G12 0ZD, United Kingdom

11. Centre for Regenerative Medicine, Institute for Regeneration and Repair, The University of Edinburgh, Edinburgh BioQuarter, EH16 4UU, United Kingdom.

* co-responding author: EEP e.patton@ed.ac.uk

Abstract

Cancer cellular heterogeneity and therapy resistance arise from metabolic and transcriptional adaptations, but how these are interconnected is less well understood. Here, we show that in melanoma, the cancer stem cell marker aldehyde dehydrogenase 1A3 (ALDH1A3) forms an enzymatic partnership with acetyl-CoA synthetase 2 (ACSS2) in the nucleus to couple high glucose metabolic flux with acetyl-histone H3 modification of neural crest lineage and glucose metabolism genes. Critically, we show acetaldehyde is a metabolite source for acetyl-histone H3 modification dependent on ALDH1A3, providing a physiologic function for this highly volatile and toxic metabolite. In a zebrafish model of melanoma residual disease, a subpopulation of ALDH1A3-high cells emerges following BRAF inhibitor treatment and targeting these with an ALDH1 suicide inhibitor, nifurtimox, delays or prevents BRAF inhibitor drug-resistant relapse. Our work reveals that ALDH1A3-ACSS2 directly coordinates local acetaldehyde-acetyl-CoA metabolism with specific chromatin-based gene regulation and represents a potential therapeutic vulnerability in melanoma.

Highlights:

- ALDH1A3-high melanomas are in a high glucose metabolic flux and neural crest stem cell dual state.
- Nuclear ALDH1A3 partners with ACSS2 to promote selective acetyl-histone H3.
- Acetaldehyde is an acetyl source for ALDH1A3 dependent histone H3 acetylation.
- ALDH1A3 is a master regulator and pharmaceutical target of melanoma heterogeneity.

Introduction

The perennial challenge in the clinical development of cancer therapies is that cancer cells frequently co-opt non-genetic mechanisms to switch between dynamic cellular states[1]. This cellular plasticity enables cancer cells to adapt and thrive under environmental pressures such as immune surveillance, nutrient deprivation, or therapy[2]. Across tumour types, numerous studies have exposed that hijacking of common developmental (foetal) lineage programs and substantial cell state heterogeneity commonly underlie tumour progression and drug resistance[3-10]. While several molecular mechanisms that can give rise to cancer heterogeneity phenotypes have been described, to make transformative progress towards curtailing tumour state transitions and enhancing treatment efficacy, a deeper understanding of the interplay between epigenetic, transcriptional, and metabolic plasticity in cancer cell biology is needed.

For many patients with advanced melanoma, novel systemic targeted and immune therapies have greatly improved prognosis[11-15]. However, a vexing obstacle towards eradicating melanoma cells and achieving long-term survival is that melanoma cell subpopulations undergo phenotypic transitions into dedifferentiated stem-like states, leading to innate or acquired drug resistance and tumour recurrence. In such subpopulations, a transcriptional state resembling that of neural crest stem cells (NCSC) and characterized by low activity of Melanocyte Inducing Transcription Factor (MITF) emerges[10,16-18]. This state then becomes enriched after therapy, which is predictive for patient outcomes [6-8,10,16,17,19-24].

In addition to transcriptional states, metabolic heterogeneity has recently come to the forefront as a mechanism influencing tumour cell survival and biology[25]. Generally considered to arise from competition for nutrients between cancer cells and tumour-associated cells in the microenvironment, metabolic heterogeneity plays a role in this phenotypic plasticity, which enables melanoma and other cancer cells to survive therapy in specific tissue and organ

environments[26-33]. While the developmental neural crest state is known to be highly sensitive to metabolic deficiencies[34-38], we lack an understanding of how metabolic states are co-opted by and coordinated with neural crest programs in melanoma.

In this study, we discover that the pan-cancer stem cell marker ALDH1A3 is a central regulator of both metabolic and stem cell transcriptional and phenotypic states in melanoma. Using melanoma cell lines and patient sample-derived low passage cells, we show that ALDH1A3 forms an enzymatic partnership with ACSS2 to consolidate a high glucose metabolic flux with a NCSC transcriptional state instructed by Transcription Factor AP-2 Beta (TFAP2B). Mechanistically, we show that ALDH1A3 activity promotes selective acetylation of histone H3, and link this to transcription of genes regulating NCSC and glucose metabolism. Critically, by tracing acetaldehyde to acetyl-histone H3, we demonstrate its role as a source for acetylated histones dependent on ALDH1A3. Our findings uncover how therapy resistance states in patients with melanoma is mediated by an actionable, high dimensional metabolic-transcriptional framework that controls melanoma stem cell plasticity.

Results

ALDH1A3^{High} melanomas are enriched for neural crest stem cell and glucose metabolic states

When patients with melanoma become resistant to MAPK inhibitor therapy, their tumour cells can upregulate expression of the stem cell marker and aldehyde dehydrogenase enzyme ALDH1 **(Figure S1A)**[39-41]. Further, in experimental mouse melanoma models, we find *Aldh1a3* expression is tightly associated with dedifferentiated, neural crest and stem cell states, which have been reported to fuel cancer growth in a cellular hierarchy[42,43] (**Figure S1B, C**).

Here, we hypothesised that the ALDH-high metabolic activity of cancer cells is the consequence of a changed transcriptional state that contributes to melanoma stemness and plasticity. To investigate this state in melanoma cells, we considered that ALDH-high activity ($ALDH^{High}$) is heterogeneous in human cell lines (**Figure S1D)**, and that the predominant ALDH activity in melanoma A375 cells is due to ALDH1A3[39]. Thus, we sorted melanoma cells for the highest and lowest ALDH activity, termed $ALDH^{High}$ and $ALDH^{Low}$, using the flourescent substrate Aldefluor (an amino acetaldehyde) **(Figure 1A)**. When we validated ALDH1A3 expression by immunocytochemistry (ICC), we were intrigued to see that ALDH1A3 was predominantly expressed in the cytosol in $ALDH^{Low}$ cells but was enriched in the nucleus of sorted $ALDH^{High}$ cells **(Figure 1B).**

We then conducted comparative RNA-sequencing (RNA-seq) of sorted $ALDH^{High}$ and $ALDH^{Low}$ cells and identified high expression of genes associated with dedifferentiated, neural crest and stem cell states (including *SOX2, SOX10, GAS7, GDNF, RXRG, TFAP2B*) in $ALDH^{High}$ cells **(Figure 1C-D; Table S1-2)**. These genes were also identified as the top marker genes associated with a less-differentiated state as measured by CytoTRACE in murine melanoma **(Figure S1E-F).** This molecular profile resembles a cell state present in drug-resistant melanomas that is predictive of poor outcomes for patients treated with both targeted and immunotherapy[7,8,44].

In contrast, $ALDH^{Low}$ cells were enriched for Interferon Regulatory Factor 1 (IRF1) target genes (*IFIT3, IL18, PDL1*), proinflammatory genes (*TAP2, PSMB9, SMAD7*) as well as MITF target pigmentation (differentiation) genes (*MLPH)* **(Figure 1C; Table S1)**. Metabolic transcriptional signatures were also highly distinct in $ALDH^{High}$ and $ALDH^{Low}$ cell populations, with glycolysis pathway genes (*PDHX, PFKFB3*) enriched in $ALDH^{High}$ cells whereas starvation response-related fatty acid metabolism genes (*ACSL5, SLC12A7*) were overrepresented in $ALDH^{Low}$ cells **(Figure**

1D). Taken together, these data show that the ALDH1A3^{High} cell state has features of both neural crest stem cell development and high glucose metabolism[18,34,45,46].

Next, we asked if the ALDH1A3-enriched cell state is present in human patient samples. To answer this question, we ranked patient samples from available datasets based on *ALDH1A3* RNA levels, and then selected *ALDH1A3-high* and *ALDH1A3-low* patients from each cohort; top and bottom 10% for TCGA[47] and Lund (population cohort of primary and metastatic melanomas[48], and top and bottom 25% for Bergen dataset (Stage IV melanomas,[49] **(Figure S1G)**. In the *ALDH1A3-high* patient samples, we found that the NCSC state was indeed enriched and coupled with an immune suppressive TGF-beta signalling signature **(Figure 1E, Figure S1H; Table S2).**

Alongside the mesenchymal genes and within the NCSC genes, we were particularly intrigued to see a strong association between *ALDH1A3* and *TFAP2B* expression, given that we have recently shown *tfap2b* marks an adult multipotent melanocyte stem cell population in zebrafish[46] **(Figure 1F)**. The *ALDH1A3-TFAP2B* association was isoform-specific: across the entire ALDH family, *TFAP2B* was only correlated with *ALDH1A3* **(Figure S1I)**. In contrast, *TFAP2A*, a related family member expressed in proliferative melanoma that shares near identical binding motif with TFAP2B **(Figure S1J)**, was negatively correlated with *ALDH1A3* **(Figure 1F)**. Given these data, we re-analysed scRNA-seq data from 15 melanoma patient samples[50] and found that *ALDH1A3* and *TFAP2B* are co-expressed in a distinct cell cluster present in a subset of samples (5 out of 14; 3 females plus 2 males) with varied mutation subtypes **(Figure 1G, Figure S1K)**. Thus, we concluded that *ALDH1A3-high* cells that are enriched for *TFAP2B-NCSC* programs are prevalent in melanoma across disease stages and genotypic subtypes.

In contrast to *ALDH-high* melanomas, and in agreement with our RNA-seq analysis **(Figure 1D),** *ALDH1A3-low* patient samples were enriched in proinflammatory genes and MITF pigmentation target genes across all three independent melanoma datasets **(Figure 1E-F, Figure S1H)**.

Conversely, analysis of RNA-seq data from pre-treatment patient samples showed that high levels of *ALDH1A3* portends worse response to immunotherapy **(Figure 1H)**. When we performed deconvolution analysis of immune cell infiltration in the *ALDH1A3-high* and *ALDH1A3-low* samples of the Lund dataset, we were intrigued to find that *ALDH1A3-low* melanomas, relative to *ALDH1A3-high* samples, were enriched in M1 macrophages, CD8+ T-cells, and plasma cells, consistent with an active ("hot") immune microenvironment (**Figure 1I, Figure S1L)**.

Collectively, our data thus far support a model in which high ALDH1A3 metabolic activity segregates with activation of *TFAP2B-NCSC*-driven developmental stem cell pathways, high expression of glucose metabolism genes, and therapy resistance, while low ALDH1A3 metabolic states have features of melanocyte differentiation, low glucose metabolism and robust immunogenicity **(Figure 1J)**.

TFAP2B promotes stemness and dedifferentiation in ALDH1A3^{High} melanoma cells

We next asked whether ALDH-associated melanoma stem cell phenotypes are directly regulated by TFAP2B. We engineered two melanoma cell lines to knock out or over-express *ALDH1A3*, both of which are bearing wildtype p53 and BRAF(V600E) mutations: the A375 line, in which *ALDH1A3* and *TFAP2B* is expressed at high levels, and the C089 line, with lower expression of both genes **(Figure S2A, B)**. Notably, analysis of human melanoma cell lines in the Cancer Cell Line Encyclopaedia[51] and EGAS00001000815[52] revealed a significant correlation between *ALDH1A3* and *TFAP2B* expression, consistent with human patient samples and sorted ALDH^{High} and ALDH^{Low} cell RNA-seq **(Figure S2B)**. Using the Aldefluor assay, we validated *ALDH1A3* knockout (KO) and overexpression (OE) as significantly shifted ALDH activity toward low and high states respectively **(Figure 2A)**. Consistently, RT-qPCR analysis revealed that loss of *ALDH1A3*

led to reduced *TFAP2B-NCSC* gene expression, concomitant with increased expression of *IRF1* and *MITF* target genes **(Figure 2B)**. By western blot analysis, we validated that loss of ALDH1A3 led to reduced TFAP2B protein levels, while ALDH1A3 overexpression upregulated TFAP2B **(Figure 2C).** Taken together, these data support that ALDH1A3 is a direct regulator of the *TFAP2B-NCSC* state.

Next, we found that loss of *ALDH1A3* led to decreased colony formation while *ALDH1A3* overexpression increased colony formation **(Figure 2D-E),** consistent with the pan-cancer effect of ALDH activity in cancer stem-like states[39,40,53]. Remarkably, loss of *TFAP2B* via CRISPR-Cas9 gene knockout in melanoma cells phenocopied *ALDH1A3* deficiency in colony assays **(Figure 2D-E, Figure S2C)**. We were unable to assess the impact of *TFAP2B* overexpression alone in colony assays due to neuronal-like differentiation, consistent with TFAP2B being a powerful transcriptional regulator of the neural crest. However, in *ALDH1A3* KO cells, overexpression of *TFAP2B* partially rescued the stemness deficiency, leading to enhanced colony formation capacity in both *ALDH1A3* and *TFAP2B* KO cells **(Figure 2D-E, Figure S2C).** These data indicate that TFAP2B is a downstream mediator of ALDH1A3 that sustains cancer stemness in melanoma.

To directly test if ALDH1A3 mediates *TFAP2B* expression in patients, we obtained low-passage human melanoma cells, termed melanoma patient-derived (MPD) 001 cells, from a patient diagnosed with BRAF-mutant melanoma **(Figure 2F)**. The clinical record for these cells shows that the patient's disease continued to progress on immune as well as targeted therapies**.** When we stably over-expressed *ALDH1A3* in the MPD001 cells, we found that *TFAP2B* gene expression was upregulated whereas expression of the melanocyte differentiation gene *TFAP2A* was downregulated **(Figure 2F)**. Moreover, these changes upon *ALDH1A3* overexpression segregated with an observable loss of pigmentation, indicative of melanocyte de-differentiation

into a stem cell-like state **(Figure 2G)**. Thus, in human melanoma cells, *ALDH1A3* expression is, in and of itself, sufficient to promote *TFAP2B* gene expression and pigmentation loss.

ALDH1A3^{High} cells use glucose while ALDH1A3^{Low} cells rely on acetate for Acetyl-CoA production

Next, we sought to explore relationships between ALDH1A3 function and the metabolic states we identified by RNA-seq in cells and in patients (**Figure 1)**. To this end, we traced $^{13}C_6$-labelled glucose in two A375-based melanoma cell models: (1) WT cells (that express high levels of *ALDH1A3*) and *ALDH1A3* knockout cells (*ALDH1A3* KO) for 24 hours labelling, and (2) non-engineered cells labelled for 12 hours, and then sorted for ALDH1A3^{High} and ALDH1A3^{Low} activity **(Figure 3A, Table S3)**. We did not find any difference in $^{13}C_6$-glucose uptake between cells having different levels of *ALDH1A3* **(Figure 3B)**, despite the increased expression of glycolysis genes in ALDH1A3^{High} cells determined by RNA-seq analysis. Instead, we observed significantly higher glucose-derived carbon flux (termed high-GF) in cells with high ALDH1A3 activity, as evidenced by an increase in ^{13}C-labelled carbon in pyruvate and TCA cycle intermediates, including citric acid, α-ketoglutaric acid, succinic acid, and malic acid **(Figure 3C-D)**. Conversely, in *ALDH1A3* KO and ALDH1A3^{Low} cells, glucose-derived carbon was converted into ketone bodies such as aceto-acetate, a metabolite synthesized from acetyl-CoA, reflecting a starvation-like metabolic state **(Figure 3E)**. Despite the lower glycolytic flux, *ALDH1A3* KO cells produced higher levels of lactate both in cells and secreted into the culture media **(Figure 3F)**. These data indicate that ALDH1A3^{High} cells are primed to execute glycolysis and oxidative phosphorylation (OXPHOS), whereas ALDH1A3^{Low} cells exist in a starvation-like state and secrete lactate. Consequently, we reasoned these states exert differential metabolic and epigenetic effects on surrounding cancer and stromal cells [54,55].

Acetyl-CoA, an acetyl group donor in biochemical reactions, is generated from pyruvate through glycolysis or through oxidation of long-chain fatty acids or certain amino acids, and its availability is known to shape both metabolic processes and epigenetic regulation[56]. Unexpectedly, we found that total levels of acetyl-CoA increased in ALDH1A3^{Low} cells despite decreased glycolysis and reduced TCA cycle flux **(Figure 3G)**. This increase in the total levels of acetyl-CoA was not derived from a glucose source, as indicated by $^{13}C_6$ glucose tracing **(Figure 3H)**. This increase was also concomitant with a significant increase in acetyl-histone H3 (~30%), as assessed by western blot and mass spectrometry analysis of histones **(Figure 3I-J; Figure S3)**. Consistently, overexpression of ALDH1A3 led to a ~15% decrease in acetyl-histone H3 **(Figure3 I-J; Figure S3)**.

Given that cancer cells under metabolic stress can support as much as half of their lipid synthesis by using acetate as a substrate[57], we hypothesized that ALDH1A3^{Low} melanoma cells supplement their carbon source by switching from glucose to acetate. When we traced $^{13}C_2$-acetate incorporation into acetyl-CoA **(Figure 3K),** we found that *ALDH1A3* KO cells had a higher percentage of acetate derived acetyl-CoA relative to control cells, both within the whole cell and in the nuclear fraction **(Figure 3L)**. Thus, we conclude that cells with high levels of ALDH1A3 rapidly metabolise glucose to generate pyruvate and to enter the TCA cycle, while cells with low ALDH1A3 activity preferentially use acetate as a carbon source for acetyl-CoA and to generate lactate from glucose **(Figure 3M)**.

Selective histone H3 acetylation depends on ALDH1A3

Previous work reported that selective histone acetylation mechanisms can regulate specific gene expression, including glucose metabolism enzymes[58,59], and that global histone repositories serve

as reservoirs for acetyl groups[60]. When extrapolating to our findings, we hypothesized that in the high ALDH1A3 state, histone acetylation by glucose-derived acetyl-CoA is coordinated with regulation of specific NCSC and metabolic genes to generate a cancer stem cell-like state. Conversely, in the low ALDH1A3 state, acetate derived acetyl-CoA is deposited on chromatin as a reservoir[60], serving as a rapid source for acetyl-coA and as a buffer for intracellular pH[61,62].

To determine how ALDH1A3-dependent histone H3 acetylation is deposited on chromatin, we performed quantitative acetyl-histone H3 chromatin immunoprecipitation (ChIP)-seq using an antibody against pan-histone H3 acetylated sites (K9 + K14 + K18 + K23 + K27). We found that the average acetyl-histone H3 signal increased by ~1.4 fold across H3-ac peak regions in *ALDH1A3* KO cells **(Figure 4A)** and identified 5129 acetyl-histone H3 peaks enriched in *ALDH1A3* KO cells, versus 2020 peaks enriched in *ALDH1A3* control cells (fold change >1, FDRq < 0.05) **(Figure 4B, C).** Notably, in control cells, the enriched acetyl-histone H3 peaks were clustering around transcription start sites (TSSs), especially within 1kb of promoters, whereas in *ALDH1A3* KO cells, the enriched acetyl-histone H3 peaks were dispersed throughout the genome, and particularly spreading into the distal intergenic region and intronic regions **(Figure 4C, D)**. This distribution effect is even more prominent when comparing the top fold-change enriched acetyl-histone H3 peaks (>2 fold) between control and *ALDH1A3* KO cells **(Figure 4C, D)**. To test the robustness of our differential acetyl-histone H3 analysis, we adjusted the average peak score of *ALDH1A3* KO samples by manually increasing the log base 2 value by +0.5, ~ log2(1.4), which is the fold change determined by western blot and mass spec **(Figure S4A, Figure S3)**. With the adjusted peak score (a more stringent analysis), enriched acetyl-histone H3 remained distributed at distal intergenic and intronic regions in the *ALDH1A3* KO, in contrast with the enriched acetyl-histone H3 in promoter areas in control cells **(Figure S4B-E).**

Next, we found that acetyl-histone H3 sites enriched in the ALDH1A3^{High} state were present in NCSC and glucose metabolism pathway genes, such as *TFAP2B* and *PDHB* (Pyruvate Dehydrogenase E1 Subunit Beta) **(Figure 4E; Table S4)**, consistent with their increased RNA expression **(Figure 1, Figure 2B).** Conversely, the loss of *ALDH1A3* was associated with a broadly dispersed acetyl-histone H3 landscape, in which subsets of elevated acetyl-histone H3 peaks were associated with proinflammatory genes, such as *CGAS* and *IRF1*, **(Figure 4D, E; Table S4),** and consistent with the increased *IRF1* RNA expression we detected in ALDH1A3^{Low} cells and patient cancers **(Figure 1, Figure 2B)**. In addition, and consistent with our prior observations **(Figure 1, Figure 3),** pathway enrichment analysis revealed that neural crest and stem cell state marker genes, neuronal signalling, TGF-beta signalling, and O-glycosylation pathways were over-represented in ALDH1A3^{High} specific acetyl-histone H3 states **(Figure S4F, G).** Pathways enriched in acetyl-histone H3 sites in ALDH1A3 KO cells also included cell-cell junction terms **(Figure S4H)**, possibly reflecting the difference in cell morphology between the rounded ALDH^{High} cells and the flat, elongated ALDH^{Low} cells **(Figure 1B).** Together, these observations support that the high-GF and NCSC transcriptional states promoted by high ALDH activity arise from selective histone H3 acetylation.

To determine the mechanistic basis of information flow from acetyl-histone H3 to transcription in cells with high ALDH1A3, we compared genes with promoters marked by high acetyl-histone H3 (total 1499) with the RNA-seq from **Figure 1C.** We found that ~19% of genes with enriched expression in ALDH1A3^{High} (total 1144) have ALDH1A3-driven acetyl-histone H3 in their promoters (215, $p<1e-10$) **(Figure 4F)**. Of these genes, and by analysing transcription motif over-representation, we identified TFAP2 transcription factor (AP-2) and Early Growth Response Factor (EGR) 1, a pioneer factor for injury-induced regeneration[63] and regulator of melanoma proliferation[64] and neuronal stem cell states[65,66], as candidate central transcriptional mediators

(Figure 4G, H). Importantly, a significant number of targets reported in ChIP-seq datasets on EGR-1 are glucose metabolism enzymes including *G6PD*, *PDHB* and *HK2* **(Table S5)**.

The discovery of the TFAP2 motif supports a likely TFAP2B feed-forward loop, acting both as a direct target and mediator of ALDH1A3-dependent gene expression, and thereby establishing and sustaining the ALDH1A3^{High} neural crest identity. The feed-forward loop is supported by i) the high similarity of binding motif across all TFAP2 members and the exclusive association of *ALDH1A3-TFAP2B* **(Figure S1J)**, ii), the fact that TFAP2B can bind chromatin regulatory regions upstream of its own gene locus[67], and iii) our work in zebrafish proposing that Tfap2b regulates a set of genes that defines the melanocyte stem cell, including *tfap2b* itself[46]. To test our prediction, we employed gene regulatory network inference (SCENIC)[68] and identified a melanoma derived *Tfap2b* regulon **(Table S5)**. When comparing AUCell expression scores of the *Tfap2b* regulon with our ALDH^{High} enriched gene set across 16,700 single melanoma cells, originating from 5 primary mouse murine lesions[43], we detected significant co-expression of both transcriptional programs **(Figure 4I).** When we intersected Tfap2b regulon genes with ALDH^{High} enriched genes, we found a striking overlap that included both NC and stem cell genes such as *SOX2*, *SOX10*, *SPRY4* and *RXRG* **(Figure 4J, Table S5)**.

ALDH1A3 forms a predicted complex with ACSS2

Nutrition, cellular metabolism, and transcription are intimately linked with epigenetic control of gene expression as exemplified in yeast where oxidation of ethanol to generate acetyl-CoA is tightly co-ordinated with histone acetylation and expression of selective growth genes[69]. When we analysed the proteomic database STRING[70] to identify protein interactions potentially forming regulatory networks with ALDH1A3, we noted that acetyl-CoA synthetase 2 (Acs2) and pyruvate

decarboxylase (Pdc) are interacting partners of aldehyde dehydrogenase (Ald6; ortholog of ALDH1) in yeast (*Saccharomyces cerevisiae*)[71] **(Figure 5A)**. The Ald6 and Acs2 interaction in yeast was intriguing given studies of neuronal stem cell differentiation and memory formation in mammals in which acetyl-CoA synthase (ACSS2) was reported to accumulate in the nucleus and generate acetyl-CoA "on-demand" from chromatin-bound acetate associated with selective histone acetylation and gene expression[72,73].

Intrigued by these relationships, we used AlphaFold-Multimer, an extension of AlphaFold2 designed to predict protein-protein interactions[74], to predict interactions between yeast Acs2 and Ald6 and between their human homologues ACSS2 and ALDH1A3 **(Figure 5B, Figure S5A)**. This approach yielded complexes for both the yeast and human pairs with moderate confidence (45% and 47%, respectively). In addition, we identified strong interface conservation between Acs2 and ACSS2 (Pearson correlation of 0.61 between buried surface area values of homologous residues, $p = 1.7 \times 10^{-66}$). In these models, Acs2/ACSS2 binds at the perimeter of the Ald6/ALDH1A3 dimerization interface in a 1:2 stoichiometry. This binding mode ensures accessibility of the Acs2/ACSS2 binding pockets for CoA and ATP and the Ald6/ALDH1A3 substrate pockets for NAD and aldehyde. Thus, both the predicted yeast and human complexes are consistent with an ACSS2-ALDH1A3 enzymatic partnership in the nucleus.

When we performed immunofluorescence on sorted ALDH1A3^{High} and ALDH1A3^{Low} melanoma cells, we found that ACSS2, along with ALDH1A3 and TFAP2B, was localized in the nucleus of sorted ALDH1A3^{High} cells but not in ALDH1A3^{Low} cells **(Figure 5C, D)**. In the presence of an ACSS2 inhibitor, we found no change in nuclear ALDH1A3 whereas nuclear TFAP2B levels were significantly reduced **(Figure S5B-C)**. We propose that the ALDH1A3-ACCS2 enzymatic partnership is not required for ALDH1A3 nuclear localisation, however we suggest ALDH1A3-dependent TFAP2B expression and nuclear localisation is mediated through ACSS2.

To investigate potential co-localisation of ACSS2 and ALDH1A3 in cells, we sorted $ALDH^{High}$ and $ALDH^{Low}$ cells and performed super-resolution imaging. We detected abundant nuclear co-localization signals of ALDH1A3-ACSS2-TFAP2B **(Figure 5E)** and a significant positive correlation between the linear signal distribution of any two of the three target proteins **(Figure 5F)**. In contrast, we saw minimal nuclear signals of ALDH1A3, ACSS2, or TFAP2B in sorted $ALDH^{Low}$ cells and the nuclear signals captured by super-resolution imaging were not overlapping **(Figure 5G-H).**

Acetaldehyde serves as an acetyl source for histone H3-acetylation

In human cells, pyruvate dehydrogenase (PDH) can generate acetaldehyde, an ALDH substrate, from pyruvate[75]. Here, we asked if acetaldehyde could serve as a source of acetyl for histone H3 in melanoma, which by extension would support that acetaldehyde metabolism in $ALDH1A3^{High}$ cancer cells impacts upon ALDH1A3-ACSS2 regulated gene activity **(Figure 6A)**. To answer this question, we traced $^{13}C_2$-labelled acetaldehyde to histones and performed mass spectrometry analysis **(Figure 6B)**. Indeed, in our two independent melanoma cell lines (control vs *ALDH1A3* KO), we consistently detected higher levels of ^{13}C-labelled acetylated histone H3 proteins in cells with high ALDH1A3 activity **(Figure 6C)**. To our knowledge, this is the first demonstration that acetaldehyde is a direct source for acetylation of histone H3.

To identify genomic loci associated with ACSS2 dependent acetylated histone peaks and increased gene expression, we designed primers targeting the *TFAP2B* gene upstream of a histone acetylation island dependent on ALDH1A3 **(Figure 6D)** and performed ACSS2 ChIP-qPCR **(Figure 6E).** We found that ACSS2 indeed binds to these regions in control cells but much less so in *ALDH1A3* knockout cells**.** Moreover, ACSS2 ChIP-qPCR also revealed higher levels of ACSS2 binding to multiple neural crest gene promoter regions in control cells comparing to

ALDH1A3 knockout, while in contrast ACSS2 binding was enriched at the *IRF1* promoter in *ALDH1A3* knockout cells.

Based on our findings thus far, we propose a model in which ALDH1A3 and ACSS2 work together in a metabolic cascade to generate a regional chromatin source of acetyl-CoA from nuclear acetaldehyde. The acetyl-CoA source is used to selectively deposit localised histone H3 acetylation and is associated with increased expression of NCSC genes (including *TFAP2B* itself) and glucose metabolism genes.

ALDH1A3 cells promote drug resistance and disease recurrence *in vivo*

We next sought to evaluate functional relationships between ALDH1A3^{High} cells, melanoma disease progression and therapy resistance. To this end, we employed the widely used BRAF^{V600E} p53 mutant zebrafish melanoma model, in which BRAF inhibitors initially generate a response that is followed by drug resistance and recurrent melanoma growth[76-78]. Here, we dissected melanomas from BRAF^{V600E} p53 mutant zebrafish that also express GFP in the melanocyte lineage *(mitfa:GFP)*, and performed AldeRed analysis (which is similar to the Aldefluor assay but uses a red fluorescent substrate) **(Figure 7A-B)**. Through FACS analysis, we identified a distinct zebrafish cell population with ALDH^{High} activity that was also low for *mitfa:GFP* signal **(Figure 7B)**, indicating that both zebrafish and human melanoma ALDH^{High} activity cells express low levels of *MITF*. When we sorted zebrafish melanoma cells with the highest and lowest ALDH activity and compared RNA expression of key marker genes **(Figure 7C-D)**, we found that *aldh1a3* was the most enriched *ALDH* family isoform in the ALDH^{High} cells **(Figure S6A)**, in agreement with our findings in human melanoma. Furthermore, in zebrafish, melanoma cells with ALDH^{High} activity

were enriched with *aldh1a3, tfap2b, sox2,* and *sox10* while melanoma cells with ALDH^{Low} activity expressed *irf1b* **(Figure 7D)**.

Next, we administered vemurafenib (BRAF inhibitor) drug pellets to adult zebrafish with melanomas[78], and upon assessing Aldh1a3 expression by IHC through disease stages (melanoma responding, stable, and progressive disease stages) **(Figure 7E, F)**, we found that clusters of Aldh1a3 cell populations increased in drug-resistant and progressing disease. Together, these data indicate that the zebrafish melanoma model faithfully models the transcriptional states of human ALDH^{High} melanoma cells and the therapy-associated increase in levels of ALDH1A3 identified in **Figures 1 and S1.**

Finally, we designed a zebrafish drug-trial to test whether killing the Aldh1a3^{High} cells that emerge during BRAF inhibitor (BRAFi) drug resistance would impact upon disease outcome **(Figure 7F-G)**. We made use of our drug pellet method to achieve consistent dosing in zebrafish[78] and administered nifuroxazide (NAZ), a 5-nitrofuran pro-drug that specifically kills ALDH1^{High} cells in human cell lines and mouse xenograft melanoma models[39]. As expected, zebrafish melanoma responded rapidly to BRAFi treatment, but then grew back quickly despite ongoing treatment **(Figure S6B)**. When we administered NAZ treatment alone, we observed a period of stable disease followed by progressive melanoma growth **(Figure S6B)**. We then tested the two drugs in combinations: (1) BRAFi followed by NAZ (BRAFi_NAZ) at day 21, and (2) BRAFi treatment alone until day 21, followed by both BRAFi + NAZ **(Figure 7G)**. Our data demonstrate that both combinations of BRAFi plus NAZ significantly improved disease outcomes and probability of survival **(Figure 7H, Figure S6B)**.

To investigate whether the improved outcomes are achieved via on-target effect of NAZ, we examined Aldh1a3 levels in NAZ-treated zebrafish melanomas in progressive, drug resistant disease and found that Aldh1a3 clusters were absent **(Figure 7F)**. Finally, although we couldn't

comprehensively investigate the host immune response to the melanoma and treatments during the drug trials (given the lack or markers for identifying zebrafish T-cell subpopulations), we successfully documented the presence of B-cell infiltration. B cells are emerging as key players in anti-tumour adaptive immune response, and there are well validated antibodies in zebrafish for IHC. In patient samples, a single MAPKi treatment significantly decreased B-cells infiltration relative to DMSO treatment[79]. Consistent with this response, we found a loss of B-cell infiltration on BRAFi regressing disease. In contrast, and as validated by staining for Cd79a, our data demonstrate that zebrafish melanomas treated with NAZ maintained high levels of B-cells, both when NAZ was administered as a single agent and in combination with BRAFi **(Figure 7I)**. Together, these results show that targeting Aldh1a3 cells delays the emergence of resistance to BRAFi through eradicating the ALDH1A3 -melanoma stem cell pool and increasing inflammation.

Discussion

Cell state heterogeneity and plasticity endow genetically identical cancer cells with the capacity to respond differently to treatment. As a result, diverse resistant cell states emerge in patients, challenging clinical strategies that target limited static states[80]. Identifying actionable core regulators of cell heterogeneity and plasticity programmes, through which multi-dimensional information and fluctuating signalling pathways converge to permit cell state dynamics, is therefore a crucial challenge of high clinical relevance[81]. Here, we uncover that nuclear ALDH1A3, functioning as a master coordinator of metabolic and transcriptional cell states, fosters therapeutic resistance in melanoma by using acetaldehyde as an acetyl donor for histones to sustain glycolytic, stem-like gene expression programs.

We refer to this coordination of metabolism and developmental lineages in melanoma as a metabolism-neural crest stem cell state axis, based on the present work and evidence from developmental biology suggesting they may be more intimately linked than previously appreciated. In embryonic mouse development, the delaminating neural crest undergoes metabolic reprogramming towards aerobic glycolysis (Warburg effect) through a mechanism known to initiate expression of transcription factors to promote the epithelial-to-mesenchymal transition (EMT)[45]. More recent evidence in the avian neural crest supports that enhanced glycolysis, mitochondrial respiration, and the pentose phosphate pathway jointly mobilize to achieve coordination of EMT and migration[82]. At later stages, during melanocyte specification, the melanocyte (hair follicle) stem cell transcription factor Yin Yang 1 (YY1) couples MITF with metabolism gene expression and is essential for melanoma development[83,84]. However, we lack an understanding of whether these metabolic states are relevant to neural crest and stem cell states in melanoma and how they crosstalk to induce melanoma drug resistance and disease recurrence.

We identified a nuclear ALDH1A3-ACSS2 mechanism that controls both a glycolytic gene expression program and a developmental NCSC program mediated by the neural crest transcriptional factor TFAP2B. Mechanistically, we provide evidence to support a model whereby nuclear ALDH1A3-ACSS2 directs a glucose-derived acetaldehyde-acetate-acetyl-CoA flux and selectively deposits histone H3 acetylation at genomic loci encoding for glycolysis and TFAP2B-NCSC genes. Cells with low ALDH1A3 preferentially take up acetate to generate a nuclear acetyl-CoA flux, which activates a MITF and IRF1 gene program, but is otherwise broadly deposited on the chromatin as an acetate reservoir**.** To help visualize these concepts, we present a detailed schematic of this model in **Figure S7A**.

Critically, we demonstrate that endogenous acetaldehyde is a metabolite utilized by ALDH1A3 to modify chromatin. This functional relationship expands the scope of acetaldehyde metabolism from solely protecting cells from aldehyde macromolecule adducts[85,86]. Endogenous and alcohol-derived acetaldehydes, as well as other aldehydes including formaldehyde, are highly reactive towards DNA, and cells employ a robust two-tier protection mechanism to protect against aldehyde-induced DNA damage[85,87-90]. Tier 1 involves aldehyde detoxification enzymes (ALDH2, ADH5) followed by Tier 2 Fanconi anaemia DNA damage repair pathways to repair aldehyde-induced DNA damage[86]. Gene variants in these pathways in people lead to a loss of protection and directly contribute to increased cancer risks, bone marrow failure and risks associated with the alcohol-exposed developing foetus[91-96]. However, recent evidence also shows exogenous alcohol can serve as an acetyl source for histones in the brain via ACSS2, near genes involved in learning and memory, and in the liver[97,98]. Although it is not known if alcohol-derived acetylated histones are generated via local ALDH enzymes and acetaldehyde, combined with our results, these findings point to a possible broader role for acetaldehyde as a metabolite source for chromatin regulatory marks. Supporting this intimate relationship between aldehydes and chromatin modification, recent evidence shows that nucleosome demethylation releases formaldehyde[99], and that histone deacetylase 3 suppresses endogenous formaldehyde production, the reaction products of which are used in one-carbon (formate) metabolism[100,101]. Thus, chromatin may be both modified by aldehydes to shape the regulatory landscape, as well as serve as a source of metabolites when required, including Acetyl-CoA and formate.

Our data here, along with recent discoveries that formaldehyde metabolism promotes differentiation of primed melanocyte stem cells[102], highlights an emerging physiological role for aldehydes as essential metabolites governing stem cell function. Pioneering studies of nuclear condensates show that the spatial clustering of active gene transcription is intimately linked with nuclear protein-metabolite condensate distribution[103], providing a possibility that subnuclear

ALDH-aldehyde metabolism may exert local control of gene expression or acetate reservoirs. It will be important for future studies to identify the external signals that promote the ALDH1A3^{High} state following drug treatment, and address whether these contribute to the cross-resistance mechanisms triggered by targeted therapies and immune checkpoint blockade. Transcriptionally, analysis of two independent HA-tagged MITF ChIP-seq experiments show HA-MITF binds the ALDH1A3 promoter region[104,105], and given that we find MITF activity is low in ALDH1A3^{High} cells, this suggests that MITF may be a repressor of ALDH1A3 transcription **(Figure S7B)**. While an increase in pyruvate can account for the source of acetaldehyde in ALDH1A3^{High} cells in the models presented here, it will be important to identify if aldehydes generated endogenously or that are available from the microenvironment (possibly even through diet or the microbiome) can impact upon on the metabolism-neural crest stem cell state axis in melanoma.

In conclusion, we present evidence that ALDH1A3 is a master regulator of a metabolism-neural crest stem cell state axis that partners with ACSS2 to modify histone acetylation by locally depositing acetaldehyde-derived acetyl groups. In cancer cell lines and low passage patient-derived melanoma cells, we show that ALDH1A3 expression alone is sufficient to shift cells towards NCSC states. Further, we show that targeting ALDH1A3-NCSC cells *in vivo* is a relevant therapeutic strategy given evidence from our pre-clinical melanoma models showing that ALDH1A3^{High} states emerge during BRAF inhibitor resistance, and that eliminating these cells extends disease-free survival. The conceptual framework we present here for melanoma may be broadly applicable to ALDH^{High} cancer stem cell subpopulations in other cancer types, as ALDH isoforms potentially cooperate with lineage specific master transcription factors (often from developmental lineages) that are co-opted to regulate tumour cell states[106].

STAR METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal anti-ALDH1A3	Abcam	Cat # ab129815; RRID: AB_2937054
Rabbit polyclonal anti-ALDH1A3	Thermo Fisher Scientific	Cat# PA5-29188, RRID: AB_2546664
Mouse polyclonal anti-TFAP2B	Santa Cruz	Cat # SC-390119x; RRID: AB_2828008
Mouse monoclonal anti-beta Actin	Abcam	Cat # ab8226; RRID: AB_306371
Rabbit polyclonal anti-acetyl-Histone H3	Millipore	Cat# 06-599, RRID: AB_2115283
Mouse monoclonal anti-Histone H3	Abcam	Cat# ab10799, RRID: AB_470239
Rabbit recombinant monoclonal anti-ACSS2	Abcam	Cat # ab133664, RRID: AB_2943489
Rabbit monoclonal anti-ACSS2 (AceCS1)	Cell Signaling Technology	Cat# 3658, RRID: AB_2222710
Rabbit monoclonal anti-CD79a	Abcam	Cat# ab16698, RRID: AB_443433
IRDye 680RD Goat polyclonal anti-Mouse IgG	LI-COR Biosciences	Cat# 925-68070, RRID: AB_2651128
IRDye 800CW Goat polyclonal anti-Rabbit IgG	LI-COR Biosciences	Cat# 925-32211, RRID: AB_2651127
Donkey anti-rabbit conjugated Alexa Fluor 488	Invitrogen	Cat# A-21206, RRID: AB_2535792
Donkey anti-mouse conjugated Alexa Fluor 568	Invitrogen	Cat# A-10037, RRID: AB_2534013
Donkey anti-rabbit conjugated Alexa Fluor 647	Invitrogen	Cat# A-31573, RRID: AB_2536183

Bacterial and virus strains		
One Shot™ Stbl3™ Chemically Competent E. coli	Invitrogen	Cat# C737303
Biological samples		
Human patient melanoma samples	Manchester Cancer Research Centre (MCRC)	Patient samples were collected with MCRC Biobank ethics application #07/H1003/161+5 and with approval for the work under MCRC Biobank Access Committee application 13_RIMA_01
Chemicals, peptides, and recombinant proteins		
DAPI (for nucleic acid staining)	Sigma-Aldrich	Cat# D9542
7-AAD (7-Aminoactinomycin D)	Eurogentec	Cat# AS-83201
TRIzol™ Reagent	Invitrogen	Cat# 15596026
RIPA Lysis buffer (10x)	Merck Millipore	Cat # 20-188
PhosSTOP™ phosphatase inhibitor	Merck Roche	Cat # 4906837001
EASYpack™ Protease Inhibitor Cocktail	Merck Roche	Cat # 5892970001
Phenylmethylsulphonyl Fluoride (PMSF)	Life Technologies	Cat # 36978
Laemmli (SDS-Sample) Buffer (Reducing, 6X)	Enzo Life Sciences	Cat # NOV-NBP- 111R
Sodium dodecyl sulfate (SDS)	Sigma-Aldrich	Cat # L3771
Sodium Butyrate	Abcam	Cat # ab120948
DNase/RNase-Free Distilled Water	Life Technologies	Cat # 10977035
LightCycler 480 SYBR Green I Master Mix	Roche	Cat # 04707516001
EGF (Epidermal Growth Factor)	Thermo Fisher Scientific	Cat # PHG0313
FGF (Fibroblast Growth Factor)	Bio-Techne R&D	Cat # 233-FB- 025/CF

B27 Supplement	Thermo Fisher Scientific	Cat # 17504044
Low Melting Temperature Agarose	Sigma-Aldrich	Cat # A4018
UltraPure™ Agarose	Invitrogen	Cat # 16500500
Restriction digestion Enzyme BbsI-HF	NEB	Cat # R3539
Restriction digestion Enzyme KpnI-HF	NEB	Cat # R3142
Restriction digestion Enzyme BsiWI-HF	NEB	Cat # R3553
CutSmart™ Buffer	NEB	Cat # M0202M
T4 DNA ligase	NEB	Cat # M0202M
T4 DNA Ligase Reaction Buffer	NEB	Cat # B0202
Monarch RNase A (20 mg/ml)	NEB	Cat # T3018
Opti-MEM™ Reduced Serum Medium	Gibco	Cat # 31985062
DMEM High Glucose	Gibco	Cat # 11965092
RPMI 1640 Medium	Gibco	Cat # 61870036
RPMI 1640 with HEPES	Gibco	Cat # 22400089
DMEM, no glucose, no glutamine, no phenol red	Gibco	Cat # A1443001
Tumor Dissociation Kit, human	Miltenyi Biotec	Cat # 130-095-929
Red blood cell lysis buffer	BioLegend	Cat # 420301
Trypsin-EDTA (10x)	Gibco	Cat # 15400054
L-Glutamine (100x)	Gibco	Cat # 25030081
Fetal Bovine Serum	Gibco	Cat # 26140079
Penicillin/Streptomycin	Gibco	Cat # 10378016
Lipofectamine 2000 Transfection Reagent	Life Technologies	Cat # 11668019
Polybrene Infection/Transfection Reagent	Merck Millipore	Cat # TR-1003-G
Puromycin	Life Technologies	Cat # A1113803

Bovine Serum Albumin	Merck Life Sciences	Cat # A9647
16% Pierce™ Methanol-free Formaldehyde	Thermo Fisher Scientific	Cat # 11586711
Tris Base	Merck Millipore	Cat # 648310-M
Dithiothreitol (DTT), reducing reagent	Abcam	Cat # ab141390
Sucrose	Merck Millipore	Cat # S0389
NP-40	Sigma-Aldrich	Cat # I8896
Trichloroacetic acid (TCA)	VWR	Cat # 89500-550
Trifluoracetic acid (TFA)	VWR	Cat # 153112E
Acetonitrile (ACN)	VWR	Cat # 83640.320
LC-MS Grade Water	Supelco	Cat # 115333.2500
Monarch® RNase A	NEB	Cat # T3018L
Proteinase K (20 mg/ml)	Thermo Fisher Scientific	Cat # EO0491
Vectashield™ Vibrance Antifade Mounting Medium	2BScientific	Cat # H-1700-10
D-Glucose (13C-6)	Merck Life Sciences	Cat # 389374
Sodium Acetate (13C-2)	CK Isotopes	Cat # CLM-381-1
FACSmax™Cell Dissociation Solution	AMS Biotechnology	Cat # T200100
Liberase TM Research Grade	Merck Roche	Cat # 5401119001
vemurafenib (PLX4032)	SelleckChem	Cat # S1267
Nifuroxazide (NAZ)	Merck Millipore	Cat # 481984
Dimethyl sulfoxide (DMSO)	Sigma Aldrich	Cat # 67-68-5
Critical commercial assays		
Aldefluor Assay (with DEAB and 2N HCl included)	STEMCELL technology	Cat# 01705
AldeRed ALDH Detection Assay	Millipore Europe	Cat # SCR150
SureBeads™ Starter Kit Protein A and G	Bio-Rad	Cat #1614833
Direct-zol RNA Miniprep Kits	ZYMO Research	Cat# 2052

SuperScript III Reverse Transcriptase	Life Technologies	Cat # 18080044
Illumina TruSeq Stranded mRNA Sample Prep Kit	Illumina	Cat # RS-122-2001
NEBNext Ultra Directional RNA library Prep Kit	NEB	Cat # E7420S
NEBNext Ultra II DNA Library Prep Kit for Illumina	NEB	Cat # E7645S
NEBNext Multiplex Oligos for Illumina® (Index Primers Set 1)	NEB	Cat # E7335S
AMPure XP beads	Beckman Coulter	Cat # A63882
QIAquick PCR Purification Kit	Qiagen	Cat # 28104
Pierce™ BCA Protein Assay Kit	Thermo Fisher	Cat # 23225
4–15% Mini-PROTEANTM TGX Precast Protein Gels	Bio-rad	Cat # 4561085DC
Trans-Blot™ Turbo Mini 0.2 µm Nitrocellulose Transfer Packs	Bio-rad	Cat # 1704158
Collagen I, Coverslip Glass 22mm Round	VWR	Cat # 734-1009
L-Lactate Assay Kit (Colorimetric/Fluorometric)	Abcam	Cat # ab65330
Zenon™ Rabbit IgG labelling kits Alexa Fluor 488	Thermo Fisher	Cat # Z25302
Zenon™ Rabbit IgG labelling kits Alexa Fluor 647	Thermo Fisher	Cat # Z25308
Deposited data		
Raw and analyzed A375 $ALDH^{High}$ and $ALDH^{Low}$ RNA-seq data	This paper	GEO: GSE243840
Raw and analyzed A375 Cas9 ALDH1A3 WT and ALDH1A3 KO ChIP-Histone acetyl-H3-seq data	This paper	GEO: GSE243834

AlphaFold multimer modelling result of ALDH1A3-ACSS2 interaction	This paper	Mendeley Data (Reserved DOI:10.17632/4hfw7rpspb.1); To be published on Model Archive upon manuscript publication
AlphaFold multimer modelling result of Ald6-Acs2 interaction	This paper	Mendeley Data (Reserved DOI:10.17632/4hfw7rpspb.1); To be published on Model Archive upon manuscript publication
Mouse melanoma single-cell RNAseq	(Karras et al., 2022)[43]	GEO: GSE207592
Patient melanoma single-cell RNAseq	(Tirosh et al., 2016)[50]	GEO: GSE72056
Human cutaneous melanoma TCGA dataset	TCGA cBioPortal; https://www.cbioportal.org/	RRID: SCR_003193
Human cutaneous melanoma Lund dataset	(Cirenajwis et al., 2015)[48]	GEO: GSE65904
Human cutaneous melanoma Bergen dataset	(Jonsson et al., 2010)[49]	GEO: GSE22155
Human melanoma pre-anti-PD-1 therapy RNA-seq correlating with Ipilimumab response	(Hugo et al., 2016)[24]	GEO: GSE78220
Tfap2b (biotin) ChIP-seq	(Ling and Sauka-Spengler, 2019)[67]	GEO: GSE125711
HA-MITF ChIP-seq	(Louphrasitthiphol P et al., 2020)[104] (Louphrasitthiphol P et al., 2023)[105]	GEO: GSE137522; GSE137776
Experimental models: Cell lines		
Human: A375 melanoma cell line	ATCC	RRID: CVCL_6233
Human: Cas9 A375 melanoma cell line	David Adams, Sanger Institute, UK	N/A

Human: Cas9 C089 melanoma cell line	David Adams, Sanger Institute, UK	N/A
HEK293T	ATCC	Cat# CRL-3216; RRID: CVCL_0063
MPD001	This paper	N/A
Experimental models: Organisms/strains		
Zebrafish: *Tg(mitfa:GFP, mitfa:BRAFV600E), p53*^{M214K}	(Patton et al. 2005)[76]	ZFIN Cat #: ZDB-TGCONSTRCT-070117-106, ZDB-ALT-050428-2
Oligonucleotides **(See also Table SX)**		
gALDH1A3: CGTCCCGGAGCAATCTGAAG	(Sanjana, Shalem and Zhang, 2014)[107]	https://www.genscript.com/tools/gRNA-library-design
gTFAP2B: GGACCGAGTGGTAGTCCCTC	(Sanjana, Shalem and Zhang, 2014)[107]	https://www.genscript.com/tools/gRNA-library-design
Recombinant DNA		
pKLV2-U6gRNA5(BbsI)-PGKpuro2A-BFP	Addgene	Cat # 67991
pKLV2-U6gRNA5(BbsI)-PGKpuro2A-mCherry	This paper	N/A
pKLV2-U6gALDH1A3-PGKpuro2A-BFP	This paper	N/A
pKLV2-U6gTFAP2B-PGKpuro2A-BFP	This paper	N/A
pKLV2-U6gRNA5(BbsI)-PGKpuro2A-ALDH1A3	This paper	N/A
pKLV2-U6gRNA5(BbsI)-PGKpuro2A-TFAP2B	This paper	N/A
psPAX2	Addgene	Cat # 12260
pMD2.G	Addgene	Cat # 12259
Software and algorithms		

R (v. 4.0.4)	R core team (2020): (https://www.R-project.org/.)	RRID:SCR_001905
R Studio (v. 1.4.1106)	RStudio Team (2020): (http://www.rstudio.com/.)	RRID:SCR_000432
ggPlot2 (v.3.4.2)	(Wickham, 2016) https://cran.r-project.org/web/packages/ggplot2/index.html	RRID:SCR_014601
MATLAB	The Math Works (2020): (http://www.mathworks.com/products/matlab/)	RRID:SCR_001622
FastQC (v.0.11.3)	Babraham Institute; (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/)	RRID: SCR_014583
STAR (v. STAR_2.5.1b)	(Dobin et al., 2013)[108]	RRID: SCR_015899
RNASeqQC (v.1.1.8.1)	(DeLuca et al., 2012)[109]	RRID: SCR_005120
htseq-count (0.6.1)	(Simon et al., 2015)[68]	RRID:SCR_011867
DESeq2 R package (v. 1.30.1)	(Love et al., 2014)[68]	RRID: SCR_015687
GSEA software (v. 4.0.3)	(Subramanian et al., 2005)[110] (http://www.broad.mit.edu/gsea/index.html)	RRID: SCR_016884
g:Profiler	(Raudvere et al., 2019)[111]	RRID:SCR_006809
Seurat (v. 4.0.0)	(Hao et al., 2021)[112]	RRID:SCR_016341
EpiProfile2.1_1Basic	(Yuan et al., 2018)[113]	N/A
Bowtie2 (bwa-0.7.9)	(Langmead and Salzberg, 2012)[68]	RRID:SCR_016368
HOMER (version 4.11)	(Heinz et al., 2010)[114]	RRID:SCR_010881
ChIPseeker R package (v. 1.26.2)	(Wang et al., 2022)[115]	RRID:SCR_021322
SAMtools (v. 1.9)	(Li et al., 2009)[116]	RRID:SCR_002105
MACS2 (v. 2.1.1.20160309)	(Feng et al., 2012)[117]	N/A

deeptools (v. 3.5.1)	(Ramírez et al., 2016)[118]	RRID:SCR_016366
DiffBind (v. 2.10.0)	(Ross-Innes et al., 2012)[119]	RRID:SCR_012918
Galaxy (usegalaxy.org)	(The Galaxy Community, 2022)[120]	RRID:SCR_006281
The UniProt Consortium	(The UniProt Consortium, 2019)[121]	RRID:SCR_002380
AlphaFold-Multimer	(Evans et al. 2021)[74]	N/A
ColabFold version 1.5.2	(Mirdita et al. 2022)[122]	N/A
GalaxyRefineComplex	(Heo, Lee, and Seok 2016)[123]	N/A
FreeSASA 2.0.3	(Mitternacht 2016)[124]	N/A
AlphaFill	(Hekkelman, de Vries, et al., 2023)[125]	N/A
UCSF ChimeraX version 1.6	(Pettersen et al., 2021)[126]	RRID:SCR_015872
Image J Fiji software (v. 1.53c)	National Institutes of Health, USA	RRID: SCR_003070
Prism 9 (v. 9.3.1) for macOS	GraphPad Software, San Diego, USA	RRID: SCR_002798
eulerr R (v. 6.1.1)	(Larsson, 2021) https://CRAN.R-project.org/package=eulerr	RRID:SCR_022753
Cytoscape (v.3.8.0)	(Otasek et al., 2019)[127]	RRID:SCR_003032
SCENIC	Aibar et al. 2017[68]	RRID:SCR_017247
AUCell	Aibar et al. 2017[68]	RRID:SCR_021327
CytoTRACE	Gulati et al. 2020[42]	RRID:SCR_022828
pheatmap	https://CRAN.R-project.org/package=pheatmap	RRID:SCR_016418
CIBERSORTx	(Chen et al., 2019)[128] (https://cibersortx.stanford.edu/)	N/A

NDP.view 2 (v. 2.9.29)	Hamamatsu.com: NDP.view2 U12388-01 (http://www.hamamatsu.com/jp/en/U12388-01.html)	N/A
FlowJo (v. 10.8.1) for macOS	Becton, Dickinson and Company (https://www.flowjo.com/)	RRID:SCR_008520

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, E. Elizabeth Patton (e.patton@ed.ac.uk).

Materials availability

Plasmids generated in this study will be made available upon request made to the lead contact and will be deposited in Addgene.

Patient derived low passage melanoma cells are available, upon MTA approval, upon request to v.pavet@beatson.gla.ac.uk and o.sansom@beatson.gla.ac.uk.

EXPERIMENTAL MODELS AND SUBJECT DETAILS

Zebrafish maintenance and husbandry

Zebrafish were maintained in accordance with UK Home Office regulations, UK Animals (Scientific Procedures) Act 1986, under project license P8F7F7E52. All experiments were approved by the Home Office and AWERB (University of Edinburgh Ethics Committee).

Zebrafish melanoma models

Zebrafish were genotyped using DNA extracted from tail fin clipped tissue by PCR to confirm the mutant allele status of *tp53*^{M214K} (referred to as *p53*$^{-/-}$ *or p53 mutant*) and *mitfa:BRAF*^{V600E} as

described in our previous publications[76]. The emergence of melanoma is usually observed in individuals aged 3- to 6-month-old. Individuals used in this study for DMSO control versus vemurafenib or Nifuroxazide drug pellets treatment were aged 5- to 6-month-old when entering the treatment scheme. Both female and male individuals were admitted into the treatment course regardless of the sex.

Human melanoma cell line culture

A375 cells were cultured in DMEM high glucose media, C089 cells were cultured in RPMI 1640 media, and patient sample derived cell line MPD001 were cultured in RPMI 1640 with 25 mM HEPES. All media were supplemented with 2 mM L-glutamine and 10% fetal calf serum, and all cells grown at 37 °C in a 5% CO_2 humidified incubator. All cells have been routinely tested for mycoplasma, with the most recent test performed on June 21, 2023.

METHOD DETAILS

Establishment and amplification of patient derived melanoma cell lines.

Melanoma Patients were managed in accordance with the ethical principles of Declaration of Helsinki and in accordance with Good Clinical Practice as defined by the International Conference on Harmonisation. All patients gave informed written consent to participate in clinical trials or EAP or EAMS. Patient samples were collected with written full-informed consent under Manchester Cancer Research Centre (MCRC) Biobank ethics application #07/H1003/161+5 and approval for the work under MCRC Biobank Access Committee application 13_RIMA_01. Tissue samples were collected from cutaneous melanoma patients at The Christie NHS Foundation Trust. For cell line MPD001, tumour fragments from the lymphatic melanoma metastases were obtained during the surgical procedure and dissociation was performed on the same day utilising the human tumour dissociation kit (Miltenyi Biotec) following manufacturer's instructions. Briefly, after mechanical dissociation with a scalpel, tissue was resuspended in RPMI media with a mixture of Kit Enzymes (H, R and A) and placed on the gentleMACS™ Dissociator, program 37°C _h_TDK_1 for 30 minutes. Next, cells were transferred to a 50 ml falcon tube and centrifuged at 300xg for 7 minutes, then resuspended and passed through a 70 µm cell strainer. This was centrifuged again (300xg for 7 minutes) and resuspended in 1X red blood cell lysis buffer (BioLegend) in deionized water, then incubated at room temperature, protected from light, for 15 minutes. After this step, the sample was centrifuged and finally resuspended and plated in a 10 cm cell culture petri dish

in RMPI medium supplemented with FBS (Gibco) (10%) and Penicillin/Streptomycin (Gibco) (1%). Cells were cultured in an incubator at 37°C with 5% CO_2. Cell lines were frozen and stored after a 2-week amplification period, freezing performed using FBS with DMSO (10%). Cryotubes were placed in an isopropanol freezing container (Nalgene® Cat. C1562-1EA) at -80°C for 24h for gentle freezing and transferred to liquid nitrogen tanks for long term storage.

Generation of human melanoma mutant cell lines

Human melanoma cell lines A375 and C089 engineered with stable expression of *Streptococcus pyogenes* Cas9 were gifted to us and used to build *ALDH1A3* mutant cells[39]. Briefly, for *ALDH1A3* knockout, the vehicle plasmid expressing gRNA (pKLV2-U6gRNA5(BbsI)-PGKpuro2ABFP, Addgene: 67991) was engineered by Golden Gate cloning using restriction enzyme BbsI-HF (NEB) to express gRNA targeting ALDH1A3 (CGTCCCGGAGCAATCTGAAG). Lentiviral particles were produced by co-transforming the HEK293T (ATCC) cells with target plasmid, the packaging plasmid psPAX2 (Addgene: 12260), and pMD2.G (Addgene: 12259), facilitated with Lipofectamine 2000 (Invitrogen) following the manufacturer's instructions. 48 hours post transfection, the 293T cell culture supernatant was collected and filtered (0.45 µm) to transfect the targeted recipient cell lines, supplemented with 10 µg/ml polybrene (Merck Millipore). 48-72 hours post transfection, the cells were split and seeded with the complete growth media containing 1 µg/ml puromycin (Life Technologies) to obtain clones with stable expression of the gRNA. Suitable single clones were validated by western blot to confirm full-length *ALDH1A3* knockout and Aldefluor assay to confirm the loss of ALDH activity before expanded for biological experiments. For *TFAP2B* knockout cells, the gRNA sequence was engineered similarly (GGACCGAGTGGTAGTCCCTC) using Golden Gate cloning. For vehicle control samples, the vehicle plasmid expressing empty gRNA were engineered to express mCherry instead of *BFP* sequence; for *ALDH1A3* and *TFAP2B* over-expression cells, the vehicle plasmid expressing empty gRNA were engineered to express *ALDH1A3* CDS instead of *BFP*. Briefly, restriction digestion enzyme KpnI-HF (NEB) and BsiWI-HF (NEB) were used to remove *BFP* sequence and create sticky ends matching the *mCherry*, *ALDH1A3* or *TFAP2B* CDS flanking sequence. The digested fragment of plasmid backbone and the target sequence were then ligated using T4 DNA ligase (NEB), following which a similar transfection and selection procedures were performed to acquire stable mutant lines.

Human melanoma cell ALDH activity measurement

The ALDH enzyme activity in human melanoma cells was measured using the Aldefluor assay kit (StemCell Technologies) following the manufacturer's instructions. In brief, melanoma cells dissociated with trypsin (Gibco) were resuspended in the Aldefluor buffer with the fluorescent bodipy-aminoacetaldehyde (BAAA) reagent included in the kits. For each experiment, a negative control vial was set up using a small aliquot out of the sample vial (100 µl out of 1 ml), supplemented with (5 µl) pan-ALDH inhibitor DEAB (diemethylaminobenzaldehyde) immediately after the resuspension. After incubation of 30 minutes at 37°C, the Aldefluor activity was measured using flow cytometry (Fortessa, BD Biosciences). For cell sorting to establish the $ALDH^{High}$ and $ALDH^{Low}$ cells, the stained cells were sorted by FACS Aria II (BD Biosciences) and the population with the highest and lowest 5% ALDH activity ($ALDH^{High}$ and $ALDH^{Low}$) were collected. All flow cytometry data were analysed using the software FlowJo. Dead cells were excluded using 7-Aminoactinomycin D (7-AAD, Eurogentec), or 4,6-Diamidine-2-phenylindole dihydrochloride (DAPI, Sigma-Aldrich).

RNA extraction and RT-qPCR

Quantitative reverse transcription PCR (RT-qPCR) assays were performed by standard protocol suitable for LightCycler® 480 Instrument (Roche). In brief, total RNA was extracted and purified from live cells using TRIzol (Invitrogen) and Direct-Zol RNA Miniprep Kits (ZYMO Research). After quality check and measurement by NanoDrop™ (ThermoFisher), 1 µg RNA for each sample was reverse-transcribed using Superscript III reverse transcriptase (Life Tchnologies). Quantitative PCR were carried out by setting up reactions using the reverse transcribed cDNA template, primers (see also **Table S6, related to STAR Oligonucleotides**) and LightCycler ® 480 SYBR Green I Master reagent, run by the program of Standard Roche Template (System II). Reads of gene ACTB (Beta Actin) were used as the internal control to calculate the relative expression values.

RNA-seq Pipeline

Libraries were prepared from 500 ng of each total-RNA sample using the TruSeq Stranded mRNA Library Kit (Illumina) according to the provided protocol and purified using AMPure XP beads (Beckman Coulter). Sequencing was performed using the NextSeq 500/550 High-Output v2 (150 cycle) Kit (# FC-404-2002) on the NextSeq 550 platform (Illumina Inc, #SY-415-1002). Libraries were combined in an equimolar pool based on the library quantification results and run across a single High-Output Flow Cell. Raw FASTQ sequence reads were quality checked using FastQC (v. 0.11.3) and aligned to the human genome (GrCh38) assembly using STAR (v. STAR_2.5.1b)

software with default parameters. The quality of the resulting alignment to the transcriptome (Ensembl annotation version GRCh38.91) was checked using RNASeqQC (v. 1.1.8.1). Raw counts of reads covering the transcriptome (Ensembl annotation version GRCh38.91) were obtained using htseq-count (0.6.1) with the "-s reverse" option. Differential expression was analysed using the DESeq2 R package (v. 1.20.0).

Pathway Enrichment Analysis

Gene set functional enrichment analysis (GSEA) identified enriched pathways at FDR < 0.05 (Kolmogorov–Smirnov test) using GSEA software (Subramanian et al., 2005). The gene expression matrix of sorted $ALDH^{High}$ versus $ALDH^{Low}$ cells or selected melanoma patient $ALDH1A3^{High}$ versus $ALDH1A3^{Low}$ groups were used to compare against literature-based datasets **(Table S1, S2)** using gene set permutation settings. For gene over-representation analysis across literature curated signature terms, genes with differential acetyl-Histone H3 in WT compared to *ALDH1A3* knockout were selected based on FDR < 0.05 (in total 1599, see also **Table S4**) as input to the g:Profiler website[111]. The gene terms with FDR < 0.05 (hypergeometric test, BH adjusted) were considered significantly over-represented, with the term size cut-off set to 1000 (token: gp__33Kp_FaEr_Mcs). For gene over-representation analysis across the REACTOME database[129], genes with acetyl-histone H3 peaks differentially enriched in WT or *ALDH1A3* knockout (see also **Table S4**) were separately tested and terms of FDR < 0.05 (hypergeometric test, BH adjusted) were considered significantly over-represented.

Human metastatic melanoma single cell data mining

Human metastatic melanoma single-cell RNA-seq expression matrix was accessed via supplementary data of Tirosh et al., 2016[50]. Cells classified as malignant tumour cells were extracted for UMAP clustering via Seurat (v.4.0.0), with *ALDH1A3* and *TFAP2B* visualised, clustering resolutions equals 0.5.

***NRAS$^{Q61K/°}$; Ink4a$^{-/-}$* murine melanoma data mining**

Single-cell RNA expression data of malignant melanoma cells, originating from 5 primary murine melanoma lesions[43], were interrogated for *Aldh1a3*, ALDHhigh_enriched_signature and Tfap2b_regulon expression activities, measured by AUCell[68]. A Tfap2b regulon is inferred by using the pySCENIC pipeline on a mouse melanoma single-cell RNA-seq data[43] with the genes extracted and visualized using Cytoscape (v.3.8.0)[130]. To map potential differentiation trajectories onto the single-cell UMAP space, we calculated for each melanoma cell its corresponding

CytoTRACE score[42], which is a measure of gene expression diversity and a surrogate for developmental potential (0<CytoTRACE score<1). CytoTRACE scores close to 1 are indicative of a less differentiated and close to 0 of a differentiated state. Besides the single cells, every gene was scored and correlated to either contribute to a dedifferentiated or differentiated state and ranked accordingly.

Profiling tumour infiltrating immune cells with CIBERSORT

Human cutaneous melanoma microarray of Lund patient cohort[48] were retrieved from GEO: GSE65904 and ranked based on *ALDH1A3* levels. Top and bottom 10% of ranked samples were selected as ALDH1A3^{High} and ALDH1A3^{Low} phenotype groups and used for tumour infiltrating immune cells profiling via CIBERSORTx[128] (https://cibersortx.stanford.edu/). The CIBERSORT derived leukocyte specific gene signature matrix LM22[68] was used to impute immune cell fractions, with the setting in absolute mode to allow direct comparison across relative differences between cell types. Tumour infiltrated immune cell type and fraction matrix of ALDH1A3^{High} and ALDH1A3^{Low} samples were then clustered and visualized using pheatmap package in R to allow interpretation.

Western Blot

Cells were detached by Trypsin (Gibco) and lysed on ice for 30 min at a cell density of 10^7 cells per ml of RIPA lysis buffer (Merck Millipore), supplemented with phosphatase and complete protease inhibitors (Merck Roche). Debris of cells were removed by centrifuge (10,000 rpm, 10 min, 4 °C). Protein concentrations were determined using the BCA assay (Thermo Fisher) and 10-20 mg protein per lane was electrophoresed on 4-15% precast gradient gels (Bio-Rad). Based on the protein concentration and sample volumes, calculated amount of Laemmli (SDS) buffer (Enzo Life Science) were added and incubated with the samples at 95 °C for 5 min before gel loading. After the electrophorese program, gels were transferred onto Turbo™ transfer membranes (Bio-Rad) using the semi-dry Turbo Transfer system (Bio-Rad). Membranes were then blocked with 5% w/v BSA/TBS for 30 min at room temperature, subsequently probed by primary antibodies with optimised dilution factors overnight at 4°C, incubation of goat anti-mouse IRDye 680- or goat anti-rabbit 800- labelled secondary antibodies (LI-COR Biosciences) and imaged using an Odyssey infrared scanner (LI-COR Biosciences).

To probe the histone proteins using western blot, an acid extraction protocol was carried out instead of RIPA lysing protocol. In brief, the cells were washed with ice-cold PBS and lysed on

ice for 10 min at a cell density of 10^7 cells per ml of Triton Extraction Buffer (TEB: PBS, 0.5% Triton X 100 (v/v), 2 mM phenylmethylsulfonyl fluoride (PMSF, Life Technologies)), supplemented with 5 mM sodium butyrate (Abcam) to retain levels of histone acetylation. The nuclei of cells were centrifuged (650 x g, 10 min, 4 °C) and washed in half the volume of TEB and centrifuged again. Pellets were re-suspended using 0.2 N HCl at a density of $4x10^7$ nuclei per ml. The histones were extracted overnight at 4°C and the debris were removed by centrifuge (650 x g, 10 min, 4 °C). 5M NaOH were added at 1/20 of the volume of the supernatant to neutralise the samples. The same steps of protein measurement and western blot were followed as described above for probing protein in the total cell lysates.

The primary antibodies used are as following: ALDH1A3 (1:10,000, Rabbit, Invitrogen), TFAP2B (1:1000, Mouse, Santa Cruz), beta-Actin (1:1000, Mouse, Invitrogen), Histone H3 (1:1000, Mouse, Abcam), acetyl-Histone H3 (1:10,000, Rabbit, Millipore).

Colony formation assay in Soft Agar

For each sample in a 6-well plate set up, 5000 cells were suspended in serum-free growth media supplemented with 10 ng/ml EGF, 10 ng/ml FGF, B27 supplement (1x, Thermo Fisher Scientific), and 0.3% low melting agarose (Sigma-Aldrich). Cells were then layered over a solid base of 0.5% low melting agarose, cultured for 18 days, optimised by the colony growth speed. Colonies (>10 cells) from 10 separate fields of each sample were then manually counted using the images captured by a Nikon DS-L3 camera system on an Eclipse TS100 microscope (Nikon) with a 4X objective.

Immunocytochemistry with fluorescence labelling and imaging

Cells were cultured on the collagen I coated coverslip glasses (VWR) for ICC assays. Samples were washed twice with PBS and fixed with 4% paraformaldehyde (PFA, Thermo Fisher Scientific) at room temperature for 15 min before permeabilized by PBST (0.1% triton x-100/PBS) for 10 min. Coverslips were then incubated in 3% BSA/PBST (w/v) for 30 min before probed with primary antibodies at appropriate dilution for 4 hours at room temperature or 4°C overnight (ALDH1A3, 1:300, Abcam; TFAP2B, 1:100; ACSS2, 1:100, Abcam). After being washed 3 times in PBST for 5 min each, samples were incubated with fluorochrome-conjugated secondary antibodies (Donkey anti-Rabbit 488, Donkey anti-Mouse 568, and Donkey anti-Rabbit 647, all 1:1000, Invitrogen) for 30 min in the dark. After 3 times of wash in PBST for 5 min each, DAPI (Sigma-Aldrich) was added to the final wash of PBST to stain the nucleus, and the coverslips

were mounted with antifade mounting medium (Vectashield, 2BScientific) before fluorescent microscope imaging using multimodal Imaging Platform Dragonfly (Andor technologies, Belfast UK). Images were acquired using a 20X or 40X lens equipped with 405, 488, 561, and 640 nm lasers built on a Nikon Eclipse Ti-E inverted microscope body with Perfect focus system (Nikon Instruments, Japan). Data were collected in Spinning Disk 25 µm pinhole mode on the Zyla 4.2 sCMOS camera using a Bin of 1 and no frame averaging using Andor Fusion acquisition software.

For ICC on flow cytometry sorted $ALDH^{High}$ and $ALDH^{Low}$ cells, Aldefluor-sorted cells were resuspended in normal culture media and allow to attach on the collagen I coated coverslip glasses (VWR) in 6-well plates for 4-6 hours before fixed to ensure the ALDH activity states.

For interrogating ALDH1A3-TFAP2B-ACSS2 co-localisation, primary antibodies raised in the same species, *i.e.* anti-ALDH1A3 and anti-ACSS2 Rabbit IgG, were pre-conjugated to different fluorescence dye (Alexa Fluor 488 and Alexa Fluor 647) separately following the manufacturer's instructions (Zenon Rabbit IgG labelling kits, Thermo Fisher) before incubation with the samples together with Mouse anti-TFAP2B primary antibody for 2 hours at room temperature. The cells were then washed three times in PBS and incubated with Donkey-anti-mouse conjugated Alexa Fluor 568 antibodies for 0.5 hour at room temperature (Invitrogen). Cells were washed, nuclear stained with DAPI (Sigma) and mounted in Vectashield® mounting media before imaging. Super-resolution images were acquired using instant Structured Illumination Microscopy (SIM)[131] with Nikon SoRa™ system. Imaging was carried out using an SR HP Plan Apo λS 100x 1.35NA Silicone lens (Nikon Instruments). The CMOS cameras used for acquisition were Teledyne Photometrics Prime 95B (Teledyne Photometrics 3440 E.Britannia Drive, Tucson AZ) and 405 / 488 / 514 / 561 / 640nm laser lines. Z-step size for Z stacks was set to 0.120 µm as required by manufacturers software. Acquisition of images and deconvolution was carried out using the (3D algorithm) Nikon NIS Elements Advanced Research software. Settings for acquisition and reconstruction were identical in all images. Line scanning signal correlation analysis were carried out in Fiji by taking the intensity from each channel across the scanning line as an individual variable and conducting Pearson correlation test.

$^{13}C_6$-glucose tracing via targeted UPLC-MRM/MS

For $^{13}C_6$-glucose tracing experiment, DMEM with no glucose, no glutamine, no phenol red (Gibco) was purchased and then supplemented with 4.5 g/L $^{13}C_6$-glucose (Merck Life Sciences), 2 mM L-glutamine, and 10% FCS (referred to from now on as the $^{13}C_6$-glucose DMEM). To trace $^{13}C_6$-

glucose in A375 $ALDH^{High}$ and $ALDH^{Low}$ cells, A375 melanoma cells were cultured in T175 flasks (Corning) until reaching 70% confluence, then incubated in $^{13}C_6$-glucose DMEM for 12 hours before dissociated for Aldefluor staining and live sorted by FACS Aria II (BD Biosciences). Every sample contains ~0.5 million sorted cells, which was immediately snap frozen in liquid nitrogen upon sorting. To trace $^{13}C_6$-glucose in A375 Cas9 control and *ALDH1A3* knockout cells, cells were seeded in 6-well plates until reaching 50% confluence, then incubated in $^{13}C_6$-glucose DMEM for 24 hours before trypsin dissociation and liquid nitrogen snap frozen. Metabolites extraction, derivation, and targeted UPLC-MRM/MS profiling the central carbon metabolites were performed as described in Han et al., 2013[132].

Histone acetylation profiling using bottom-up mass spectrometry

Histone extraction and derivatization workflow was optimised based on Sidoli et al., protocol[133]. Briefly, nuclei from live attached melanoma cells were isolated by 5-minute incubation on ice with NIB buffer (15 mM Tris, 60 mM KCl, 15 mM NaCl, 5 mM $MgCl_2$, 1 mM $CaCl_2$, 250 mM sucrose, pH adjusted to 7.5, supplemented with 1 mM DTT, 0.5 mM PMSF, 0.05 µM NaF, 0.05 µM $NaVO_4$ and 10 mM sodium butyrate, 0.1% NP-40) followed by scraping. Nuclei were spin down by centrifuging at 1000 rcf for 5 min at 4 °C. The nuclei pellets were washed twice by NIB without NP-40. Histone proteins were extracted by adding 0.2 M H_2SO_4 in a 1:4 ratio of nuclear pellet to H_2SO_4 and incubated for 4 hours at 4 °C. Histones were then precipitated by adding 1:3 v/v Trichloroacetic acid (TCA) and incubated overnight at 4 °C. For chemical derivatization, precipitated histone proteins were air-dried by vacuum centrifuge and rinsed with ice-cold acetone prior to four rounds of propionylation, with the last two rounds of propionylation carried out on histone peptides post trypsin digestion. Propionlated peptides were transferred to C18 staging tips for desalting and eluted using 80% acetonitrile (ACN) with 0.1% trifluoracetic acid (TFA) before LC-MS analysis.

Peptides resulting from all digestions were separated by nanoscale C18 reverse-phase liquid chromatography using an UltiMate 3000 RSLCnano system coupled online to an Orbitrap Fusion Lumos Tribrid mass spectrometer (Lumos) (all Thermo Fisher Scientific). HPLC buffers- 0.1% formic acid in HPLC-grade water (buffer A); 0.1% formic acid in HPLC-grade acetonitrile (buffer B) were prepared. HPLC method was programmed as follows: from 0 to 30% buffer B in 30 min, from 30 to 100% B for the next 5 min and at isocratic 100% B for 8 min. the flow rate was set to 250-300 nl/min.

Acquisitions were carried out in data independent acquisition mode (DIA) using Tune application 3.5.3890 (Thermo Scientific). A nanoelectrospray ion source (Sonation) was used for ionisation in positive mode. Chromatography was carried out at a flow rate of 250-300 nl/min using 50 cm fused silica emitters (CoAnn Technologies) packed in house with reverse phase Reprosil Pur Basic 1.9 µm (Dr. Maisch GmbH). The emitter was heated to 50°C using a column oven (Sonation), and an Active Background Ion Reduction Device (ABIRD) was used to decrease air contaminants signal level. Peptides were eluted with a 60-minute two-step gradient, over a total run time of 90 minutes. A full scan was acquired at a resolution of 60,000 at 200 m/z, over mass range of 300-1100 m/z, followed by a DIA scan. All precursors were fragmented using 15 consecutive windows with 50 Da width, allowing for a 1 m/z overlap, covering a mass range from 349.5 to1100.5 m/z. Higher energy collisional dissociation fragmentation spectra were recorded at 15,000 resolution at 200 m/z. All ions were fragmented using normalised collision energy of 28%, for a maximum injection time of 54 ms, or a normalised AGC target of 1000%.

Data analysis was performed on MATLAB using the EpiProfile2.1_1Basic package (DOI: 10.1021/acs.jproteome.8b00133) using label-free settings (nsource=1) or for histone H3 acetylation and 13C2-acetyl incorporation analysis the C13 on acetylation group (nsource=3). Isotopic correction was performed by MATLAB as implemented in EpiProfile.

Acetyl-CoA extraction and LC-MS analysis

To extract acetyl-CoA from whole cell lysates and nuclear department, samples were collected as described by Trefely et al., protocol[134]. Briefly, live attached melanoma cells were isolated by 5-minute incubation on ice with NIB buffer (15 mM Tris, 60 mM KCl, 15 mM NaCl, 5 mM $MgCl_2$, 1 mM $CaCl_2$, 250 mM sucrose, pH adjusted to 7.5, supplemented with 1 mM DTT and 0.1% NP-40) followed by scraping. 10% of each sample were removed and quenched in 1 ml ice-cold 10% TCA as total lysates. Nuclei were spun down by centrifuging at 600 rcf for 5 min at 4 °C. The nuclei pellets were washed twice by NIB without NP-40 and then quenched in 1 ml 10% TCA. Acetyl-CoA was extracted using solid phase extraction to remove TCA. Extraction cartridges (Oasis® HLB 1cc (30 mg)) were conditioned by running 1000 µL of methanol (MeOH) following by 1000 µL of water (H_2O). On the cartridges, 100 µL of samples were then loaded, washed with 1000 µL of H_2O and eluted with 500 µL of MeOH. The extracts were then dried under nitrogen and reconstituted in 80% ACN: 20% H_2O (20 mM ammonium carbonate 0.1% ammonium hydroxide solution 25%).

Samples were analysed on a Dionex UltiMate 3000 LC System (Thermo Scientific, Waltham, Massachusetts, EUA) coupled to a Q Exactive Orbitrap Mass Spectrometer (Thermo Scientific, Waltham, Massachusetts, EUA) operating in negative polarity with scan range from 806 to 815 m/z. Chromatographic separation was achieved using a ZIC®-pHILIC 150 x 2.1 mm column (Merck Millipore Sigma, Burlington, Massachusetts, EUA) at 45°C using a gradient starting from 20% buffer A (20 mM ammonium carbonate 0.1% ammonium hydroxide solution 25%), and 80% B (acetonitrile) to 80% buffer A, 20% buffer B at 9.5 minutes and reconditioning the column to the initial condition until 14.5 minutes. Mass spectrometry data were processed using Skyline[135] on a targeted fashion by matching accurate mass and retention time with standard.

A375 acetyl-Histone H3 ChIP-Seq

The control and *ALDH1A3* KO melanoma cells were cultured to 80% confluency and harvested by dissociation with trypsin in PBS/EDTA. Cells were resuspended in PBS and immediately fixed in 1% formaldehyde in PBS for 10 minutes at room temperature. The fixation was terminated by adding glycine and incubated for an additional 5 minutes. After being washed in cold PBS, cell pellets were resuspended in 150 µl chilled lysis buffer (1% SDS, 10mM EDTA, 50 mM Tris-HCl pH8.1, 1x protease inhibitor cocktail, 1x PhosSTOP phosphatase inhibitors (Roche), 5 mM sodium butyrate (Sigma) and fresh 1mM DTT) and supplemented with 850 µl 1% Triton X IP dilution buffer (1% Triton X, 20mM Tris-HCl pH8.1, 150mM NaCl, 2mM EDTA, 1x protease inhibitor cocktail (Roche), 1x PhosSTOP phosphatase inhibitors (Roche), 1mM DTT, 5 mM sodium butyrate and 1 mM PMSF) and incubated on ice for 10 min. Lysed cells were sonicated on ice for 8x 30 sec on / 30 sec off burst cycles with a probe sonicator (SoniPrep150) in a chilled ice-water bath (12 Amplitude) to yield chromatin fragments ranging between 200 – 800 bp in length. Sheared chromatin was centrifuged at 16,000xg for 10 minutes at 4°C and the soluble supernatant transferred to new tubes. Each 500 µl chromatin samples were supplemented with 5 µl (5mg/ml) BSA. 10% of the input was stored and the rest was used for the immunoprecipitation.

Rabbit anti-acetyl-Histone H3 Antibody (Millipore) were pre-bound to magnetic SureBeads (BioRad) in 10% w/v BSA in PBS according to the manufacturer's instructions for approximately 1h at 4°C with rotation, following which free antibody was removed with 3 washes of cold 10% w/v BSA in PBS. Chromatin and proteinG beads were combined and incubated over night at 4°C (at a ratio of 500 ng of bead bound antibody per 1 million cell equivalents of chromatin). Samples were washes at 4°C with rotation through the following series: 2 times in 1% Triton X IP dilution buffer, 2 times with ChIP wash A (50mM HEPES ph7.9, 500mM NaCl, 1mM EDTA, 1% Triton X-

100, 0.1% Na-deoxycholate, 0.1% SDS. 1x protease inhibitor cocktail, 1x PhosSTOP phosphatase inhibitors (Roche) and fresh 1mM DTT) and 2 times with ChIP wash B (20mM Tris pH 8.0, 1mM EDTA, 250mM LiCl, 1% NP-40, 0.1% Na-deoxycholate, 1x protease inhibitor cocktail, 1x PhosSTOP phosphatase inhibitors (Roche) and fresh 1mM DTT). Finally, the samples were washed with TE (1mM EDTA, 10mM Tris pH8.0). The samples were resuspended in TE and supplemented with preheated 37°C Extraction buffer (0.1M NaHCO3 and 1% SDS), vortexed and incubated for 15 minutes at 37°C on a vibrating platform.

The pH of the extracted chromatin was adjusted by adding 6µl 2M Tris-HCl pH6.8 following which both the ChIP and input samples were incubated with 20µg RNAse A (NEB) at 65°C for 1 hour. Cross-links were reversed and the protein degraded by the addition of 20µg Proteinase K and incubation at 65°C for 6-8 h. Following removal of the magnetic SureBeads from the ChIP samples, DNA was purified using a Qiagen PCR cleanup kit following manufacturer's instructions. DNA libraries were prepared using NEBNext® Ultra™ II DNA Library Prep Kit for Illumina® and NEBNext® Multiplex Oligos for Illumina® (Index Primers Set 1) following manufacturer's instructions. Sequencing was performed using the NextSeq 500/550 High-Output v2.5 (150 cycle) Kit (#20024907) on the NextSeq 550 platform (Illumina Inc, #SY-415-1002). PhiX Control v3 (Illumina, #FC-110-3001) was spiked into the library pool at a concentration of ~1% to enable troubleshooting in the event of any issues with the run.

ChIP-seq data analysis pipeline

Basecall data produced by the NextSeq 550 is automatically uploaded to BaseSpace, a cloud-based data management and analysis service provided by Illumina. Here it is converted into FASTQ files and mapped to was mapped to the human genome (GRCh38) using bowtie2 (bwa-0.7.9) applications directly accessible through BaseSpace. BAM files were uploaded to the open access bioinformatic community platform Galaxy (the public server at usegalaxy.org) for downstream analysis[120].

Briefly, mapped regions (due to fragment processing) extended beyond the end of the chromosomes were removed using SAMtools with chromosome lengths for GRCh38. MACS2 calling broadpeak algorithm was used to identify histone H3 acetylation regions, with each ChIP BAM file paired with the input control, cut off -- mfold 5 50, band width --bw 500, FDR --qvalue 0.05. Call peak results (gapped peaks) were used for differential binding (histone acetylation in

this case) analysis between A375 WT and ALDH1A3 knockout cells using the DiffBind[119]. R package ChIPseeker[136] was used for adjacent gene annotation and genomic distribution analysis. To visualise the histone H3 acetylation sites on the genome browser UCSC[137] mapped reads were converted to bigWig files (.bw) using the bamCoverage (deepTools2) algorithm[118], with replicants for each condition merged.

ACSS2 ChIP qPCR

ACSS2 bound chromatin were pulled down using rabbit anti-ACSS2 antibody (Cell Signaling Technology) with the same ChIP protocol described above in A375 acetyl-Histone H3 ChIP-seq, ChIP-ACSS2 DNA and input control were purified using the QIAquick PCR Purification Kit (Qiagen) and used for qPCR with the same LightCycler 480 SYBR Green (Roche) system as described above in the qPCR method for RT-qPCR. The ChIP qPCR primer sequences are listed in **Table S6**.

AlphaFold Multimer modelling of ALD6-ACS2 and ALDH1A3-ACSS2 interaction

Canonical sequences of yeast ALD6 (UniProt accession: P54115) and ACS2 (P52910) and human ALDH1A3 (P47895) and ACSS2 (Q9NR19) were retrieved from UniProt[121]. Disordered N-termini of the human proteins, the first 24 residues of each, were removed to facilitate complex prediction. AlphaFold-Multimer[74] predictions were performed with LocalColabFold (https://github.com/YoshitakaMo/localcolabfold), running ColabFold version 1.5.2[122] on a single 350GB NVIDIA A100 GPU. Both yeast and human sequences were run as dimers of ALD6/ALDH1A3 and monomers of ACS2/ACSS2 because of sequence length limitations. We used templates available in the Protein Data Bank[138] and the “mmseqs2_uniref” option for the –msa-mode flag. For both yeast and human complexes, we generated 3 models with 5 recycles and excluded models that were incompatible with the tetrameric structure of ALDH1A3. Confidence is 45% for yeast and 47% for the human complexes, calculated as $0.2 \times pTM + 0.8 \times ipTM$[74]. The highest ranking models of each complex were further refined with GalaxyRefineComplex[123], using default settings of protocol 2. Solvent accessible surface was calculated at residue level with FreeSASA 2.0.3[124] with the buried surface area defined as the difference in solvent accessible surface area between the monomer and the complex. Homologous residues were determined via sequence alignment with MUSCLE[139]. Conservation of ACS2/ACSS2 interface with ALD6/ALDH1A3 was determined by calculating the Pearson correlation between the buried surface area values of homologous residues. ATP/CoA binding pocket of ACSS2 was visualised by structural alignment of AlphaFill-optimized[125] protein-ligand

complexes (ATP donor: PDB ID 5k8f; CoA donor: PDB ID 3gpc) to the AlphaFold-Multimer predicted model. Visualisation of protein structures was performed with UCSF ChimeraX version 1.6[126]. The PDB models of ACSS2-ALDH1A3 and Acs2-Ald6 will be released on ModelArchive upon the publication of this manuscript and is currently available for reviewers via Mendeley Data (Reserved DOI:10.17632/4hfw7rpspb.1).

Zebrafish melanoma ALDH activity measurement

Zebrafish melanoma live cells were dissociated from freshly dissected tumour samples as described in Travnickova et al., 2019[10]. Cell suspensions were sized to 1 million cell count per ml and incubated with AldeRed (ThermoFisher Technologies) following the instruction by the manufacturer. After incubation of 1 hour at 28°C, the AldeRed activity was measured using flow cytometry (Fortessa, BD Biosciences). For cell sorting to establish the zebrafish $ALDH^{High}$ and $ALDH^{Low}$ cells, the stained cells were sorted by FACS Aria II (BD Biosciences) and the population with the highest and lowest 5% ALDH activity ($ALDH^{High}$ and $ALDH^{Low}$) were collected, as is described for AldeFluor guided ALDH subpopulation selection.

Zebrafish drug pellet treatment

To perform drug treatment on adult zebrafish bearing melanoma, we produced fish bite-size drug pellets as described in our previous publication[78] and fed single-housed individual fish with fish food agar pellets containing DMSO (Sigma Aldrich), vemurafenib (SelleckChem), and/or Nifuroxazide (Merck Millipore) once per day. Zebrafish under the drug treatment procedure were fed daily in the AM and early PM with artemia, and then fed the drug pellets in the late PM (6-8 PM). Zebrafish actively sought for and consumed the drug pellets voluntarily without any handling. Zebrafish under continuous drug treatment were imaged one day pre-treatment and once every week to track tumour size change.

Imaging of adult zebrafish tumour and size measurement

Zebrafish were briefly anesthetised (Tricaine in PBS 1:10,000 concentration) for no longer than 10 min per session and fully recovered in fresh system water. Brightfield images were taken for each fish positioned on both sides. Images of fish lesions were captured at the same magnification scale every week using a Nikon COOLPIX5400 camera attached to a brightfield microscope (Nikon SMZ1500). The size of each lesion was quantified by using the manual field selection in Fiji on each tumour image, then compared to the matching pre-treatment lesion to calculate the

relative percentage change. Lesions that could be observed from both sides of the fish were measured by combining the area number averaged from both sides.

Zebrafish histology and IHC quantification

Zebrafish melanoma samples were collected, fixed, and processed as described in our earlier publications[10,17,140]. The slides of Haematoxylin and Eosin staining were imaged using a Hamamatsu NanoZoomer SlideScanner, and the images were processed using NDP.3 software. Aldh1a3 expression was assessed using Rabbit polyclonal anti-ALDH1A3 primary antibody (1:200, Abcam); Cd79a expression was probed using Rabbit monoclonal anti-CD79a (1:200, Abcam). Following secondary fluorescent antibody incubation (Donkey anti-mouse conjugated Alexa Fluor 568, Donkey anti-Rabbit conjugated Alexa Fluor 488 or 647, Invitrogen), nuclei were stained with DAPI dye (1:1000, Life Technologies). Stained tissue slides were mounted with antifade mounting medium (Vectashield, 2BScientific) before fluorescent microscope imaging using multimodal Imaging Platform Dragonfly (Andor technologies, Belfast UK). Similar to ICC imaging, images were acquired using a 20X lens equipped with 405, 488, 561, and 640 nm lasers built on a Nikon Eclipse Ti-E inverted microscope body with Perfect focus system (Nikon Instruments, Japan). Data were collected in Spinning Disk 25 µm pinhole mode on the Zyla 4.2 sCMOS camera using a Bin of 1 and no frame averaging using Andor Fusion acquisition software. Standard deviation intensity (STD) projection of a confocal z-stack was performed in Fiji to allow intensity quantification and cell subpopulation assessment.

Quantification and Statistical Analysis

All statistical methods used in the paper are described in the figure legends and, where indicated, additional details are provided in the method details. Definitions of sample size, measures of centre and dispersion, and precision measures are also indicated in figure legends. Statistics were computed using R and GraphPad Prism. When appropriate, corrections for multiple comparisons were implemented as indicated in the figure legends.

CRedit Author contributions

Conceptualization: YL, AC, EEP

Data curation: YL, ZK, JAM, AvK

Formal Analysis: YL, MB, FR, PL

Funding acquisition: FR, AvK, JAM, OS, RI, EEP

Investigation: YL, MB, FR, AC, ZK

Methodology: YL, JT, MB, ZK, JAM, LM, AvK

Project administration: YL, EEP

Resources: YL, JT, PGM, RM, VSP, OS, EEP

Supervision: AvK, JAM, VSP, RI, EEP

Validation: YL, RI

Visualization: YL, MB, FR, EEP

Writing – original draft: YL, MB, RI, EEP

Writing – review & editing: YL, JT, AvK, RI, EEP

Declaration of interests

Richard Marais is a Founder, Director and the CSO of Oncodrug Ltd, which has a drug discovery programme targeting ALDH1A3.

Inclusion and diversity

One or more of the authors of this paper self-identifies as an underrepresented ethnic minority in their field of research or within their geographical location. One or more of the authors of this paper self-identifies as a gender minority in their field of research.

Acknowledgements

We thank James Chen (Stanford University), Kevin Myant (CRUK Scotland, Edinburgh) and Wendy Bickmore (MRC Human Genetics Unit) for helpful discussions. We thank Sergio Lilla (Cancer Research UK Beatson Institute, CRUK Scotland Centre) for MS proteomic analysis (funded by Cancer Research UK Beatson Institute Advanced Technology Facilities grant no. A17196) related with **Figure 6C**; Richard Clark and the Wellcome Trust Clinical Research Facility (WTCRF) for the next generation sequencing service; Matthew Pearson (MRC Human Genetics Unit) for assistance with super resolution imaging; Graeme Grimes (MRC Human Genetics Unit)

and the IGC bioinformatic core for the RNA-seq analysis pipeline related with **Table S1**. We thank Faith Robison (University of Washington in St. Louis) for helping us to establish histone mass spectrometry related with **Figure S3**; Jun Han (Metabolomics Innovation Centre, University of Victoria, Canada) for $^{13}C_6$-glucose traced central carbon metabolites mass spectrometry assays related with **Table S3**; David Adams (Sanger Institute) for the human melanoma cell lines A375 and C089 engineered with stable expression of Cas9. We thank Craig Nicol and Uta Mackensen for the graphical assistance; Christina Lilliehook for editing assistance; the MRC HGU zebrafish facility; the MRC HGU Flow Cytometry facility and Technical Service staff; the MRC HGU imaging facility and Edinburgh Super-Resolution Imaging Consortium.

R.S.I. is supported by an MRC Career Development award (MR/S007644/1) and the Simons Initiative for the Developing Brain (SFARI - 529085); AVK is funded by Wellcome Trust (Multiuser Equipment 208402/Z/17); ZK is funded by Melville Trust for Cancer Research (Studentship M00109.0001/TZH/MHR); FR is funded by Melanoma Research Alliance and the Wolfgang & Gertrud Boettcher Foundation; JAM is funded by the European Research Council (101001169). The generation of patient derived cell lines was supported by funds from the Wellcome Trust (100282/Z/12/Z) to RM. EEP is funded by the Medical Research Council (MC_UU_00035/13), the European Research Council (ZF-MEL-CHEMBIO-648489), Melanoma Research Alliance and Rosetrees Trust (MRA Awards 687306, 917226). This work was supported by the Cancer Ressarch UK Scotland Centre (CTRQQR-2021\100006).

References

1. Marine, J.C., Dawson, S.J., and Dawson, M.A. (2020). Non-genetic mechanisms of therapeutic resistance in cancer. Nat Rev Cancer *20*, 743-756. 10.1038/s41568-020-00302-4.

2. Rambow, F., Marine, J.C., and Goding, C.R. (2019). Melanoma plasticity and phenotypic diversity: therapeutic barriers and opportunities. Genes Dev *33*, 1295-1318. 10.1101/gad.329771.119.

3. Baggiolini, A., Callahan, S.J., Montal, E., Weiss, J.M., Trieu, T., Tagore, M.M., Tischfield, S.E., Walsh, R.M., Suresh, S., Fan, Y., et al. (2021). Developmental chromatin programs determine oncogenic competence in melanoma. Science *373*, eabc1048. 10.1126/science.abc1048.

4. Boshuizen, J., Vredevoogd, D.W., Krijgsman, O., Ligtenberg, M.A., Blankenstein, S., de Bruijn, B., Frederick, D.T., Kenski, J.C.N., Parren, M., Bruggemann, M., et al. (2020).

Reversal of pre-existing NGFR-driven tumor and immune therapy resistance. Nat Commun *11*, 3946. 10.1038/s41467-020-17739-8.

5. Kaufman, C.K., Mosimann, C., Fan, Z.P., Yang, S., Thomas, A.J., Ablain, J., Tan, J.L., Fogley, R.D., van Rooijen, E., Hagedorn, E.J., et al. (2016). A zebrafish melanoma model reveals emergence of neural crest identity during melanoma initiation. Science *351*, aad2197. 10.1126/science.aad2197.

6. Landsberg, J., Kohlmeyer, J., Renn, M., Bald, T., Rogava, M., Cron, M., Fatho, M., Lennerz, V., Wolfel, T., Holzel, M., and Tuting, T. (2012). Melanomas resist T-cell therapy through inflammation-induced reversible dedifferentiation. Nature *490*, 412-416. 10.1038/nature11538.

7. Lee, J.H., Shklovskaya, E., Lim, S.Y., Carlino, M.S., Menzies, A.M., Stewart, A., Pedersen, B., Irvine, M., Alavi, S., Yang, J.Y.H., et al. (2020). Transcriptional downregulation of MHC class I and melanoma de- differentiation in resistance to PD-1 inhibition. Nat Commun *11*, 1897. 10.1038/s41467-020-15726-7.

8. Marin-Bejar, O., Rogiers, A., Dewaele, M., Femel, J., Karras, P., Pozniak, J., Bervoets, G., Van Raemdonck, N., Pedri, D., Swings, T., et al. (2021). Evolutionary predictability of genetic versus nongenetic resistance to anticancer drugs in melanoma. Cancer Cell *39*, 1135-1149 e1138. 10.1016/j.ccell.2021.05.015.

9. Shaffer, S.M., Dunagin, M.C., Torborg, S.R., Torre, E.A., Emert, B., Krepler, C., Beqiri, M., Sproesser, K., Brafford, P.A., Xiao, M., et al. (2017). Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance. Nature *546*, 431-435. 10.1038/nature22794.

10. Travnickova, J., Wojciechowska, S., Khamseh, A., Gautier, P., Brown, D.V., Lefevre, T., Brombin, A., Ewing, A., Capper, A., Spitzer, M., et al. (2019). Zebrafish MITF-Low Melanoma Subtype Models Reveal Transcriptional Subclusters and MITF-Independent Residual Disease. Cancer Res *79*, 5769-5784. 10.1158/0008-5472.CAN-19-0037.

11. Curti, B.D., and Faries, M.B. (2021). Recent Advances in the Treatment of Melanoma. N Engl J Med *384*, 2229-2240. 10.1056/NEJMra2034861.

12. Atkins, M.B., Lee, S.J., Chmielowski, B., Tarhini, A.A., Cohen, G.I., Truong, T.G., Moon, H.H., Davar, D., O'Rourke, M., Stephenson, J.J., et al. (2023). Combination Dabrafenib and Trametinib Versus Combination Nivolumab and Ipilimumab for Patients With Advanced BRAF-Mutant Melanoma: The DREAMseq Trial-ECOG-ACRIN EA6134. J Clin Oncol *41*, 186-197. 10.1200/JCO.22.01763.

13. Larkin, J., Chiarion-Sileni, V., Gonzalez, R., Grob, J.J., Rutkowski, P., Lao, C.D., Cowey, C.L., Schadendorf, D., Wagstaff, J., Dummer, R., et al. (2019). Five-Year Survival with Combined Nivolumab and Ipilimumab in Advanced Melanoma. N Engl J Med *381*, 1535-1546. 10.1056/NEJMoa1910836.

14. Robert, C., Grob, J.J., Stroyakovskiy, D., Karaszewska, B., Hauschild, A., Levchenko, E., Chiarion Sileni, V., Schachter, J., Garbe, C., Bondarenko, I., et al. (2019). Five-Year Outcomes with Dabrafenib plus Trametinib in Metastatic Melanoma. N Engl J Med *381*, 626-636. 10.1056/NEJMoa1904059.

15. Robert, C., Ribas, A., Schachter, J., Arance, A., Grob, J.J., Mortier, L., Daud, A., Carlino, M.S., McNeil, C.M., Lotem, M., et al. (2019). Pembrolizumab versus ipilimumab in advanced melanoma (KEYNOTE-006): post-hoc 5-year results from an open-label, multicentre, randomised, controlled, phase 3 study. Lancet Oncol *20*, 1239-1251. 10.1016/S1470-2045(19)30388-2.

16. Rambow, F., Rogiers, A., Marin-Bejar, O., Aibar, S., Femel, J., Dewaele, M., Karras, P., Brown, D., Chang, Y.H., Debiec-Rychter, M., et al. (2018). Toward Minimal Residual Disease-Directed Therapy in Melanoma. Cell *174*, 843-855 e819. 10.1016/j.cell.2018.06.025.

17. Travnickova, J., Muise, S., Wojciechowska, S., Brombin, A., Zeng, Z., Young, A.I.J., Wyatt, C., and Patton, E.E. (2022). Fate mapping melanoma persister cells through regression and into recurrent disease in adult zebrafish. Dis Model Mech *15*. 10.1242/dmm.049566.

18. Diener, J., and Sommer, L. (2021). Reemergence of neural crest stem cell-like states in melanoma during disease progression and treatment. Stem Cells Transl Med *10*, 522-533. 10.1002/sctm.20-0351.

19. Chauvistre, H., Shannan, B., Daignault-Mill, S.M., Ju, R.J., Picard, D., Egetemaier, S., Varaljai, R., Gibhardt, C.S., Sechi, A., Kaschani, F., et al. (2022). Persister state-directed transitioning and vulnerability in melanoma. Nat Commun *13*, 3055. 10.1038/s41467-022-30641-9.

20. Emert, B.L., Cote, C.J., Torre, E.A., Dardani, I.P., Jiang, C.L., Jain, N., Shaffer, S.M., and Raj, A. (2021). Variability within rare cell states enables multiple paths toward drug resistance. Nat Biotechnol *39*, 865-876. 10.1038/s41587-021-00837-3.

21. Kong, X., Kuilman, T., Shahrabi, A., Boshuizen, J., Kemper, K., Song, J.Y., Niessen, H.W.M., Rozeman, E.A., Geukes Foppen, M.H., Blank, C.U., and Peeper, D.S. (2017). Cancer drug addiction is relayed by an ERK2-dependent phenotype switch. Nature *550*, 270-274. 10.1038/nature24037.

22. Strub, T., Ghiraldini, F.G., Carcamo, S., Li, M., Wroblewska, A., Singh, R., Goldberg, M.S., Hasson, D., Wang, Z., Gallagher, S.J., et al. (2018). SIRT6 haploinsufficiency induces BRAF(V600E) melanoma cell resistance to MAPK inhibitors via IGF signalling. Nat Commun *9*, 3440. 10.1038/s41467-018-05966-z.

23. Du, K., Wei, S., Wei, Z., Frederick, D.T., Miao, B., Moll, T., Tian, T., Sugarman, E., Gabrilovich, D.I., Sullivan, R.J., et al. (2021). Pathway signatures derived from on-treatment tumor specimens predict response to anti-PD1 blockade in metastatic melanoma. Nat Commun *12*, 6023. 10.1038/s41467-021-26299-4.

24. Hugo, W., Zaretsky, J.M., Sun, L., Song, C., Moreno, B.H., Hu-Lieskovan, S., Berent-Maoz, B., Pang, J., Chmielowski, B., Cherry, G., et al. (2016). Genomic and Transcriptomic Features of Response to Anti-PD-1 Therapy in Metastatic Melanoma. Cell *165*, 35-44. 10.1016/j.cell.2016.02.065.

25. Fischer, G.M., Vashisht Gopal, Y.N., McQuade, J.L., Peng, W., DeBerardinis, R.J., and Davies, M.A. (2018). Metabolic strategies of melanoma cells: Mechanisms, interactions

with the tumor microenvironment, and therapeutic implications. Pigment Cell Melanoma Res *31*, 11-30. 10.1111/pcmr.12661.

26. Fischer, G.M., Jalali, A., Kircher, D.A., Lee, W.C., McQuade, J.L., Haydu, L.E., Joon, A.Y., Reuben, A., de Macedo, M.P., Carapeto, F.C.L., et al. (2019). Molecular Profiling Reveals Unique Immune and Metabolic Features of Melanoma Brain Metastases. Cancer Discov *9*, 628-645. 10.1158/2159-8290.CD-18-1489.

27. Liu, S., Dharanipragada, P., Lomeli, S.H., Wang, Y., Zhang, X., Yang, Z., Lim, R.J., Dumitras, C., Scumpia, P.O., Dubinett, S.M., et al. (2023). Multi-organ landscape of therapy-resistant melanoma. Nat Med *29*, 1123-1134. 10.1038/s41591-023-02304-9.

28. Piskounova, E., Agathocleous, M., Murphy, M.M., Hu, Z., Huddlestun, S.E., Zhao, Z., Leitch, A.M., Johnson, T.M., DeBerardinis, R.J., and Morrison, S.J. (2015). Oxidative stress inhibits distant metastasis by human melanoma cells. Nature *527*, 186-191. 10.1038/nature15726.

29. Shen, S., Faouzi, S., Souquere, S., Roy, S., Routier, E., Libenciuc, C., Andre, F., Pierron, G., Scoazec, J.Y., and Robert, C. (2020). Melanoma Persister Cells Are Tolerant to BRAF/MEK Inhibitors via ACOX1-Mediated Fatty Acid Oxidation. Cell Rep *33*, 108421. 10.1016/j.celrep.2020.108421.

30. Tasdogan, A., Faubert, B., Ramesh, V., Ubellacker, J.M., Shen, B., Solmonson, A., Murphy, M.M., Gu, Z., Gu, W., Martin, M., et al. (2020). Metabolic heterogeneity confers differences in melanoma metastatic potential. Nature *577*, 115-120. 10.1038/s41586-019-1847-2.

31. Ubellacker, J.M., Tasdogan, A., Ramesh, V., Shen, B., Mitchell, E.C., Martin-Sandoval, M.S., Gu, Z., McCormick, M.L., Durham, A.B., Spitz, D.R., et al. (2020). Lymph protects metastasizing melanoma cells from ferroptosis. Nature *585*, 113-118. 10.1038/s41586-020-2623-z.

32. Pathria, G., Scott, D.A., Feng, Y., Sang Lee, J., Fujita, Y., Zhang, G., Sahu, A.D., Ruppin, E., Herlyn, M., Osterman, A.L., and Ronai, Z.A. (2018). Targeting the Warburg effect via LDHA inhibition engages ATF4 signaling for cancer cell survival. EMBO J *37*. 10.15252/embj.201899735.

33. Arner, E.N., and Rathmell, J.C. (2023). Metabolic programming and immune suppression in the tumor microenvironment. Cancer Cell *41*, 421-433. 10.1016/j.ccell.2023.01.009.

34. Bhattacharya, D., Khan, B., and Simoes-Costa, M. (2021). Neural crest metabolism: At the crossroads of development and disease. Dev Biol *475*, 245-255. 10.1016/j.ydbio.2021.01.018.

35. Santoriello, C., Sporrij, A., Yang, S., Flynn, R.A., Henriques, T., Dorjsuren, B., Custo Greig, E., McCall, W., Stanhope, M.E., Fazio, M., et al. (2020). RNA helicase DDX21 mediates nucleotide stress responses in neural crest and melanoma cells. Nat Cell Biol *22*, 372-379. 10.1038/s41556-020-0493-0.

36. Sporrij, A., and Zon, L.I. (2021). Nucleotide stress responses in neural crest cell fate and melanoma. Cell Cycle *20*, 1455-1467. 10.1080/15384101.2021.1947567.

37. White, R.M., Cech, J., Ratanasirintrawoot, S., Lin, C.Y., Rahl, P.B., Burke, C.J., Langdon, E., Tomlinson, M.L., Mosher, J., Kaufman, C., et al. (2011). DHODH modulates transcriptional elongation in the neural crest and melanoma. Nature *471*, 518-522. 10.1038/nature09882.

38. Johansson, J.A., Marie, K.L., Lu, Y., Brombin, A., Santoriello, C., Zeng, Z., Zich, J., Gautier, P., von Kriegsheim, A., Brunsdon, H., et al. (2020). PRL3-DDX21 Transcriptional Control of Endolysosomal Genes Restricts Melanocyte Stem Cell Differentiation. Dev Cell *54*, 317-332 e319. 10.1016/j.devcel.2020.06.013.

39. Sarvi, S., Crispin, R., Lu, Y., Zeng, L., Hurley, T.D., Houston, D.R., von Kriegsheim, A., Chen, C.H., Mochly-Rosen, D., Ranzani, M., et al. (2018). ALDH1 Bio-activates Nifuroxazide to Eradicate ALDH(High) Melanoma-Initiating Cells. Cell Chem Biol *25*, 1456-1469 e1456. 10.1016/j.chembiol.2018.09.005.

40. Ginestier, C., Hur, M.H., Charafe-Jauffret, E., Monville, F., Dutcher, J., Brown, M., Jacquemier, J., Viens, P., Kleer, C.G., Liu, S., et al. (2007). ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. Cell Stem Cell *1*, 555-567. 10.1016/j.stem.2007.08.014.

41. Luo, Y., Dallaglio, K., Chen, Y., Robinson, W.A., Robinson, S.E., McCarter, M.D., Wang, J., Gonzalez, R., Thompson, D.C., Norris, D.A., et al. (2012). ALDH1A isozymes are markers of human melanoma stem cells and potential therapeutic targets. Stem Cells *30*, 2100-2113. 10.1002/stem.1193.

42. Gulati, G.S., Sikandar, S.S., Wesche, D.J., Manjunath, A., Bharadwaj, A., Berger, M.J., Ilagan, F., Kuo, A.H., Hsieh, R.W., Cai, S., et al. (2020). Single-cell transcriptional diversity is a hallmark of developmental potential. Science *367*, 405-411. 10.1126/science.aax0249.

43. Karras, P., Bordeu, I., Pozniak, J., Nowosad, A., Pazzi, C., Van Raemdonck, N., Landeloos, E., Van Herck, Y., Pedri, D., Bervoets, G., et al. (2022). A cellular hierarchy in melanoma uncouples growth and metastasis. Nature *610*, 190-198. 10.1038/s41586-022-05242-7.

44. Liu, D., Lin, J.R., Robitschek, E.J., Kasumova, G.G., Heyde, A., Shi, A., Kraya, A., Zhang, G., Moll, T., Frederick, D.T., et al. (2021). Evolution of delayed resistance to immunotherapy in a melanoma responder. Nat Med *27*, 985-992. 10.1038/s41591-021-01331-8.

45. Bhattacharya, D., Azambuja, A.P., and Simoes-Costa, M. (2020). Metabolic Reprogramming Promotes Neural Crest Migration via Yap/Tead Signaling. Dev Cell *53*, 199-211 e196. 10.1016/j.devcel.2020.03.005.

46. Brombin, A., Simpson, D.J., Travnickova, J., Brunsdon, H., Zeng, Z., Lu, Y., Young, A.I.J., Chandra, T., and Patton, E.E. (2022). Tfap2b specifies an embryonic melanocyte stem cell that retains adult multifate potential. Cell Rep *38*, 110234. 10.1016/j.celrep.2021.110234.

47. Cancer Genome Atlas, N. (2015). Genomic Classification of Cutaneous Melanoma. Cell *161*, 1681-1696. 10.1016/j.cell.2015.05.044.

48. Cirenajwis, H., Ekedahl, H., Lauss, M., Harbst, K., Carneiro, A., Enoksson, J., Rosengren, F., Werner-Hartman, L., Torngren, T., Kvist, A., et al. (2015). Molecular stratification of metastatic melanoma using gene expression profiling: Prediction of survival outcome and benefit from molecular targeted therapy. Oncotarget *6*, 12297-12309. 10.18632/oncotarget.3655.

49. Jonsson, G., Busch, C., Knappskog, S., Geisler, J., Miletic, H., Ringner, M., Lillehaug, J.R., Borg, A., and Lonning, P.E. (2010). Gene expression profiling-based identification of molecular subtypes in stage IV melanomas with different clinical outcome. Clin Cancer Res *16*, 3356-3367. 10.1158/1078-0432.CCR-09-2509.

50. Tirosh, I., Izar, B., Prakadan, S.M., Wadsworth, M.H., 2nd, Treacy, D., Trombetta, J.J., Rotem, A., Rodman, C., Lian, C., Murphy, G., et al. (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science *352*, 189-196. 10.1126/science.aad0501.

51. Barretina, J., Caponigro, G., Stransky, N., Venkatesan, K., Margolin, A.A., Kim, S., Wilson, C.J., Lehar, J., Kryukov, G.V., Sonkin, D., et al. (2012). The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature *483*, 603-607. 10.1038/nature11003.

52. Ranzani, M., Kemper, K., Michaut, M., Krijgsman, O., Aben, N., Iyer, V., Wong, K., Roumeliotis, T.I., Velasco-Herrera, M.D.C., Nsengimana, J., et al. (2017). A screen for combination therapies in BRAF/NRAS wild type melanoma identifies nilotinib plus MEK inhibitor as a synergistic combination. bioRxiv, 195354. 10.1101/195354.

53. Vassalli, G. (2019). Aldehyde Dehydrogenases: Not Just Markers, but Functional Regulators of Stem Cells. Stem Cells Int *2019*, 3904645. 10.1155/2019/3904645.

54. Rabinowitz, J.D., and Enerback, S. (2020). Lactate: the ugly duckling of energy metabolism. Nat Metab *2*, 566-571. 10.1038/s42255-020-0243-4.

55. Torrini, C., Nguyen, T.T.T., Shu, C., Mela, A., Humala, N., Mahajan, A., Seeley, E.H., Zhang, G., Westhoff, M.A., Karpel-Massler, G., et al. (2022). Lactate is an epigenetic metabolite that drives survival in model systems of glioblastoma. Mol Cell *82*, 3061-3076 e3066. 10.1016/j.molcel.2022.06.030.

56. Guertin, D.A., and Wellen, K.E. (2023). Acetyl-CoA metabolism in cancer. Nat Rev Cancer *23*, 156-172. 10.1038/s41568-022-00543-5.

57. Schug, Z.T., Vande Voorde, J., and Gottlieb, E. (2016). The metabolic fate of acetate in cancer. Nat Rev Cancer *16*, 708-717. 10.1038/nrc.2016.87.

58. Etchegaray, J.P., and Mostoslavsky, R. (2016). Interplay between Metabolism and Epigenetics: A Nuclear Adaptation to Environmental Changes. Mol Cell *62*, 695-711. 10.1016/j.molcel.2016.05.029.

59. Hsieh, W.C., Sutter, B.M., Ruess, H., Barnes, S.D., Malladi, V.S., and Tu, B.P. (2022). Glucose starvation induces a switch in the histone acetylome for activation of gluconeogenic and fat metabolism genes. Mol Cell *82*, 60-74 e65. 10.1016/j.molcel.2021.12.015.

60. Ye, C., and Tu, B.P. (2018). Sink into the Epigenome: Histones as Repositories That Influence Cellular Metabolism. Trends Endocrinol Metab *29*, 626-637. 10.1016/j.tem.2018.06.002.

61. McBrian, M.A., Behbahan, I.S., Ferrari, R., Su, T., Huang, T.W., Li, K., Hong, C.S., Christofk, H.R., Vogelauer, M., Seligson, D.B., and Kurdistani, S.K. (2013). Histone acetylation regulates intracellular pH. Mol Cell *49*, 310-321. 10.1016/j.molcel.2012.10.025.

62. Kurdistani, S.K. (2014). Chromatin: a capacitor of acetate for integrated regulation of gene expression and cell physiology. Curr Opin Genet Dev *26*, 53-58. 10.1016/j.gde.2014.06.002.

63. Gehrke, A.R., Neverett, E., Luo, Y.J., Brandt, A., Ricci, L., Hulett, R.E., Gompers, A., Ruby, J.G., Rokhsar, D.S., Reddien, P.W., and Srivastava, M. (2019). Acoel genome reveals the regulatory landscape of whole-body regeneration. Science *363*. 10.1126/science.aau6173.

64. Schmidt, K., Carroll, J.S., Yee, E., Thomas, D.D., Wert-Lamas, L., Neier, S.C., Sheynkman, G., Ritz, J., and Novina, C.D. (2019). The lncRNA SLNCR Recruits the Androgen Receptor to EGR1-Bound Genes in Melanoma and Inhibits Expression of Tumor Suppressor p21. Cell Rep *27*, 2493-2507 e2494. 10.1016/j.celrep.2019.04.101.

65. Sun, Z., Xu, X., He, J., Murray, A., Sun, M.A., Wei, X., Wang, X., McCoig, E., Xie, E., Jiang, X., et al. (2019). EGR1 recruits TET1 to shape the brain methylome during development and upon neuronal activity. Nat Commun *10*, 3892. 10.1038/s41467-019-11905-3.

66. Baek, J., Lopez, P.A., Lee, S., Kim, T.S., Kumar, S., and Schaffer, D.V. (2022). Egr1 is a 3D matrix-specific mediator of mechanosensitive stem cell lineage commitment. Sci Adv *8*, eabm4646. 10.1126/sciadv.abm4646.

67. Ling, I.T.C., and Sauka-Spengler, T. (2019). Early chromatin shaping predetermines multipotent vagal neural crest into neural, neuronal and mesenchymal lineages. Nat Cell Biol *21*, 1504-1517. 10.1038/s41556-019-0428-9.

68. Aibar, S., Gonzalez-Blas, C.B., Moerman, T., Huynh-Thu, V.A., Imrichova, H., Hulselmans, G., Rambow, F., Marine, J.C., Geurts, P., Aerts, J., et al. (2017). SCENIC: single-cell regulatory network inference and clustering. Nat Methods *14*, 1083-1086. 10.1038/nmeth.4463.

69. Cai, L., Sutter, B.M., Li, B., and Tu, B.P. (2011). Acetyl-CoA induces cell growth and proliferation by promoting the acetylation of histones at growth genes. Mol Cell *42*, 426-437. 10.1016/j.molcel.2011.05.004.

70. Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N.T., Morris, J.H., Bork, P., et al. (2019). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res *47*, D607-D613. 10.1093/nar/gky1131.

71. Gavin, A.C., Aloy, P., Grandi, P., Krause, R., Boesche, M., Marzioch, M., Rau, C., Jensen, L.J., Bastuck, S., Dumpelfeld, B., et al. (2006). Proteome survey reveals modularity of the yeast cell machinery. Nature *440*, 631-636. 10.1038/nature04532.

72. Alexander, D.C., Corman, T., Mendoza, M., Glass, A., Belity, T., Wu, R., Campbell, R.R., Han, J., Keiser, A.A., Winkler, J., et al. (2022). Targeting acetyl-CoA metabolism attenuates the formation of fear memories through reduced activity-dependent histone acetylation. Proc Natl Acad Sci U S A *119*, e2114758119. 10.1073/pnas.2114758119.

73. Mews, P., Donahue, G., Drake, A.M., Luczak, V., Abel, T., and Berger, S.L. (2017). Acetyl-CoA synthetase regulates histone acetylation and hippocampal memory. Nature *546*, 381-386. 10.1038/nature22405.

74. Evans, R., O'Neill, M., Pritzel, A., Antropova, N., Senior, A., Green, T., Žídek, A., Bates, R., Blackwell, S., Yim, J., et al. (2022). Protein complex prediction with AlphaFold-Multimer. bioRxiv, 2021.2010.2004.463034. 10.1101/2021.10.04.463034.

75. Liu, X., Cooper, D.E., Cluntun, A.A., Warmoes, M.O., Zhao, S., Reid, M.A., Liu, J., Lund, P.J., Lopes, M., Garcia, B.A., et al. (2018). Acetate Production from Glucose and Coupling to Mitochondrial Metabolism in Mammals. Cell *175*, 502-513 e513. 10.1016/j.cell.2018.08.040.

76. Patton, E.E., Widlund, H.R., Kutok, J.L., Kopani, K.R., Amatruda, J.F., Murphey, R.D., Berghmans, S., Mayhall, E.A., Traver, D., Fletcher, C.D., et al. (2005). BRAF mutations are sufficient to promote nevi formation and cooperate with p53 in the genesis of melanoma. Curr Biol *15*, 249-254. 10.1016/j.cub.2005.01.031.

77. Travnickova, J., and Patton, E.E. (2021). Deciphering Melanoma Cell States and Plasticity with Zebrafish Models. J Invest Dermatol *141*, 1389-1394. 10.1016/j.jid.2020.12.007.

78. Lu, Y., and Patton, E.E. (2022). Long-term non-invasive drug treatments in adult zebrafish that lead to melanoma drug resistance. Dis Model Mech *15*. 10.1242/dmm.049401.

79. Haas, L., Elewaut, A., Gerard, C.L., Umkehrer, C., Leiendecker, L., Pedersen, M., Krecioch, I., Hoffmann, D., Novatchkova, M., Kuttke, M., et al. (2021). Acquired resistance to anti-MAPK targeted therapy confers an immune-evasive tumor microenvironment and cross-resistance to immunotherapy in melanoma. Nat Cancer *2*, 693-708. 10.1038/s43018-021-00221-9.

80. Goyal, Y., Busch, G.T., Pillai, M., Li, J., Boe, R.H., Grody, E.I., Chelvanambi, M., Dardani, I.P., Emert, B., Bodkin, N., et al. (2023). Diverse clonal fates emerge upon drug treatment of homogeneous cancer cells. Nature. 10.1038/s41586-023-06342-8.

81. Bai, X., Fisher, D.E., and Flaherty, K.T. (2019). Cell-state dynamics and therapeutic resistance in melanoma from the perspective of MITF and IFNgamma pathways. Nat Rev Clin Oncol *16*, 549-562. 10.1038/s41571-019-0204-6.

82. Nekooie-Marnany, N., Fodil, R., Féréol, S., Depp, M., Motterlini, R., Foresti, R., Duband, J.-L., and Dufour, S. (2022). Glucose oxidation and nutrients availability drive neural crest development. bioRxiv, 2022.2009.2005.506657. 10.1101/2022.09.05.506657.

83. Varum, S., Baggiolini, A., Zurkirchen, L., Atak, Z.K., Cantu, C., Marzorati, E., Bossart, R., Wouters, J., Hausel, J., Tuncer, E., et al. (2019). Yin Yang 1 Orchestrates a Metabolic Program Required for Both Neural Crest Development and Melanoma Formation. Cell Stem Cell *24*, 637-653 e639. 10.1016/j.stem.2019.03.011.

84. Li, J., Song, J.S., Bell, R.J., Tran, T.N., Haq, R., Liu, H., Love, K.T., Langer, R., Anderson, D.G., Larue, L., and Fisher, D.E. (2012). YY1 regulates melanocyte development and function by cooperating with MITF. PLoS Genet *8*, e1002688. 10.1371/journal.pgen.1002688.

85. Hodskinson, M.R., Bolner, A., Sato, K., Kamimae-Lanning, A.N., Rooijers, K., Witte, M., Mahesh, M., Silhan, J., Petek, M., Williams, D.M., et al. (2020). Alcohol-derived DNA crosslinks are repaired by two distinct mechanisms. Nature *579*, 603-608. 10.1038/s41586-020-2059-5.

86. Wang, M., Dingler, F.A., and Patel, K.J. (2022). Genotoxic aldehydes in the hematopoietic system. Blood *139*, 2119-2129. 10.1182/blood.2019004316.

87. Garaycoechea, J.I., Crossan, G.P., Langevin, F., Mulderrig, L., Louzada, S., Yang, F., Guilbaud, G., Park, N., Roerink, S., Nik-Zainal, S., et al. (2018). Alcohol and endogenous aldehydes damage chromosomes and mutate stem cells. Nature *553*, 171-177. 10.1038/nature25154.

88. Garaycoechea, J.I., Crossan, G.P., Langevin, F., Daly, M., Arends, M.J., and Patel, K.J. (2012). Genotoxic consequences of endogenous aldehydes on mouse haematopoietic stem cell function. Nature *489*, 571-575. 10.1038/nature11368.

89. Langevin, F., Crossan, G.P., Rosado, I.V., Arends, M.J., and Patel, K.J. (2011). Fancd2 counteracts the toxic effects of naturally produced aldehydes in mice. Nature *475*, 53-58. 10.1038/nature10192.

90. Rosado, I.V., Langevin, F., Crossan, G.P., Takata, M., and Patel, K.J. (2011). Formaldehyde catabolism is essential in cells deficient for the Fanconi anemia DNA-repair pathway. Nat Struct Mol Biol *18*, 1432-1434. 10.1038/nsmb.2173.

91. Chen, C.H., Kraemer, B.R., and Mochly-Rosen, D. (2022). ALDH2 variance in disease and populations. Dis Model Mech *15*. 10.1242/dmm.049601.

92. Oberbeck, N., Langevin, F., King, G., de Wind, N., Crossan, G.P., and Patel, K.J. (2014). Maternal aldehyde elimination during pregnancy preserves the fetal genome. Mol Cell *55*, 807-817. 10.1016/j.molcel.2014.07.010.

93. Dingler, F.A., Wang, M., Mu, A., Millington, C.L., Oberbeck, N., Watcham, S., Pontel, L.B., Kamimae-Lanning, A.N., Langevin, F., Nadler, C., et al. (2020). Two Aldehyde Clearance Systems Are Essential to Prevent Lethal Formaldehyde Accumulation in Mice and Humans. Mol Cell *80*, 996-1012 e1019. 10.1016/j.molcel.2020.10.012.

94. Hira, A., Yabe, H., Yoshida, K., Okuno, Y., Shiraishi, Y., Chiba, K., Tanaka, H., Miyano, S., Nakamura, J., Kojima, S., et al. (2013). Variant ALDH2 is associated with accelerated progression of bone marrow failure in Japanese Fanconi anemia patients. Blood *122*, 3206-3209. 10.1182/blood-2013-06-507962.

95. Wang, M., Brandt, L.T.L., Wang, X., Russell, H., Mitchell, E., Kamimae-Lanning, A.N., Brown, J.M., Dingler, F.A., Garaycoechea, J.I., Isobe, T., et al. (2023). Genotoxic aldehyde stress prematurely ages hematopoietic stem cells in a p53-driven manner. Mol Cell *83*, 2417-2433 e2417. 10.1016/j.molcel.2023.05.035.

96. Pontel, L.B., Rosado, I.V., Burgos-Barragan, G., Garaycoechea, J.I., Yu, R., Arends, M.J., Chandrasekaran, G., Broecker, V., Wei, W., Liu, L., et al. (2015). Endogenous Formaldehyde Is a Hematopoietic Stem Cell Genotoxin and Metabolic Carcinogen. Mol Cell *60*, 177-188. 10.1016/j.molcel.2015.08.020.

97. Mews, P., Egervari, G., Nativio, R., Sidoli, S., Donahue, G., Lombroso, S.I., Alexander, D.C., Riesche, S.L., Heller, E.A., Nestler, E.J., et al. (2019). Alcohol metabolism contributes to brain histone acetylation. Nature *574*, 717-721. 10.1038/s41586-019-1700-7.

98. Kriss, C.L., Gregory-Lott, E., Storey, A.J., Tackett, A.J., Wahls, W.P., and Stevens, S.M., Jr. (2018). In Vivo Metabolic Tracing Demonstrates the Site-Specific Contribution of Hepatic Ethanol Metabolism to Histone Acetylation. Alcohol Clin Exp Res *42*, 1909-1923. 10.1111/acer.13843.

99. Shen, X., Wang, R., Kim, M.J., Hu, Q., Hsu, C.C., Yao, J., Klages-Mundt, N., Tian, Y., Lynn, E., Brewer, T.F., et al. (2020). A Surge of DNA Damage Links Transcriptional Reprogramming and Hematopoietic Deficit in Fanconi Anemia. Mol Cell *80*, 1013-1024 e1016. 10.1016/j.molcel.2020.11.040.

100. Wit, N., Gogola, E., West, J.A., Vornbaumen, T., Seear, R.V., Bailey, P.S.J., Burgos-Barragan, G., Wang, M., Krawczyk, P., Huberts, D., et al. (2023). A histone deacetylase 3 and mitochondrial complex I axis regulates toxic formaldehyde production. Sci Adv *9*, eadg2235. 10.1126/sciadv.adg2235.

101. Burgos-Barragan, G., Wit, N., Meiser, J., Dingler, F.A., Pietzke, M., Mulderrig, L., Pontel, L.B., Rosado, I.V., Brewer, T.F., Cordell, R.L., et al. (2017). Mammals divert endogenous genotoxic formaldehyde into one-carbon metabolism. Nature *548*, 549-554. 10.1038/nature23481.

102. Brunsdon, H., Brombin, A., Peterson, S., Postlethwait, J.H., and Patton, E.E. (2022). Aldh2 is a lineage-specific metabolic gatekeeper in melanocyte stem cells. Development *149*. 10.1242/dev.200277.

103. Schede, H.H., Natarajan, P., Chakraborty, A.K., and Shrinivas, K. (2023). A model for organization and regulation of nuclear condensates by gene activity. Nat Commun *14*, 4152. 10.1038/s41467-023-39878-4.

104. Louphrasitthiphol, P., Siddaway, R., Loffreda, A., Pogenberg, V., Friedrichsen, H., Schepsky, A., Zeng, Z., Lu, M., Strub, T., Freter, R., et al. (2020). Tuning Transcription Factor Availability through Acetylation-Mediated Genomic Redistribution. Mol Cell *79*, 472-487 e410. 10.1016/j.molcel.2020.05.025.

105. Louphrasitthiphol P, L.A., Pogenberg V, Picaud S, Schepsky A, Friedrichsen H, Zeng Z, Lashgari A Thomas B, Patton EE, Wilmanns M, Filippakopoulos P, Lambert J-P,

Steingrímsson E, Mazza D and Goding CR (in press). Acetylation reprograms MITF target selectivity and residence time. Nature Communications.

106. Reddy, J., Fonseca, M.A.S., Corona, R.I., Nameki, R., Segato Dezem, F., Klein, I.A., Chang, H., Chaves-Moreira, D., Afeyan, L.K., Malta, T.M., et al. (2021). Predicting master transcription factors from pan-cancer expression data. Sci Adv *7*, eabf6123. 10.1126/sciadv.abf6123.

107. Sanjana, N.E., Shalem, O., and Zhang, F. (2014). Improved vectors and genome-wide libraries for CRISPR screening. Nat Methods *11*, 783-784. 10.1038/nmeth.3047.

108. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21. 10.1093/bioinformatics/bts635.

109. DeLuca, D.S., Levin, J.Z., Sivachenko, A., Fennell, T., Nazaire, M.D., Williams, C., Reich, M., Winckler, W., and Getz, G. (2012). RNA-SeQC: RNA-seq metrics for quality control and process optimization. Bioinformatics *28*, 1530-1532. 10.1093/bioinformatics/bts196.

110. Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A *102*, 15545-15550. 10.1073/pnas.0506580102.

111. Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., and Vilo, J. (2019). g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). Nucleic Acids Res *47*, W191-W198. 10.1093/nar/gkz369.

112. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., 3rd, Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell *184*, 3573-3587 e3529. 10.1016/j.cell.2021.04.048.

113. Yuan, Z.F., Sidoli, S., Marchione, D.M., Simithy, J., Janssen, K.A., Szurgot, M.R., and Garcia, B.A. (2018). EpiProfile 2.0: A Computational Platform for Processing Epi-Proteomics Mass Spectrometry Data. J Proteome Res *17*, 2533-2541. 10.1021/acs.jproteome.8b00133.

114. Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell *38*, 576-589. 10.1016/j.molcel.2010.05.004.

115. Wang, Q., Li, M., Wu, T., Zhan, L., Li, L., Chen, M., Xie, W., Xie, Z., Hu, E., Xu, S., and Yu, G. (2022). Exploring Epigenomic Datasets by ChIPseeker. Curr Protoc *2*, e585. 10.1002/cpz1.585.

116. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078-2079. 10.1093/bioinformatics/btp352.

117. Feng, J., Liu, T., Qin, B., Zhang, Y., and Liu, X.S. (2012). Identifying ChIP-seq enrichment using MACS. Nat Protoc *7*, 1728-1740. 10.1038/nprot.2012.101.

118. Ramirez, F., Ryan, D.P., Gruning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dundar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res *44*, W160-165. 10.1093/nar/gkw257.

119. Ross-Innes, C.S., Stark, R., Teschendorff, A.E., Holmes, K.A., Ali, H.R., Dunning, M.J., Brown, G.D., Gojis, O., Ellis, I.O., Green, A.R., et al. (2012). Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature *481*, 389-393. 10.1038/nature10730.

120. Galaxy, C. (2022). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update. Nucleic Acids Res *50*, W345-W351. 10.1093/nar/gkac247.

121. UniProt, C. (2023). UniProt: the Universal Protein Knowledgebase in 2023. Nucleic Acids Res *51*, D523-D531. 10.1093/nar/gkac1052.

122. Mirdita, M., Schutze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., and Steinegger, M. (2022). ColabFold: making protein folding accessible to all. Nat Methods *19*, 679-682. 10.1038/s41592-022-01488-1.

123. Heo, L., Lee, H., and Seok, C. (2016). GalaxyRefineComplex: Refinement of protein-protein complex model structures driven by interface repacking. Sci Rep *6*, 32153. 10.1038/srep32153.

124. Mitternacht, S. (2016). FreeSASA: An open source C library for solvent accessible surface area calculations. F1000Res *5*, 189. 10.12688/f1000research.7931.1.

125. Hekkelman, M.L., de Vries, I., Joosten, R.P., and Perrakis, A. (2023). AlphaFill: enriching AlphaFold models with ligands and cofactors. Nat Methods *20*, 205-213. 10.1038/s41592-022-01685-y.

126. Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H., and Ferrin, T.E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. Protein Sci *30*, 70-82. 10.1002/pro.3943.

127. Otasek, D., Morris, J.H., Boucas, J., Pico, A.R., and Demchak, B. (2019). Cytoscape Automation: empowering workflow-based network analysis. Genome Biol *20*, 185. 10.1186/s13059-019-1758-4.

128. Chen, B., Khodadoust, M.S., Liu, C.L., Newman, A.M., and Alizadeh, A.A. (2018). Profiling Tumor Infiltrating Immune Cells with CIBERSORT. Methods Mol Biol *1711*, 243-259. 10.1007/978-1-4939-7493-1_12.

129. Gillespie, M., Jassal, B., Stephan, R., Milacic, M., Rothfels, K., Senff-Ribeiro, A., Griss, J., Sevilla, C., Matthews, L., Gong, C., et al. (2022). The reactome pathway knowledgebase 2022. Nucleic Acids Res *50*, D687-D692. 10.1093/nar/gkab1028.

130. Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res *13*, 2498-2504. 10.1101/gr.1239303.

131. Azuma, T., and Kei, T. (2015). Super-resolution spinning-disk confocal microscopy using optical photon reassignment. Opt Express *23*, 15003-15011. 10.1364/OE.23.015003.

132. Han, J., Gagnon, S., Eckle, T., and Borchers, C.H. (2013). Metabolomic analysis of key central carbon metabolism carboxylic acids as their 3-nitrophenylhydrazones by UPLC/ESI-MS. Electrophoresis *34*, 2891-2900. 10.1002/elps.201200601.

133. Sidoli, S., Bhanu, N.V., Karch, K.R., Wang, X., and Garcia, B.A. (2016). Complete Workflow for Analysis of Histone Post-translational Modifications Using Bottom-up Mass Spectrometry: From Histone Extraction to Data Analysis. J Vis Exp. 10.3791/54112.

134. Trefely, S., Huber, K., Liu, J., Noji, M., Stransky, S., Singh, J., Doan, M.T., Lovell, C.D., von Krusenstiern, E., Jiang, H., et al. (2022). Quantitative subcellular acyl-CoA analysis reveals distinct nuclear metabolism and isoleucine-dependent histone propionylation. Mol Cell *82*, 447-462 e446. 10.1016/j.molcel.2021.11.006.

135. MacLean, B., Tomazela, D.M., Shulman, N., Chambers, M., Finney, G.L., Frewen, B., Kern, R., Tabb, D.L., Liebler, D.C., and MacCoss, M.J. (2010). Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics *26*, 966-968. 10.1093/bioinformatics/btq054.

136. Yu, G., Wang, L.G., and He, Q.Y. (2015). ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics *31*, 2382-2383. 10.1093/bioinformatics/btv145.

137. Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, D. (2002). The human genome browser at UCSC. Genome Res *12*, 996-1006. 10.1101/gr.229102.

138. Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2000). The Protein Data Bank. Nucleic Acids Res *28*, 235-242. 10.1093/nar/28.1.235.

139. Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res *32*, 1792-1797. 10.1093/nar/gkh340.

140. Lister, J.A., Capper, A., Zeng, Z., Mathers, M.E., Richardson, J., Paranthaman, K., Jackson, I.J., and Elizabeth Patton, E. (2014). A conditional zebrafish MITF mutation reveals MITF levels are critical for melanoma promotion vs. regression in vivo. J Invest Dermatol *134*, 133-140. 10.1038/jid.2013.293.

141. Hugo, W., Shi, H., Sun, L., Piva, M., Song, C., Kong, X., Moriceau, G., Hong, A., Dahlman, K.B., Johnson, D.B., et al. (2015). Non-genomic and Immune Evolution of Melanoma Acquiring MAPKi Resistance. Cell *162*, 1271-1285. 10.1016/j.cell.2015.07.061.

Figure Legends

Figure 1: ALDH1A3^{High} melanomas are enriched for neural crest stem cell and glucose metabolic states

A. ALDH^{High} and ALDH^{Low} melanoma cell subpopulations. The Aldefluor assay quantifies ALDH activity in live cells by measuring the accumulated fluorescence from converted BODIPY-aminoacetaldehyde (BAAA) to BODIPY-aminoacetate (BAA-). Highest and lowest Aldefluor activity, termed ALDH^{High} and ALDH^{Low} subpopulations, were isolated by FACS sorting. Negative control: DEAB, a pan-ALDH inhibitor.

B. ALDH1A3 cellular localisation and levels. Immunocytochemistry (ICC) staining by fluorescence antibody labelling ALDH1A3 in A375 sorted ALDH^{High} and ALDH^{Low} cells. Scale bar = 50 µm. Quantification of cells enriched with nuclear ALDH1A3: N=2 biological repeat experiments, quantified image fields n=11 for ALDH^{High} and n= 9 for ALDH^{Low}, mean±s.d.; non-paired Kolmogorov-Smirnov t-test, ****$p<0.0001$. Cells showing minimal cytoplastic content outside DAPI-stained regions (indicated with an asterisk*) were determined as unfit following FACS and thus excluded from quantification.

C. Differential gene expression in ALDH^{High} and ALDH^{Low} melanoma cells. Volcano plot showing differentially expressed genes in ALDH^{High} (red) and ALDH^{Low} (blue) cells. DESeq2-analysis was performed using RNA-seq data of 3 biological replicants and cut-off thresholds using fold-change >1 and p.adj < 0.05. See also **Table S1**.

D, E. Significantly enriched pathway analysis in melanoma cells and patient samples. Dot plot of pathway analysis showing significantly enriched terms **(D)** in ALDH^{High} and ALDH^{Low} cells and **(E)** in ALDH1A3^{High} and ALDH1A3^{Low} patient samples. Dot sizes represent -$\log_{10}$ FDR q-value (Benjamini–Hochberg test) and colours indicate normalised enrichment score (NES). Patient groups are defined in **Figure S1E**. See also **Table S1, S2**.

F. *ALDH1A3* gene network heatmap. Heatmap of TCGA melanoma patient samples with RNA expression of *TFAP2B, RXRG* and mesenchymal genes, as well as *MITF, TFAP2A*, and their target genes ranked by *ALDH1A3* level. Colour index filling scaled by row (each gene).

G. *ALDH1A3* and *TFAP2B* scRNA-seq cluster in metastatic patient samples. UMAP feature plot showing expression level of *ALDH1A3* and *TFAP2B* in scRNA-seq of reanalysed patient samples of metastatic melanomas[50]. Red circle highlights a *ALDH1A3* and *TFAP2B* cluster. Gene expression is visualised by a colour index from blue (low) to red (high).

H. *ALDH1A3* expression in immune therapy treatment responders. Ipilimumab treatment responders (complete and partial) versus non-responders (progressive disease) in *ALDH1A3*^{High} and *ALDH1A3*^{Low} patient subgroups in GSE78220 cohort[24]. ALDH1A3^{High} patients were defined by having *ALDH1A3* RNA levels higher than the rounded median reads of the cohort (>9 FPKM) and ALDH1A3^{Low} patients having *ALDH1A3* less than the rounded median reads (<9 FPKM).

I. *ALDH1A3* expression in immune-hot and immune-cold melanomas. Immune-hot versus immune-cold samples in *ALDH1A3*^{High} and *ALDH1A3*^{Low} patient subgroups in the Lund melanoma cohort. ALDH1A3^{High} patients were defined by having *ALDH1A3* RNA levels higher than the rounded median reads of the cohort. Immune-hot versus immune-cold subtypes are defined by immune cell infiltration deconvolution analysis as shown in **Figure S1J**.

J. Schematic of new model of ALDH1A3 activity. High levels of ALDH1A3 Aldefluor activity are associated with nuclear localization and a high glucose metabolism and TFAP2B-NC stem cell dual state. Low levels of ALDH1A3 Aldefluor activity are associated with a differentiated, immunogenic, and starved-like state.

Figure 2: TFAP2B promotes stemness and dedifferentiation in ALDH1A3^{High} melanoma cells

A. ALDH1A3 melanoma cell models. Aldelfuor activity and quantification of *ALDH1A3* knockout (KO), over-expression (OE), and vehicle control cells (A375 and C089 cells). n=3, mean±s.d.; one-way ANOVA with Tukey's test correction for multiple comparisons.

B. The ALDH1A3-neural crest stem cell gene signature. RT-qPCR results of signature genes in control versus *ALDH1A3* KO cells (A375 cells, 3 bio-replicants with 3 technical replicates).

C. TFAP2B protein levels are dependent on ALDH1A3. Western blot and quantification of ALDH1A3, TFAP2B protein expression in ALDH1A3 control, KO and OE cells. n=3, mean±s.d.; one-way ANOVA with Tukey's test for multiple comparisons.

D, E. TFAP2B rescues ALDH1A3 activity in colony assays. Representative images and quantification of colonies formed by control, *ALDH1A3* KO, *ALDH1A3* OE, *TFAP2B* KO, and combined *ALDH1A3* KO with *TFAP2B* OE conditions. n=3 per condition, mean±s.d. one-way ANOVA with Tukey's test for multiple comparisons.

F. ALDH1A3 overexpression promotes TFAP2B expression in low-passage patient cells. Melanoma tissue was acquired from patients and maintained as low passage melanoma patient derived (MPD) 001 cells. RT-qPCR of *ALDH1A3, TFAP2A, TFAP2B* expression in control cells and cells over-expressing *ALDH1A3*. n=3, mean±s.d.; Multiple non-paired t-test corrected with Holm-Sidak's method.

G. ALDH1A3 overexpression leads to less pigmentation. Cell pellets showing brown melanin in control and ALDH1A3 overexpressing MPD001 cells.

*P<0.05; **P<0.01; ***P<0.001; ****P<0.0001.

Figure 3. ALDH1A3^{High} cells use glucose while ALDH1A3^{Low} cells rely on acetate for Acetyl-CoA production

A-E. Schematic of $^{13}C_6$ glucose tracing experiment. Central carbon metabolites were profiled by UPLC-MRM/MS in panel (**B**) Glucose (**C**) Pyruvate (**D**) TCA cycle metabolites: citrate, alpha-ketoglutarate, succinic acid, malic acid, and (**E**) aceto-acetate. (n=3 for WT vs ALDH1A3 knockout cells; n=5 for ALDH^{High} vs ALDH^{Low} cells. Multiple paired t-test corrected with Holm-Sidak's method.)

F. *ALDH1A3* KO cells generate and secrete lactate. Intracellular lactate measured by targeted UPLC-MRM/MS. (n=3, Multiple paired t-test corrected with Holm-Sidak's method). Secreted lactate levels were measured by a colorimetric lactate assay kit and normalised to standards. (n=3, unpaired student t-test.)

G, H. *ALDH1A3* KO cells generate more acetyl-CoA but not from glucose. Intracellular glucose w/o 13C_6-labelling and **(H)** total acetyl-CoA measured by targeted UPLC-MRM/MS. (n=3, Multiple paired t-test corrected with Holm-Sidak's method.

I, J. Western blot analyses of pan-acetyl-histone H3 (Acetyl-K9+K14+K18+K23+K27), and total histone H3 protein levels in lysates of A375 and C089 cell lines with vehicle control (WT), *ALDH1A3* knockout (KO) or *ALDH1A3* overexpression (OE) respectively, with total histone H3 probed as loading control and **(J)** quantification (n=3, mean±s.d.; one-way ANOVA with Tukey's correction for multiple comparisons. *P<0.05; ***P<0.001.)

K, L. Schematic of $^{13}C_2$ acetate tracing experiment. $^{13}C_2$ acetate tracing experiment designed for ^{13}C-incoporation profiling by HPLC in (**L**) whole-cell and nuclear Acetyl-CoA. n=5, two-way ANOVA corrected with Holm-Sidak's method.)

M. Schematic of ALDH1A3 and metabolic states. ALDH1A3 acts through two different sources of Acetyl (Ac)-CoA production: high glucose flux generates high levels of pyruvate, leading to Ac-CoA in ALDH^{High} cells, while ALDH^{Low} cells preferentially uptake acetate as an Ac-CoA source.

*P<0.05; **P<0.01; ***P<0.001; ****P<0.0001.

Figure 4: Selective histone H3 acetylation depends on ALDH1A3.

A. Collective acetyl-histone H3 reads abundance at peak regions measured by acetyl-histone H3 ChIP-seq. (Wilcoxon signed-rank test. ****P<0.0001).

B. Volcano plot of differentially acetylated histone H3 peaks in WT control and ALDH1A3 knockout (highlighted by magenta, fold change >1, FDR.q < 0.05).

C. Heatmap of ChIP-seq histone pan-acetyl-histone H3 peaks aligned by gene body extended by 20% upstream transcription starting site (TSS) and downstream transcription termination site (TTS) (bottom panels), with average profile showing read count frequency distribution (upper panels) and **(D)** genomic annotations by bar plot (top panels) among A375 control (Ctrl_All), ALDH1A3 knockout (KO_All), significantly enriched peaks in A375 control (Ctrl_Enrich, Ctrl_Enrich>2 folds), and significantly enriched peaks in ALDH1A3 knockout (KO_Enrich, KO_Enrich>2 folds).

E. Gene tracks of histone pan-acetyl-histone H3 ChIP-seq reads in control versus *ALDH1A3* KO cells. Significant changes in histone H3 acetylation are highlighted (arrows).

F. Venn diagram of overlapping genes between upregulated genes by RNA-Seq in sorted $ALDH^{High}$ cells **(Figure 1C)**, and genes enriched in control cells versus *ALDH1A3* knockout cells by histone H3 acetylation in **B**. P values by Fisher's exact test.

G. Transcription factor binding motif over-representation analysis of the Venn diagram overlapping genes in **F**. Enrichment score by g:Profiler (version e109_eg56_p17_1d3191d) with g:SCS multiple testing correction method applying statistical significance threshold of 0.05 [111].

H. Binding motif of human EGR1 and TFAP2B from JASPAR database.

I. Scatter plot shows the correlation of $ALDH^{High}$ enriched gene signature and Tfap2b regulon activities (AUCell score) in murine melanoma cells (*NRAS*$^{Q61K/°}$;*Ink4a*$^{-/-}$)[43].

J. SCENIC-inferred Tfap2b regulatory network using a murine mouse scRNA-seq dataset[43], with Tfap2b target genes in green and $ALDH^{High}$ enriched Tfap2b target genes highlighted in yellow. Tfap2b regulon genes were mapped to human homologs to allow comparison and visualization. See also **Table S5**.

Figure 5: ALDH1A3 forms a predicted complex with ACSS2

A. ALDH1A3 functional association. STRING functional protein association analysis between yeast Ald6, Acs2, and Pdc1, orthologous to human ALDH1A3, ACSS2, and PDH.

B. AlphaFold multimer modelling of human ALDH1A3 and ACSS2. Proteins visualised as a complex in surface model (left) and ribbon model (right). The binding sites of the aldehyde substrate (retinaldehyde) and co-factor (NAD+) were created by structural alignment of the AlphaFold-Monomer predicted structure to PDB ID 5fhz, a published tetramer ALDH1A3 structure. The ATP/CoA binding pocket of ACSS2 was visualised by structural alignment of AlphaFill-optimized protein-ligand complexes (ATP donor: PDB ID 5k8f; CoA donor: PDB ID 3gpc).

C-D. Subcellular expression of ALDH1A3, ACSS2, and TFAP2B. **(C)** ICC staining by fluorescence antibody probing ALDH1A3 (yellow), ACSS2 (green), TFAP2B (red) in sorted A375 $ALDH^{High}$ and $ALDH^{Low}$ cells. DAPI (blue). Scale bar = 10 µm. **(D)** Fluorescence signal intensity quantification of nuclear TFAP2B and ACSS2 in ICC images. N=2 biological repeat experiments, n = 55 quantified single cells for $ALDH^{High}$ TFAP2B, n= 46 single cells for $ALDH^{Low}$ TFAP2B; n = 26 single cells for $ALDH^{High}$ ACSS2, n= 20 single cells for $ALDH^{Low}$ ACSS2, (represented as individual dots), mean±s.d., unpaired non-parametric Komogorov-Smirnov t-test. ****P<0.0001.

E-H. ALDH1A3, ACSS2 and TFAP2B co-localise in the nucleus. Structured illumination microscopy (SIM) of ALDH1A3 (magenta), TFAP2B (orange), ACSS2 (green), and DAPI (blue) in A375 melanoma FACS sorted **(E)** $ALDH^{High}$ cells and **(G)** $ALDH^{Low}$ cells. Scale bar = 1 µm. Complex co-localization signals were abundant (arrows) and are indicated as highlighted dots in the zoomed image in **E**. Low signals of nuclear TFAP2B are highlighted in zoomed image in **G**. **(F)** Intensity plot profile of the line scan across two co-localisation hotspots in zoomed image of **F**. Signal overlap peaks are indicated by arrows. **(H)** Intensity plot profile of the line scan across the two TFAP2B signal spots in zoomed image of **G**. TFAP2B signal peaks are indicated by arrows, where no ALDH1A3 or ACSS2 signals are present.

Figure 6: Acetaldehyde serves as an acetyl source for histone H3-acetylation

A. Schematic summarising the acetyl group transfer from pyruvate to histone mediated by ALDH1A3 metabolising acetaldehyde.

B. Overview of $^{13}C_2$ acetaldehyde tracing experimental design**.** Control and *ALDH1A3* KO cells (A375, C089) were treated with $^{13}C_2$ acetaldehyde for two-hours and the carbon flow from acetaldehyde to histone traced by HPLC-MS/MS.

C. Acetaldehyde is an acetyl source across multiple lysine residues on Histone H3. $^{13}C_2$-acetyl groups were traced from acetaldehyde into lysine acetylation at K9, K14, K18, K23, and K27. *ALDH1A3* knockout cells treated with $^{13}C_2$-acetaldehyde showed less $^{13}C_2$-acetyl groups at K9, K14 and K23 residues. n=5 per cell line; P values by two-way ANOVA with Sidak's multiple comparisons test.

D-E. ALDH1A3 directs gene-specific ACSS2 chromatin binding. **(D)** Schematics showing the primer design for ChIP-qPCR for ACSS2 binding upstream of TFAP2B, with a local histone H3 acetylation island controlled by ALDH1A3 (from **Figure 4**). **(E)** ACSS2 ChIP-qPCR results in control versus *ALDH1A3* knockout cells. N=3 biological repeat experiments in both A375 and C089 cells; each with n=3 technical replicates and normalised to IgG pull-down control before heatmap scaling and plotting. Unpaired Mann-Whitney test with Holm-Sidak's method.

Figure 7. ALDH1A3^{High} subpopulations promote melanoma drug resistance in vivo

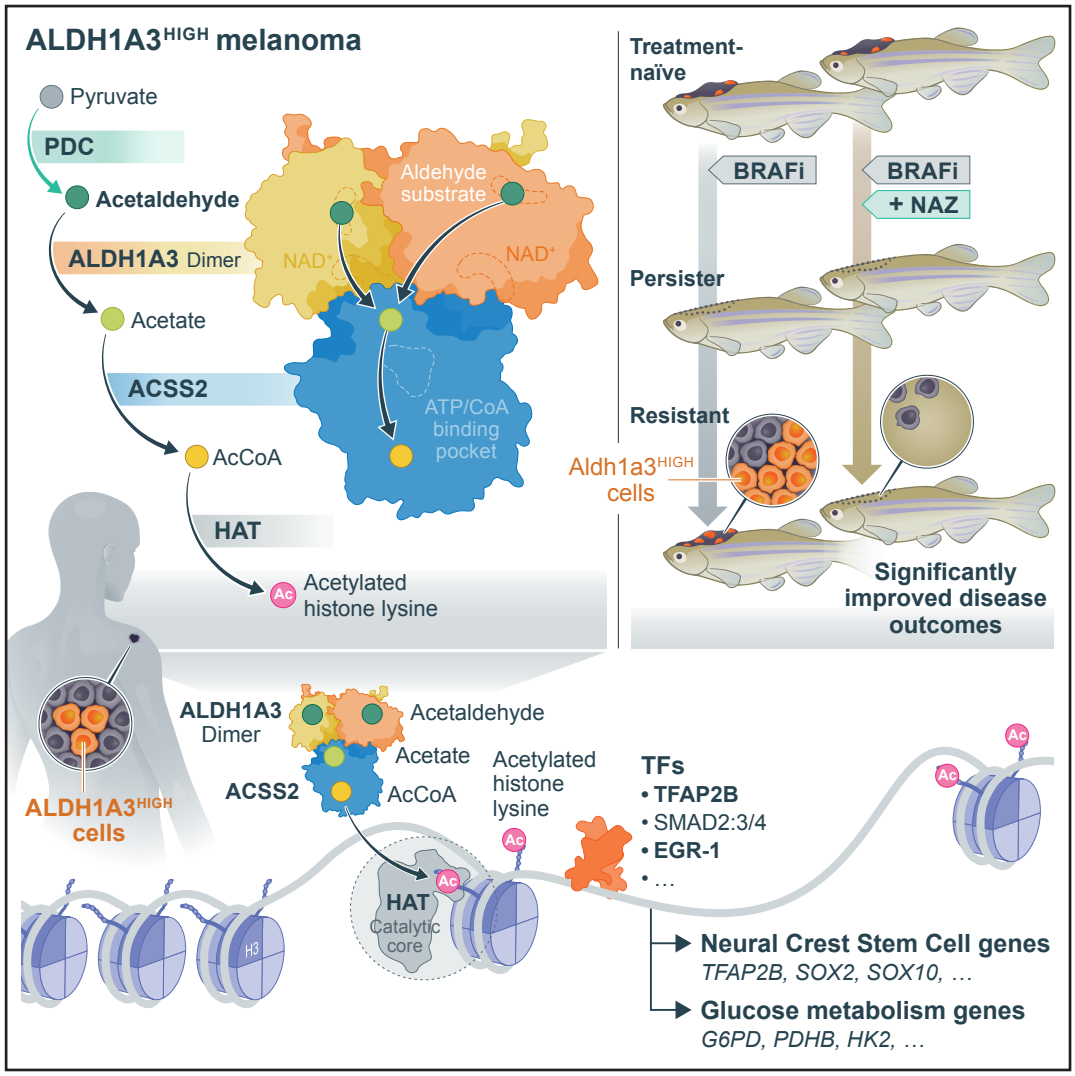
A. Zebrafish BRAF^{V600E} melanoma model. Mesoscopic images of a zebrafish with melanoma (arrow highlighted). *BRAF^{V600E}* and *GFP* are expressed in the melanocyte lineage, by the *mitfa* promoter, and the zebrafish is mutant for p53.

B-D. ALDH^{High} cells in zebrafish melanoma. Dissected zebrafish melanoma tissues were dissociated into single cell suspensions and processed with the AldeRed assay. A cell population with high ALDH activity and correspondingly low levels of *mitfa:GFP* is indicated (red circle). DEAB is an ALDH inhibitor and serves as a negative control. (**C**) Flow cytometry sorting of ALDH^{High} and ALDH^{Low} zebrafish melanoma cells. (**D**) RT-qPCR quantification showing significantly higher *sox2, sox10 and tfap2b* expression in sorted ALDH^{High} versus ALDH^{Low} zebrafish melanoma cells. (n=3 bio-replicates each with 3 technical replicates, Multiple paired t-test corrected with Holm-Sidak's method).

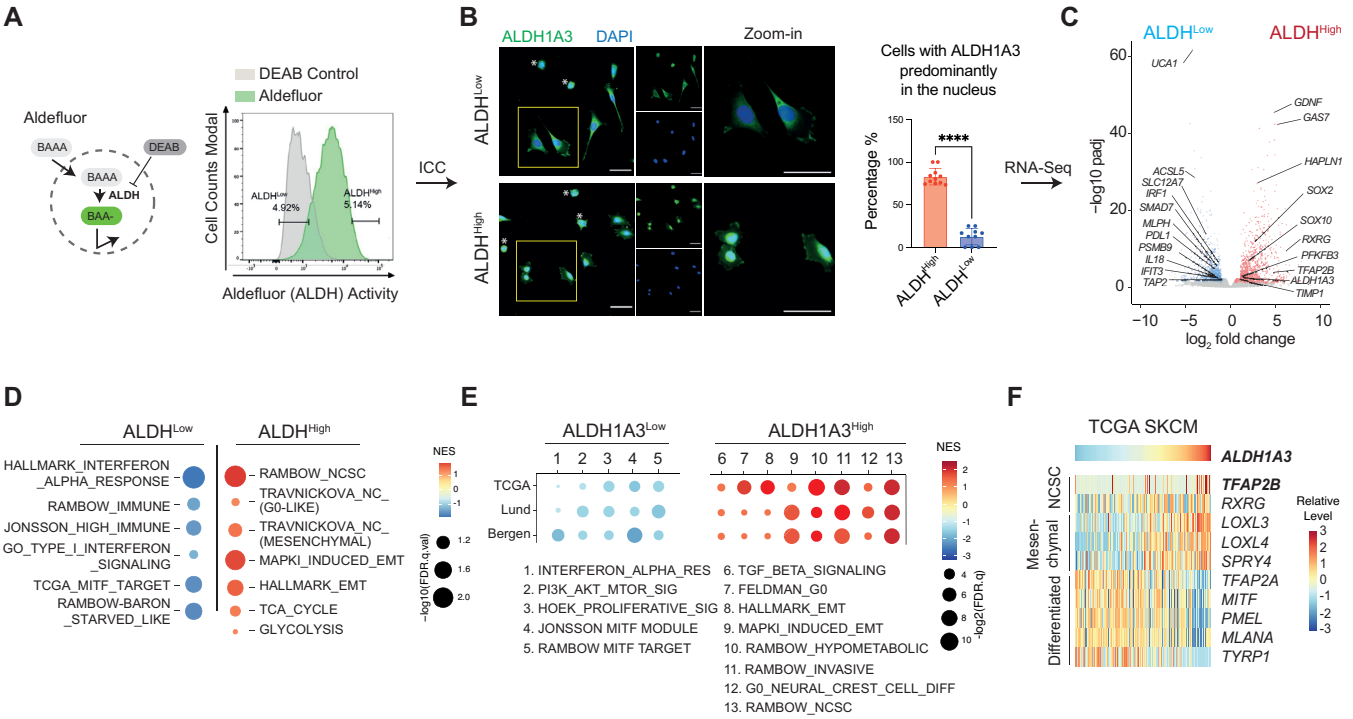
E-F. Aldh1a3 in zebrafish model of BRAF-inhibitor regression and recurrent disease. **(E)** Schematic of *BRAF^{V600E} p53* mutant zebrafish with melanomas were non-invasively fed with 200 mg/kg/day vemurafenib-containing food pellets leads to melanoma regression, drug resistance and disease recurrence. **(F)** Aldh1a3 immunohistochemistry in treatment naïve (DMSO), BRAFi and/or NAZ treatment responding disease, and in drug-resistant recurrent disease. Aldh1a3 IHC in melanoma recurrent disease treated with indicated NAZ alone or in combination with BRAFi shows Aldh1a3 on-target efficacy of nifuroxazide. One-way ANOVA with Tukey's test correction for multiple comparisons. *P<0.05; ***P<0.001; ****P<0.0001.

G-H. Zebrafish combination drug trial to target ALDH^{High} cells. **(G)** Long-term drug pellet treatment design. **(H)** Kaplan Meier survival curves of zebrafish melanomas under different drug treatment shown in **G**: DMSO (Black): daily DMSO control treatment. BRAFi (Orange): 200 mg/kg/day vemurafenib treatment. NAZ (Blue): daily 150 mg/kg/day Nifuroxazide treatment. BRAFi_NAZ (Magenta): 3-week treatment of 200 mg/kg/day vemurafenib followed by daily 150 mg/kg/day Nifuroxazide treatment. BRAFi+NAZ (Green): 3-week treatment of 200 mg/kg/day vemurafenib followed by the combination of 200 mg/kg/day vemurafenib and 150 mg/kg/day Nifuroxazide treatment. The weekly tumour size change of each fish was calculated and an event of tumour size increase over 20% is considered as the end of progression free survival period. Log-rank tests, $p < 0.001$.

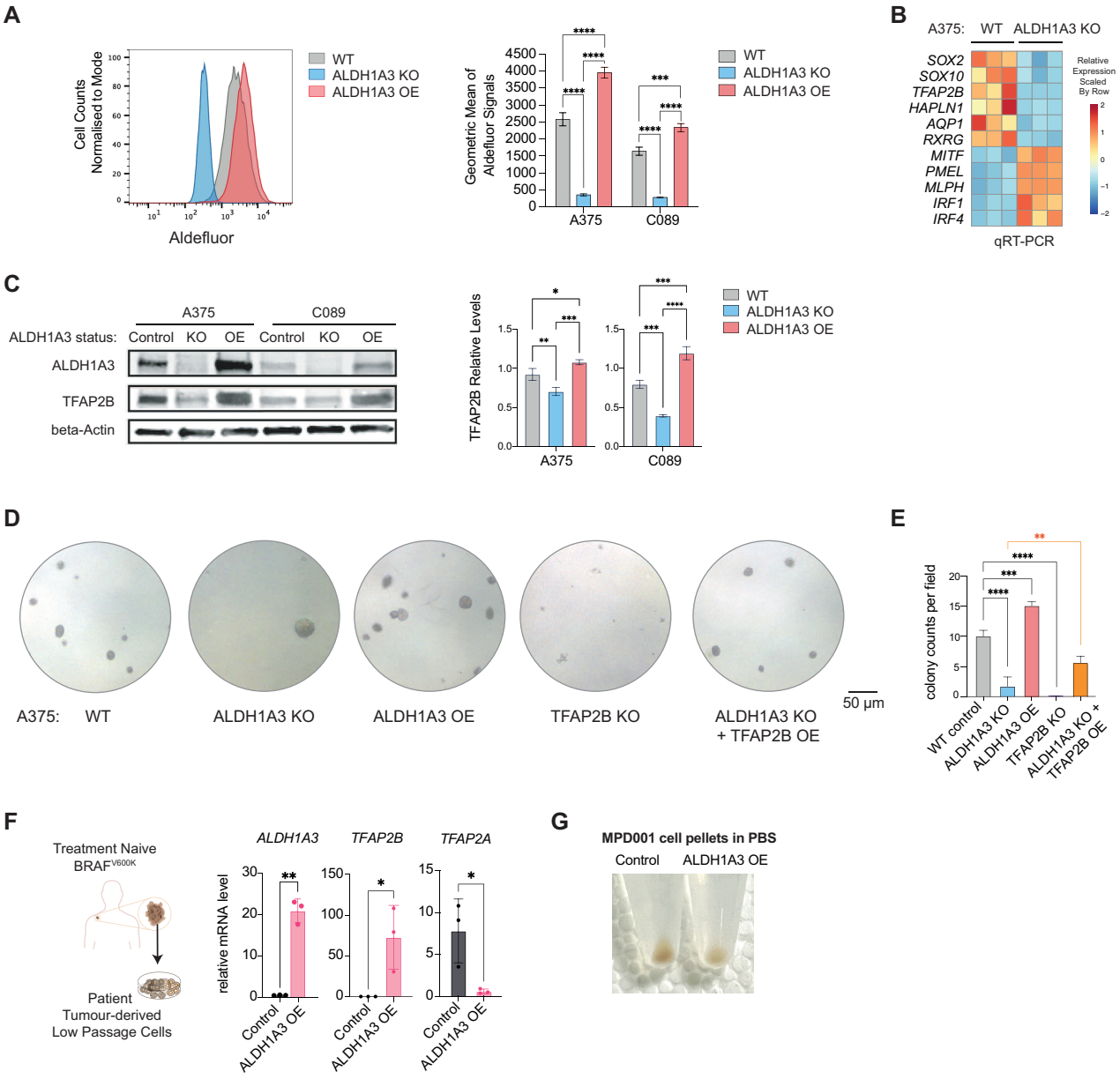
I. Melanoma immune response following drug treatment. Representative immunofluorescence staining of melanomas with zebrafish B-cell marker Cd79a at 2-week point on-drug treatment. Matched H&E staining indicate melanoma region of interest "**M**". Quantification of Cd79a staining between samples from DMSO (N=3 fish, n=6 lesions), BRAFi (N=3 fish, n=6 lesions), NAZ (N=2 fish, n=6 lesions), and BRAFi+NAZ (N=4 fish, n=6 lesions) treated animals. One-way ANOVA with Tukey test correction for multiple comparisons. ***P<0.001.



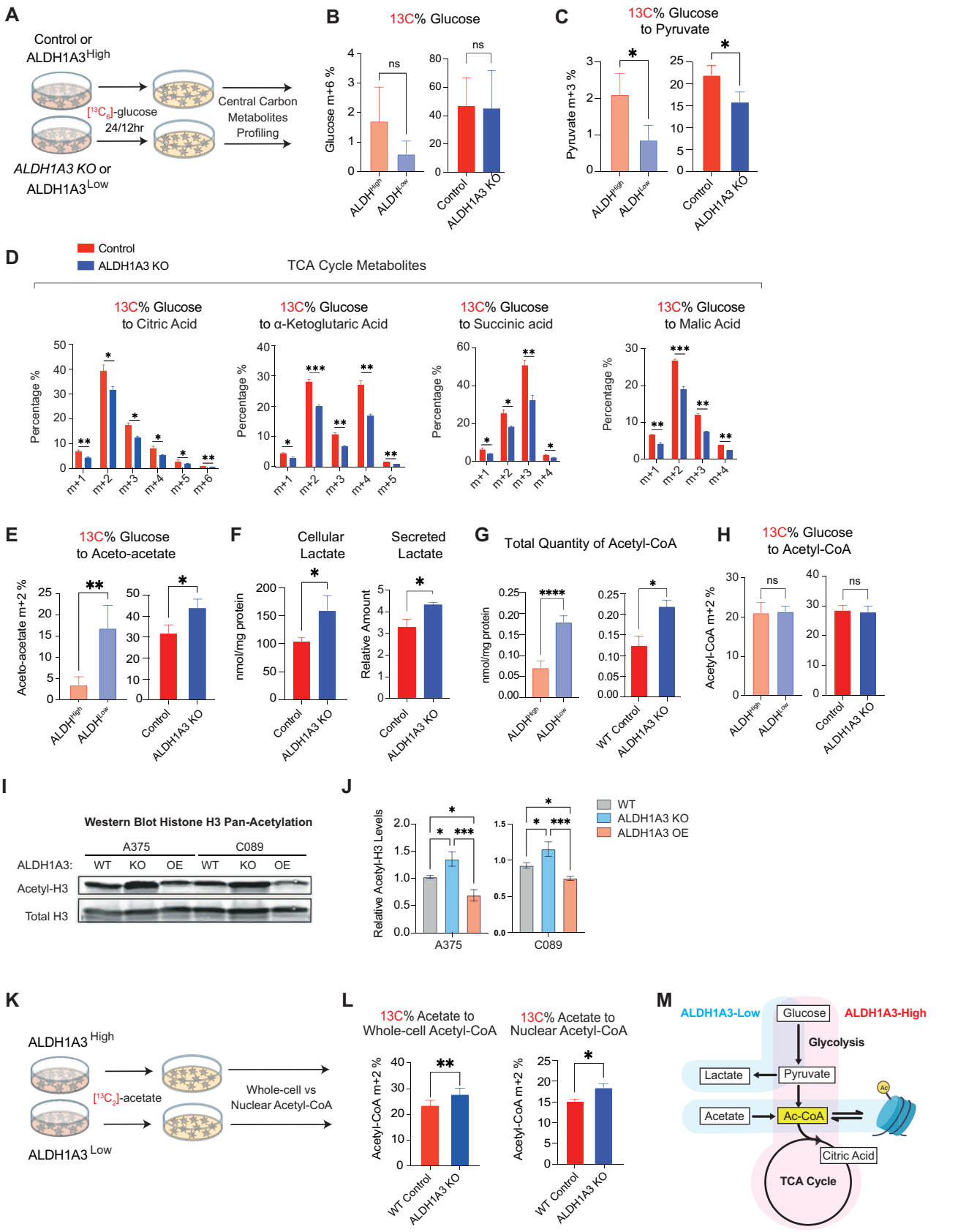
Graphical Abstract



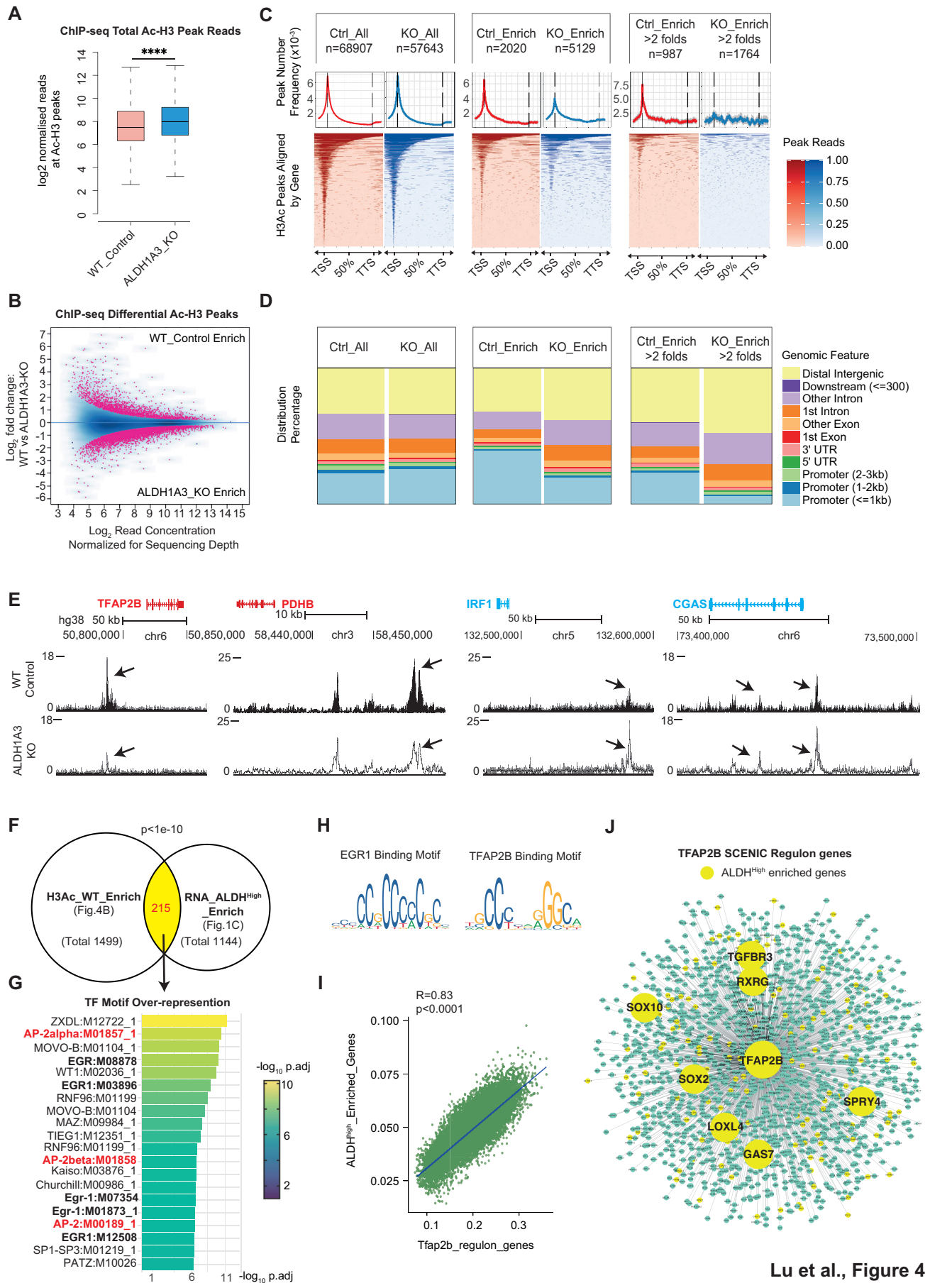
Lu et al., Figure 1



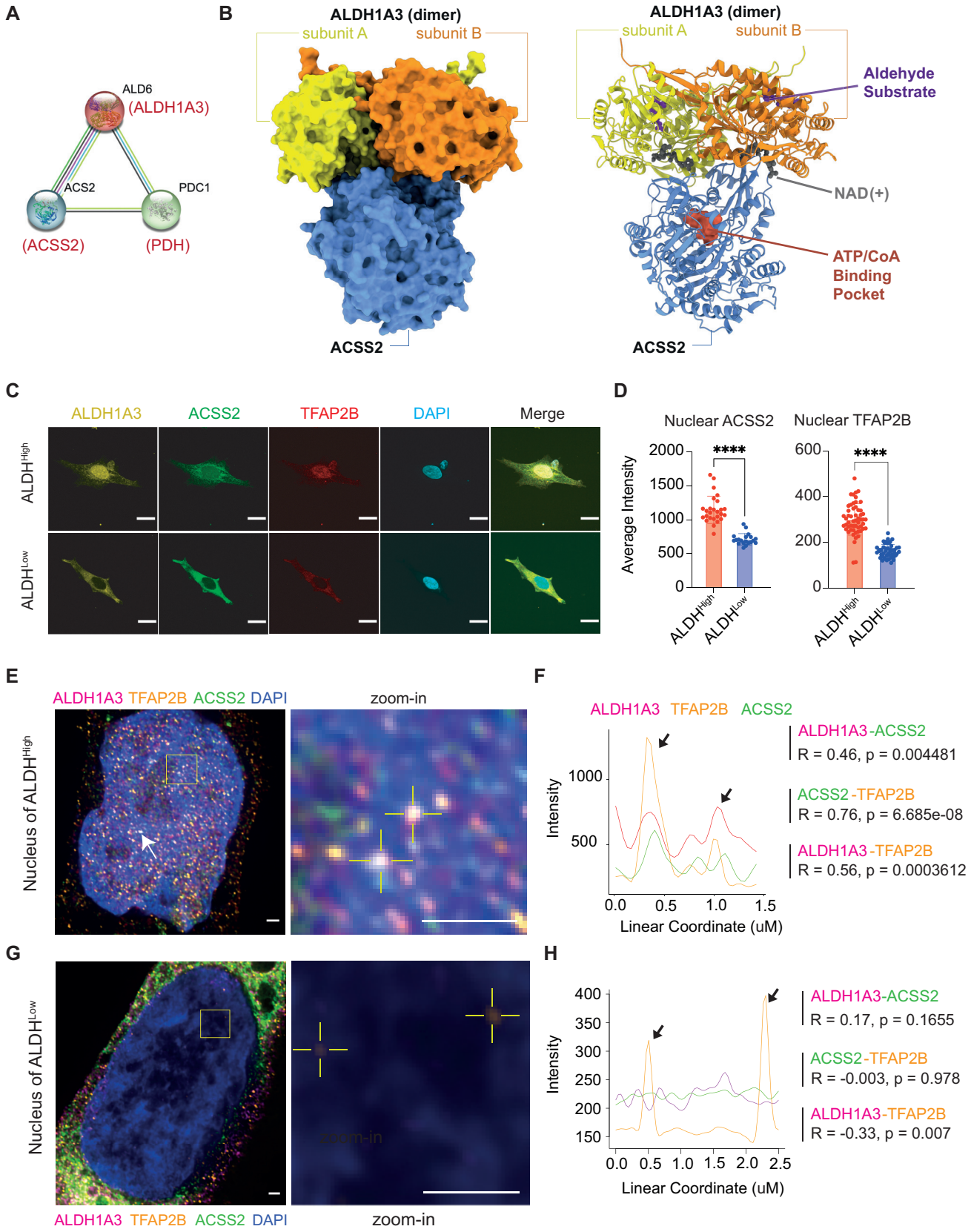
Lu et al., Figure 2



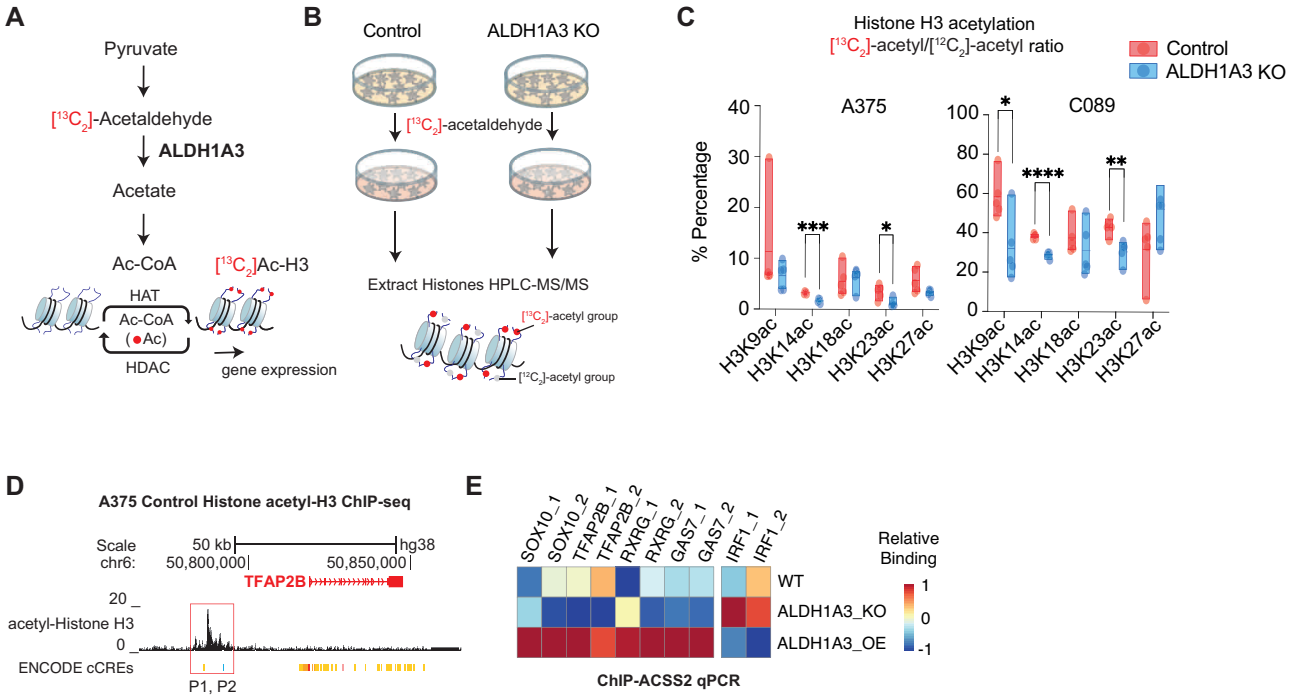
Lu et al., Figure 3



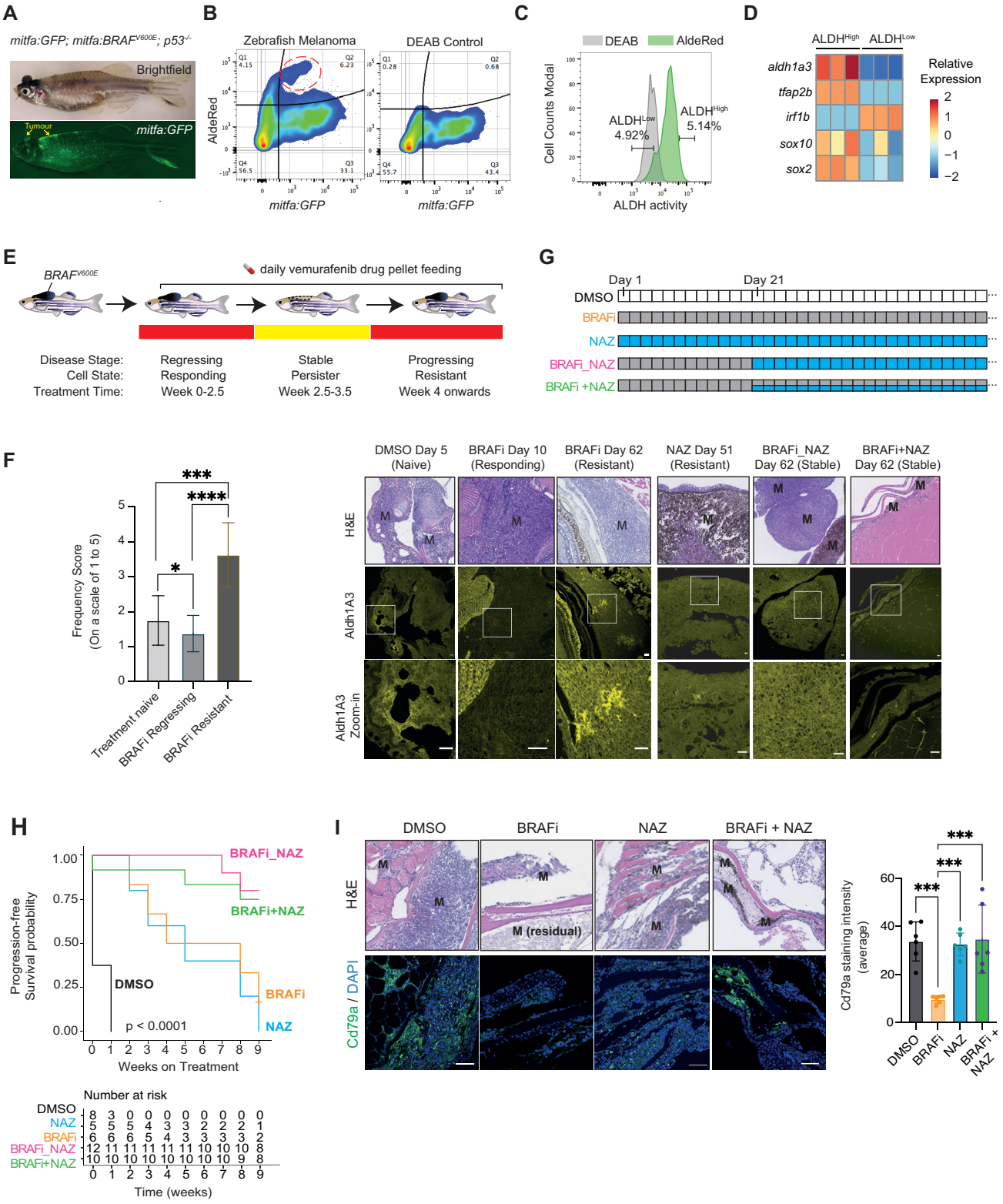
Lu et al., Figure 4



Lu et al., Figure 5



Lu et al., Figure 6



Lu et al., Figure 7

Supplementary Figure Legends

Supplementary Figure 1

A. Pair-matched patient samples (GSE65185)[141] ranked by *ALDH1A3* fold enrichment comparing post MAPKi resistance to pre-treatment biopsies. The average of *ALDH1A3* was used for comparison when multiple MAPKi-resistant tumour biopsies were taken from the same patient.

B. Violin plot depicts *Aldh1a3* RNA expression in different murine melanoma cell states, including neural crest (NC)-like (cluster 1) and stem-like states (cluster 4)[43].

C. UMAP projection of murine melanoma cells colored by CytoTRACE score (gene expression diversity score, less differentiated = score close to 1 and more differentiated = score close to 0) and Aldh1a3 expression. Cluster numbers that relate to **B** are indicated.

D. Microscope image overlaying brightfield and GFP channel shows the heterogeneous ALDH activity among A375 melanoma cells as measured by the Aldefluor assay on (adhered) live cells. Scale bar = 20 µm.

E. Murine melanoma genes from **B, C** ranked by decreasing correlation coefficient with CytoTRACE score (genes conferring a less differentiated phenotype close to 1 and genes conferring a differentiated phenotype close to -1).

F. Heatmap showing RT-qPCR results validating differentially expressed genes identified from RNA-seq on 3 bio-replicants of independently sorted A375 $ALDH^{High}$ and $ALDH^{Low}$ cells. Relative expression of each gene was averaged across 3 technical replicants and normalised to beta-actin loading control before heatmap plotting. For heatmap visualisation, each row (each gene) was scaled for the colour index filling.

G. Patient samples ranked by *ALDH1A3* expression, with the top and bottom 10% ranking samples of TCGA and Lund datasets defined as $ALDH1A3^{High}$ and $ALDH1A3^{Low}$ group respectively, while the top and bottom 25% ranking samples of Bergen datasets were defined as $ALDH1A3^{High}$ and $ALDH1A3^{Low}$ groups due to fewer patient numbers in the Bergen cohort.

H. Heatmap showing gene expression fold enrichment levels comparing the $ALDH1A3^{High}$ to $ALDH1A3^{Low}$ patient groups. Colorimetric values normalized by column (each cohort).

I. Bar plot showing spearman correlation value of each *ALDH* isoform with *TFAP2B* in TCGA dataset. Colour filling scaled by FDR.q.value. Only *ALDH1A3* shows significant positive correlation with TFAP2B.

J. DNA Binding motif sequences of TFAP2 family members from JASPAR database.

K. Mutation and clinical background patient samples table from Tirosh et al., 2016[50]. The cluster of *ALDH1A3-TFAP2B* consists of cells from highlighted patient entries.

L. Lund bulk RNA-seq dataset immune cell infiltration deconvolution analysis using CIBERSORTx algorithm based on LM22 profile. Heatmap showing unsupervised clustering of the Lund dataset based on immune cell infiltration profile divided the entire cohort into 3 subgroups, with an immune-hot group defined by a high percentage of M1 macrophage and active lymphocytes infiltration, while an immune-cold group is defined by a high percentage of M2 macrophage and low levels of active lymphocytes infiltration.

Supplementary Figure 2

A. Schematics of plasmids used to generate wildtype (WT), *ALDH1A3* knockout (KO), and *ALDH1A3* overexpression (OE) cell lines.

B. Scatter plot showing the distribution of *ALDH1A3* and *TFAP2B* expression over different human melanoma cell lines. RNAseq data (FPKM) from a large panel of human melanoma cell lines (45 from Cancer Cell Line Enclopedia and 20 from study EGAS00001000815) were retrieved and plotted. The level of *ALDH1A3* and *TFAP2B* expression in the A375 and C089 melanoma cell line are indicated. Pearson correlation analysis showed a significant positive correlation between *ALDH1A3* and *TFAP2B* expression across all melanoma cell lines.

C. Western blot of TFAP2B levels in *TFAP2B* knockout cells and *ALDH1A3* knockout cells overexpressing *TFAP2B.* Beta-Actin is presented as a loading control.

Supplementary Figure 3

Volcano plot of histone H3 peptide modification quantification measured by bottom-up histone HPLC-MS/MS. (n=5, FDR q values by two-way ANOVA corrected with Sidak's multiple comparison method.) Acetyl-histone H3 lysine residuals with significant upregulation in *ALDH1A3 KO* group comparing to the control are highlighted in **Bold**.

Supplementary Figure 4

A. Schematics of adjusted (right Y axis, highlighted in red) differential peak analysis on acetylated histone H3 peaks in WT control compared to *ALDH1A3* knockout cells. Here, we have manually increased log base 2 peak score by +0.5 in *ALDH1A3* KO samples to account for the potential biased loss of peak score during normalisation.

B. Heatmap of adjusted differentially acetylated histone H3 peaks aligned by the gene body and extended by 20% upstream transcription starting site (TSS) and downstream of the transcription termination site (TTS) (bottom panels). Read count frequency distribution is shown in an average profile (upper panels).

C. Genomic annotations by bar plot among adjusted enriched peaks in A375 control (Ctrl_Enrich_Adj, Ctrl_Enrich_Adj>2 folds), and in *ALDH1A3* knockout cells (KO_Enrich_Adj, KO_Enrich_Adj>2 folds).

D-E. Overlay of the differentially enriched acetyl-histone H3 peaks aligned by gene body from **Figure 4C** and **(E)** from **Figure S4B.** Both showed reduced proportion of TSS-centring peaks in ALDH1A3 KO cells, coupled with increased proportion of peaks towards the TTS and extended gene downstream regions as highlighted by arrows.

F-H. Significant Pathway Enrichment terms with literature curated gene lists **(F).** Gene Ontology pathway datasets enriched in control **(G)** and *ALDH1A3* knockout **(H)** cells. Benjamini-Hochberg test. FDR <0.05, Benjamini-Hochberg test.

Supplementary Figure 5

A. AlphaFold multimer modelling of yeast Ald6 and Acs2 proteins visualised in surface model (left) and ribbon model (right).

B-C. ICC staining by fluorescence antibody for ALDH1A3 (yellow, in **B**), and TFAP2B (red, in **C**) in sorted A375 $ALDH^{High}$ and $ALDH^{Low}$ cells. DAPI (blue). Scale bar = 20 µm. Fluorescence signal intensity quantification of nuclear **(B)** ALDH1A3 and **(C)** TFAP2B in ICC images. n = 60 single cells for ALDH1A3 and n=19 single cells for TFAP2B quantification (represented as individual dots), mean±s.d., unpaired non-parametric Komogorov-Smirnov t-test. ****P<0.0001.

Supplementary Figure 6

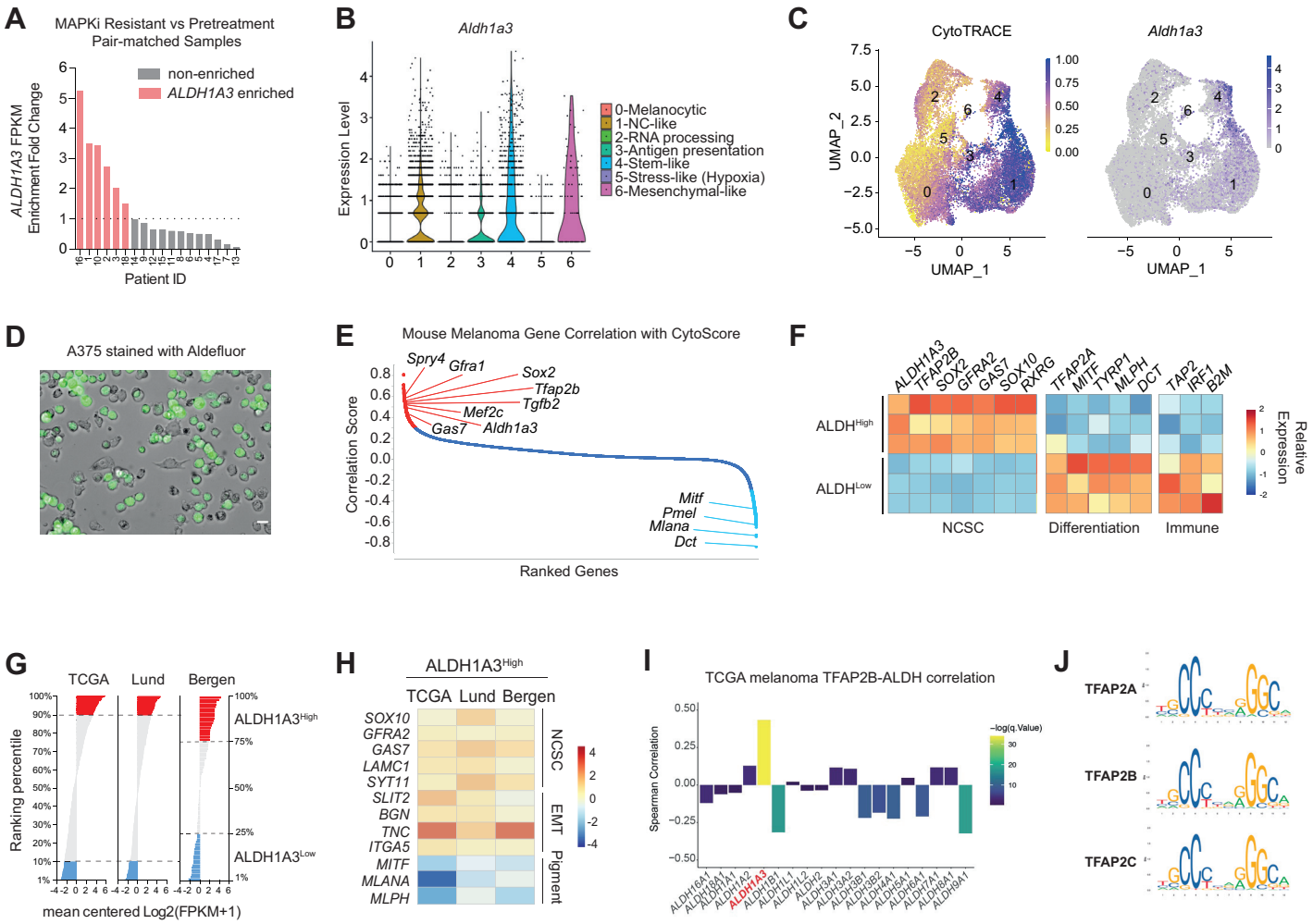
A. RT-qPCR quantification shown by a heat rose map of *aldh1a3* expression enriched with the highest folds in zebrafish $ALDH^{High}$ versus $ALDH^{Low}$ melanoma cells. (n=3 bio-replicates each with 3 technical replicates, Multiple paired t-test corrected with Holm-Sidak's method).

B. Spaghetti summary plot related to **Figure 7H**, showing tumour volume changes during drug trial design as listed in **Figure 7G**. DMSO: melanomas n=8, fish N=5; BRAFi: melanomas n=5, fish N=4; NAZ: melanomas n=6, fish N=4. BRAF_NAZ: melanomas n=12, fish N=5; BRAFi + NAZ: melanomas n=10, fish N=4. Each dot on the plot represents one melanoma lesion. Drug pellet treatment is colour coded as shown in **G**: DMSO (Black): daily DMSO control treatment. BRAFi (Orange): 200 mg/kg/day vemurafenib treatment. NAZ (Blue): daily 150 mg/kg/day Nifuroxazide treatment. BRAFi_NAZ (Magenta): 3-week treatment of 200 mg/kg/day vemurafenib followed by daily 150 mg/kg/day Nifuroxazide treatment. BRAFi+NAZ (Green): 3-week treatment of 200 mg/kg/day vemurafenib followed by the combination of 200 mg/kg/day vemurafenib and 150 mg/kg/day Nifuroxazide treatment.

Supplementary Figure 7

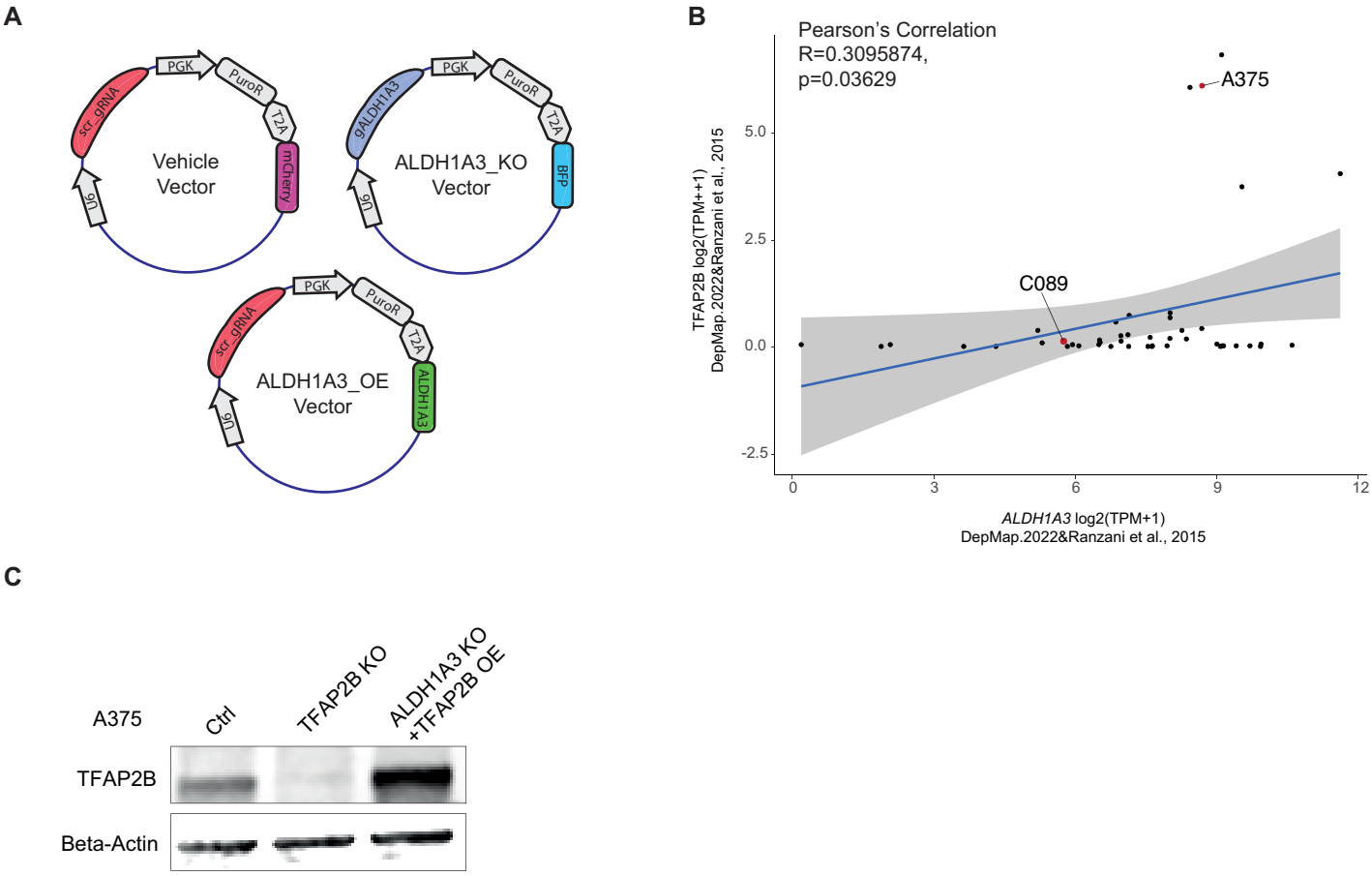
A. Schematic model of ALDH1A3 orchestrating the metabolic, epigenetic, and transcriptional cell states of melanoma.

B. UCSC genome browser snapshot of MITF ChIP-seq multiWig tracks at the ALDH1A3 locus from two experiments (Top, GSE137522; Bottom, GSE137776). Each track is an overlay of an input control and 2 independent HA-MITF ChIP-seq experiments in the human 501mel melanoma cell line engineered to inducibly express HA-tagged MITF. Inputs are shown in Green Mist colour RGB (200,200,150).

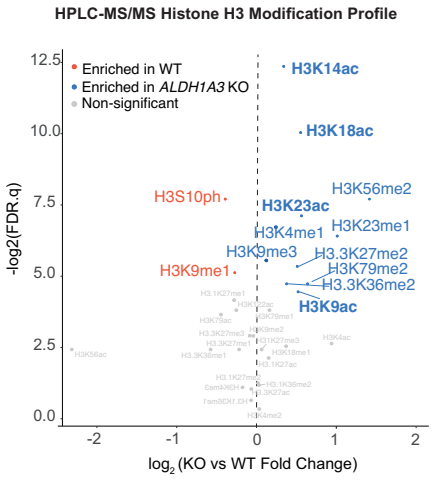


Sample ID	Age/sex	Mutation status	Pre-operative treatment	Site of resection
Melanoma_53	77/F	Wild-type	None	Subcutaneous back lesion
Melanoma_59	80/M	Wild-type	None	Femoral lymph node
Melanoma_60	69/M	BRAF V600K	Trametinib, ipilimumab	Spleen
Melanoma_65	65/M	BRAF V600E	None	Paraspinal intramuscular
Melanoma_67	58/M	BRAF V600E	None	Axillary lymph node
Melanoma_71	79/M	NRAS Q61L	None	Transverse colon
Melanoma_78	73/M	NRAS Q61L	WDVAX, ipilimumab + nivolumab	Small bowel
Melanoma_79	74/M	Wild-type	None	Axillary lymph node
Melanoma_80	86/F	NRAS Q61L	None	Axillary lymph node
Melanoma_81	43/F	BRAF V600E	None	Axillary lymph node
Melanoma_82	81/M	Wild-type	None	Axillary lymph node
Melanoma_84	67/M	Wild-type	None	Acral primary
Melanoma_88	54/F	NRAS Q61L	Tremelimumab + MEDI3617	Cutanoues met
Melanoma_89	67/M	n/a	None	Axillary lymph node
Melanoma_94	54/F	Wild-type	IFN, ipilimumab + nivolumab	Iliac lymph node

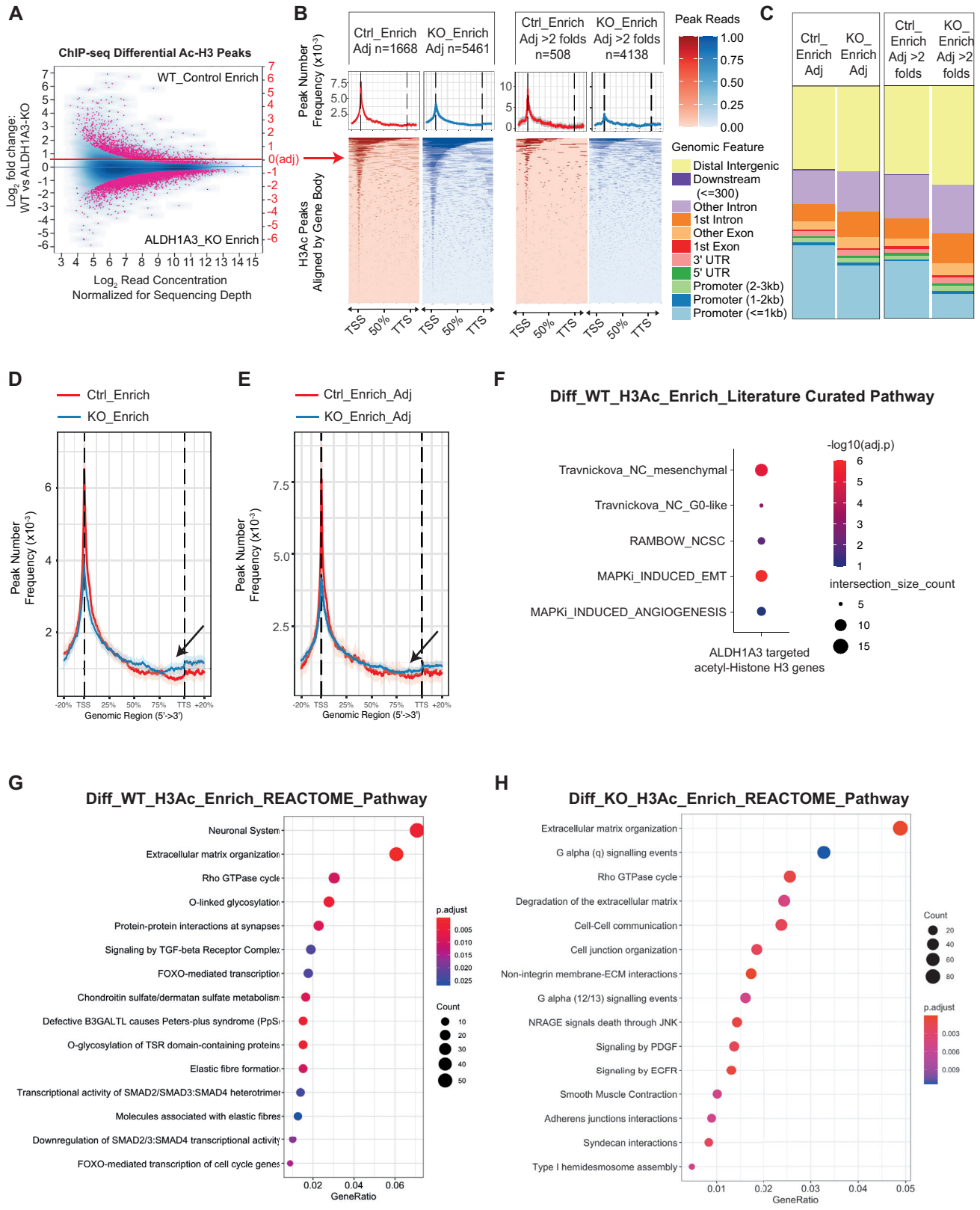
Lu et al., Supplementary Figure 1



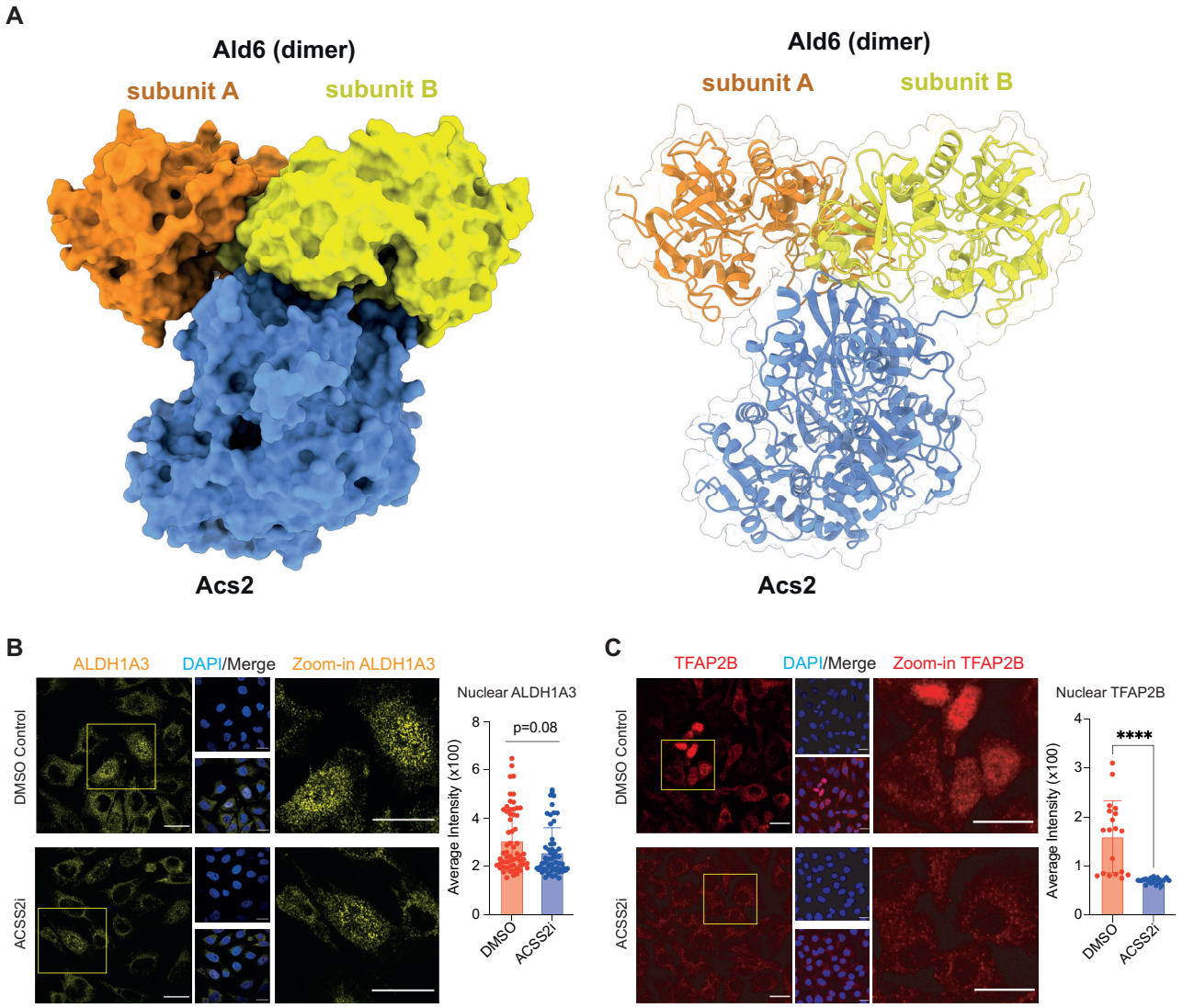
Lu et al., Supplementary Figure 2



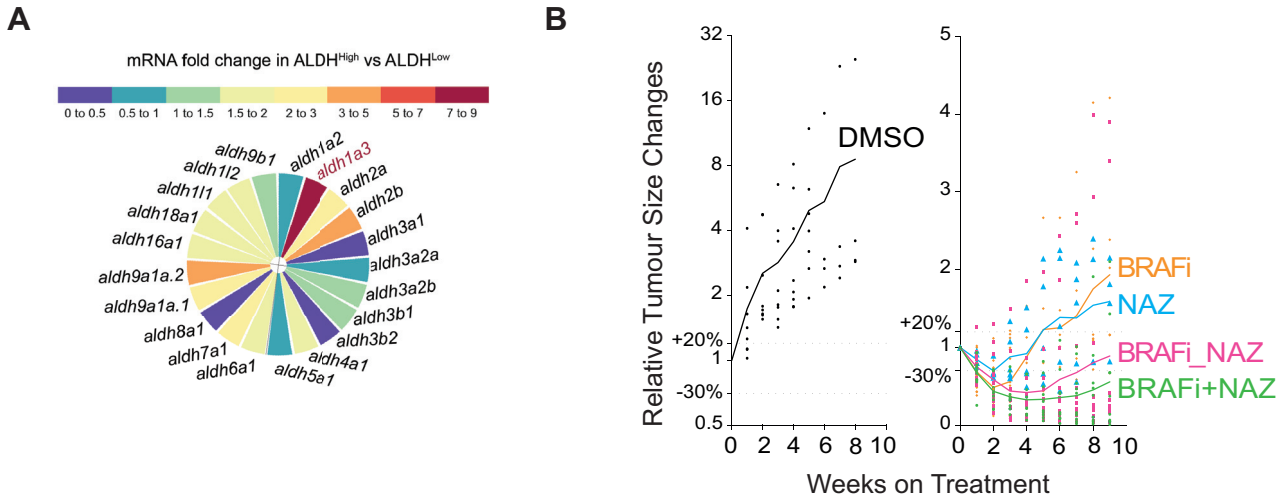
Lu et al., Supplementary Figure 3



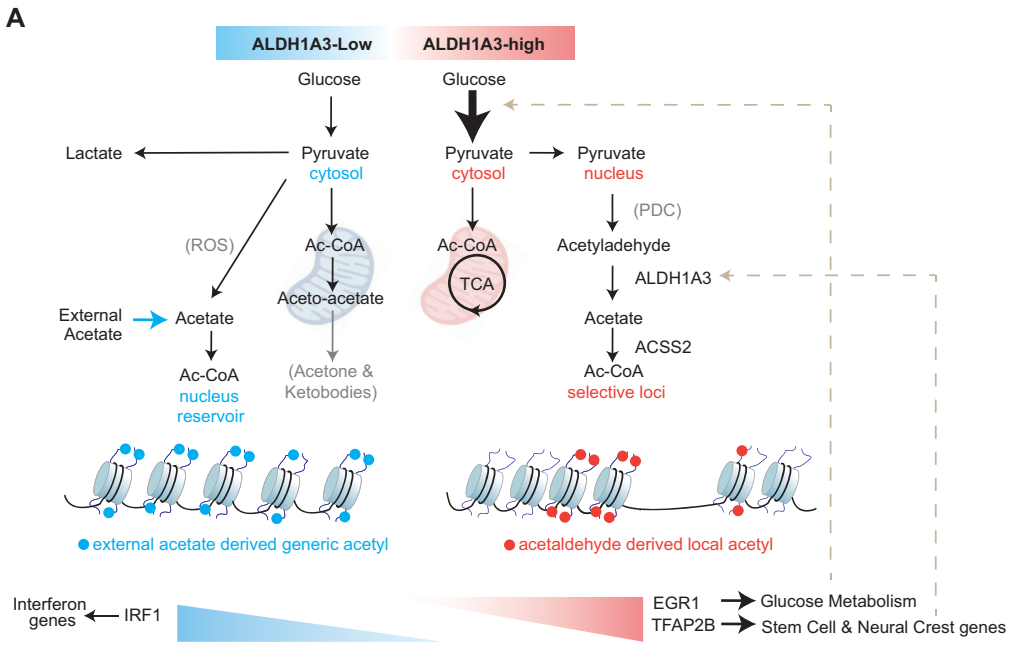
Lu et al., Supplementary Figure 4

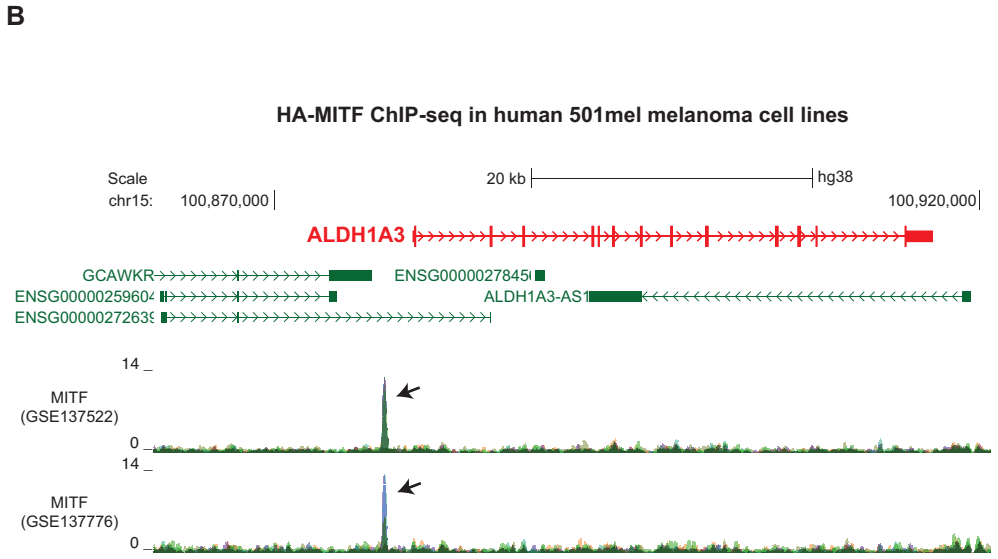


Lu et al., Supplementary Figure 5



Lu et al., Supplementary Figure 6





Lu et al., Supplementary Figure 7