Aldh2 dependent formaldehyde metabolism fuels purine demands in melanocyte stem cells

Hannah Brunsdon[1,2], Alessandro Brombin[1,2], Samuel Peterson[3], John H. Postlethwait[3], and E. Elizabeth Patton[1,2*]

1. MRC Human Genetics Unit, Institute of Genetics and Cancer, The University of Edinburgh, Western General Hospital Campus, Crewe Road, Edinburgh, EH4 2XU, UK.
2. Cancer Research UK Edinburgh Centre, Institute of Genetics and Cancer, The University of Edinburgh, Western General Hospital Campus, Crewe Road, Edinburgh, EH4 2XU, UK.
3. Institute of Neuroscience, University of Oregon, Eugene OR 97403 USA

* Corresponding author: e.patton@ed.ac.uk

ABSTRACT

Aldehyde-processing enzymes are viewed as essential clearing agents that rapidly deactivate harmful aldehydes. In the bone marrow, two specific enzymes, aldehyde dehydrogenase (ALDH) 2 and alcohol dehydrogenase (ADH) 5, were previously reported to protect hematopoietic stem cells from endogenous formaldehyde accumulation. Unexpectedly, we found that melanocyte stem cells (McSCs) in zebrafish depend on formate, an Aldh2-generated reaction product, to drive regeneration. Activated McSCs require Aldh2 (but not Adh5) to generate differentiated progeny, and by using scRNA-sequencing analysis, we identified a *de novo* purine biosynthesis program that is uniquely present in activated McSCs. Consistent with formate serving as one-carbon units for nucleotide biosynthesis, we found that purine supplementation (but not pyrimidine supplementation) was able to restore melanocyte regeneration in the absence of Aldh2. This work shows that Aldh2 enzymes generate reaction products that are needed to meet metabolic demands in regeneration.

MAIN

Biomolecules required for all cellular functions are continuously generated through anabolism, or biosynthesis. Essential for metabolism is disassembly of biomaterials into smaller building blocks, however this process can also generate reactive molecules that cause cellular damage. For instance, aldehydes, a class of highly reactive metabolite intermediates, can easily diffuse through cell membranes and form cytotoxic adducts and crosslinks with DNA, lipids and proteins, and are therefore subject to rapid clearing mechanisms[1-3]. As a first line of defence, widely expressed aldehyde-processing enzymes that metabolize aldehydes or aldehyde conjugates stand ready, including Aldehyde dehydrogenase (ALDH) enzymes, a family with 19 members, and Alcohol dehydrogenase (ADH) 5[4]. The importance of aldehyde detoxification in human biology is exemplified by the genetic

variants of *ALDH2* in the human population, such as the single nucleotide polymorphism r671 in *ALDH2* (c.1510G>A; p.E504K; ALDH2*2), which confers loss-of-function in 560 million people, mainly of East Asian origin[5, 6]. Carriers of the r671 *ALDH2* polymorphism can experience adverse reactions to acetaldehyde from exogenous alcohol consumption and are at risk for a range of diseases including osteoporosis, cardiovascular disease, neurodegeneration, and Fanconi Anemia[7-13].

Much of the toxicity from aldehydes can be attributed to metabolites such as acetaldehyde and formaldehyde that cause mutations and chromosomal rearrangements by direct damage to DNA[2, 3]. Recent work shows that a two-tier protection mechanism in cells defends against aldehyde-induced DNA crosslinks: first, aldehydes are cleared by enzymes, such as ALDH2 and ADH5, and second, replication-coupled DNA damage response pathways repair crosslinks and remove adducts[2, 3, 14-17]. This work emphasizes the nature of aldehyde toxicity and homeostatic clearance, primarily investigated in the hematopoietic stem cell (HSC) compartment. However, other work proposes more varied roles for aldehydes, namely that by-products generated by aldehyde-detoxification enzyme reactions also sustain essential downstream cellular metabolic processes[18-20]. For example, in mammalian cells, formate, generated by ADH5 metabolism of a formaldehyde-glutathione conjugate, can serve as a carbon source in the one-carbon (1C) cycle that supports nucleotide synthesis[19, 20]. In addition, one-carbon metabolism, compartmentalized within different cell types and organs, is becoming more broadly recognized as a physiological process impacting on cell states and associated with disease[21]. Thus, toxic aldehydes can be converted to metabolites essential for life. What is yet unknown is how the reaction products of aldehyde metabolism by ALDH2 and ADH5 contribute to physiological processes in specific cells and tissues in the context of a whole animal.

To learn how ALDH2 functions in stem cells other than HSCs and in an intact animal, we set out to study the zebrafish McSC population in melanocyte regeneration. During zebrafish embryonic development, melanocytes that originate directly from the neural crest generate lateral stripes along the body. In contrast, in adult animals, McSCs that reside at the dorsal root ganglion (DRG) niche

regenerate and form the adult stripes[22-27] **(Fig. 1a)**. McSCs are multi-potent, and give rise to glia and multiple pigment cell types to generate the adult pigmentation pattern[28, 29]. Recently, we identified a developmental *tfap2b*+ McSC population that we found to be distinct within neural crest and pigment cell lineages[29] and found *aldh2* gene paralogs were expressed in these cells **(Fig. S1)**. Here, we set out to investigate the function of Aldh2 in McSCs. To this end, we used the ALDH2 inhibitor (ALDH2i) CVT-10216 in a melanocyte regeneration assay that is dependent on a temperature sensitive allele (*mitfa*^{vc7}) of the master melanocyte transcription factor MITF [30]. In this model, fish embryos are grown at higher temperatures to deplete Mitfa activity, which prevents embryonic melanocyte development from the neural crest. When the water temperature is lowered to a level permissive for restoring Mitfa activity, melanocytes are regenerated from McSCs[30] **(Fig. 1a)**. In zebrafish embryos grown in the presence of CVT-10216, we did not detect any discernible effects on embryonic melanocyte development. However, melanocyte regeneration from McSCs was significantly delayed in ALDH2i-treated embryos, suggesting that Aldh2 primarily functions in McSCs **(Fig. 1b, S1).**

CVT-10216 is reported to have a >40-fold selectivity for ALDH2 over other ALDH enzymes[5], however, to confirm this specificity in zebrafish, we generated an *aldh2.1* - *aldh2.2* double mutant line by CRISPR-Cas9, henceforth referred to as *aldh2-/-*. The genetic similarity between these two paralogs made generating specific *aldh2* mutants difficult, so we created a double null mutant instead by designing guide RNAs to excise a large intergenic region between the tandem duplicate genes **(Fig. S1)**. In keeping with the ALDH2i experiments, *aldh2-/-* mutants generated embryonic melanocytes, yet were defective in melanocyte regeneration from the McSC compartment **(Fig. 1c)**. We noticed that after multiple rounds of breeding of our *aldh2-/-* mutants, the melanocyte regeneration phenotype was lessened, coupled with transcriptional upregulation of other *aldh* enzyme family members, suggesting some plasticity in *aldh* expression in regeneration (**Fig. S1)**. Therefore, we confirmed these results in *aldh2.1 and aldh2.2* paralog knockdown experiments with morpholino oligonucleotides **(Fig. S1).** We found that the embryonic melanocytes in *aldh2-/-* mutants were

defective for the dopaminergic camouflage response, a neuronally regulated innate behavior, reflecting the function for Aldh2 in dopamine metabolism[31]. This phenotype recapitulates our previous data with Daidzin, another ALDH2i, and provides confidence that the *aldh2-/-* mutants are defective for Aldh2 activity[32] **(Fig. S1)**.

To investigate whether Aldh2 activity impacts directly upon the McSCs, we employed a *Tg(mitfa:GFP)* transgenic line that was previously shown to mark McSCs[24, 29]. Following ALDH2i treatment in regenerating embryos, we observed a significant loss of GFP+ expression in McSCs at the niche (**Fig. 1d**). One interpretation of this result is that McSCs are depleted in the absence of Aldh2. Alternatively, McSCs may be present but expressing only low (or no) *mitfa:GFP* under conditions of ALDH2 inhibition.

In the earliest stages of embryonic development, McSCs that emerge from the neural crest maintain a neural crest identity at the niche, but lose this identity by day 3[24, 29]. Given our results in ALDH2i-treated regenerating embryos, we postulated that regenerative (activated) McSCs would re-express neural crest identity markers in addition to *mitfa*. To assess this hypothesis, we employed a double transgenic line *Tg(mitfa:GFP; crestin:mCherry)* in which *mCherry* is expressed from the promoter of *crestin*[29, 33], a neural crest gene, and applied this to a second, independent regeneration assay. In this assay, the pro-drug MoTP kills differentiated embryonic melanocytes, and melanocytes are regenerated from the McSC compartment[34]. Following MoTP washout, McSCs expressed both mCherry and GFP in control animals **(Fig. 1e)**. The intensity of GFP+ was heterogeneous between McSC clusters in control embryos, but all McSCs expressed mCherry indicating that McSCs re-express a neural crest identity in regeneration. Upon ALDH2i treatment, and as seen in **Fig. 1d**, we again observed a specific and strong reduction of GFP in McSCs. However, this time, mCherry+ McSCs were still clearly visible. Thus, McSCs re-express a neural crest identity during regeneration and require Aldh2 to increase expression of *mitfa* and generate melanoblasts.

Using live confocal imaging of McSCs to capture this process over time, we performed an MoTP regeneration assay and observed cells expressing high levels of *mitfa:GFP*+ emerging from McSCs and migrating dorsally in control embryos **(Fig. 1f; Movie 1)**. In contrast, the McSC niches in ALDH2i-treated embryos had little discernible cell movement, with very little *mitfa:GFP* expression **(Fig. 1f; Movie 2)**. Taken together, these data show that there are at least two distinct cell states within the regenerative McSC niche (*mitfa-low* and *mitfa-high*) and that Aldh2 is required for activated McSCs to increase *mitfa* expression and generate migratory progeny.

Thus far, we had visually captured activated McSCs (*crestin*+ *mitfa-low*) uncoupled from emerging progeny (*crestin*+ *mitfa-high*), and discovered a novel role for Aldh2 in this process. Next, we went on to investigate the transcriptional signatures of these cell populations by scRNA-sequencing to ascertain how they might be affected by Aldh2 deficiency. To this end, we designed a scRNAseq analysis of a MoTP melanocyte regeneration experiment in which double transgenic *mitfa:GFP; crestin:mCherry* embryos were treated with DMSO or CVT-10216 **(Fig. 2a)**. We identified 24 clusters of transcriptionally distinct cell populations by comparing the top 30 variably expressed genes, generating UMAPs featuring expression of known lineage-defining NC genes, and mapping the cluster identities from two recent zebrafish scRNA publications onto our data[35, 36] **(Fig. 2b, c; Fig. S2; Tables S1, 2)**.

As *crestin:mCherry* is expressed in a wide range of neural crest-derived cell populations[33], we captured both pigment cell lineages and cells of the neural lineage. The expression of *mitfa* and *dct* in clusters 7 and 11 suggested that these are late and early melanoblast (Mb) populations respectively. Cells in clusters 2, 6 and 12 expressed *crestin*, but low *mitfa*, and contained a mix of markers consistent with McSC identity[29]. *aldh2.2* and *aldh2.1* were expressed across multiple clusters, but were particularly enriched in regenerating pigment clusters including melanoblasts **(Fig. 2c)**. Relating the above cluster identities to our imaging analyses, we propose that the *crestin*+ *mitfa-low* McSCs are within clusters 2, 6 and 12 and that the *crestin*+ *mitfa-high* McSCs and progeny (and remaining

embryonic melanoblasts) are within clusters 7 and 11 **(Fig. 2d)**. The predicted cell cycle phase shows clusters 11 (*mitfa-high*) and 12 (*mitfa-low*) to be in S and G2/M, and may reflect the cycling McSCs we observe during regeneration **(Fig. 1f, 2d)**.

Next, we analysed the dataset by drug treatment condition. Overall, we found that Aldh2 inhibition did not substantially change cell or cluster identity **(Fig. 2b)**. However, the proportions of cells within some clusters differed significantly between treatment conditions **(Fig. 2e)**. Specifically, we detected a higher proportion of *crestin+ mitfa-low* cells (clusters 2,6,12), and a lower proportion of *crestin+ mitfa-high* cells (cluster 7) after ALDH2i. This population shift is consistent with our imaging experiments, in which we detected fewer *mitfa:GFP* expressing cells at the McSC niche **(Fig. 1d-f)**, and suggestive of a block in McSC differentiation.

To understand the physiological and mechanistic implications of the ALDH2-dependent *mitfa-high* to *mitfa-low* McSC transition, we performed differential expression analysis with the control dataset between *crestin+ mitfa-low* cells and *crestin+ mitfa-high* cells **(Table S3)**. Overall, *mitfa-high* cells (clusters 7,11) were enriched for pigmentation programs and melanoma-related terms, whereas *mitfa-low* cells (clusters 2,6,12) were enriched for essential metabolic pathways, including the 1 Carbon (THF) cycle, the TCA cycle, and *de novo* purine biosynthesis **(Fig. 2f)**, suggesting that regenerative McSCs acquire metabolic requirements distinct from those of melanoblasts.

Next, to understand why McSCs require ALDH2 activity to generate progeny, we performed differential expression analyses between controls and ALDH2i-treated cell populations **(Fig. 2g), Tables S4-6)**. Within the ALDH2i treated *crestin+ mitfa-low* cell population, *de novo* purine synthesis was again significantly upregulated **(Fig. 2g-i)**, suggesting that McSCs “blocked” by ALDH2i are starved for purines. Because we found no ALDH2i-dependent change in *de novo* purine synthesis or glucose metabolism genes in cells from either clusters 7,11 (melanoblast) or another pigment cell cluster requiring purine synthesis for pigmentation (cluster 9; iridophores)[37], this pattern was specific to *crestin+ mitfa-low* cells and not a general effect of drug treatment. Taken together, these analyses

support a mechanism in which regenerative McSCs require Aldh2 for metabolic rewiring to generate progeny.

Next, we wanted to identify the Aldh2 substrate in McSCs. We reasoned that aldehyde substrates in melanocyte regeneration would be toxic if supplied in excess, and that toxicity would increase in *aldh2-/-* mutant embryos. Hence, we screened known ALDH2 substrates for sensitivity in zebrafish development overall and specifically in the context of melanocyte regeneration **(Fig. 3a; Fig. S3)**. Unexpectedly, we found that *aldh2-/-* embryos were resistant to acetaldehyde and propionaldehyde, widely recognized ALDH2 substrates. In contrast, however, *aldh2-/-* embryos were sensitive to formaldehyde. Notably, low doses of exogenous formaldehyde (that had no other apparent effect on the fish) impaired melanocyte regeneration, and this response was significantly stronger in *aldh2-/-* zebrafish mutants **(Fig. 3b, Fig. S3)**. These data indicate that formaldehyde, but not other aldehydes, is an important Aldh2 substrate in the McSC compartment.

Recent studies show that ALDH2 and ADH5 function together to clear endogenous formaldehyde during HSC differentiation to prevent immune depletion in mouse and induced pluripotent stem cells (iPSCs), as well as in patients with biallelic *ALDH2* and *ADH5* mutations[38-40] **(Fig. 3c)**. Mice lacking both ALDH2 and ADH5 develop leukemia and have shorter lifespans, and in spite of active DNA repair, bone marrow-derived progenitors acquire a formaldehyde-associated mutation signature that resembles human cancer mutation signatures associated with aging[38]. To address if Adh5 can function in melanocyte regeneration and compensate for Aldh2, we generated an *adh5-/-* mutant line by CRISPR-Cas9 **(Fig. 3d)**. We found that the *adh5-/-* mutant was highly sensitive to exogenous formaldehyde treatment, indicating that, like in mammals, formaldehyde is an Adh5 substrate in zebrafish **(Fig. 3e)**. However, *adh5* loss had no effect on melanocyte regeneration, and did not enhance the regeneration defects in *aldh2-/-* mutants or ALDH2i-treated embryos **(Fig. 3f, g)**. Thus, despite the shared formaldehyde substrate and conservation across species, Aldh2 has a

unique function for formaldehyde metabolism in McSC differentiation, and Adh5 does not compensate for Aldh2 in this cell lineage.

One explanation for the Aldh2-deficient regeneration phenotype is that accumulation of endogenous formaldehyde causes McSC toxicity. However, we believe this to be unlikely given our experimental data; i) our observations while imaging over time showed no evidence of McSC loss, ii) following ALDH2i treatment, *crestin+ mitfa-low* McSCs were present in our scRNA-seq analysis, even at relatively higher numbers, and iii) the McSC block by ALDH2i treatment was reversible following washout **(Fig. S4)**. These findings led us to hypothesize that the reaction products of formaldehyde metabolism are required for McSCs differentiation but not for survival. To test this hypothesis, we performed a regeneration assay in CVT-10216 treated embryos in the presence or absence of formate, and found that formate supplementation fully restored melanocyte regeneration **(Fig. 3h)**. At the cellular level, formate even fully rescued *crestin+ mitfa-high* expression at the niche, whilst having no discernible effect on *crestin+ mitfa-low* cells **(Fig. 3i)**.

Formate is a carbon donor for the 1C cycle, and we found the McSC metabolic switch identified here reminiscent of cell state transitions reported for naïve to primed murine stem cells, that depend on 1C cycling and nucleotide biosynthesis[41], and formate overflow mechanisms that induce a metabolic shift from low to high adenine nucleotide levels in human cancer cell lines and mouse cancer models[42]. Taken together, our data suggest that regenerative McSCs depend on formate and 1C cycling to transition from a neural crest to a melanoblast cell state. To test this hypothesis, we used the dihydrofolate reductase inhibitor methotrexate (Mtx) to inhibit 1C metabolism **(Fig. 4a)**. Mtx had no effect on the embryonic melanocyte lineage but its inhibitor function is easy to validate in zebrafish embryos; wild-type embryos treated with Mtx lack pigmentation in xanthophores and iridophores, both of which require functional 1C metabolism for pigment synthesis[37] **(Fig. 4b, S4)**. In the McSC lineage, we found that Mtx treatment caused melanocyte regeneration defects that were significantly

exacerbated in *aldh2-/-* mutants **(Fig. 4c, d)**. These data indicate that zebrafish McSCs have metabolic requirements that require functional 1C metabolism.

Given the upregulation of *de novo* purine metabolism genes in McSCs, we next set out to examine purine nucleotide supplementation in regeneration. In the presence of ALDH2i, we found that exogenously provided purine nucleotides rescued the melanocyte regeneration defect in a dose-dependent manner **(Fig. 5a)**. This effect was not simply a consequence of providing embryos with an additional energy source in the form of ATP, because purine ribonucleosides were also capable of rescuing melanocyte regeneration **(Fig. 5b)**. However, pyrimidine supplementation did not rescue melanocyte regeneration, demonstrating that this effect does not reflect a general requirement for all nucleotides. Next, we explored the specificity of this rescue using confocal imaging, and found that purine, but not pyrimidine, supplementation selectively rescued *mitfa:GFP* expression at the McSC niche after ALDH2i treatment **(Fig. 5c, d)**. Hence, McSCs have a specific requirement for purine nucleotides to generate progeny **(Fig. 5e)**.

While all cells require nucleotides as fundamental building blocks, and for energy and signaling, the neural crest is especially sensitive to nucleotide depletion[43], which has direct metabolic consequences in rare disease and melanoma. For instance, patients with Miller syndrome, a rare genetic neurocristopathy affecting face and limb development, have mutations in dihydroorotate dehydrogenase (DHODH), the rate-limiting enzyme for pyrimidine *de novo* biosynthesis[43, 44]. In zebrafish, expression of a neural crest program defines melanoma initiation, and these cancers are sensitive to leflunomide, a DHODH inhibitor[33, 45]. Similarly, in mouse, a metabolic gene program driven by the transcription factor Yin Yang 1, a neural crest stem cell regulator, is essential for neural crest lineages, and its loss of function causes hypoplasia and prevents initiation of melanoma[46]. In these contexts, nucleotide levels may directly influence the transcriptional response, as we and others have shown for the neural crest and McSC[47, 48]. What we discovered here is that regenerative McSCs have a select requirement for purine nucleotides, findings that may point to purine nucleotide functions

unrelated to transcription or DNA replication. For instance, purine nucleotides have an ancient function as neurotransmitters that activate purinergic receptors, and as such can regulate neural stem and progenitor cells, and melanocyte-keratinocyte communication in human skin[49, 50]. Hence, purine nucleotides could facilitate McSC communication with DRG niche cells (of which we know very little) and with peripheral nerves that are used as migratory routes for melanoblast progenitors[23, 24]. Given that neural crest and McSCs programs re-emerge in melanoma[29, 33, 45, 46, 51] our findings may be relevant to understanding the metabolic reprogramming in melanomas, such as the dependency on formate metabolism during melanoma metastasis[52-54].

How stem cells generate progeny is a fundamental question in regenerative medicine. Here, we show that McSCs have a neural crest identity in regeneration and a metabolic demand for purines to generate progeny. This purine requirement is supplied by formate, the reaction product of ALDH2-dependent formaldehyde metabolism. Formaldehyde is abundant in the blood (>40 µM) and can arise from demethylation reactions from histones and nucleic acids[38, 39]. Based on our data in **Fig. 3**, we suggest that an as yet unknown, endogenous formaldehyde source is active in melanocyte regeneration. Mechanistically, our work identifies an unanticipated lineage-specific requirement for Aldh2 in both protection from genotoxicity and supply of essential metabolites in McSCs. This could mean that in individuals with mutations in ALDH2, both aldehyde cytotoxicity and depletion of aldehyde derived metabolites could result in the clinical disease features.

ACKNOWLEDGEMENTS

We are grateful to Cameron Wyatt and the IGC Zebrafish Facility for zebrafish management and husbandry, Elisabeth Freyer and the IGC FACS/10x facility, Ann Wheeler and the IGC Imaging Facility for supporting the imaging experiments, Richard Clarke at the Genetics Core ECRF for sequencing, and Jana Travnickova for sharing R code and expertise. JHP supported by US National Institutes of Health grants R01 OD011116 and R24 OD018555. EEP is funded by MRC HGU Program

(MC_UU_00007/9), the European Research Council (ZF-MEL-CHEMBIO-648489), and Melanoma Research Alliance (687306).

AUTHOR CONTRIBUTIONS

Conceptualization: EEP, HB; Methodology: HB; Software: HB; Validation: HB; Formal analysis: HB, AB; Investigation: HB, AB, SP, JHP; Resources: SP, JHP, EEP; Writing original draft: EEP, HB; Writing review and editing: HB, AB, JHP, EEP; Visualisation: HB, EEP; Supervision: JHP, EEP; Funding acquisition: JHP, EEP.

COMPETING INTERESTS STATEMENT

The authors declare no competing interests.

RESOURCES

Data and code availability

scRNA-seq experiment data have been submitted to GEO (GSE183868). A private access token is available for reviewers. Previously published sequencing data that was reanalyzed here are available from GEO: GSE131136[36], NCBI SRA: PRNJNA56410[35], and GEO: GSE178364[29] (in revision).

METHODS

Fish husbandry, fish lines

Zebrafish were maintained in accordance with UK Home Office regulations, UK Animals (Scientific Procedures) Act 1986, amended in 2013, and European Directive 2010/63/EU under project license 70/8000 and P8F7F7E52. All experiments were approved by the Home Office and AWERB (University of Edinburgh Ethics Committee). Fish stocks used were: wild-type AB, *mitfa*^{vc7} [30, 55], *Tg(mitfa:GFP)* [24], *Tg(crestin:mCherry)*[33], *aldh2-/-* (this study), and *adh5-/-* (this study). Combined transgenic and mutant lines were generated by crossing. Adult fish were maintained at ~28.5°C under 14:10 light-dark cycles. Embryos were kept at either 24°C, 28.5°C or 32°C and staged according to the reference table provided by Kimmel and colleagues [56].

Genotyping

Whole embryos or fin clips from adult fish were genotyped by resuspending tissue in DirectPCR® DNA-Tail solution (Viagen), and heating samples to 56°C for 2 hours, then 84°C for 20 minutes. Primers used for genotyping can be found in **Table S7**.

CRISPR-Cas9 mutant line generation

sgRNAs (**Table S7**) were synthesized using the EnGen® sgRNA Synthesis Kit, *S. pyogenes* (New England Biolabs) according to manufacturer's instructions. CRISPR-Cas9 knock-out lines were generated as previously described[57]. Briefly, 200 ng/µl sgRNAs targeting exon 3 of *aldh2.1*

(GCCAGAGATGCCTTTAAGCT) and exon 3 of *aldh2.2* (GCCAGAGATGCCTTTAAGCT) were co-injected with Cas9 mRNA into zebrafish embryos at the 1 cell stage. An allele was recovered which was the result of a large deletion between *aldh2.1* and *aldh2.2*, creating a gene fusion and single base-pair insertion at the fusion site. This introduced an adjacent frameshift mutation and premature stop codon. 200 ng/µl sgRNA targeting exon 3 of *adh5* (CTCAGTGGAAGTGACCCCGAG) was co-injected with recombinant 300 ng/µl Cas9 protein (SBI). These F0 fish were raised to adulthood, and outcrossed with WT fish to obtain progeny that were screened for presence of indels through PCR amplification of a 600bp region surrounding the target site, and digestion of the amplicon using T7 endonuclease (New England Biolabs). Outcrossed F1 fish that contained a 25bp deletion were isolated and raised to adulthood.

Morpholino injection

Standard control morpholinos and translation blocking morpholinos were sourced from Genetools LLC, based off previously published sequences for *aldh2.1* (ZDB-MRPHLNO-120517-2*)* and *aldh2.2* (ZDB-MRPHLNO-120517-3)[58]. 2-6 ng of each morpholino was injected into sibling *mitfa*^{vc7} embryos at the 1-2 cell stage.

Imaging

Images of embryos immobilized with MS:222 and 1.5% LMP agarose were acquired using a 0 20X/0.75 lens on the multimodal Imaging Platform Dragonfly (Andor technologies, Belfast UK) equipped with 405, 445, 488, 514, 561, 640 and 680nm lasers built on a Nikon Eclipse Ti-E inverted microscope body with Perfect focus system (Nikon Instruments, Japan). Data were collected in Spinning Disk 40 µm pinhole mode on the Zyla 4.2 sCMOS camera using a Bin of 1 and no frame averaging using Andor Fusion acquisition software. Z stacks were collected using the Nikon TiE focus drive. Multiple positions were collected using a Sigma-Koki Stage (Nikon Instruments Japan). Data

were visualized and analyzed using Imaris (Oxford Instruments, v. 9.7.0) or Image J Fiji software (v. 1.53c).

Whole zebrafish embryos fixed in 4% PFA/PBST were imaged with a Leica MZFLIII fluorescence stereo microscope with a 1x objective fitted with a Qimaging Retiga Exi CCD camera (Qimaging, Surrey, BC, Canada). Image capture was performed using Micromanager (Version 1.4).

To quantify the area of *GFP* or *mCherry*-expressing cells within niches, homozygous *Tg(mitfa:GFP)* fish were outcrossed with non-fluorescent fish to obtain embryos with similar levels of transgene expression. The McSC compartment was imaged at the same magnification, within the same anatomical area, and with consistent laser power and other imaging settings between individual samples and biological replicates. In Fiji, a maximum projection Z-stack of images was cropped to only include McSC compartment cells (typically containing 6-7 compartments per image) and converted to a binary image. Consistent threshold settings were applied, and the total GFP+ area measured in pixels^2 and divided by the number of somites visible in the field of view.

Melanocyte regeneration assays

If using the *mitfa*^{vc7} regeneration model line, embryos were kept in a 32°C incubator from 0-72hpf to repress the developmental melanocyte lineage. Embryos were then moved to a 24°C incubator to allow regeneration over a period of 48 hours. When using chemical methods for regeneration, 150 μM 4-(4-Morpholinobutylthio)phenol (MoTP) (Sigma) was added to embryos kept at 28.5°C from 24hpf onwards. MoTP was washed out to allow regeneration between 72 and 120 hpf. After fixation, embryos were imaged and melanocytes counted within a set region with the Image J CellCounter plugin.

Camouflage response assays

The camouflage response assay was performed as described previously[32]. 5 dpf wild type or *aldh2-/-* mutant embryos were placed in a dark place for 15 minutes to standardize their light exposure. These embryos were split into cohorts which were either placed under a lamp or kept in the dark for 1.5

hours. The embryos were then moved to the opposite light condition for a further 45 minutes, during which time melanin dispersed or contracted depending on light exposure. This was repeated once or twice more when assessing the embryos ability to learn to adapt to changing light conditions. Afterwards, embryos were then briefly anaesthetized in MS-222 and fixed in 4% PFA. Embryos were imaged dorsally at a fixed magnification. Melanin coverage was measured with Image J Fiji, by outlining a predetermined region of the head, converting the image to an 8-bit binary image with a uniform threshold, and then measuring the area of black pixels.

Small molecule inhibitor and rescue experiments

Unless otherwise stated, 10 µM CVT-12016 (Sigma-Aldrich) or equimolar Dimethyl Sulphoxide DMSO (Sigma-Aldrich) was added to embryos at 24hpf after manual or pronase-assisted (Sigma) dechorionation and refreshed every 24 hours. Embryos were arrayed in 6-well tissue culture plates with 10-15 embryos per well. For formate supplementation assays, 25 µM sodium formate (Sigma) was added. For nucleotide supplementation assays, 400 µM of AMP, UMP, GMP, IMP or TMP were added to embryos, or 200 µM of dA, dG, dU or T. 4-HNE (range of concentrations in ethanol) (Sigma) and Mtx (Sigma) (range of concentrations in DMSO) were added at 24hpf and refreshed every 24 hours.

Aldehyde treatments

Stock solutions of fresh acetaldehyde and formaldehyde were made in a fume hood just before use. Various aldehyde concentrations were added to embryos kept in screw cap centrifuge tubes to limit aldehyde evaporation, and embryos scored for survival after 48 hours.

RNA extraction and RT-qPCR

Samples to be processed for RT-qPCR were collected at the required stage and frozen on dry ice. RNA was extracted from frozen tissues with the Qiagen RNeasy Mini kit according to manufacturer's instructions. RNA was quantified and quality checked using a Nanodrop 2000c (Thermo Scientific).

500 µg of RNA was used as input for Reverse Transcription using Superscript™ III reverse transcriptase (Invitrogen) and an Oligo$(DT)_{15}$ primer (Promega). RT-qPCR was performed with Sybr Green® Lightcycler Green I Master mix (Roche), using a Lightcycler 480 instrument and associated software. *β-actin* was used as a housekeeping control (**Table S7**). Gene expression fold changes were found using the delta-delta ct method.

Single cell sequencing experimental setup and sequencing

24hpf *Tg(mitfa:GFP; crestin:mCherry)* and were divided into groups of ~500 embryos and treated with MoTP, and co-treated with either 10 µM CVT-10216 or equimolar DMSO. At 64hpf, MoTP was washed out, and embryos left to regenerate for 8 hours. Embryos were anaesthetized in MS-222 and trunks dissected, and a cell suspension of each treatment condition obtained as previously described[59]. Samples were sorted on a FACSAria2 SORP instrument (BD Biosciences UK) as previously described[29] but stage-matched non-fluorescent AB embryos also treated with MoTP used as a control to enable gating of *mCherry* and *GFP* fluorescence. 10,000 fluorescent cells were collected in 100ul of 0.04% BSA/PBS in LoBind tubes. Single cell libraries were prepared using the Chromium Single Cell 3' GEM, Library & Gel Bead Kit v3 (10x Genomics).

The samples were sequenced on a Nextseq 2000 using a P2 flow cell on a 100 cycle run. ~2.97M reads passed quality filters for CVT-10216 treated, and ~1.87M reads for DMSO-treated, however due to the greater number of cells processed in the CVT-10216 sample, the mean reads per cell were fairly equal (37,405-CVT vs 34,832-DMSO).

Bioinformatics analyses

Aldh2.2 expression within developmental melanocytes was visualized using the recent Brombin et al. GEO #: GSE178364 scRNA-seq dataset[29].

For this study, FASTQ files were generated using CellRanger (v.3.1.0, 10x Genomics) mkfastq function with default settings and -qc parameter and aligned to the zebrafish STAR genome index

using gene annotations from Ensembl GRCz11 release 94 with manually annotated entries for *GFP* and *mCherry*. Libraries were aggregated (CellRanger aggr pipeline) to generate a gene-barcode matrix. Gene matrices (13360 total, DMSO-5,394, CVT-7966), barcodes and features were uploaded to R (v. 4.0.5) and standard quality control filtering performed with the Scater package (v. 1.18.6)[60]. Only cells with total features >1000, $\log_{10}$ total counts >3.0 and mitochondrial gene counts <20% were considered as high quality and kept for further analysis (DMSO-4488 CVT-6795). The dimensionality of the combined dataset was visualized with Elbow and JackStraw plots before running linear dimensional reduction. Louvain clustering was then performed using the FindNeighbors and FindClusters functions (dims=50, resolution=0.5) in Seurat (v. 4.0.3)[61]. Data were projected onto 2D spaces using Uniform Manifold Approximation and Projection (UMAP) using the same dimensions as above. Cluster-specific genes were identified using the FindAllMarkers and FindMarkers function in Seurat with default parameters (**Table S1,2**).

Cluster calling was done after detection of published marker genes for specific cell types and by making unbiased pairwise comparisons based on gene overdispersion against published datasets GEO #: GSE131136[36] and NCBI SRA #: PRNJNA56410[35] using the scMap package (v.1.12.0)[62] and between the datasets presented in this paper. Plots were generated either using Seurat or ggplot2 (v.3.3.5)[63]. Prediction of cell cycle phase was performed with Seurat, using canonical cell cycle markers described in Tirosh et al,[64].

For DE analyses, scRNA-seq data were first corrected for zero-inflated counts by using the ZINB-WaVE package (v. 1.12.0) with default parameters[65]. Then, the DEseq2 package (v. 1.30.1)[66] was used to generate genelists of significantly (p.adj < 0.05) upregulated and downregulated genes (raw data in **Tables S3-6**). Pathway analyses were performed using the clusterProfiler package (v. 3.18.1)[67], using GO, KEGG, Reactome[68] or Literature-based genelists[69,70]. GSEA analysis was performed using GSEA software (v. 4.1.0) with genelists generated through DeSeq2, using the "RunGSEAPreranked" function.

Statistics

Statistical details of experiments and n numbers can be found in figure legends. Statistics and plots were generated using GraphPad Prism 7 (v. 7.0e) and R. Unless otherwise stated, experiments were replicated at least three times (N=3 biological replicates), with 10-15 embryos per condition.

REFERENCES

1. O'Brien, P.J., Siraki, A.G. & Shangari, N. Aldehyde sources, metabolism, molecular toxicity mechanisms, and possible effects on human health. *Crit Rev Toxicol* **35**, 609-662 (2005).
2. Garaycoechea, J.I. *et al.* Alcohol and endogenous aldehydes damage chromosomes and mutate stem cells. *Nature* **553**, 171-177 (2018).
3. Pontel, L.B. *et al.* Endogenous Formaldehyde Is a Hematopoietic Stem Cell Genotoxin and Metabolic Carcinogen. *Mol Cell* **60**, 177-188 (2015).
4. Jackson, B. *et al.* Update on the aldehyde dehydrogenase gene (ALDH) superfamily. *Hum Genomics* **5**, 283-303 (2011).
5. Chen, C.H., Ferreira, J.C., Gross, E.R. & Mochly-Rosen, D. Targeting aldehyde dehydrogenase 2: new therapeutic opportunities. *Physiol Rev* **94**, 1-34 (2014).
6. Chen, C.H. *et al.* Novel and prevalent non-East Asian ALDH2 variants; Implications for global susceptibility to aldehydes' toxicity. *EBioMedicine* **55**, 102753 (2020).
7. Harada, S., Agarwal, D.P. & Goedde, H.W. Aldehyde dehydrogenase deficiency as cause of facial flushing reaction to alcohol in Japanese. *Lancet* **2**, 982 (1981).
8. Brooks, P.J., Enoch, M.A., Goldman, D., Li, T.K. & Yokoyama, A. The alcohol flushing response: an unrecognized risk factor for esophageal cancer from alcohol consumption. *PLoS Med* **6**, e50 (2009).
9. Masaoka, H. *et al.* Aldehyde dehydrogenase 2 (ALDH2) and alcohol dehydrogenase 1B (ADH1B) polymorphisms exacerbate bladder cancer risk associated with alcohol drinking: gene-environment interaction. *Carcinogenesis* **37**, 583-588 (2016).
10. Matsuo, K. *et al.* The aldehyde dehydrogenase 2 (ALDH2) Glu504Lys polymorphism interacts with alcohol drinking in the risk of stomach cancer. *Carcinogenesis* **34**, 1510-1515 (2013).
11. Chang, J.S., Hsiao, J.R. & Chen, C.H. ALDH2 polymorphism and alcohol-related cancers in Asians: a public health perspective. *J Biomed Sci* **24**, 19 (2017).
12. Takeuchi, F. *et al.* Genome-wide association study of coronary artery disease in the Japanese. *Eur J Hum Genet* **20**, 333-340 (2012).
13. Hiura, Y. *et al.* A genome-wide association study of hypertension-related phenotypes in a Japanese population. *Circ J* **74**, 2353-2359 (2010).
14. Garaycoechea, J.I. *et al.* Genotoxic consequences of endogenous aldehydes on mouse haematopoietic stem cell function. *Nature* **489**, 571-575 (2012).
15. Hodskinson, M.R. *et al.* Alcohol-derived DNA crosslinks are repaired by two distinct mechanisms. *Nature* **579**, 603-608 (2020).
16. Langevin, F., Crossan, G.P., Rosado, I.V., Arends, M.J. & Patel, K.J. Fancd2 counteracts the toxic effects of naturally produced aldehydes in mice. *Nature* **475**, 53-58 (2011).
17. Rosado, I.V., Langevin, F., Crossan, G.P., Takata, M. & Patel, K.J. Formaldehyde catabolism is essential in cells deficient for the Fanconi anemia DNA-repair pathway. *Nat Struct Mol Biol* **18**, 1432-1434 (2011).
18. Jacobson, M.K. & Bernofsky, C. Mitochondrial acetaldehyde dehydrogenase from Saccharomyces cerevisiae. *Biochim Biophys Acta* **350**, 277-291 (1974).
19. Burgos-Barragan, G. *et al.* Mammals divert endogenous genotoxic formaldehyde into one-carbon metabolism. *Nature* **548**, 549-554 (2017).
20. Bae, S., Chon, J., Field, M.S. & Stover, P.J. Alcohol Dehydrogenase 5 Is a Source of Formate for De Novo Purine Biosynthesis in HepG2 Cells. *J Nutr* **147**, 499-505 (2017).
21. Ducker, G.S. & Rabinowitz, J.D. One-Carbon Metabolism in Health and Disease. *Cell Metab* **25**, 27-42 (2017).

22. Budi, E.H., Patterson, L.B. & Parichy, D.M. Embryonic requirements for ErbB signaling in neural crest development and adult pigment pattern formation. *Development* **135**, 2603-2614 (2008).

23. Budi, E.H., Patterson, L.B. & Parichy, D.M. Post-embryonic nerve-associated precursors to adult pigment cells: genetic requirements and dynamics of morphogenesis and differentiation. *PLoS Genet* **7**, e1002044 (2011).

24. Dooley, C.M., Mongera, A., Walderich, B. & Nusslein-Volhard, C. On the embryonic origin of adult melanophores: the role of ErbB and Kit signalling in establishing melanophore stem cells in zebrafish. *Development* **140**, 1003-1013 (2013).

25. Hultman, K.A. *et al.* Defects in ErbB-dependent establishment of adult melanocyte stem cells reveal independent origins for embryonic and regeneration melanocytes. *PLoS Genet* **5**, e1000544 (2009).

26. Irion, U. & Nusslein-Volhard, C. The identification of genes involved in the evolution of color patterns in fish. *Curr Opin Genet Dev* **57**, 31-38 (2019).

27. Kelsh, R.N. & Barsh, G.S. A nervous origin for fish stripes. *PLoS Genet* **7**, e1002081 (2011).

28. Singh, A.P. *et al.* Pigment Cell Progenitors in Zebrafish Remain Multipotent through Metamorphosis. *Dev Cell* **38**, 316-330 (2016).

29. Brombin, A. *et al.* Tfap2b specifies an embryonic melanocyte stem cell that retains adult multi-fate potential. *bioRxiv*, 2021.2006.2018.448859 (2021).

30. Johnson, S.L., Nguyen, A.N. & Lister, J.A. mitfa is required at multiple stages of melanocyte differentiation but not to establish the melanocyte stem cell. *Dev Biol* **350**, 405-413 (2011).

31. Yao, L. *et al.* Inhibition of aldehyde dehydrogenase-2 suppresses cocaine seeking by generating THP, a cocaine use-dependent inhibitor of dopamine synthesis. *Nat Med* **16**, 1024-1028 (2010).

32. Zhou, L. *et al.* ALDH2 mediates 5-nitrofuran activity in multiple species. *Chem Biol* **19**, 883-892 (2012).

33. Kaufman, C.K. *et al.* A zebrafish melanoma model reveals emergence of neural crest identity during melanoma initiation. *Science* **351**, aad2197 (2016).

34. Yang, C.T. & Johnson, S.L. Small molecule-induced ablation and subsequent regeneration of larval zebrafish melanocytes. *Development* **133**, 3563-3573 (2006).

35. Farnsworth, D.R., Saunders, L.M. & Miller, A.C. A single-cell transcriptome atlas for zebrafish development. *Dev Biol* **459**, 100-108 (2020).

36. Saunders, L.M. *et al.* Thyroid hormone regulates distinct paths to maturation in pigment cell lineages. *Elife* **8** (2019).

37. Ng, A., Uribe, R.A., Yieh, L., Nuckels, R. & Gross, J.M. Zebrafish mutations in gart and paics identify crucial roles for de novo purine synthesis in vertebrate pigmentation and ocular development. *Development* **136**, 2601-2611 (2009).

38. Dingler, F.A. *et al.* Two Aldehyde Clearance Systems Are Essential to Prevent Lethal Formaldehyde Accumulation in Mice and Humans. *Mol Cell* **80**, 996-1012 e1019 (2020).

39. Mu, A. *et al.* Analysis of disease model iPSCs derived from patients with a novel Fanconi anemia-like IBMFS ADH5/ALDH2 deficiency. *Blood* **137**, 2021-2032 (2021).

40. Shen, X. *et al.* A Surge of DNA Damage Links Transcriptional Reprogramming and Hematopoietic Deficit in Fanconi Anemia. *Mol Cell* **80**, 1013-1024 e1016 (2020).

41. Chandrasekaran, S. *et al.* Comprehensive Mapping of Pluripotent Stem Cell Metabolism Using Dynamic Genome-Scale Network Modeling. *Cell Rep* **21**, 2965-2977 (2017).

42. Oizel, K. *et al.* Formate induces a metabolic switch in nucleotide and energy metabolism. *Cell Death Dis* **11**, 310 (2020).

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

44. Ng, S.B. *et al.* Exome sequencing identifies the cause of a mendelian disorder. *Nat Genet* **42**, 30-35 (2010).

45. White, R.M. *et al.* DHODH modulates transcriptional elongation in the neural crest and melanoma. *Nature* **471**, 518-522 (2011).

46. Varum, S. *et al.* Yin Yang 1 Orchestrates a Metabolic Program Required for Both Neural Crest Development and Melanoma Formation. *Cell Stem Cell* **24**, 637-653 e639 (2019).

47. Johansson, J.A. *et al.* PRL3-DDX21 Transcriptional Control of Endolysosomal Genes Restricts Melanocyte Stem Cell Differentiation. *Dev Cell* **54**, 317-332 e319 (2020).

48. Santoriello, C. *et al.* RNA helicase DDX21 mediates nucleotide stress responses in neural crest and melanoma cells. *Nat Cell Biol* **22**, 372-379 (2020).

49. Lee, E.J. *et al.* Critical Role of ATP-P2X7 Axis in UV-Induced Melanogenesis. *J Invest Dermatol* **139**, 1554-1563 e1556 (2019).

50. Ulrich, H., Abbracchio, M.P. & Burnstock, G. Extrinsic purinergic regulation of neural stem/progenitor cells: implications for CNS development and repair. *Stem Cell Rev Rep* **8**, 755-767 (2012).

51. Marie, K.L. *et al.* Melanoblast transcriptome analysis reveals pathways promoting melanoma metastasis. *Nat Commun* **11**, 333 (2020).

52. Fischer, G.M. *et al.* Metabolic strategies of melanoma cells: Mechanisms, interactions with the tumor microenvironment, and therapeutic implications. *Pigment Cell Melanoma Res* **31**, 11-30 (2018).

53. Patton, E.E. *et al.* Melanoma models for the next generation of therapies. *Cancer Cell* **39**, 610-631 (2021).

54. Piskounova, E. *et al.* Oxidative stress inhibits distant metastasis by human melanoma cells. *Nature* **527**, 186-191 (2015).

55. Zeng, Z., Johnson, S.L., Lister, J.A. & Patton, E.E. Temperature-sensitive splicing of mitfa by an intron mutation in zebrafish. *Pigment Cell Melanoma Res* **28**, 229-232 (2015).

56. Kimmel, C.B., Ballard, W.W., Kimmel, S.R., Ullmann, B. & Schilling, T.F. Stages of embryonic development of the zebrafish. *Dev Dyn* **203**, 253-310 (1995).

57. Sorlien, E.L., Witucki, M.A. & Ogas, J. Efficient Production and Identification of CRISPR/Cas9-generated Gene Knockouts in the Model System Danio rerio. *J Vis Exp* (2018).

58. Ma, A.C., Chung, M.I., Liang, R. & Leung, A.Y. A DEAB-sensitive aldehyde dehydrogenase regulates hematopoietic stem and progenitor cells development during primitive hematopoiesis in zebrafish embryos. *Leukemia* **24**, 2090-2099 (2010).

59. Manoli, M. & Driever, W. Fluorescence-activated cell sorting (FACS) of fluorescently tagged cells from zebrafish larvae for RNA isolation. *Cold Spring Harb Protoc* **2012** (2012).

60. McCarthy, D.J., Campbell, K.R., Lun, A.T. & Wills, Q.F. Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics* **33**, 1179-1186 (2017).

61. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587 e3529 (2021).

62. Kiselev, V.Y., Yiu, A. & Hemberg, M. scmap: projection of single-cell RNA-seq data across data sets. *Nat Methods* **15**, 359-362 (2018).

63. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. (Springer-Verlag New York, 2016).

64. Tirosh, I. *et al.* Single-cell RNA-seq supports a developmental hierarchy in human oligodendroglioma. *Nature* **539**, 309-313 (2016).

65. Risso, D., Perraudeau, F., Gribkova, S., Dudoit, S. & Vert, J.P. Publisher Correction: A general and flexible method for signal extraction from single-cell RNA-seq data. *Nat Commun* **10**, 646 (2019).

66. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).

67. Yu, G., Wang, L.G., Han, Y. & He, Q.Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284-287 (2012).

68. Yu, G. & He, Q.Y. ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. *Mol Biosyst* **12**, 477-479 (2016).

69. Martens, M. *et al.* WikiPathways: connecting communities. *Nucleic Acids Res* **49**, D613-D621 (2021).

70. Travnickova, J. *et al.* Zebrafish MITF-Low Melanoma Subtype Models Reveal Transcriptional Subclusters and MITF-Independent Residual Disease. *Cancer Res* **79**, 5769-5784 (2019).

FIGURE LEGENDS

Figure 1: **Aldh2 is required for regenerative McSCs to generate progeny**

a. Schematic of the melanocyte lineages in zebrafish development. Embryonic melanocytes and McSCs both develop directly from the neural crest but McSCs go on to generate late-stage, regenerative, and adult melanocytes. On the right, confocal Z-stack depicting McSCs expressing *mitfa:GFP* located at the dorsal root ganglia (DRG) and melanoblasts (Mb) on the motor neurons. Neural tube and DRG are marked by *nbt:dsRed* expression.

b. ALDH2 inhibitor (CVT-10216) delays melanocyte regeneration. Representative images of wild type embryos treated with or without CVT-10216 during development (embryonic melanocytes) or in an McSC regeneration assay. Each data point represents an embryo. Scale bar = 500 μm. **** $p<0.0001$. One-way ANOVA with Tukey's multiple comparisons test. N=4.

c. *aldh2-/-* mutants are defective in melanocyte regeneration. Schematic of CRISPR-Cas9 strategy to target *aldh2.1* and *aldh2.2* with excision sites between scissor symbols **(see Fig. S1)**. Wild type or *aldh2-/-* embryos in normal development, or a McSC regeneration assay. **** $p<0.0001$. Unpaired two-tailed t-test performed to calculate statistical significance. N=3.

d. ALDH2 inhibitor (CVT-10216) causes loss of *mitfa:GFP* expression in McSCs. Representative confocal stack images of *Tg(mitfa:GFP);mitfa*^{vc7} McSCs at the niche after 24 hours regeneration with or without CVT-10216 treatment. The average *mitfa:GFP* niche area μm^2/somite (as depicted in **Fig 1a**) was quantified per embryo (one data point) and plotted on a box plot. McSC fluorescence was lost in CVT-10216 treated embryos, yet dorsal stripe epithelial (epi) *GFP*+ melanoblasts remained. Scale bars are 50 μm, N=3 with >5 embryos imaged per repeat.**** $p<0.0001$. Unpaired, two-tailed t-test.

e. McSCs maintain neural crest identity when treated with ALDH2 inhibitor (CVT-10216). Confocal stack images of McSC niches in CVT-10216 treated embryos after 6 hours washout of MoTP. McSCs with very low to no GFP signal are arrowed. N=2, >5 embryos used per condition, representative images are shown. Scale bars = 50 μm

f. Time lapse stills of individual regenerating McSCs at the niches. *Tg(mitfa:GFP; crestin:mCherry)* embryos with or without CVT-10216 were imaged 2 hours post-MoTP washout. In a control embryo, an McSC undergoes cell division and a new *mitfa:GFP-high* cell migrates upwards towards the epidermis (see **Movie 1**). In a CVT-10216 treated embryo, *mitfa:GFP* expression is not visible, and migration is not observed (see **Movie 2**). Scale bars = 20 μm.

Figure 2: Regenerative McSCs express a *de novo* purine nucleotide transcriptional program

a. Experimental design for the scRNA-seq experiment to capture the McSCs in regeneration.

b. UMAPs of *Tg(crestin:mCherry, mitfa:GFP)* positive cells after clustering, split by drug treatment. Mb = melanoblasts, Xanth = xanthophores, Irid = Iridophores, Mb/I = mixed melanoblast/iridophore.

c. UMAPs of both DMSO and CVT-10216 treated cells with color change from grey (negative) to purple based on $\log_2$ expression of *aldh2.1* and *aldh2.2* in pigment lineages and *crestin* (neural crest), *tfec* (melanophore/iridophore progenitors), *mitfa* (early melanoblasts) and *dct* (late melanoblasts).

d. Proposed relation of imaged McSCs to scRNA-seq clusters, using an example niche from **Fig. 1f** (scale bar 20 μm). We predict *crestin+ mitfa-high* cells (green arrow/box) are represented in clusters 7, 11, and *crestin+ mitfa-low* cells (magenta arrow/box) in clusters 2,6,12. UMAPs of these clusters (top) and their predicted cell cycle phase (bottom) are shown.

e. The proportion of total cells within each cluster compared between treatment conditions. The $\log_{10}$ percentage difference of numbers of cells in the CVT-10216 treated clusters compared to DMSO equivalents was plotted in a bar chart, with solid colour filled bars indicating a significant difference in proportions (Chi squared test).

f. Dot-plot of pathway analysis showing selection of significantly upregulated GO (G), KEGG (K), Reactome (R) and Literature-based (L) terms in clusters 2,6,12 compared to 7,11, and vice versa. Dot size represents observed/expected ratio, and colour adjusted p-value (Benjamini–Hochberg test).

g. As **f**, but significant enrichment of pathways in CVT-10216 treated cells relative to DMSO from clusters 2,6,12 (*crestin+ mitfa-low)*, clusters 7,11 (*crestin+ mitfa-high)*, and cluster 9 (predicted iridophores).

h. Enrichment plot of *de novo* purine biosynthesis signature upregulated in clusters 2,6,12 in CVT-10216 treated embryos compared to control, (NES -1.18, FDR <25%) Kolganov Smirnov test. Leading edge genes are listed.

i. Schematic diagram of *de novo* purine biosynthesis, with genes encoding enzymes significantly upregulated in the CVT-10216 dataset from **g** and **h** shown in red.

Figure 3: Selective requirement for Aldh2, but not Adh5, in formaldehyde metabolism during melanocyte regeneration

a. Table of known Aldh2 substrates and their effects on *aldh2-/-* embryos (See **Fig. S3**).

b. Melanocyte regeneration is sensitive to formaldehyde and this effect is stronger in *adlh2-/-* mutants. Images and quantification of melanocytes in zebrafish embryos in a *mitfa*^{vc7} regeneration assay. Melanocyte counts were normalised to the mean of respective control, each dot represents a single embryo. ** p<0.0021, **** p<0.0001. Ordinary One-way ANOVA with Tukey's multiple comparisons. N=3.

c. Schematic diagram of formaldehyde metabolism by Adh5 (cytosol) and Aldh2 (mitochondria).

d. Schematic diagram showing *adh5-/-* CRISPR-Cas9 mutant, with sgRNA target site in exon 3 and alignment to WT sequence showing a deletion of 25bp.

e. Sensitivity of *adh5-/-* embryos to increasing concentrations of formaldehyde from 24hpf for 24 hours, and surviving embryos quantified. N=5. ***p<0.0002 **** p<0.0001 Two-way ANOVA with Sidak's multiple comparisons test, error bars indicate SE.

f. MoTP regeneration assay on *aldh2-/-*, *adh5-/-* mutant embryos and embryos from an incross of *adh5+/-;aldh2-/-* fish (embryos genotyped after counting). *** p<0.0002, **** p<0.0001, ns not significant.

g. MoTP Regeneration assay on wild type and *adh5-/-* mutants treated +/- CVT-10216. N=3. *** p<0.0002, ** p<0.0021, ns: not significant. One way ANOVA with Tukey's multiple comparisons.

h. Formate rescues the ALDH2i melanocyte regeneration phenotype. Representative images of a regeneration assay where control or CVT-10216-treated embryos were supplemented with 25 μM sodium formate. P value: **** p<0.0001, ns= not significant. Kruskall-Wallis test with Dunn's multiple comparisons. N=3.

i. Formate rescues the McSC differentiation deficiency. An MoTP assay on *Tg(mitfa:GFP;crestin:mCherry)* embryos treated with +/- CVT-10216 +/- 25 μM sodium formate from 24hpf. MoTP was washed out at 72hpf, and embryos imaged confocally at 74hpf. N=2, >5 embryos imaged per replicate. Scale bars are 25 μm. Single channel images of *crestin:mCherry* expression (magenta) and *mitfa:GFP* expression (green) are shown alongside merged images.

Figure 4: The McSC lineage is sensitive to disruption of the 1-Carbon cycle

a. Schematic of 1C metabolism and proposed function for Aldh2 supply of formate through formaldehyde metabolism (based on Burgos-Barragan et al 2017[19]). Tetrahydrofolate (THF) combines with formate to make 10-formyl-THF, which directly provides two carbons to make purine nucleosides (inosine, adenosine, and guanosine) which can then be converted into nucleotides such as AMP. Serine also donates carbon units to THF as a methyl group to make 5-10,methylene-THF, which in turn donates carbon for DNA/histone methylation, as well in pyrimidine synthesis.

b. Mtx treatment has no effect on embryonic melanocytes. Zebrafish embryos (wildtype and *aldh2-/-*) treated with or without Mtx at 24 hpf for 48 hr. N=3.

c. Melanocyte regeneration is sensitive to disruption of the 1C cycle. Representative images of control and *aldh2-/-* mutants +/- Mtx treatment in a *mitfa*^{vc7} regeneration assay are shown. To compare regeneration counts between genotypes, the melanocyte count at each dose was normalised to its respective genotype DMSO control. Each dot represents a single embryo. *** $p<0.0002$, **** $p<0.0001$. Ordinary One-way ANOVA performed with Tukey's multiple comparisons test. N=3.

d. McSCs are sensitive to disruption of the 1C cycle. Confocal Z-stacks of *mitfa:GFP* McSCs in a *mitfa*^{vc7} regeneration assay, in control or *aldh2-/-* embryos treated with or without Mtx. Scale bars are 50 μm. N= 2 biological repeats, with >5 embryos imaged per repeat. Quantification of GFP+ niche area/somite of embryos treated with Mtx is shown. *** $p<0.0002$, **** $p<0.0001$. Ordinary One-way ANOVA with Tukey's multiple comparisons.

Figure 5: Purine, but not pyrimidine, nucleotides rescue Aldh2 requirements in regeneration

a. Purine nucleotides rescue Aldh2 deficient melanocyte regeneration. Melanocyte regeneration assay in *mitfa*^{vc7} embryos +/- CVT-10216 plus increasing concentrations of purine nucleotides (IMP,GMP and AMP cocktail). N=3. Melanocyte counts are normalized to the untreated control. Each dot represents a single embryo. Error bars represent SE. ** $p<0.0021$, *** $p<0.0002$, **** $p<0.0001$, ns= not significant. Ordinary One-way ANOVA performed with Tukey's multiple comparisons test with a single pooled variance.

b. Purine, but not pyrimidine nucleosides, rescue Aldh2 deficient melanocyte regeneration. Melanocyte regeneration assay on *mitfa*^{vc7} embryos +/- CVT-10216 and supplemented with deoxyadenosine (dA), deoxguanosine (dG) purine nucleosides, or deoxyuridine (dU) or thymidine (T) pyrimidine nucleosides (200 µM) N=3. **** $p<0.0001$, ns= not significant. Error bars represent SE. Ordinary One-way ANOVA performed with Tukey's multiple comparisons test with a single pooled variance.

c. Purine nucleotides rescue McSC differentiation in ALDH2 inhibitor treated embryos. Representative confocal Z-stacks of *Tg(mitfa:GFP;crestin:mCherry)* embryos treated with MoTP +/- CVT-12016, as well as 400 µM AMP/GMP purine nucleotides, or 400 µM UMP or Thymidine pyrimidine nucleotides.

d. Quantification of *crestin:mCherry* and *mitfa:GFP* niche areas from **c.** Each dot represents the sum of the GFP niche area/ number of somites in view in one embryo. ****:$p<0.0001$, ns: not significant. Error bars represent SE. Ordinary One-way ANOVA performed with Tukey's multiple comparisons test with a single pooled variance.

e. Proposed model for Aldh2 control of the McSC lineage. Regenerating McSCs start expressing *crestin* and low levels of *mitfa*. Next, McSCs increase their metabolic demands for purine nucelotides to express high levels of *mitfa* and generate progeny. This metabolic demand is met by Aldh2, that metabolises endogenous formaldehyde into formate, which is then used in the 1C cycle to fuel the production of purine nucleotides. McSCs undergo cell division to generate progeny, which migrate away from the niche to the epidermis. ALDH2i (CVT-10216) delays the progression of the activated McSC to generate progeny in regeneration.

MOVIES

Movie 1: **Time-lapse video of regenerative McSCs**

Time-lapse video of a *Tg(crestin:mCherry;mitfa:GFP)* embryo during McSC regeneration (DMSO control). Embryos were treated with MoTP to kill differentiated melanocytes and initiate melanocyte regeneration. McSCs were followed for over 14 hours (post MoTP washout). McSCs are *crestin+ mitfa-low*, but then during or shortly after cell division, a cell strongly expresses GFP+ to become a *mitfa-high* cell, which then leaves the McSC compartment and migrates towards the epidermis. McSCs generate progeny.

Movie 2: **Time-lapse video regenerative McSCs with ALDH2i.**

Time-lapse video of a *Tg(crestin:mCherry;mitfa:GFP)* embryo during McSC regeneration in the presence of ALDH2i. Embryos were treated with CVT-10216 to inhibit Aldh2, and co-treated with MoTP to kill differentiated melanocytes, and initiate melanocyte regeneration. McSCs were followed for over 14 hours (post MoTP washout, but in the presence of ALDH2i). McSCs do not generate progeny.

SUPPLEMENTARY DATA

Supplementary Figures

Figure S1: *aldh2* expression in the McSC lineage, and generation of an *aldh2-/-* mutant line

Figure S2: Identification of transcriptionally distinct scRNA-seq clusters during McSC regeneration

Figure S3: Results of aldehyde screen on *aldh2-/-* mutant zebrafish

Figure S4: Recovery of melanocyte regeneration following ALDH2i removal

Supplementary Tables

Table S1: scRNA-seq: top 30 cluster markers

Table S2: scRNA-seq: metrics, clustering information and cell states

Table S3: Differential expression analysis of *crestin+ mitfa-low* vs *crestin+ mitfa-high* cells

Table S4: Differential expression analysis of *crestin+ mitfa-low* cells, DMSO vs CVT-10216 treated

Table S5: Differential expression analysis of *crestin+ mitfa-high* cells, DMSO vs CVT-10216 treated

Table S6: Differential expression analysis of iridophore cluster 9, DMSO vs CVT-10216 treated

Table S7: Oligonucleotide sequences

SUPPLEMENTARY FIGURE LEGENDS

Figure S1: Aldh2 expression in the McSC lineage, and generation of an *aldh2-/-* mutant line

a. UMAP of scRNA-seq data from Brombin et al,[29] with McSCs in red. UMAPs of these isolated McSCs showing $\log_2$ expression of specific genes, *tfap2b, mitfa*, *aldh1* and *aldh2* genes with color change from grey (negative) to purple.

b. Phylogenetic tree showing the relationship between human *ALDH* and zebrafish *aldh* genes from the ALDH1 and ALDH2 families.

c. sgRNAs targeting exon 3 of *aldh2.1* and *aldh2.2* were co-injected with Cas9. This caused deletion of the intergenic region and creation of a fusion transcript, with a base insertion and frame shift generating a premature stop codon (PTC). To confirm this deletion, RT-qPCR was performed with primers targeting Primer site 1, which should persist in the truncated fusion transcript, and Primer site 2 within the intergenic region. Relative expression of mutant fusion transcript in *aldh2-/-* mutants is shown relative to WT, and normalized to *β-actin*. N=3 biological replicates. Error bars represent SE.

d. A camouflage response assay on WT or *aldh2-/-* embryos. Representative images are shown of embryos after adaptation to dark or light surroundings. Melanin coverage within the red outlined area was quantified. N=3 biological replicates. Error bars represent SD. P values indicated. Ordinary One-way ANOVA performed with Tukey's multiple comparisons test.

e. RT-qPCR showing quantification of wild type *aldh2.2* expression levels relative to other zebrafish *aldh* genes in 72hpf WT or *aldh2-/-* mutant embryos. *ß-actin* was used as a housekeeping control. N=3 biological replicates. Error bars show SE. Asterisks mark *aldh* genes upregulated >1.5-fold in *aldh2-/-* mutants compared to wild type.

f. Regeneration assay on *mitfa*^{vc7} embryos injected with 6 ng of standard control morpholino (CMO), or morpholinos against *aldh2.1, aldh2.2* or combination. Regenerated melanocytes are quantified. Each dot represents a single embryo. N=3 biological replicates. * $p<0.0332$, **** $p<0.0001$, ns: not significant. One-way ANOVA performed with Tukey's multiple comparisons test.

Figure S2: Identification of transcriptionally distinct scRNA-seq clusters during McSC regeneration

a. Heatmap showing top 5 cluster-defining genes per selected clusters.

b. UMAP of the combined dataset showing gene expression of non-pigment clusters marked by *pou3f1* (Schwann Cell Progenitors), *mcamb* and *elavl3* marking NC-derived neural, *mvp17* marking iridophores, and *aox5* marking xanthophores.

c. UMAP of this scRNA seq data mapped with cell identity annotation from Farnsworth et al (2020) and Saunders et al (2019).

Figure S3: Results of aldehyde screen on *aldh2-/-* mutant zebrafish

a. Representative images of 72hpf WT and *aldh2-/-* mutant embryos treated with 0.75 mM acetaldehyde with or without CVT-10216. Unexpectedly, Aldh2 loss or deficiency confers resistance to acetaldehyde. N=3.

b. Representative images of 72hpf WT and *aldh2-/-* mutant embryos treated with 25 µM 4-HNE, showing whole body sensitivity in the mutant.

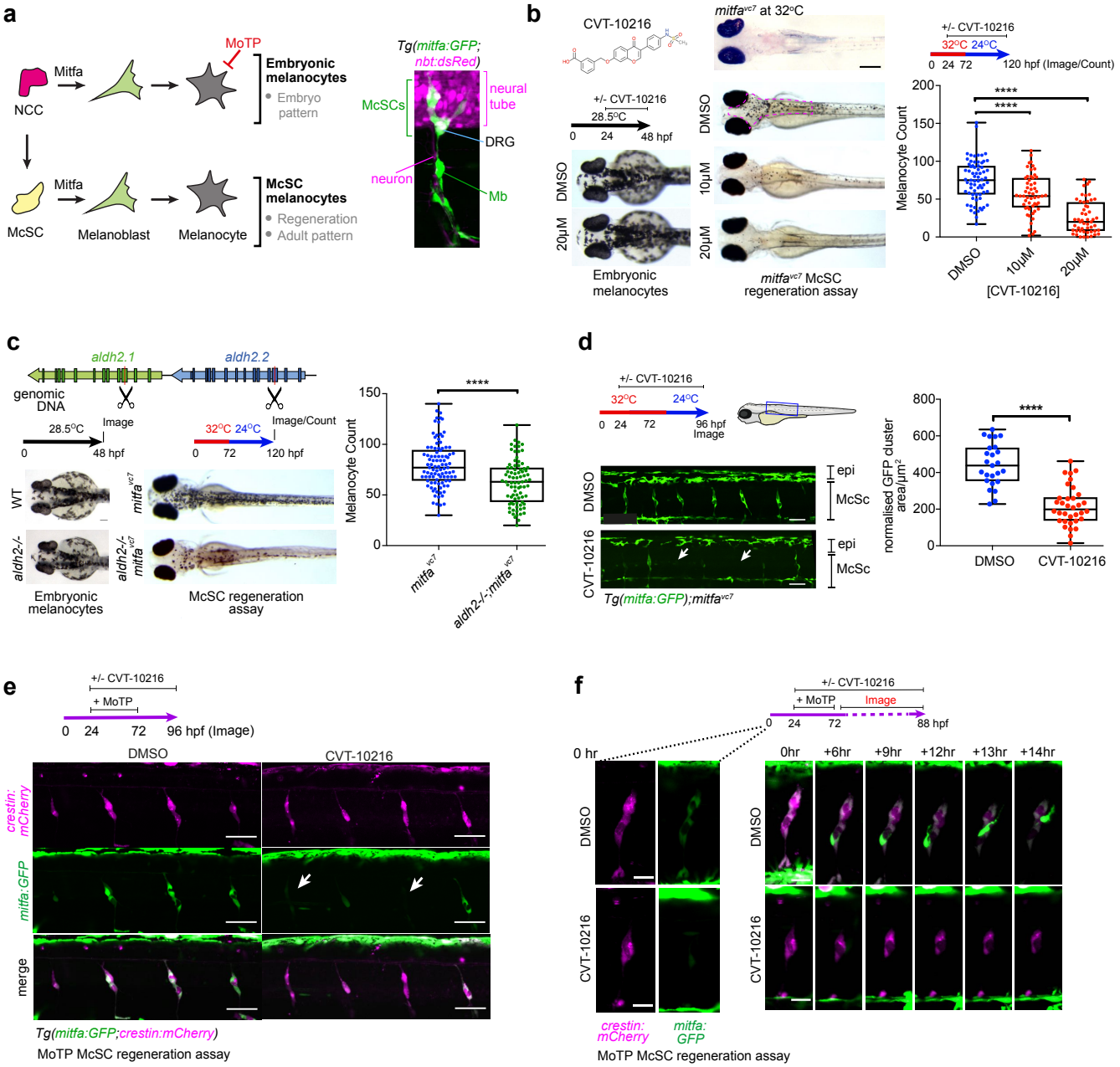
c. *mitfa*^{vc7} melanocyte regeneration assay and subsequent quantification of embryos treated with increasing doses of 4-HNE, showing no significant difference between controls and *aldh2-/-; mitfa*^{vc7} embryos in terms of reduction in regeneration potential after 4-HNE treatment. For comparison between genotypes, melanocyte numbers were normalized to the average untreated condition for each genotype. Each dot represents a single embryo. N=2. ** $p<0.0021$, **** $p<0.0001$, ns not significant. One way ANOVA with Tukey's multiple comparisons.

d. Survival percentage of WT and *aldh2-/-* mutant embryos treated with various concentrations of formaldehyde. N=6. Error bars are SE of the mean. ** $p<0.0021$. A two-way ANOVA with Sidak's multiple comparisons test.

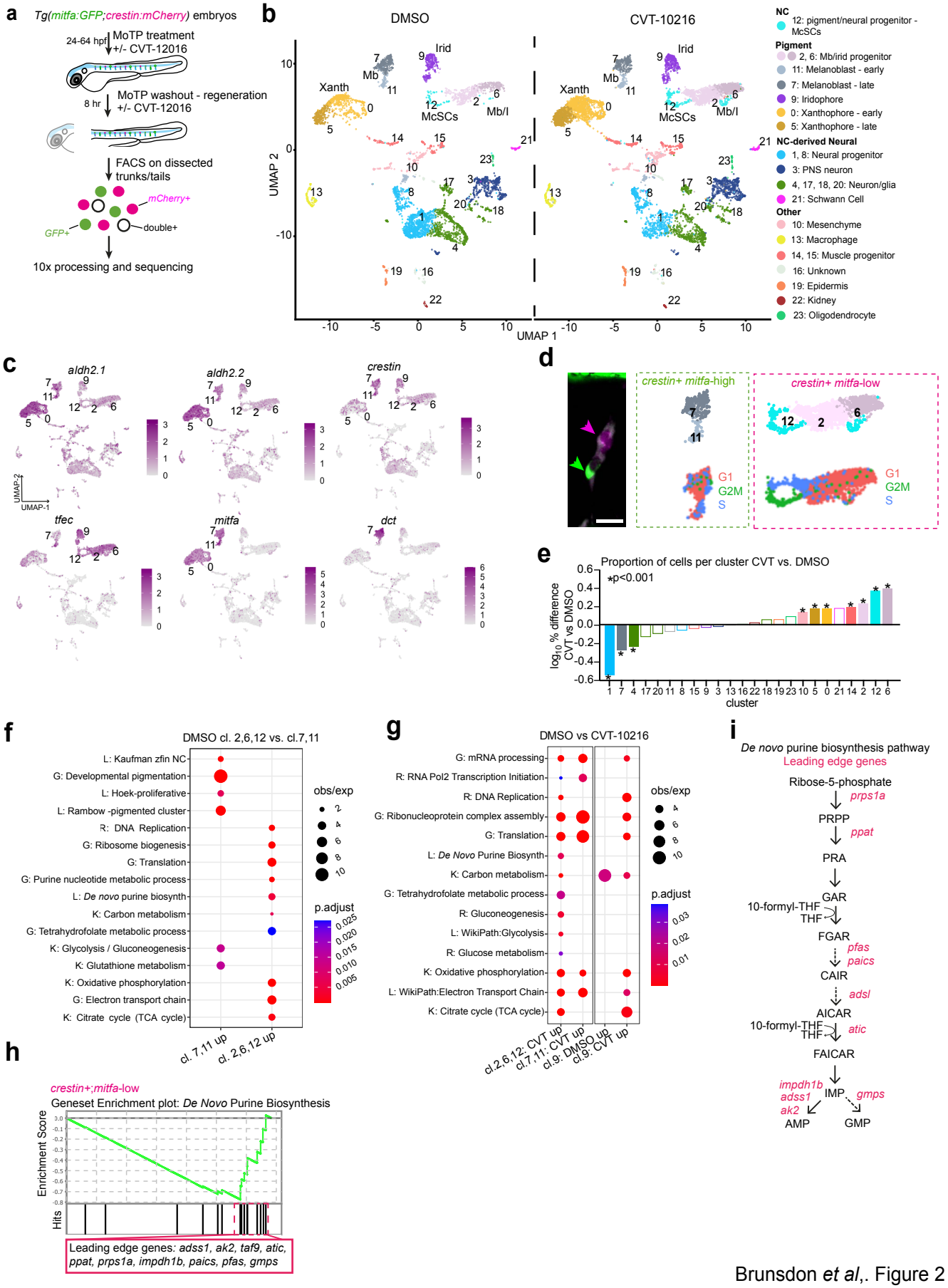
Figure S4: Recovery of melanocyte regeneration following ALDH2i removal

a. Extended regeneration assay on *mitfa*^{vc7} embryos treated with CVT-10216 from 24-120hpf. After washout, larvae were imaged at a number of time points to monitor recovery/continuation of melanocyte regeneration. Representative images are shown from 5 embryos per condition.

b. Mtx treated embryos (96 hpf) have lost reflective iridophore pigments (clearly observed in the eye) and yellow xanthophore pigments.



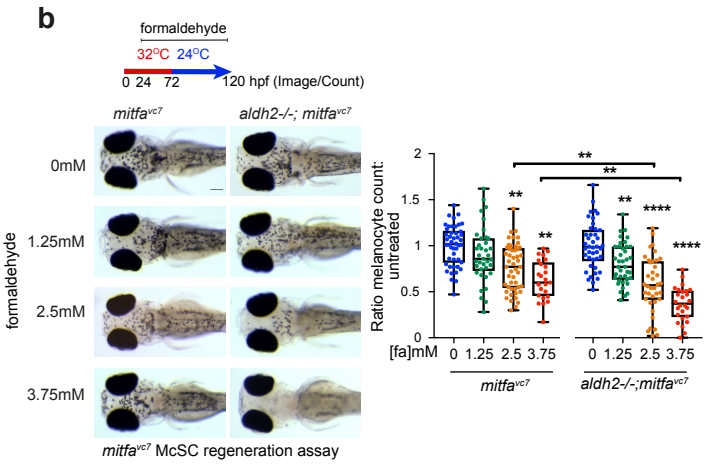
Brunsdon *et al*,. Figure 1

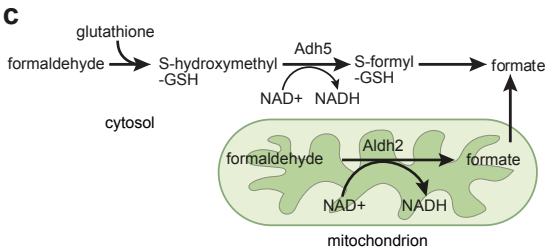


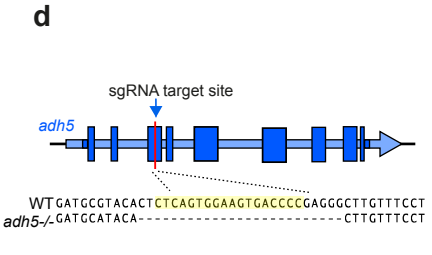
Brunsdon *et al*,. Figure 2

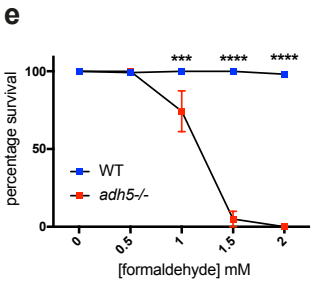
a

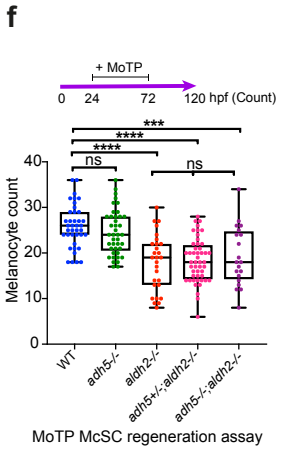
	ALDH2 substrate → product	aldh2-/- embryo phenotype
Short chain aldehydes	Formaldehyde → Formate	High dose: whole embryo sensitivity Low dose: McSC regeneration delay
	Acetaldehyde → Acetate	Whole embryo resistance
	Propionaldehyde → Propionate	Whole embryo resistance
Dopamine metabolism	DOPAL → DOPAC	Defective dopaminergic camouflage response
Oxidative stress	Malondialdehyde → Malonic acid semialdehyde	Not tested
	4-HNE → 4-HNA	Whole embryo sensitivity

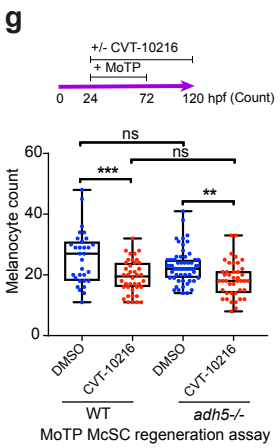


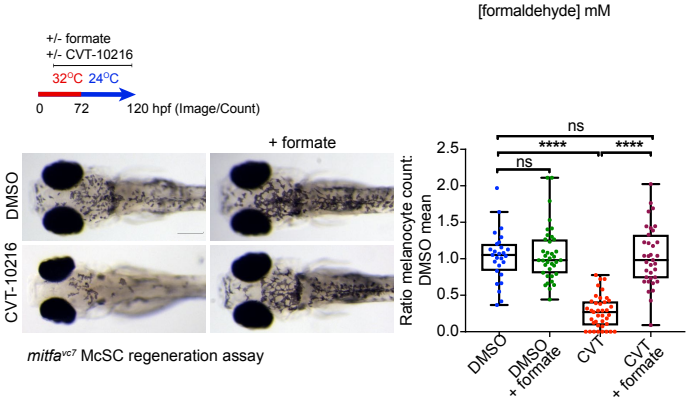


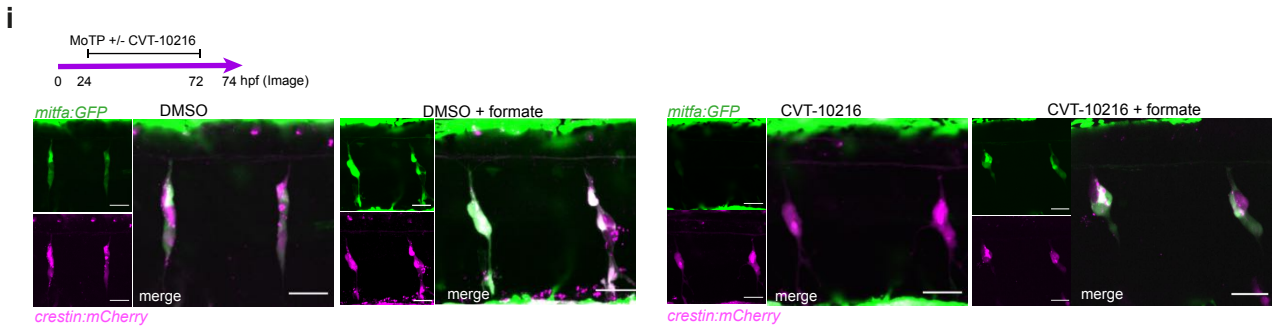




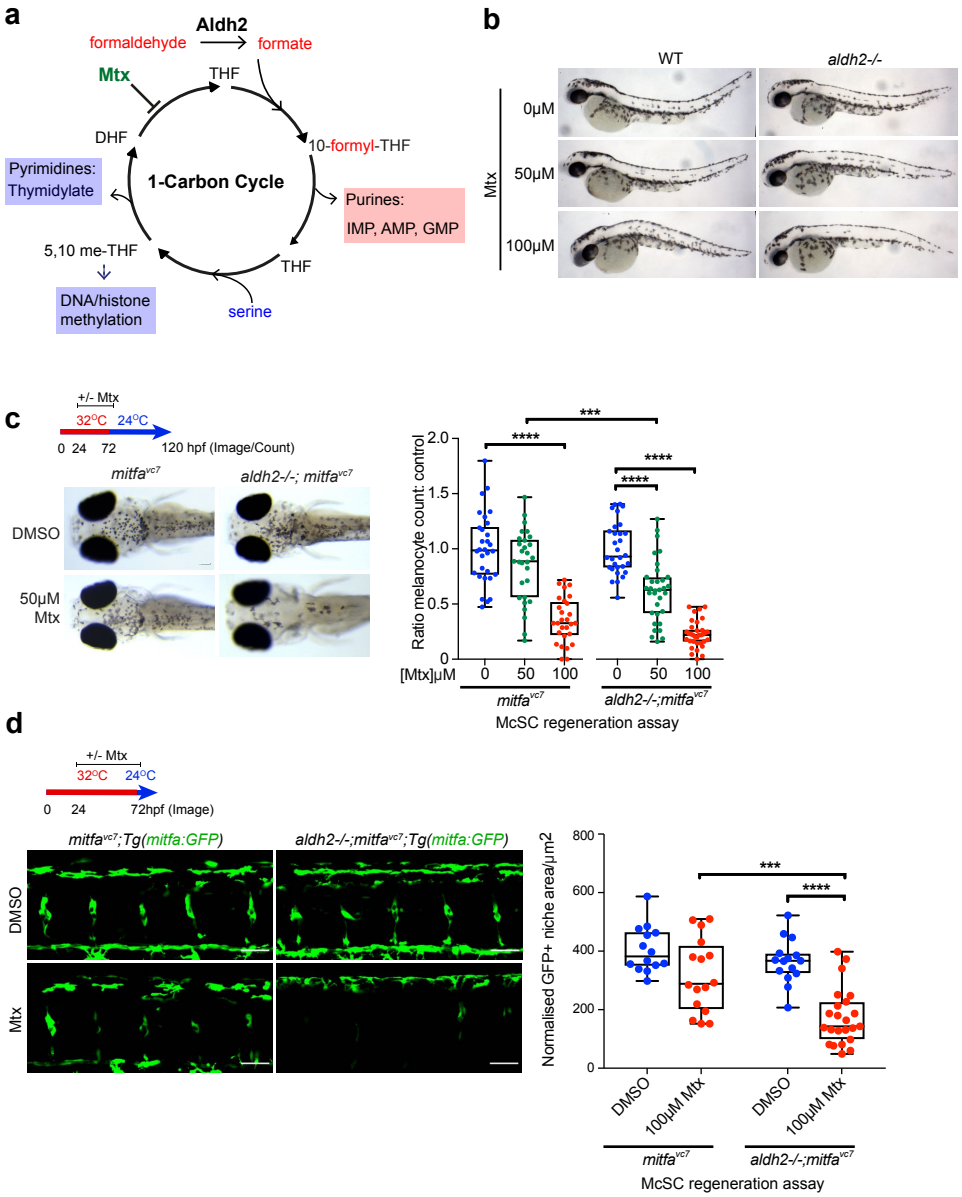




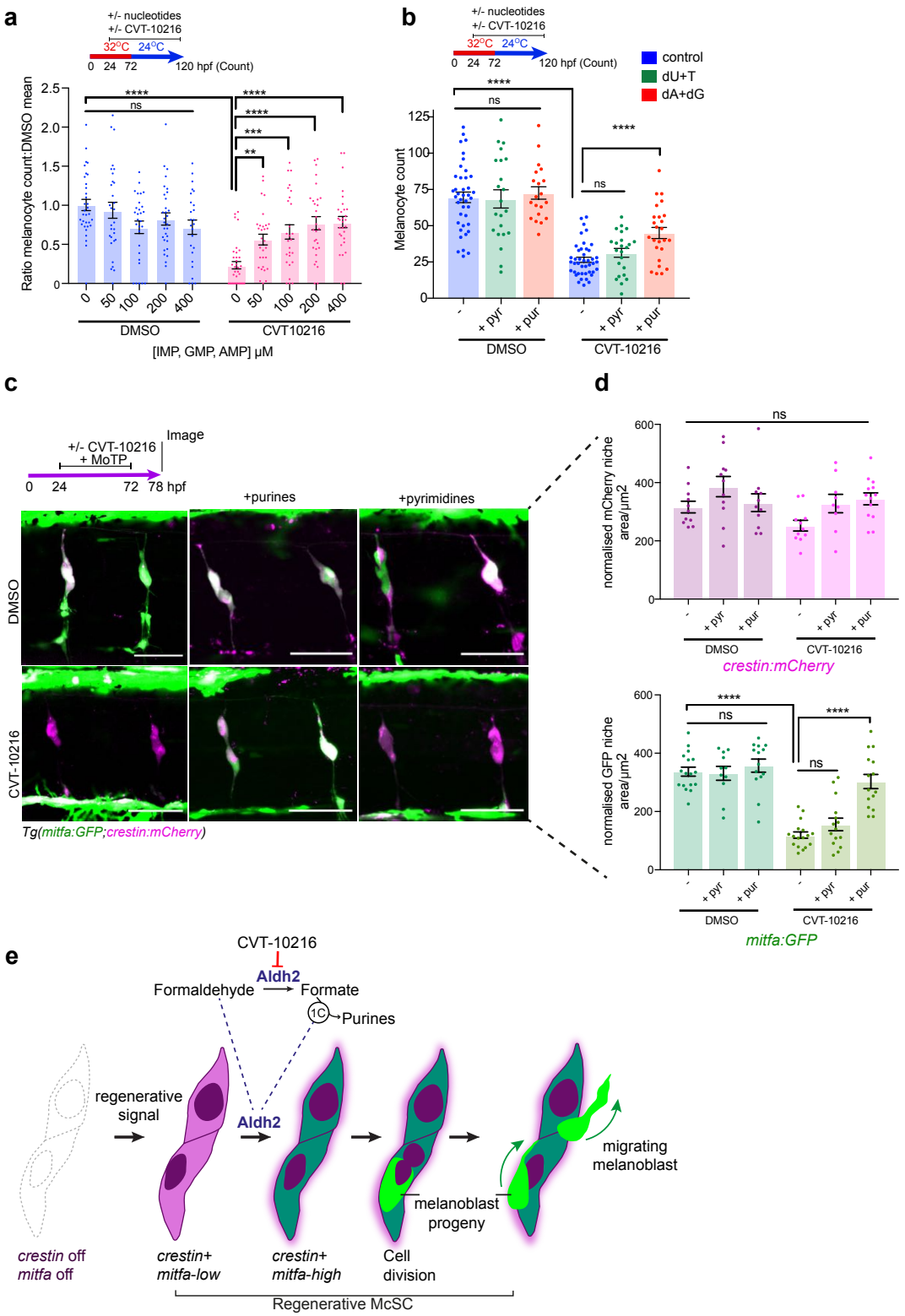




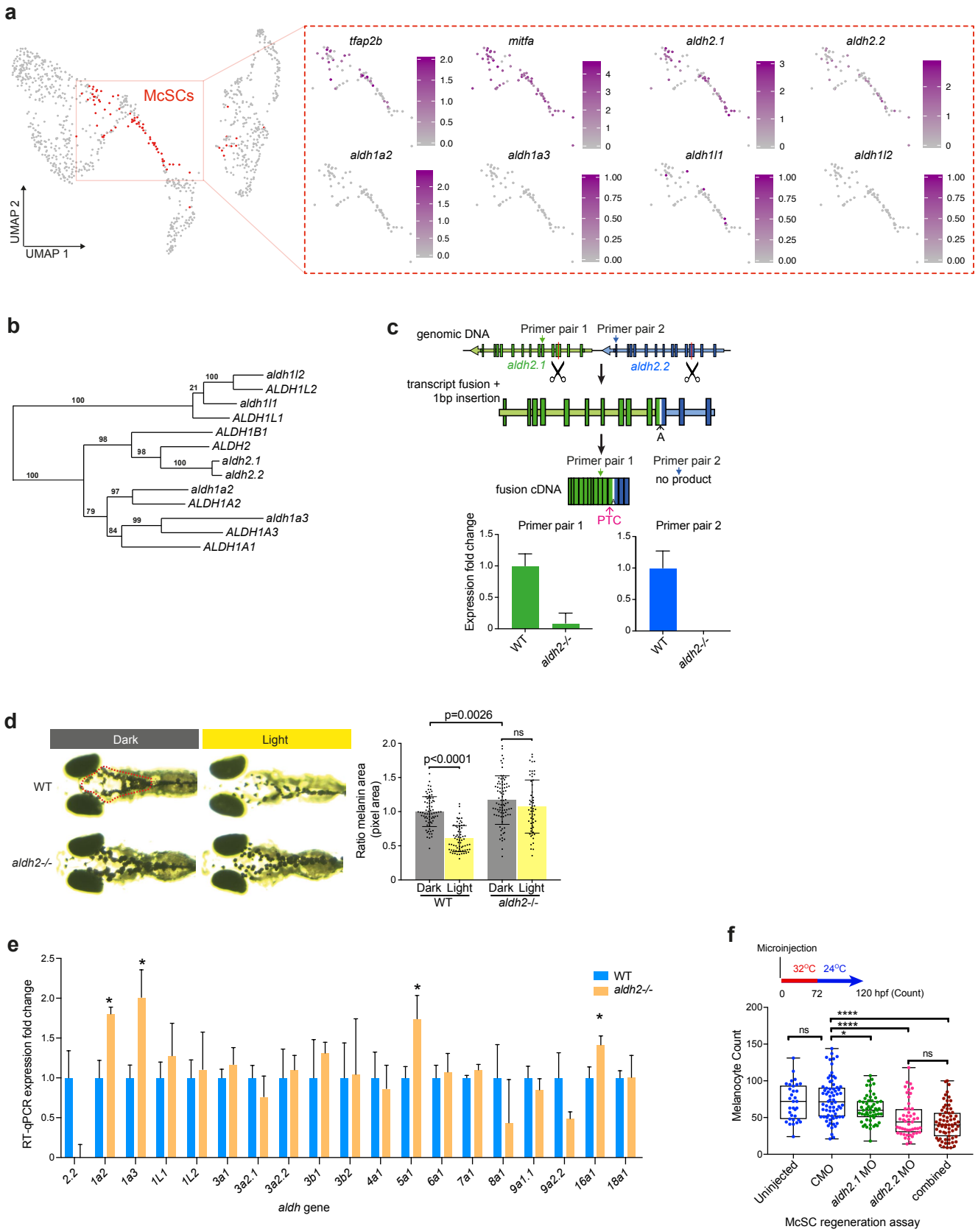
Brunsdon *et al*,. Figure 3



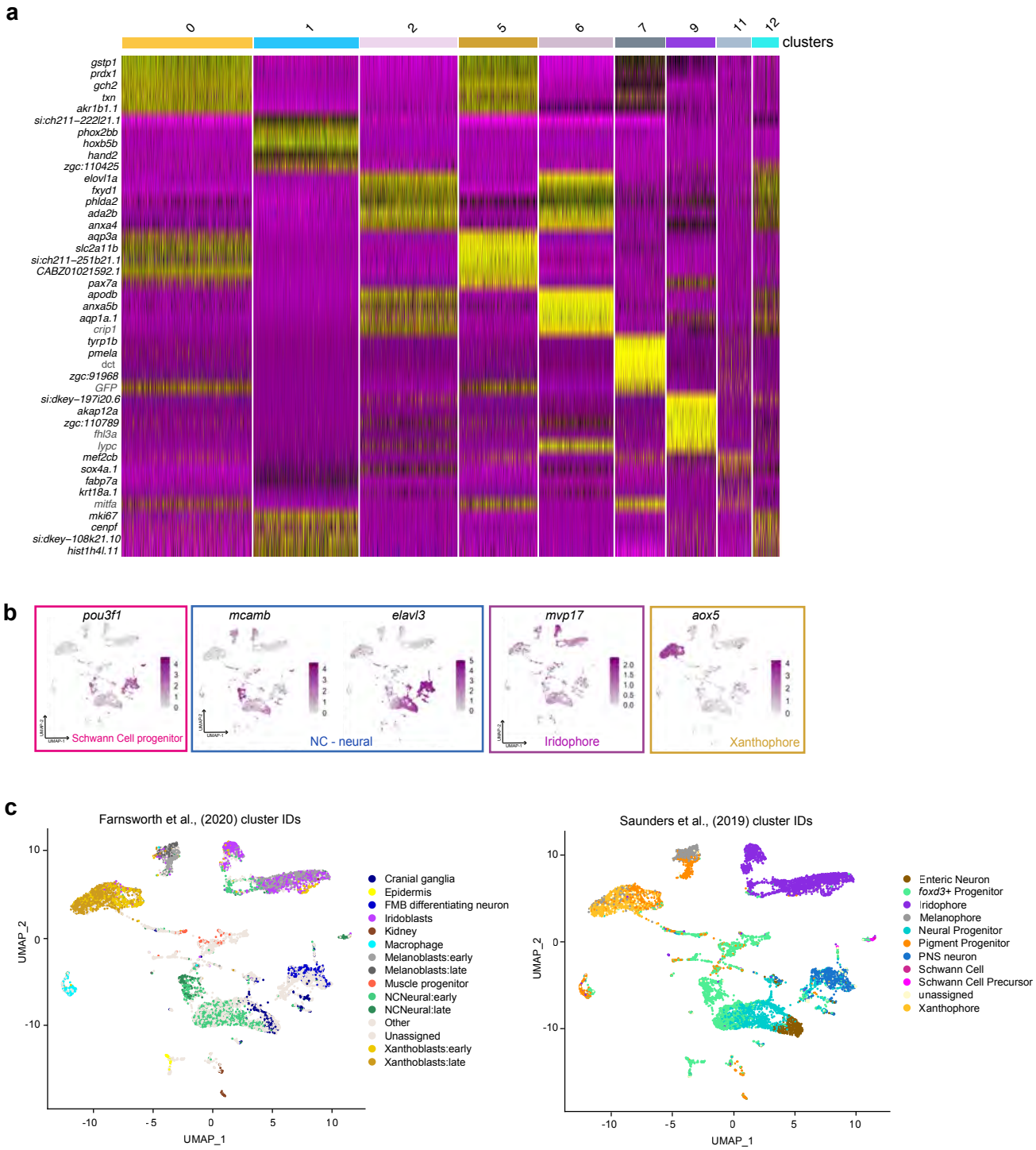
Brunsdon *et al*,. Figure 4



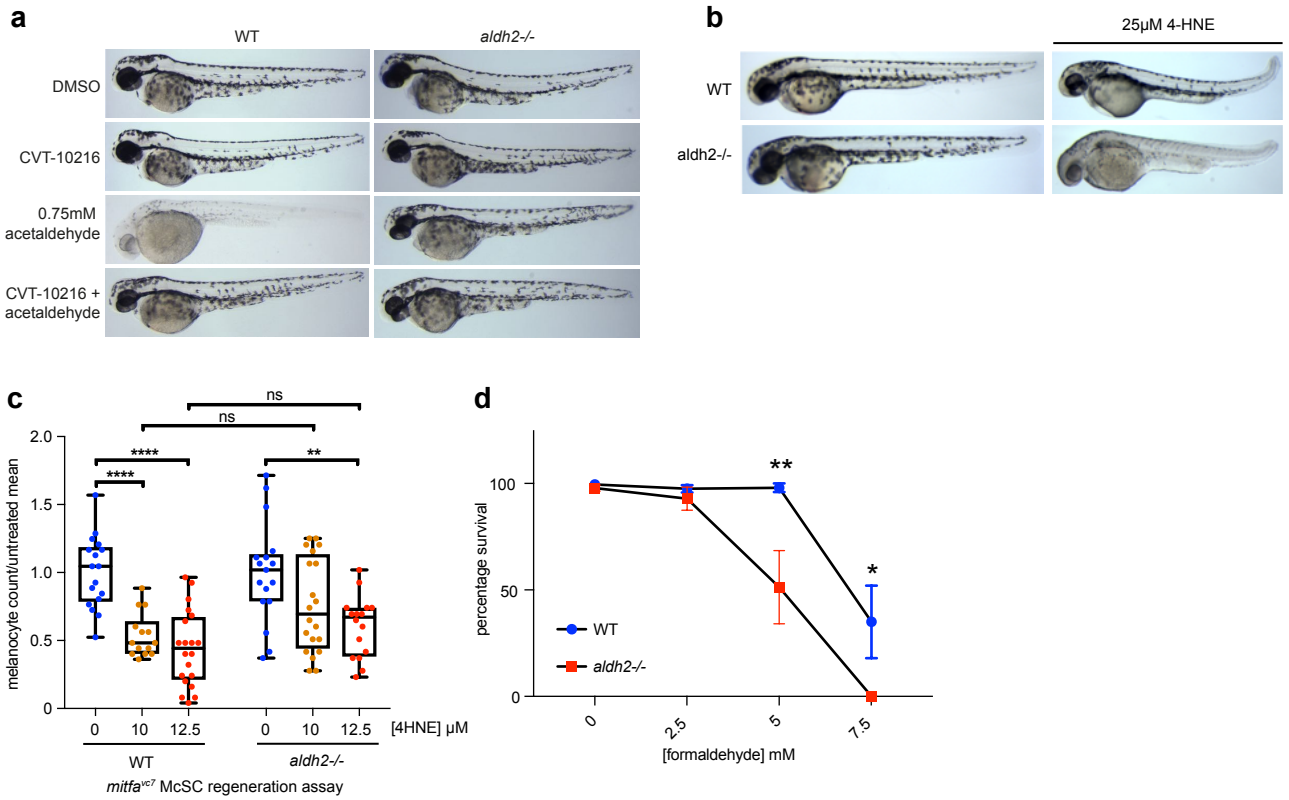
Brunsdon et al,. Figure 5



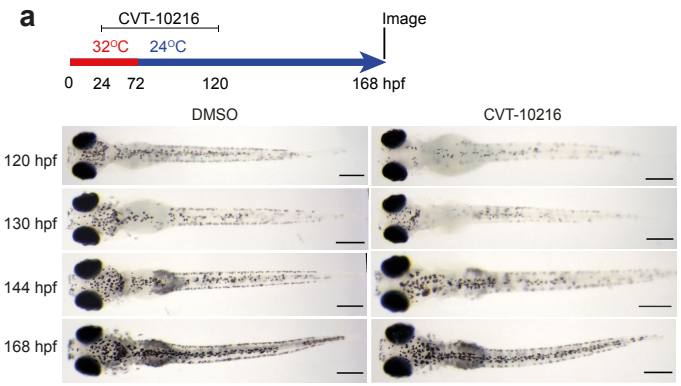
Brunsdon *et al*,. Figure S1

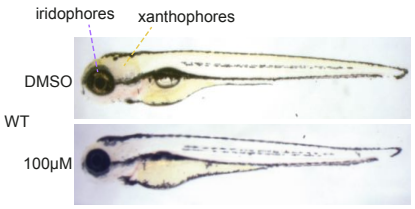


Brunsdon *et al*,. Figure S2



Brunsdon *et al*,. Figure S3





Brunsdon *et al*,. Figure S4