Comprehensive single-cell transcriptomic analysis of embryonic melanoblasts uncovers lineage-specific mechanisms of melanoma metastasis and therapy resistance

Vishaka Gopalan[1], Chi-Ping Day[2], Eva Pérez-Guijarro[2], Sung Chin[3], Jessica Ebersole[3], Cari Smith[3], Mark Simpson[2], Antonella Sassano[2], Maira Alves Constantino[2], Emily Wu[2], Howard H. Yang[2], Maxwell P. Lee[2], Sridhar Hannenhalli[1,C.], Glenn Merlino[2, C.], Kerrie L. Marie[2, 4, C.]

1. Cancer Data Science Laboratory, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA.
2. Laboratory of Cancer Biology and Genetics, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA
3. Frederick National Laboratory for Cancer Research, National Cancer Institute, National Institutes of Health, Frederick, MD, USA.
4. Division of Molecular and Cellular Function, School of Biological Sciences, Faculty of Biology, Medicine and Health, Manchester Academic Health Sciences Centre, University of Manchester, M13 9PT Manchester, United Kingdom
C. Corresponding Authors: kerrie.marie@manchester.ac.uk; sridhar.hannenhalli@nih.gov; merlinog@dc37a.nci.nih.gov

Abstract

Across cancers, tumor cells can resemble embryonic cell states that may allow them to metastasize and evade therapies. Melanoma is a cancer of the melanocyte that exhibits a wide range of transcriptional states characterized by alterations in embryonic melanocyte gene expression patterns. How these states and their functions are related to the embryonic precursors of melanocytes, the melanoblasts, is unknown. Here, we present the first high-resolution single-cell RNA-seq profiles of embryonic melanocytic lineages in mice. We discover a diverse array of transcriptional cell states in this lineage and confirm, for the first time at the single-cell level, that melanocytes arise from Schwann-cell precursors (SCPs), a highly plastic cell population, via a newly described intermediate mesenchymal-like state. Via novel computational strategies to map these developmental cell states to metastatic melanoma, we find that SCP-resembling tumors are associated with exclusion of the immune cells and non-response to immune checkpoint blockade. In contrast, a higher mesenchymal profile underlies immune dysfunction and resistance to BRAF-inhibition therapy. We also carry out the first time-resolved single-cell RNA-seq study of early melanoma metastatic colonization, demonstrating that melanoma cells activate a SCP program transiently during early metastatic colonization. Finally, we discover a hybrid lineage state that resembles multiple melanocytic lineages simultaneously and is enriched in melanoma cells during metastatic seeding and in therapy resistance. Our work reveals that the lineage-specific mechanisms underlie melanoma progression/evolution, including early metastatic colonization and therapeutic resistance.

Introduction

Multiple cancers have been observed to "de-differentiate" and to re-express embryonic genes[1]. This co-opting of embryonic cell states in cancer, including melanoma, is usually tied with plasticity, metastatic progression and therapy resistance[2]. Due to this link with poor outcome, efforts have been underway to understand developmental states in melanoma. This is proving an attractive avenue to uncover novel targets for therapy, because: 1) these targets may be evolutionarily conserved and shared across multiple tumors, 2) targeting embryonic lineage-specific biology may result in reduced side effects in adults, 3) these may provide alternative methods to target immune "cold" tumors (that do not trigger a strong immune response). In recent years, investigation of embryonic lineage has begun to gain traction amongst cancer biologists.

Melanoma is a cancer of the melanocytes, the pigment-producing cells known to protect against harmful Ultraviolet (UV) irradiation in the skin. However, melanocytes also populate non-UV irradiated areas, such as palms of hands, soles of feet, eyes and mucosal membranes. In embryos, the precursor of melanocytes, melanoblasts, develop from Neural Crest Cells (NCCs), an intriguing population of cells renowned for being highly migratory and multipotent[3, 4]. For these reasons, they have long fascinated developmental biologists. Collectively, NCCs migrate through almost every organ of the developing body — the skin being the largest organ of the body, thus melanoblasts need to migrate extensively for successful colonization[4]. In fact, the highly invasive nature of melanoma has been attributed to this innate capacity for migration within the melanocytic lineage. NCCs are also a highly multipotent, giving rise to not only melanocytes, but also peripheral nerves, Schwann cells, glial cells, bones, muscle and cartilage[1, 2]. Therefore, in lineage terms, melanocytes are thought to be closely related to many diverse tissues[5]. Melanoma is a highly plastic cancer, which can transition between cell states to adapt to challenging stresses, such as during metastases or following therapy[2, 6]. How melanoma's plastic nature is related to the multipotency of NCCs or the melanocytic lineage is unknown, nor if there are conserved routes of differentiation or responses to extracellular stimuli that melanoma cells hijack to achieve cell state switching.

In the mouse embryo, NCCs have been shown to be specified to the melanocytic lineage as early as Embryonic day (E) 8 in the Migration Staging Area (MSA)[7]. This specification is then reinforced by environmental signals as melanoblasts receive as they migrate to their final location. Dopachrome tautomerase (Dct) is a known melanocytic lineage gene and is involved in melanin biosynthesis. Transgenic reporter mice, which detect expression from the Dct promoter, develop labeled melanoblasts by E9.5[8]. They migrate while differentiate as a heterogeneous cell population until several days after birth of the pups. Two lineages of melanoblasts have been described to date in the embryonic trunk: 1) Dorsolateral migrating neural crest cells, and 2) Ventrally migrating and nerve-associated Schwann Cell Precursors, which once denervated, produce a 2nd wave of melanocyte development[4, 5]. At any given time melanoblasts face different environmental stresses at different stages in their migration or colonization of the skin. Moreover, melanoblast states may vary depending upon which lineage they came from. To understand the full spectrum of melanoblast development the transcriptome of melanoblasts must be profiled at a single-cell level – not yet accomplished in great detail owing to recent technological challenges. We hypothesized that many functionally distinct melanoblast

subpopulations exist during embryonic development and that some of these subpopulations represent advantageous biology that can be co-opted by melanoma cells during disease progression.

To date, a high-resolution characterization of embryonic melanoblast subpopulations has not been reported. Here, we address these limitations by using a developmentally synchronized, inbred mouse model, coupled with a specific lineage reporter of both immature and mature melanoblasts, the *i*Dct-GFP transgenic mouse[8]. Using single-cell transcriptomic (scRNA-seq) analysis here we profile developing mouse melanoblasts to a resolution greater than was previously known. We thus identified transcriptomically distinct subpopulations. Using signatures from the identified cell states we then developed a Developmental Gene Module (DGM) classification system, as an effective tool to characterize melanocytic lineage gene expression within complex large transcriptomic datasets.

In prior work, melanoma tumors have been subtyped based on either their transcriptional signatures (such as overall differentiation state)[1, 2, 6, 9, 10] or tumor microenvironment (TME) status[11, 12]. While melanoma cell states have been defined by phenotype (invasive/proliferative)[13], assumed lineage[1, 2, 6] and developmental age (fetal/neonatal/adult)[9, 10], their association with function and prognosis is yet unclear. We hypothesized that T-cell response may be driven by lineage-specific signaling, because prior work has shown expression of melanosome genes to be positively correlated with patient response to immune checkpoint inhibitors (ICI)[14]. Studies to address this hypothesis are required for deriving biomarkers or pathways that predict therapeutic response and identify targets. Up until now, our ability to fully comprehend how lineage-specific signaling contributes to melanoma progression has been hampered by a lack of comprehensive characterization of neural-crest derived melanoblasts. With our new DGM classification tool and developmental insights this question can now be addressed.

By applying DGM classification system, we have been able to identify developmental states associated with therapeutic resistance in ICI and targeted therapies. In particular, we uncover a new population of melanoblasts that express mesenchymal-like (Mes) signatures, that are enriched in ECM remodeling genes. Gene expression signatures from this newly characterized melanoblast population is active in melanoma metastatic outgrowth in the lung and correlated with vemurafenib targeted therapy resistance. Finally, our analyses revealed an unexpected population of progenitor cells that could simultaneously exhibit features of multiple melanoblast lineages, termed hybrid states. These hybrid states were found to be enriched in melanoma, particularly during early metastatic seeding, implying the mechanism of plasticity.

Results

ScRNA-seq uncovers diverse subpopulations of mammalian melanoblasts with distinct phenotypes

We hypothesized that multiple embryonic melanoblast subtypes exist, and that further characterization of these would provide unique insights into melanoma. To address this, we used a mouse melanoblast lineage tracing mouse model (*i*Dct-GFP), which labels melanocytes from immature embryonic stages through to mature adult development. Immunohistochemical analysis of melanoblasts with an anti-GFP antibody confirmed GFP^+ cells were localized to regions known to be populated by melanoblasts in embryogenesis (Fig. 1a-b)[15, 16]. This includes melanoblasts in the dermis, epidermis and hair follicles, along peripheral nerves and surrounding the dorsal root ganglion (DRG) – as reported by Adameyko and colleagues[5]. We also noted melanoblasts clustered in collagen-rich regions and surrounding vasculature, which has only previously been reported of embryonic melanoblasts populating the choroid (eye)[17]. Using scRNA-seq we queried melanoblast subpopulations at two developmentally relevant embryonic stages (Fig. 1c): 1) E11.5 is enriched for melanoblasts transmigrating from the dermis to the epidermis, across the basement membrane[16]. This is relevant to melanoma, in which cells also invade through the basement membrane as the tumor progresses. 2) E15.5 is enriched for melanoblasts colonizing hair follicles, which could be analogous to metastatic colonization of melanoma[16]. Importantly, cells were isolated from multiple embryos (3 embryos at E11.5 and 4 embryos at E15.5) across different litters per timepoint, thus accounting for any embryo or litter-specific effects.

Uniform Manifold Approximation and Projection (UMAP) analysis of data revealed up to 15 distinct subpopulations of cells across both developmental stages, supporting our hypothesis that multiple embryonic melanoblast subtypes exist (Fig. 1d). To characterize these subpopulations, we cross-referenced our data with that of the Mouse Organogenesis Cell Atlas (MOCA) dataset (Fig. S1a). This analysis corroborated the presence of melanocytic (Mel) cells and also of Schwann Cell Precursors (SCPs), a known multipotent progenitor from which melanocytes can develop. These are well-documented cell types in the embryonic melanocyte lineage and thus verifies our approach. However, our study additionally revealed multiple variations within these known melanocyte lineage populations. For example, Mel cells from stage E11.5 are different from those at E15.5. Both groups of Mel cells express the melanocytic genes, *Dct* and *Pmel*, known to be important in early melanocyte specification, as would be expected (Fig. 1e). Late Mel cells however also upregulated *Tyr* expression, a gene encoding a pigmentation enzyme known to be associated in late melanocytic specification. Moreover, late Mel cells also upregulate *Sox4* expression, which is a neural crest transcription factor associated with the Peripheral Nervous System (PNS)[18]. This suggests these late Mel cells are derived from a nerve-associated SCP population, as has been previously shown[5]. Interestingly, a subpopulation of immune cells was identified at E15.5 (Fig. 1d-e). Based on the expression of known immune-specific genes (Cx3cr1, Mrc1; Fig. 1e) and confirmed alignment of reads from GFP construct, we deduce that this population represents genuine embryonic macrophages that engulfed GFP^+ melanoblasts. An implication of this is that E15.5 is the first stage in this study in which the immune cells are in contact with melanoblasts at an appreciable level.

In addition to described melanoblast populations, our data highlights new melanoblast cell states. Broadly, these new populations can be classified as either neural (including Neural Progenitor Cells (NPC) and neural tube, in E11.5 exclusively) or mesenchymal-like (Mes). Neural

subtypes are characterized by expression of classical transcriptional regulators of neural differentiation, including Neurog1, *Pou4f1* and *Neurod1* (Fig. 1e-f), and upregulation of *Tubb3* expression, a gene encoding a tubulin subtype that is expressed exclusively in neurons[19]. They are also highly enriched in transcriptional and translational processing genes, which is suggestive of translational reprogramming (Fig. 1f). Finally, NPC1 is enriched for cell cycle genes suggesting a mitotically active neural-derived population (Fig. 1f).

Our work also highlights four ECM-enriched mesenchymal-like populations (Mes Early at E11.5 and Mes 1-3 at E15.5), characterized by expression of the mesenchymal transcription factors (Fig. 1e-f), *Twist1* and *Prrx1* and Transforming Growth Factor beta (*Tgfb*) pathway genes. At E15.5, mesenchymal-like cells upregulate the Epithelial-to-Mesenchymal (EMT) transcription factor, *Zeb1,* and the Endoplasmic Reticulum (ER) Stress transcription factor, *Creb3l1* (cAMP responsive element binding protein 3; Fig. 1e-f). While these three Mes populations are related, they have distinct features. Most noticeably, Mes-1 cells have a gene expression pattern characteristic of a mesenchymal stem cell/ osteogenic progenitor cell state. This includes upregulation of *Sox4*, osteopontin (*Spp1*), Matrixmetalloprotease genes (*Mmp13*, *Mmp14*, *Mmp9*) and *Runx1* and *Runx2*. Mes-3 expresses bone morphogenetic protein (*Bmp*) genes and the family of *Col4a* genes, which encode the basal lamina, a component of the basement membrane. This is the first time a subpopulation of mesenchymal-like melanoblasts has been described.

SCP cells are defined by expression of Neural Growth Factor Receptor (*Ngfr*; Fig. 1e), which in melanoma has been associated with tumor dormancy and resistance to therapies[20]. As expected, SCP cells in this study are enriched for expression of the tyrosine kinase receptor, *Erbb3*, which is a cell surface receptor for neuregulin-1 and has been documented as a key mediator of SCP fate[5]. Likewise, expression of Endothelin-B receptor (*Ednrb*) is observed, which is a key determinant that drives differentiation of melanocytes and the enteric nervous system from the neural crest[21]. As with the other lineages, SCPs also exist in multiple forms, each intrinsically different. SCP1 co-expresses some melanocytic genes, such as *Rab38*, a gene whose protein product is required for packaging and trafficking of melanin in melanosoMes-2. SCP1 also upregulates *Notch1* gene expression which is known to maintain immature melanoblasts and is responsible for survival and regulation of proper differentiation in this population[22, 23]. In contrast, SCP2 shows expression of the EMT transcription factor, *Zeb1* and upregulation of the serine protease inhibitor, *Serpine2*, which is known to remodel the ECM and is expressed in many tumors[24]. Our study thus characterized new and existing melanoblast subpopulations to a resolution greater than was previously known. We next asked how/if these lineage markers are expressed/contribute to melanoma progression. To address this, we first needed to develop a tool to accurately measure melanoblast lineage cell states in large datasets.

Construction of a Developmental Gene Module (DGM) classification tool

We developed a method to apply the knowledge of mouse melanoblast subpopulations gained in this study to enhance understanding of human melanoma progression. To do this, we generated gene signatures of each melanoblast subpopulation detected by Seurat. These

signatures consisted of ~50-200 genes that are (a) significantly up-regulated in a given cluster reltive to other clusters, (b) do not contain cell cycle (or correlated) genes, and (c) co-expressed more strongly within the lineage represented by the cluster compared to clusters of other lineages (see Methods). We have termed this collective gene signature resource as, "Developmental Gene Modules" (hereafter referred to as DGMs; Table S1). The DGM classification tool can be evaluated within melanocyte/melanoma datasets to understand if the relative levels of melanoblast lineage gene expression can reveal new biomarkers and/or mechanisms of development and disease. We follow two distinct gene set activity normalization strategies, one for single-cell and one for bulk RNA-seq datasets (see Methods), where the resulting activity scores represent the extent of co-expression of genes within a DGM. We did not derive a DGM for the macrophage population as it was not part of the melanoblast lineage.

To validate both the DGMs and our method of scoring their activities, we computed DGM activities in an independent dataset of human melanocyte scRNA-seq data[9] of adult, neo-natal and fetal melanocytes. DGMs were widely and dynamically expressed between human melanocyte subpopulations at different life stages (Fig. 1g), demonstrating that DGMs are indeed both relevant and informative to human melanocyte development. The most widely expressed DGMs were unsurprisingly those corresponding to classical "melanocytic" subpopulations. DGMs that correspond to immature, early melanoblasts, such as the SCP lineage, were more active in human fetal melanocytes, as would be expected (Fig. 1g). We found mesenchymal-like DGM activity was strongly enriched in a sub-cluster of human fetal melanocytes, associated with sub-cluster 10 in the study by Belote et al (see Fig. 1h), a population believed to be Melanocyte Stem Cells (MSCs)[9]. The cells of this sub-cluster also had upregulated SCP activity but low Mel activity. Importantly, the Neural lineage DGMs, which were only found in very immature melanoblasts in mice (E11.5), had low overall activity within human melanocytes. This is likely because very few, if any, of the isolated human melanocytes were equally as immature as our isolated mouse embryonic melanoblasts (human fetal week 9 versus mouse Embryonic day 11.5), corroborating the specificity and sensitivity of the DGMs and validate our DGM activity normalization scheme.

DGM-labelling and cell fate prediction reveals SCPs and Mesenchymal-like cells with multipotent states

Palantir, a computational tool that calculates cell fate probabilities, works on the assumption that cell states and cell fate decisions are a continuous process (rather than bimodal decisions)[25]. This has proven to be a highly effective tool to predict lineage and cell fate choices in development, such as during hematopoiesis[25]. We therefore employed Palantir to represent complex cell fate probabilities taken from scRNA-seq data as a simplified 2-dimensional plot (diffusion embedding). We carried out diffusion embedding separately on E11.5 and E15.5 cells (Fig. 2a-b), identifying cell states within the plot that are most likely to represent terminal states and arranging cells in pseudotime between these states within one timepoint. In E11.5 embryos, NPC2, notochord and early melanocytes are identified by Palantir as terminal states. We also learned that in E15.5 embryos, a neural population (sensory), a mesenchymal-like population (Mes-3) and melanocytic cells represented terminal states. To complement this visual

representation, we used Palantir's transition probability matrix that calculates a quantitative probability metric of a cell transitioning from a given cluster (or state) towards other cell states (Fig. 2c-d). In E15.5 embryos we expected to see evidence of SCPs transitioning into melanocytic cells (as has been previously documented[5]). This expectation bears out in Palantir transition probabilities, thus validating the approach and quality of our scRNA-seq data. However, in addition to this we were surprised to observe a population of mesenchymal-like melanoblasts (Mes 2) could be derived from SCPs and was predicted to transition into a melanocytic state or towards a SCP state (Fig. 2b,d). This data further confirms that mesenchymal-like melanoblasts identified in our study are genuine members of the melanocytic lineage, although we do not yet understand their role in embryonic development.

To make accurate cell fate predictions, the Palantir algorithm computes an entropy score for each cell. This entropy score measures a cell's potential to differentiate towards multiple different states, where a higher entropy represents a greater multipotent potential. The embeddings in Fig. 2a-b reveal that cells in each time-point exist along a continuum of differentiation. A given choice of clustering parameters divides this continuum into discrete clusters or states. For instance, within the SCP.1 cluster one would expect some SCP cells to express more melanocytic genes (have a higher Mel activity) than others. To detect such cells, we computed the active DGMs in each cell (see Methods), and classified cells as mesenchymal (Mes), melanocytic (Mel), Neural or Schwann cell-precursor (SCP) if at least one DGM belonging to a given lineage was active. Crucially, this methodology allowed for the possibility that a given cell expresses DGMs from multiple lineages – we refer these as cells in 'hybrid' state.

We found that the fraction of cells in a hybrid state was significantly higher among SCP cells compared to other clusters (Fig. 2e). We also found rare melanoblasts (21 cells in E11.5, 67 cells in E15.5) that expressed DGMs of all four lineages, which we denote as Mel/Mes/Neural/SCP cells and refer to as the quad state. The frequency of quad state cells was greater than expected by chance (Fig. S1b). The quad state is overrepresented amongst SCP cells in both time-points (Fig. 2F; Two-sided Fisher test, E11.5: OR = 11.68, $p = 1.57 \times 10^{-6}$, E15.5: OR = 2.41, p = 0.003). SCPs are known to be multipotent neural crest progenitor cells and thus should have the greatest entropy[26]. However, the SCPs in this study are already committed to the melanocytic lineage (express GFP from the Dct promoter), therefore we queried what the level of entropy would be in these cells and if quad state cells truly measured entropy of the melanocytic lineage. Cells with an active SCP gene expression (SCP^+), and quad state cells, had a significantly higher entropy than in SCP^- cells (Fig. 2g). In contrast, Neural+, Mes+, or Mel+ cells do not have a consistently higher entropy than Neural-, Mes- or Mel- cells (Fig. S1c). These data indicate that SCP population has a high potential to differentiate (entropy) towards other states. Expression of the SCP DGM alone is able to distinguish between high- and low-entropy cells. However, the quad state represents a rare multi-potent SCP sub-population that cannot solely be defined by its entropic potential, suggesting that hybrid states may be a new phenomenon worthy of further study in melanocytic development and melanoma.

A SCP-like melanoma cell state is associated with T cell exclusion in patient tumors

We asked if our newly devised DGM classification system could be used to analyze transcriptomes of cutaneous melanoma (SKCM) tumors profiled in The Cancer Genome Atlas (TCGA) project[27]. Firstly, we used human orthologs of genes in each DGM (Table S2) to calculate normalized activity scores of all DGMs (DGM activity) in RNA-sequencing data from patient tumors. DGM activity is the mean expression of DGM genes normalized using expression of a random subset of DGM genes (see Methods) For each DGM, we plotted the frequency of patient samples that scored highly for the DGM (activity score > 0) (Fig. 3a). Melanocytic DGMs are active in nearly 80% of all melanomas. SCP, Mesenchymal, and NPC1 and NT-Early DGMs are active in 20-40% of all samples. We confirmed that DGMs are expressed in patient tumors and next asked if DGM scoring could improve on the prognostic or subtyping abilities of other methodologies.

While the TME is a major determinant in melanoma patient prognosis, we do not understand if any tumor cell autonomous cues could drive the changes in TME observed throughout melanoma patients. We therefore asked if DGM activity are associated with T-cell response. Here we employed Tumor Immune Dysfunction Exclusion (TIDE) scoring[28], a transcriptional deconvolution method of tumor subtyping, to evaluate T-cell infiltration and functionality in tumors. We calculated DGM activity and TIDE scores within all melanomas of the TCGA SKCM cohort[27]. Spearman's Rho correlation analysis revealed that tumors with a high SCP DGM activity had a high T cell exclusion score (Fig. 3b). Conversely, tumors with high Mes DGM activities had a high T cell dysfunction score. This result prompted us to test if a combination approach of DGM-TIDE subtyping was predictive of survival in patients. Based on combined DGM-TIDE activity scores, a Principal Component Analysis (PCA) dimension reduction method was used and k-means clustering applied to find four patient tumor subtypes (Fig. 3c-d; Fig S2a): 1. Mel/Neural; 2. Mel/Dysfunction; 3. SCP/Neural/Exclusion; and 4. SCP/Mes/Dysfunction-Exclusion. Kaplan Meier plots of patient survival coupled with Cox-regression analyses showed that this subtyping system was a robust predictor of patient prognosis over and above the lineage scoring system derived by Tsoi et al (Figs. 3e, S2a-f, S3a-b). Subtype 3 (SCP/Neural/Exclusion) had the worst overall survival prognosis, whereas subtype 2 (Mel/Dysfunction) had the best patient prognosis. Notably, overall survival was indistinguishable between subtypes 1 and 4 even though both subtypes had dramatically different DGM profiles and despite subtype 1 displaying no obvious T-cell phenotype. This suggests that different mechanisms are at play between these two tumor subtypes.

We next asked how our DGM classification improved upon known gene markers of melanoma progression/ patient outcome. In the literature, *AXL* and *KDELR3-001* expression and loss of *SOX10* expression are associated with invasiveness[10, 29, 30], while *MITF* expression is shown to drive proliferation/invasiveness[6, 13]. NGFR and the Melanocytic Plasticity Signature (MPS) are associated with immune therapy resistance[14, 20]. None of these markers were directly associated with one tumor subtype over another (Fig. S3c-e). However, we did observe a significant enrichment of SCP gene expression in BRAF-mutated melanomas vs. all other mutational subtypes (Fig. 3f). In summary, DGM classification revealed that Tumor Immune Exclusion is strongly associated with SCP DGM activity. Moreover, DGM classification can link functional and

prognostic features (TME and survival) with real lineage-derived melanocytic phenotypes to help reveal how tumor cell autonomous features can drive melanoma progression.

SCP and Mesenchymal DGMs underlie resistance to ICI and BRAF-targeted therapy

Given the association between high SCP activity and immune exclusion, we reasoned that DGM activity would be predictive of response to ICI therapy. We obtained re-processed RNA-seq data[31] obtained from tumors of responders and non-responders to anti-CTLA4 therapy[32] and anti-PD1 therapy[33, 34] and calculated DGM activities in samples in these cohorts. Across datasets we found that SCP activity was up-regulated in pre-treatment samples of non-responders vs responders (Fig. 4a), which was in contrast to the melanocytic DGM that was up-regulated in responders. Higher activity of the melanocytic DGM was also predictive of longer overall survival and progression free survival, while higher expression of the SCP2 DGM was predictive of shorter overall survival and progression free survival (Fig. 4b-c, S4a-b). These findings support our hypothesis that adoption of SCP lineage state by melanoma cells is associated with immune evasion and thus drive resistance to ICI therapy.

Encouraged by these findings, we next asked if our DGM classification could inform other clinically relevant therapies. An "undifferentiated-neural-crest-like state" was implicated in melanoma resistance to vemurafenib[1, 6]. However, this developmental state was coined by a post-hoc interpretation of differentially expressed genes in melanoma, without the knowledge of how these gene signatures were expressed in the embryonic melanocytic lineage. Similarly, recent work uncovered a stress-like gene signature[35], seemingly distinct from melanocytic and neural crest signatures, whose induction in melanoma cells leads to vemurafenib resistance. Our methodology enables us to pinpoint the exact lineage contribution to these clinically relevant states. Thus, by cross correlating these previously identified signatures with our embryonic melanocytic lineage we can begin to fill the above knowledge gaps. Using correlative analysis in patient-derived melanoma cell lines in Tsoi et al we found that an "undifferentiated-neural-crest-like state" is strongly correlated with melanoma cells with both high Mes DGM activity and Early-neural/SCP DGM activity (Fig. 4d). We also found that the stress-like gene signature consists of genes that are up-regulated in the mesenchymal clusters (Mes-1, Mes-2, Mes-Early) and the Notochord cluster from our melanoblast data (Fig. 4e). Therefore, we predict Mes DGM activity is necessary for vemurafenib resistance. We investigated DGM activity over time in the vemurafenib-sensitive $BRAF^{V600E}$-mutated WM989 melanoma cell line. We found that Mes DGM activity increases over time following vemurafenib treatment (Fig. 4f) and that EGFR-high cells ($EGFR^{+}$), which have greater vemurafenib resistance[33], have higher activity of the Mes DGMs than unsorted cells. We therefore show that a mesenchymal-like DGM activity is enriched in vemurafenib resistant $BRAF^{V600E}$-mutated cells. We also found an increase in SCP DGM activity following vemurafenib treatment in $EGFR^{+}$ cells. Likewise, SCP activity transiently increased in two vemurafenib-resistant melanoma cell lines, M229 and M397, as analyzed by Tsoi et. al[1] (Fig. 4g). The spike in SCP activity precedes an increase in NPC and NT-Early DGM activity in M229 cells and an increase in Mes DGM activity in M397 cells. This demonstrates cross-over between

mechanisms of resistance to Vemurafenib and ICI treated cells, with an activation of SCP DGMs preceding the terminal DGM activity of cells.

Since the DGMs of SCP and/or Mes mediate the de novo resistance to ICI in melanoma, we investigated which DGMs may induce acquired resistance after ICI treatment. We therefore analyzed single-cell RNA-seq data of malignant cells from patient melanomas[12] (Fig. 4h). The malignant cells in this study came from either untreated tumors of unknown response to ICI therapy, or from ICI post-progression lesions. Survival data is not available for these samples. Therefore, while we are not able to fully assess risk factors in these cells, we are able to ask what state human melanoma cells exist in and what common changes occur following ICI therapy. We see that almost all of our DGMs are significantly active (normalized gene set activity >1) in malignant melanoma cells from patient tumors, showing that embryonic melanoblast gene expression is re-acquired in malignant cells (Fig. 4h). All main categories of DGMs are represented in patient cells (Mel, SCP, Mes, Neural) with Mel and SCP DGMs being particularly enriched. Active Mes DGM activity confirms that individual malignant cells retain the ability to remodel their microenvironment in a similar manner to their embryonic melanoblast predecessors.

Remarkably, we found a near-uniform up-regulation of more immature DGMs (derived from E11.5 melanoblasts) and a down-regulation of more differentiated DGMs (derived from E15.5) in post-treatment vs. treatment naive malignant cells. This suggests an overall trend of de-dedifferentiation towards an earlier embryonic time-point in ICI non-responders. Notably, the neural DGMs share expression of mRNA splicing, translation and protein folding gene sets (Fig. 1f, S5a). This result is in-line with work by Goding and colleagues, who showed that melanoma cells leverage translational rewiring to adapt to periods of stress or starvation[36, 37]. The near-uniform up-regulation of more immature DGMs suggests these processes may be required for melanoma cells to adopt an early-embryonic immature melanoblast-like state, as can be modeled by E11.5 melanoblasts in mouse. While SCP2 DGM activity is strongly predictive of poor patient progression from treatment naive bulk RNA-seq studies, SCP2 activity is significantly lower in post-treatment melanoma cells than treatment naive cells. This suggests a difference in the biomarkers/mechanisms at play that determine prognostic potential and those that are initiated in response to ICI insult. To fully understand how developmental cell states contribute to melanoma progression one must be able to map these states over time under periods of selective pressure. Nonetheless, coupled with the predictive power of SCP signature activity in untreated patients we suggest that SCP-high activity predisposes cells to undergo efficient and effective rewiring upon therapeutic stress. We thus see an SCP-high state underlying resistance to ICI and a Mes-high state underlying resistance to BRAF inhibition, although a transitory SCP-high state is consistent with/potentially required for a switch to a Mes-high phenotype.

scRNA-seq uncovers dynamic DGM evolution during metastatic outgrowth

The results so far suggest that melanoma tumors re-express melanoblast developmental states in response to therapeutic insult. We hypothesized that re-expression of these states is a mechanism of adaptation in melanoma cells. This prompted a temporal analysis of

developmental cell state contribution to melanoma progression under selective pressure. We reasoned that early metastatic colonization and outgrowth in a foreign microenvironment requires melanoma cell adaptation. We therefore sought to understand how lineage-derived cell states, determined by DGM classification, dynamically contribute to metastatic progression. We curated a scRNA-seq dataset of mouse melanoma (B2905[14]/M4-BRN2) cells over time during lung metastatic outgrowth in mice. B2905/M4 cells are genetically heterogeneous[14] and the M4-BRN2 cell line derivative has been cycled through mice twice to enrich for cells of high metastatic potential[38]. We have previously shown that B2905 (model M4) cells of similar mutational status can have differential expression of melanocytic lineage genes[39], indicating that this model undergoes non-genetic plastic cell state switching. It is therefore an ideal model to study the contribution of cell-state switching between DGMs in metastatic outgrowth. We confirmed that lung sections had increased tumor burden over time and presented with multiple phenotypes (Fig. 5a-g). These include amelanotic and pigmented metastatic nodules, as well as lesions with a heterogeneous mixture of both pigmented and nonpigmented neoplastic cells. We validated that some metastatic nodules in this immunocompetent model had a component of tumor infiltrating lymphocytes. Anti-Sox10 immunolabelling revealed both melanotic and amelanotic lesions had Sox10$^+$ cells and that Sox10 expression was heterogeneous. Some amelanotic lesions could be confirmed to be anti-Sox10 negative (Fig. 5a-g).

eGFP$^+$ metastatic cells were injected into the tail vein of nude mice and then isolated from lungs 1-, 2- and 3-weeks post injection for scRNA-seq analysis (Fig. 5a-i). Gene expression of single cells were scored using the DGM classification system. We used a multi-pronged approach to confirm malignant cells in our data. First, we used inferCNV to estimate the number of Copy Number Variations (CNVs) in each cluster (Fig. S6a), where a higher CNV load is representative of a malignant cell. Second, we counted the number of reads aligning to the eGFP construct, to identify putative malignant clusters. Third, we derived markers of cell populations in healthy mouse lung biopsies from the Tabula Muris single-cell RNA-seq dataset[40] and scored each cluster in our data for the expression of these genes (Fig. S6b). Out of 9 cell clusters, one malignant cluster was excluded for being enriched in apoptotic signatures. Another cluster expressed macrophage marker genes and had an intermediate CNV score, potentially representing macrophages that have engulfed a malignant cell. One intriguing cluster expressed endothelial marker genes and had an intermediate CNV score. This is suggestive of a possible mixture of metastatic and normal cells within this cluster, both with an endothelial phenotype. This cluster may therefore represent metastatic cells that have transdifferentiated to an endothelial subtype, which has been noted in an earlier study of disseminated melanoma cells[41]. While this is a notable feature of this model, it does not appear to be linked to co-option of melanocytic lineage biology and therefore was not pursued further in this study.

We scored each malignant cell for the activity of all DGMs (see Methods) and analyzed their expression over time (Fig. 5j, S6c, d). Any DGM whose normalized activity is above 1 in a cell is considered active in that cell. Virtually all cells had active Mel and Mes DGMs (Mes-1, Mes-2 and Mes-3) across all time-points, with an overall increase in Mes activity from the 1-week to 2-week time-points (Fig. 5j, S6d). This suggests that Mes DGM activity is enriched upon metastatic progression in our model, and that this is in-line with a previously described “undifferentiated

state"[2, 29] that is AXL^{Hi}, MITF^{Low} and SOX10^{Low}. Thus, our system is an excellent model to understand how "undifferentiated" (herein termed "mesenchymal-like" or "Mes" for short) melanomas are regulated during metastatic colonization and outgrowth.

Of the Neural lineage DGMs only one (NT-Early) was significantly expressed in M4 metastatic melanoma cells (Fig. 5j). NT-Early is a gene expression signature derived from an E11.5 melanoblast subcluster with a neural tube-like cell state (Fig. 1d, S5a). The NT-Early DGM showed high activity at the 1-week time-point, with a reduction in activity over time, suggesting it is required specifically during early metastatic colonization/adaptation to a new microenvironment. NT-Early genes are strongly enriched for proteostasis genes (Cct, Psma families), translational control genes (Eif family, Ybx1) which cumulatively represent translational and transcriptional reprogramming (Fig. S5a, S7f). These processes are active as early as 1 week post metastatic colonization, suggesting their role during melanoma cell adaptation in metastasis. This is consistent with our findings showing neural lineages upregulated post-treatment in the Jerby-Arnon ICI-treated patient data (Fig. 4h). Selective protein synthesis (translational reprogramming) has recently been described as a mechanism of metastatic adaptation in other models[37]. NT-Early DGM expression was concomitant with expression of SCP DGMs, which had high activity at 1-week post metastatic colonization, with a reduction over time. This suggests a requirement for a progenitor-like state and translational reprogramming for early metastatic melanoma adaptation to a foreign microenvironment. Our data from early metastasis yet again highlights that high SCP activity correlates with adaptation of melanoma, in this instance at the metastatic site. We know that in development, SCPs are a nerve-associated multipotent population, we therefore posit that co-option of SCP gene expression is a lineage-specific mechanism to enhance plastic adaptability. To address this, we need to measure the multipotent potential of developmental cell states during metastatic progression.

Hybrid DGM states regulate the tumor microenvironment during metastatic evolution

Next we asked if we could measure multipotency of the lineage-derived cell states and if so, how this would change over time during metastatic colonization. To do this, we use the DGM hybrid state labeling defined in our developmental studies (Fig. 2). Through this measurement we can define if a single cell is significantly expressing features of multiple lineages. We suggest that such a hybrid state may represent multipotentency potential. In normal embryonic development, hybrid states make up a minority of all cells (Fig. 2e, f). Likewise in development we observed a rare population of quad-state cells that significantly express features of all four main lineages (Neural, Mel, Mes and SCP) and were overrepresented amongst SCPs (Fig. 2f). However, when hybrid state labeling was applied to M4-BRN2 malignant cells extracted from lung metastases, we found the **majority** of these cells existed in a hybrid DGM state (Fig. 6a). We next assessed whether DGM hybrid mapping could reveal single-cell evolution of developmental cell states during metastatic colonization.

Analysis of single cell transcriptomic data of lung metastatic colonization over time (following M4-BRN2 intravenous injection) revealed that the majority of cells at 1-week time-point were in

quad state, with a monotonic decrease in the fraction of quad-state cells over time. This was accompanied by an increase in the fraction of tri-state (Mel/Mes/SCP) cells (Fig. 6a). This reciprocal relationship is suggestive of a selective advantage for potentially multipotent/adaptable quad state (Mel/Mes/Neural/SCP) cells during early metastatic colonization. A strong translational/transcriptional reprogramming gene signature defines the quad state, and these cells are more melanocytic than their tri-state counterparts (Fig. 6b, S7b). Over time during metastatic outgrowth a tri-state, Mel/Mes/SCP phenotype, is selected for. A high mesenchymal-like/matrix remodeling gene signature defines the tri-state (Fig. 6d, S7b). As mentioned previously, Mes DGM expression is a defining feature of this metastatically enriched melanoma model and is also represented in human patient tumors. It has never previously been appreciated how mesenchymal-like/ "undifferentiated" cell states coexist alongside other lineage-derived phenotypes.

We asked if/how hybrid states contribute to tumor cell evolution over time. To calculate the likely transitions occurring between malignant cells in different hybrid states over time, we applied the Waddington-OT algorithm[42] (Fig. 6c). With this analysis we could confirm that the quad- and tri-states were the two key transcriptional states responsible for metastatic evolution in our model. Quad- and tri-state cells tended to transition towards one another over time, whereas other cell states were predicted to transition into either quad- or tri-states (Fig. 6c). We next asked if DGM hybrid state mapping could provide any additional insights into key pathways responsible for metastatic melanoma evolution. We used DGM hybrid mapping as a method of stratifying metastatic cells and performed Gene Set Enrichment Analysis (GSEA) on all quad-state and tri-state cells at each time point (Fig. 6d). This analysis revealed a shift in inflammatory signaling over time between hybrid states. At 1-week post metastatic colonization quad state cells upregulated Interferon gamma (IFNγ) pathways and the Interleukin-6 (IL6)/ Signal transducer and activator of transcription 3 (Stat3) pathways. IFNg is an inflammatory cytokine with a pleiotropic pro and anti-tumor role in tumor progression[8, 43, 44]. By 2-weeks post metastatic colonization all cells have upregulated genes associated with EMT gene expression, which is strongly upregulated in tri-state cells alongside myogenesis GSEA pathway gene. Mesenchymal/myogenesis pathways are consistent with activated Transforming Growth Factor beta (TGFb) signaling, a known activator of proinflammatory immune response[45]. It follows therefore that by 3-weeks post injection, tri-state malignant cells have strongly enhanced IL2-Stat5 signaling and non-canonical nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) signaling. These signaling pathways are known to promote a pro-inflammatory environment, which in tumors stimulates tumor cell proliferation and cell survival signals[45]. Thus, analysis of hybrid DGM states in metastases have revealed an evolution from immune evasive to pro-inflammatory tumor cell characteristics that is associated with a switch of quad-state cells to tri-state cells.

Targetable pathways in hybrid states

Next we asked if developmental cell states are sensitive to changes in IFNγ signaling and if this could be related to either pro- or anti-inflammatory phenotypes. To address this, we measured DGM activity in RNA expression data from melanocytes that had been exposed to UVB 6 days prior[8]. These melanocytes were previously reported to have upregulated an IFNγ gene expression

signature and exhibited a pro-tumorigenic immunoevasion phenotype[8]. We reasoned that if SCP gene expression mediates immunoevasion in the melanocytic lineage and if this was in-part regulated by IFNγ then we would observe SCP gene expression to be elevated in UVB-treated melanocytes vs. controls (UVA and no UV). Indeed this was the case (Fig. 6e; Fig. S8a). We next asked if direct IFNγ treatment of cell lines could evoke a significant response in DGM gene expression. To test this, we used two datasets following 6h and 3-week treatment of IFNγ[44, 46]. We found that IFNγ treatment resulted in dramatic changes in DGM activity and that this DGM expression change was abrogated by knockout of genes in the IFNγ signaling pathway (Fig. 6f; Fig. S8b). Owing to the sensitivity of DGM state switching to IFNγ, we asked if inflammatory signaling played a role in embryonic melanoblast development. We performed GSEA across all melanoblast single cell clusters identified in Figure 1 (Table 1, Table S3). Remarkably, inflammatory signaling was significantly enriched in all the mesenchymal-like clusters (Mes-1, Mes-2, Mes-3) and in the SCP1 cluster, both from E15.5, but in no other melanocytic lineage. E11.5 had significantly negative regulation of inflammatory signalling, corroborating that we isolated no macrophages at this time point. Thus our data points to inflammatory signalling in specifically Mes and SCP phenotypes, which we have also shown to be the most most relevant phenotypes in melanoma progression and resistance to therapy. Our data therefore links IFNγ regulation of developmental cell states to melanoma progression.

To address if hybrid states are relevant to human melanomas, we evaluated DGM activity in single-cell RNA-seq profiles of sixteen CCLE melanoma cell lines[47] (Fig. 6g; top panel) and in single-cell RNA-seq biopsies of human melanoma samples[12] (Fig. 6g; bottom panel). We found that quad- and tri-states were frequent in melanoma cell lines (in vitro), with a wide variation in the fraction of quad-state cells across cell lines and a much higher fraction of more differentiated Mel/Neural cells than in patient tumors. Analysis of patient tumor cells revealed that virtually all cells in each tumor expressed at least one active DGM, with quad state cells being the most prevalent across patients. Likewise, quad- and tri- state expression were most prevalent in a patient-derived xenograft model of vemurafenib resistance[6] (Fig. S8c). While there are differences in sequencing platforms between both the in vivo datasets (SMART-seq) and the in vitro cell line dataset (10X), these observations, coupled with our in vivo data (10X), suggest that quad-state cells are likely more prevalent *in vivo*. We computed the relative regulon activity in quad-state cells vs. other hybrid states in both mouse melanoma metastases (M4) and human tumors and highlighted the regulons that are highly differentially active (z-score > 0) in quad-state cells shared across both datasets (Fig. 6h). A correlation between human and mouse cells suggests that the quad state is regulated by the same mechanisms in both species. This analysis revealed a mixture of Mes-specific (such as *Twist1*, *Creb3l1*, *Foxc2*, *Nfatc4*), SCP-specific (*Sox4*, *Foxp1*, *Dlx1*) and Neural.-specific (*Rfx4*, *Hoxa7*) regulons, with other regulons not being specific to any one lineage. *Xbp1* is up-regulated in response to ER stress[48], which is a crucial aspect of melanoma progression[10, 48]. We suggest that future efforts to treat plastic adaptation in melanoma should target both quad-state cells (a putative highly adaptable, immune-evasive population) and their transition to mesenchymal-like pro-inflammatory cells. Regulons identified in this study represent potential targetable pathways for therapeutic intervention.

Discussion

Our study reinforces and refines, at a single-cell transcriptomic level, century-old observations and modern systems-biology notions of similarities between cancer and development[49]. Numerous studies have shown that diverse malignant tumors express embryonic genes, including melanoma[10], glioblastoma[50], oligodendroglioma[51], neuroblastoma[52], lung adenocarcinoma[53] and Wilms' tumors[54]. Melanoma is a particularly plastic tumor. Cell state switching between "neural crest-like", "undifferentiated" and "melanocytic" phenotypes is instrumental in disease progression. However, until now there was no comprehensive mapping of how these transcriptional cell states are expressed in the embryonic melanocytic lineage, nor an understanding of whether they are requisite for specific cell functions or fates. Our work pinpoints the specific neural crest states, and combinations thereof, that can be observed in patient tumors and early metastatic colonization and are strongly associated with therapy response (Fig. 6i-j).

We show the embryonic melanocytic lineage is more diverse than was previously realized. Multiple populations of melanoblast cell types within this lineage were discovered, each with both shared and unique gene expression. For example, three distinct subgroups of SCP-like melanoblasts were discovered which likely represent distinct SCP developmental stages and cell fate decision points. The discovery of mesenchymal-like melanoblast subpopulations in embryonic development is particularly relevant to understanding melanoma plasticity. This is because melanocytic melanomas frequently adopt mesenchymal-like phenotypes (sometimes referred to as "undifferentiated" to promote aggressive metastases. Previously this phenotype was thought to require the melanoma cell to exit the melanocytic lineage and the external/internal mechanisms that regulated this were little understood.

Cancer is a system in which many of the usual checks and balances are deregulated, thus confounding investigation of the mechanisms that drive phenotypes. In contrast, development is a highly orchestrated system. We can take advantage of this to learn what pathways initiate and establish a mesenchymal-like phenotype from the developmental system. Interestingly, in our study, mesenchymal-like phenotypes are mostly observed in more developed, E15.5 embryos, thereby challenging the prevailing notion in melanoma that equates mesenchymal-like with more "undifferentiated". Our work shows that embryonic mesenchymal-like melanoblasts may develop from a multipotent SCP progenitor population and may even differentiate into pigment-producing melanocytes. This is in line with recent findings in melanoma showing that "neural-crest-like" melanoma cells induced by Notch3 signaling (via an interaction with endothelial cells) is a precursor to a mesenchymal-like melanoma phenotype (pre-EMT)[55]. These seminal findings however leave a lot to be understood, what is the true relation of "neural-crest-like" signatures to the melanocytic lineage and how clinically relevant are these findings.

Our work pinpoints a neural crest gene signature derived from a specific subpopulation of nerve-associated melanoblast SCPs. This gene signature is transiently upregulated post metastatic

colonization and post vemurafenib therapy. Melanoma tumors have previously been prognostically subtyped based on TME status[11, 12, 28], leaving the question open on what tumor cell autonomous factors correspond to immune exclusion. Our joint DGM-TIDE classification system reveals the correlation between the co-expression of SCP genes and the exclusion of T cells from the TME. The role of the SCP state in immune exclusion is further supported by our analysis of multiple anti-CTLA4 and anti-PD1 treatment cohorts where high SCP expression is predictive of poor response and overall survival.

We take developmental subtyping[56] of melanoma one step further, finding that certain melanoma cells can express genes marking multiple melanoblast lineages simultaneously. Such hybrid states also occur (with much lower frequency) in normal development, primarily amongst E15.5 SCPs, which we also find to be the most plastic in terms of their differentiation potential. The quad state is an extreme instance of the hybrid state. We show the quad state to be of high clinical significance as it is a key feature of both early metastatic colonization, vemurafenib-resistant and ICI-resistant tumors. We suggest that the quad state represents a measure of melanocytic lineage-specific multipotency. Tumor cells characterized as being in the quad state therefore may have a greater potential to adopt multiple developmental phenotypes and ergo could have greater plastic potential. Plastic potential is a key factor for tumor cell adaptation to challenges and is not yet effectively targeted in the clinic[2]. Our study highlights quad state regulons that are consistent between mouse and human melanoma, and we suggest these regulons would make effective therapeutic targets.

Our data strongly implicates inflammatory signaling as a key mediator of the quad state and its subsequent cell state transitions. In particular, IFNγ signaling is upregulated in early metastatic quad-state cells prior to their switch to a mesenchymal-like status (Fig. 6j). IFNγ has a paradoxical role as both an anti-tumorigenic and pro-tumorigenic agent[43]. Presence of IFNγ in a tumor microenvironment can increase immune targeting to that region; however exposure of tumor to IFNγ is known to stimulate immune evasive capabilities in tumor cells. We show that UVB-irradiated, immune evasive melanocytes that upregulate IFNγ signaling also upregulate SCP signaling in parallel. Added to our data demonstrating that SCP DGM expression is associated with tumor immune exclusion and because SCP DGM activity is likewise upregulated across all cells at 1-week post injection vs. later timepoints, we suggest that SCP cell states could be important mediators of IFNγ signaling, immune evasion and hybrid states in melanoma. This is corroborated by the fact that in development, the only melanoblast lineages with upregulated inflammatory signaling are the SCP and Mes lineages (Fig. 6i). Moreover, quad-state cells were mostly clustered with SCP-like melanoblasts, suggesting that this multipotent hybrid-state is a feature of the SCP lineage (Fig. 2f, 6i). Erbb3 is specifically expressed in SCP lineages and Stat3 is known to be activated by Erbb3/Her3[57], signifying a potential melanoblast lineage-induced mechanism of immune cell evasion.

In summary, our data and comprehensive new analyses of published data demonstrates that inflammatory signals strongly impact developmental phenotypes. The exact nature of IFNγ regulation in relation to complex systems such as metastasis, plasticity, and therapy is yet to be

clarified and requires comprehensive longitudinal studies in combination with our newly described DGM and hybrid state classification methods. In conclusion we have generated rich datasets and developed informatic tools to fully map the contribution of fundamental, conserved embryonic melanoblast lineage pathways to melanoma progression. We explore, for the first time, how single malignant cells can simultaneously co-opt multiple melanoblast pathways, and how these hybrid states contribute to tumor evolution. We have demonstrated how these data/tools can be used to gain new insight into mechanisms of melanoma plasticity.

Figure legends

Figure 1. Subpopulations of mammalian melanoblasts with distinct phenotypes
a, b. Saggital sections of E15.5 iDct-GFP embryos in skin (a) and at the dorsal root ganglion (b). Left panel, H&E, Middle panel, Masson's trichrome staining, right panel, anti-GFP antibody (green). hf, hair follicles, d, dermis, e, epidermis, bv, blood vessels. Melanoblasts in collagen-rich regions (arrowhead). (b) Melanoblasts (arrows) localise margin of the forming dorsal root ganglion (drg) capsule. Scale bar, 100 μm.
c. Schematic of melanoblast stages.
d. UMAP clustering of melanoblast sc-RNA-seq data (10x Genomics).
e. Expression of selected genes in melanoblast clusters.
f. GO enrichment analysis of differentially up-regulated genes in melanoblast clusters.
g. Normalized DGM activity scores in human melanocytes from Belote et al.
h. Normalized DGM activity scores in the MSC subcluster from Belote et al versus all other melanocytes.

Figure 2. Hybrid developmental cell states are multipotent
a, b. Diffusion-embeddings of melanoblast transcriptomes from E11.5 (A) and E15.5 (B) time-points.
c, d. Palantir-computed transition probability matrix of cells between different cell types seen in E11.5 (C) and E15.5 (D). Each entry denotes the median probability of cells transitioning from the cell type along the row towards the cell type shown in the column. Frequency of different hybrid states across clusters in each lineage.
e. Frequency of hybrid state cells amongst the four melanoblast lineages across E11.5 and E15.5 time-points.
f. UMAP highlighting quad state cells amongst melanoblasts at E11.5 (left) and E15.5 (right)
g. Entropy of quad-state cells (Mel/Mes/Neural/SCP) compared with cells without an active SCP DGM (SCP-) and those with an active SCP DGM (SCP+)

Figure 3. A SCP-like melanoma cell state is associated with T cell exclusion in patient tumors
a. Fraction of cutaneous melanoma samples in TCGA cohort with an active DGM.
b. Spearman correlation coefficient between DGM activities and TIDE exclusion and dysfunction scores.
c. Schematic of analysis strategy.
d. DGM activity variation across different DGM/TIDE subtypes in TCGA.

e. Kaplan Meier of DGM/TIDE subtypes in TCGA.
f. DGM activity variation across TCGA samples split by mutation status of BRAF, NF1, and RAS genes.

Figure 4. SCP and Mesenchymal DGMs underlie resistance to ICI and BRAF-targeted therapy
a. Coefficient corresponding to response term in linear model fit of DGM expression versus response status. Positive/Negative values indicate that the DGM is more active in non-responders/responders. ** and * indicate p-values less than 0.01 and 0.05, respectively.
b. Hazard ratio of expression of different DGMs in progression-free survival (PFS, red) and overall survival (OS, blue) during immune checkpoint blockade after pooling multiple cohorts. * indicates a p-value < 0.05.
c. Kaplan-Meier curves of overall survival of multiple ICB cohorts[32-34] corresponding to Mel, SCP.2 and Mes-Early DGMs.
d. Spearman correlation coefficient between neural crest signatures from Tsoi et. al with DGMs. Activities of both sets of signatures were evaluated in panel of vemurafenib-resistant and sensitive cell lines described in Tsoi et. al (GSE80829).
e. Median Log-FC of mouse orthologs of signature gene sets described in Baron et. al, Cell Systems, 2020 amongst E11.5 and E15.5 cells.
f. Time-course of DGM activity of SCP and Mes lineage DGMs in Shaffer et. al data. Each point is a median of 3-4 replicates (variation was low and thus error bars are not visible).
g. Time- course of DGM activity of SCP, Neural, Mel, Mes lineage DGMs in Tsoi et. al data.
h. Normalized gene set activities of DGMs amongst malignant cells from unpaired pre-treatment and non-responding post-treatment melanoma patients receiving anti-PD1 ICB therapy in Jerby-Arnon et. al cohort.

Figure 5. Developmental cell states over time in melanoma metastases

a-g. Characterization of multiple pulmonary metastatic phenotypes from two week to 22-day post tail vein injection of M4-BRN2 mouse melanoma cells, representative examples. Bars = 100 μm.
a-b. Melanotic and amelanotic lesions, H&E stain.
b. Heterogeneous lesions with both pigmented and nonpigmented neoplastic cells and a component of tumor infiltrating lymphocytes (*).
c.-g. anti-SOX10 Immunohistochemistry reveals a heterogeneous pattern of nuclear immuno-labeling including some but not all lesions.
c. Example metastatic nodules range from intensely pigmented, to heterogeneous mixtures of pigmented and nonpigmented tumor cells in some nodules (middle lesion), and relatively amelanotic nodules with distinct anti-SOX10 positive nuclei (lower nodule).
d-e. The same tumor nodule in serial sections is represented after anti-SOX10 without (d) and with (e) melanin bleaching to remove pigment.
f-g. Metastatic pulmonary nodules from the same lung section revealing anti-SOX10 negative (f) and immunopositive (g) lesions.
h. Schematic of experimental set up.
i. UMAP of all cells across time-points, obtained after batch correction and integration.

j. Normalized activities of selected DGMs across all malignant cells over time.

Figure 6. Hybrid DGM states regulate the tumor microenvironment during metastatic evolution
a. Change in frequency of cells in quad- and tri-states over time. Other hybrid states with a small number of cells are shown in gray.
b. Normalized activities of selected DGMs across all cells divided into quad-states (Mel/Mes/Neural/SCP) and tri-states (Mel/Mes/SCP).
c. Transition probabilities between hybrid states over time as computed by Waddington-OT.
d. Gene set enrichment analysis of differentially expressed genes amongst quad-state and tri-state cells at each time-point. Gene sets with a significant enrichment (q-value < 0.2) are shown in shades of red (up-regulated) or blue (down-regulated).
e. SCP1 DGM activity score of UVA/UVB treated melanocytes from Zaidi et al.
f. DGM activity score of melanoma cells treated +/- IFNγ for 6h from Grasso et al.
g. Fraction of cells in different hybrid states across tumors in Jerby-Arnon et. al cohort (bottom) and panel of CCLE cell lines (top).
h. Z-scores of median regulon activities of quad-state cells in Jerby-Arnon et. al cohort and M4 metastasis data. Top regulons are highlighted based on their Z-scores, with the colors indicating the development lineages in which regulons are specifically active (Teal: No lineage specificity, Green: Mesenchymal, Purple: SCP, Red: Melanocytic).
i. Schematic showing embryonic melanoblasts of different subpopulations, the existence of quad state cells in development and inflammatory signaling is restricted to SCP and Mes phenotypes. SCP, Schwann Cell Precursor; Mel, melanocytic' Mes, mesenchymal-like.
j. Schematic of hyrbrid states in tumor evolution. Quad state cells are required for early metastatic survival and following IFNγ (interferon gamma) signaling upregulate their mesenchymal-like properties.

Table 1. GSEA analysis of up/down-regulated inflammatory pathways in embryonic melanoblast clusters

Supplementary Figure 1.

a. Fraction of cells in each cluster (along rows) of the mouse organogenesis cell atlas (MOCA) that contain an active DGM (indicated along columns).
b. Negative logarithm of frequency of cells in different hybrid states. The red bar indicates the frequency of different hybrid states as expected by chance (see Methods), while the blue bar indicates the actual frequencies.

c. Entropy of quad-state cells (Mel/Mes/Neural/SCP) compared with entropy of cells split into two sub-populations based on whether they contain an active DGM from the lineage indicated.

Supplementary Figure 2.
a. DGM/TIDE subtyping of TCGA SKCM patient samples.
b-c. Hazard ratios and Kaplan Meier of DGM/TIDE subtypes in Cox regression of progression free survival in TCGA, after controlling for sex, age and tumor purity as covariates.
d-e. Hazard ratios and Kaplan Meier of DGM/TIDE subtypes in Cox regression of overall survival in TCGA, after controlling for sex, age and tumor purity as covariates.

f. Tumor purity scores of DGM/TIDE subtypes 1-4.

Supplementary Figure 3.
a, b. Hazard ratio of Tsoi et.al subtypes in Cox regression of overall survival (left) and progression-free survival (right), along with sex, age and purity as covariates.
c. Spearman correlation coefficient between expression of melanoma differentiation marker genes and DGM activity in TCGA SKCM tumors.
d. Variation in activity of all DGMs across DGM/TIDE subtypes.
e. Expression of melanoma differentiation markers, and melanocytic plasticity signature (MPS)[14] scores of different DGM/TIDE subtypes.

Supplementary Figure 4.
a, b. Kaplan-Meier curves of overall survival of multiple ICB cohorts[32-34] stratified by DGM activity.

Supplementary Figure 5.
a. String plots (https://string-db.org) of key pathways in the NT early (neural tube-like) DGM.

Supplementary Figure 6.
a. Boxplot of number of inferred gene-level CNVs (amplifications + deletions) across cells in each cluster.
b. Z-scores of AUCell scores of marker gene sets derived from cells in healthy mouse lungs (along columns) in different clusters from M4 metastasis data (along rows).
c. Fraction of cells with an active DGM across sub-clusters of malignant cells.
d. Normalized DGM activities across all malignant cells over time

Supplementary Figure 7.
a. GSEA pathway analysis of pathways enriched in neural DGMs.
b. Normalized activities of selected DGMs across all cells over time and divided into quad-states (Mel/Mes/Neural/SCP) and tri-states (Mel/Mes/SCP).

Supplementary Figure 8.
a. DGM activity scores of UVA/UVB treated melanocytes from Zaidi et al.
b. DGM activity score change of melanoma cells after treatment with IFNg and TNFa from Kim et al.

Supplementary Figure 9.
a. Hybrid state frequency measured in single cells across melanoma cell lines (CCLE) from Kinker et al.
b. Hybrid state frequency measured in single cells across human melanoma patient tumors from Jerby Arnon et al.
c. Frequency of all hybrid states across time-points (measured in days) of individual cells in a PDX outgrowth of a vemurafenib-resistant melanoma tumor.

Methods

Mouse models

Melanomas were derived from the following mouse melanoma model: M4, C57BL/6 male background, with Hgf-tg transgenic allele. UV was used as the tumor-inducing carcinogen; M4 mice were treated at postnatal day 3. Cell lines derived from this model (B2905) were then cycled through mice via intracardiac injection and isolated from the brain and cultured. This was done twice to make the M4-BRN2 highly metastatic derivative. Experimental metastasis studies were performed using a filtered, single-cell suspension in phosphate-buffered saline (PBS). 1x10^6 M4-BRN2 cells were injected in 100-µl volume into the tail vein of 6–8-week-old Glowing-Head mice (Frederick National Laboratory for Cancer Research) x C57BL/6N mice (Charles River, Frederick National Laboratory for Cancer Research) making a F1 cross. Lungs were perfused with PBS through the circulation prior to removal for sample preparation. Lungs were removed from mice at the timepoints indicated.

Melanoblasts were isolated from the Dct-rtTA; tre-H2B-GFP (*i*Dct-GFP; Frederick National Laboratory for Cancer Research) mouse model. Embryonic development was timed based on the number of days post coitum. Pregnant females and were placed on a doxycycline-enriched diet to activate expression of GFP.

Single Cell – Sample Preparation, Sorting and Oligo-Labeled Antibody Staining

For melanoblast preparation: multiple litters were used per time point (3 litters at E11.5 and 4 litters at E15.5) and one embryo per litter was used for melanoblast isolation. Trunk melanoblasts were isolated from a preparation that excluded the head, arms, legs and internal organs, but retained the dermis, epidermis, spinal cord and associated nerve and muscle tissue. Enzymatic disassociation of melanoblasts was as follows: tissue was cut finely and digested with an incubation of 30 minutes at 37°C with a digestion media of RPMI media and 200 U/mL Liberase TL (Roche). Following digestion tissue was kept cold on ice. Physical disassociation of melanoblasts was performed using 50µm green medicons (BD Bioscience), E11.5 and E15.5 embryos processed on the medimachine for 1.5 mins and 2 mins respectively. Cells resuspended in a media/10% fetal bovine serum/Dnase I and spun at 300xg for 5 mins. Cells were washed in a solution of PBSA (PBS and 1% Bovine Serum albumin) and filtered in a 30µm filter read for cell sorting. Cells were incubated for 30 minutes with LIVE/DEAD Violet stain (thermofisher) as per manufactuer's instructions, or were stained with DAPI immediately prior to sorting.

For Lung metastasis preparation: 3-4 lungs were used per time point (1-, 2-, 3-weeks post injection). Each lung was perfused with PBS via the circulation and cut into small pieces. A solution of 0.5 µg/µl Liberase TL (Roche) in media (RPMI/DMEM) was added to the tissue chunks in a Miltenyi C tube, chunks were processed on the Miltenyi tumor dissociator then incubated at 37°C for 30-60 min. The chunks were processed on the tumor dissociator once more then transferred to a solution of RPMI containing 20% FBS and 0.5-mg/ml DNase (Sigma). Cells were filtered using a 70µm filter and the pellet resuspended in ACK lysis buffer for 5-10 minutes. Cells

were washed, filtered and stained with LIVE/DEAD Violet stain (thermofisher) as per manufactuer's instructions, or were stained with DAPI immediately prior to sorting.

Oligo-labeled antibody staining was done on both melanoblasts and melanom cells using the TotalSeq™-A Antibodies and Cell Hashing with 10x Single Cell 3' Reagent Kit (Biolegend). The HashTag Oligonucleotidess used for the analysis are mHTO01, mHTO02, mHTO04, and mHTO05 TotalSeqA from BioLegend Staining was performed as per the manufacturer's guidelines using TruStain FcX™ PLUS (anti-mouse CD16/32) (BioLegend) as a blocking agent. For cell hashing antibodies we used 0.5μg per 1million cells, followed by 3 washes prior to 10x Genomics sequencing.

Fluorescence-activated cell sorting (FACS)

iDct-GFP-positive melanoblasts/melanocytes and eGFP-positive melanoma cells were sorted on a BD FACSAria IIu (BD Biosciences) cell sorter. Cells were initially identified on forward scatter (FSC) and side scatter (SSC). Cell doublets and aggregates were excluded using light scatter pulse width. Cells were sorted based on GFP expression and viability as determined by staining with DAPI or LIVE/DEAD Violet stain (ThermoFisher). Embryos of the same developmental age that were heterozygous for the TRE-H2B-GFP gene but lacked the Dct-rtTA gene or derived from the FVB (for melanoblasts) / C57BL6 (for melanoma studies) wildtype strain were used as negative controls. Due to low sample cell number, reanalysis of sorted cells was not usually done, but representative post-sort analyses confirmed that initial frequencies of 0.6–4.4% GFP-positive cells were enriched up to 97% following the sort.

Single Cell – Partitioning and Library Preparation

Prepared antibody-label multiplexed single cell suspensions were washed with cold PBS with 0.04% BSA by centrifugation at 300g and gentle resuspension in fresh buffer. Cell counts and viability measurement were performed on each cell suspension using an automated cell counter with propidium iodide and acridine orange dyes (LunaFL, Logos Biosystem). Cell concentrations were adjusted for loading onto the 10x Genomics Chromium platform using the 3' v3 gene expression chemistry (10x Genomics) targeting approximately 6,000 cells per sample. Single cell RNA-Seq libraries were prepared according to vendor recommendations.

Single Cell – Sequencing

Sequencing was performed on the Illumina NextSeq 550 (for embryonic melanoblasts samples) or NextSeq 2000 (for melanoma metastasis samples), with separate sequencing runs for gene expression libraries and for associated antibody-oligo libraries. Single cell gene expression libraries were sequenced with a 28bp read to identify cell barcodes and unique molecular indices, an 8bp read for sample indices, and a 98bp read to identify cDNA insert. For antibody-oligo libraries, a shorter insert read was used. Samples were multiplexed for sequencing based on experimental groups, and reads were combined from multiple sequencing runs to achieve a target greater than 50,000 reads per cell on average for gene expression and at least 10,000 reads per cell for antibody-oligo libraries for all samples. >4000 melanoblasts were isolated collectively from the multiple embryo donors at each stage. Sequencing depth was excellent, >71,000 mean reads per cell, yielding >12,000 median Unique Molecular Identifiers (UMI) per cell and identifying > 3,700 median genes per cell.

Single Cell – Data Processing

Data was processed using the 10x Genomics cellranger pipeline to demultiplex reads and then align reads to a GRCh38 reference for scRNA-Seq data (cellranger v3.1.0 against reference refdata-cellranger-mm10-3.0.0 (for embryonic melanoblasts samples) and cellranger v5.0.1 against refdata-gex-mm10-

2020-A (for melanoma metastasis samples). UMI-adjusted aligned reads were used to generate a single cell barcode and gene expression matrix and antibody-oligo count matrix that was used for downstream analysis.

Processing of embryonic melanoblast data :

Reads were aligned to a custom CellRanger (v6.1.0) reference generated using the mm10 genome assembly with the sequence of the eGFP construct added as a separate chromosome. Cells with more than 15% of reads aligning to mitochondrial genes were discarded. For each embryonic time-point, read counts were normalized using the NormalizeData function in Seurat (PMID: 31178118) (v4.0.1) with default parameters. The top 10000 most variable features were chosen to carry out PCA, and the first 50 PCs were used to cluster cells using the FindNeighbors function (at a resolution of 0.1) and to generate a UMAP embedding of cells. DoubletFinder (PMID: 30954475) was used (with parameters pN = 0.25, and the first 50 PCs) to detect putative doublet cells that were then removed from further analyses. To derive regulons, we ran SCENIC (PMCID: PMC5937676) with its default parameter settings. We normalized single-cell regulon activity scores as described in the section on gene set activity normalization, and retained regulons with a mean normalized activity score of 0.1 for further analyses. To derive cluster-specific regulons, we averaged normalized gene set activity scores across cells in each cluster and z-scored the mean activity scores across clusters. The top ten differentially active regulons (with a mean activity z-score > 1) in each cluster were chosen as cluster-specific regulons.

We ran Palantir (PMID: 30899105) at each embryonic time-point separately based on the PCA coordinates of each cell as computed by Seurat. Diffusion maps were computed using the first five diffusion components and 30 nearest neighbors. A notochord cell and an SCP.2 cell were set as the start cell, with the exact cell in each cluster chosen such that it had the highest activity of its respective DGM. At each time-point, Palantir yielded an N x N transition probability matrix where N is the number of cells in each time-point. For each pair of clusters, we computed the median probability of transition across cells in both clusters for constructing an overall cell type transition matrix.

Derivation of developmental gene modules :

We merged cells across E11.5 and E15.5 time-points and inferred differentially expressed genes for each cluster using the FindAllMarkers function in Seurat, which returned genes with an average log2-fold change of at least 0.25. The MAST test was used to estimate p-values with the cell cycle score of each cell being considered as a co-variate. Genes with an adjust p-value less than 0.1 were retained. Amongst the up-regulated genes in each cluster C, we construct a gene set G_C by discarding all cell cycle genes along with any genes whose expression has a Spearman correlation coefficient of at least 0.6 with cell cycle genes.

We then performed a lineage-based correlation filtering to retain co-expressed genes within each cluster C as follows. We grouped clusters into four lineages : SCP – SCP.1, SCP.2, SCP.Early, Mes

– Mes-1, Mes-2, Mes-3, Mes-Early, Mel – Mel, Mel-Early, Neural – NPC1, NPC2, Notochord, NT-Early, Sensory (immune cluster is not considered for further analysis). For a given cluster C belonging to a lineage L, we compute the co-expression score (Spearman correlation) between every gene pair in G_C across cells in C, and repeat this process for every other cluster C'. If the maximum co-expression score of a gene pair across clusters in L is less than the maximum observed in other lineages, then that gene pair is discarded. Amongst genes in the retained gene pairs, we retain at most 200 genes (based on their average log2-fold change) as the DGM of that cluster.

Normalization of gene set activity scores :

Single-cell RNA-seq data : We used AUCell [PMID: 28991892] to compute raw gene set activity scores for a given collection of gene sets (DGMs or otherwise). While AUCell can perform thresholding to determine cells where a gene set is significantly active, we followed a different normalization approach that better accounts for changes in raw activity scores due to variation in library sizes between batches of cells [PMID: 34049974]. Briefly, we bin the mean expression of all genes in a given batch of cells into 20 bins. For each foreground gene set, we select 100 control gene sets from the union of genes in all sets such that the bin-wise distribution of genes in each control set matches the foreground. We then score the activity of all 100 control gene sets, and divide the activity of the foreground gene set in each cell with the 95th percentile of control gene set activity scores to obtain the normalized activity of the foreground gene set. Normalized activities greater than 1 represent an active gene set.

Bulk RNA-seq data : For a given cohort of bulk RNA-seq samples, count data is transformed into a log-TPM scale. The mean log-TPM score of all foreground gene sets is first computed. Each gene in the collection is assigned to one of 20 bins based on its mean expression. A set of 100 control gene sets are constructed consisting of genes from the union of all gene sets as described above. The mean expression of each control gene set is computed, and the 95^{th}percentile is subtracted from the activity of the foreground gene set to obtain the normalized activity score.

Estimating background distribution of hybrid states :

As each cell is a vector of DGM activities swith each pair of DGM activities bearing some correlation with each other, we fit a multivariate normal distribution to each cluster of cells. We then generated N random DGM activity vectors from this distribution, where N is the number of cells in the cluster, and labelled each vector by our hybrid state classification. This is done separately for each cluster, and the frequency of cells in each hybrid state is calculated. This entire process is repeated 100 times in order to obtain the mean and standard deviation of the frequency of each hybrid state, which is then used to compute the p-value of the observed frequency of each hybrid state.

Processing of metastatic single-cell RNA-seq data :

Reads were aligned to a custom CellRanger (v6.1.0) reference generated using the hg38 human genome assembly with the sequence of the eGFP construct added as a separate chromosome. Cells with more than 15% of reads aligning to mitochondrial genes were discarded. For each

metastatic time-point, read counts were normalized using the NormalizeData function in Seurat (PMID: 31178118) (v4.0.1) with default parameters. DoubletFinder (PMID: 30954475) was used (with parameters pN = 0.25, and the first 30 PCs) to detect putative doublet cells that were then removed from further analyses. The SCTransform method (PMID: 31870423) was used to perform integration and batch correction (based on first 20 PCs with k.score = 20 and k.weight = 20). Clustering was performed on the integrated data with a resolution of 0.2, and based on the first 30 PCs.
We performed three steps of computation to separate malignant from non-malignant cells. First, we ran inferCNV (inferCNV of the Trinity CTAT Project. https://github.com/broadinstitute/inferCNV) to infer the number of gene-level deletions and amplifications of cells in each cluster. We then scored each cluster using AUCell with signature genes of every cell type in the healthy mouse lung derived from droplet scRNA-seq data from the Tabula Muris project (PMID: 30283141). Finally, we counted the number of eGFP reads from each cluster. Clusters which had a high CNV burden, did not score highly on markers of healthy lung cell types and contained eGFP reads were declared as malignant cells.
Waddington-OT (PMID: 30712874) was run to compute likely transitions between all cells across time-points. Transport maps were computed by setting epsilon = 0.05, lambda1=1, lambda2=50 and growth_iters=10. The transport map at the 3 week time-point consisted of a matrix of size $(N_1+N_2)xN_3$, where N_1, N_2 and N_3 were the numbers of cells at the one-,two- and three-week time-points. For each pair of hybrid states (S_1,S_2) at time-points (t_1,t_2), we computed the median probability of transition between all cells in S_1 to all cells in S_2 for visualizing transitions between hybrid states.

Histology and Immunolabelling:
All histology was undertaken by Histoserv, Inc. (Germantown, USA). FFPE slides were deparaffinized in graded alcohols and brought to water. For slides that required bleaching, bleaching was performed in a heated bath with 3% H2O2 in PBS for 30' at 60°C. The slides were then retrieved at 100°C for 20 minutes. After retrieval, slides were blocked with a peroxidase block, incubated with the primary antibody, and detected with an HRP-conjugated goat-anti-rabbit secondary. HRP was then visualized with either DAB or Vina Green chromogens. For multiplex slides, the staining was repeated on the same slides, beginning with another round of retrieval, following the same detection protocol, and completed by visualizing the staining with an alternative chromogen. Upon staining completion, slides were dehydrated and cleared in rapid changes of graded alcohols and xylene (or clearite for Vina Green) and mounted with permanent mounting medium. Primary antibodies, concentrations used, and hosts: Sox10 - Cellmarque #383R, 1:300; Rabbit host; Col VI - Novus #NB120-6588, 12ug/ml; Rabbit host; GFP - Genetex #20290, 1:400; Rabbit host; Zeb1 - Cell Signaling #70512, 1:100; Rabbit host.

Acknowledgments

This research was supported by the NCI Intramural Research Program of the NIH (all authors), NCI Director's Innovation award (KLM), The Wellcome Trust [204796/Z/16/Z] (KLM) and The University of Manchester FBMH Dean's Prize award (KLM). We thank the NCI LGI Flow Cytometry Core: Cell sorting

and flow cytometry analyses were performed at the NIH, NCI LGI Flow Cytometry Core supported by funds from the Center for Cancer Research. We acknowledge support from CCR Single Cell Analysis Facility for single cell RNA-Seq captures, library preparation, sequencing and primary data processing was funded by FNLCR Contract 75N91019D00024. This work utilized the computational resources of the NIH HPC Biowulf cluster. (http://hpc.nih.gov). We thank Histoserv, Inc. (Germantown, USA) for expert advise and histology and immunolabeling of samples. We acknowledge Kun Wang for providing re-processed immune checkpoint inhibition RNA-seq data and single-cell RNA-seq data, and Rachel Belote for single-cell RNA-seq data of human melanocytes. We thank Andrew Macdonald, Constantinos Stathopoulos, Adam Hurlstone, and Bill Pavan for useful discussions. This work used the computational resources of the NIH High Performance Computing Biowulf cluster.

References

1. Tsoi, J. *et al.* Multi-stage Differentiation Defines Melanoma Subtypes with Differential Vulnerability to Drug-Induced Iron-Dependent Oxidative Stress. *Cancer Cell* **33**, 890-904 e895 (2018).
2. Rambow, F., Marine, J.C. & Goding, C.R. Melanoma plasticity and phenotypic diversity: therapeutic barriers and opportunities. *Genes Dev* **33**, 1295-1318 (2019).
3. Soldatov, R. *et al.* Spatiotemporal structure of cell fate decisions in murine neural crest. *Science* **364** (2019).
4. Mayor, R. & Theveneau, E. The neural crest. *Development* **140**, 2247-2251 (2013).
5. Adameyko, I. *et al.* Sox2 and Mitf cross-regulatory interactions consolidate progenitor and melanocyte lineages in the cranial neural crest. *Development* **139**, 397-410 (2012).
6. Rambow, F. *et al.* Toward Minimal Residual Disease-Directed Therapy in Melanoma. *Cell* **174**, 843-855 e819 (2018).
7. Li, A. *et al.* Rac1 drives melanoblast organization during mouse development by orchestrating pseudopod- driven motility and cell-cycle progression. *Dev Cell* **21**, 722-734 (2011).
8. Zaidi, M.R. *et al.* Interferon-gamma links ultraviolet radiation to melanomagenesis in mice. *Nature* **469**, 548-553 (2011).
9. Belote, R.L. *et al.* Human melanocyte development and melanoma dedifferentiation at single-cell resolution. *Nat Cell Biol* **23**, 1035-1047 (2021).
10. Marie, K.L. *et al.* Melanoblast transcriptome analysis reveals pathways promoting melanoma metastasis. *Nat Commun* **11**, 333 (2020).
11. Bagaev, A. *et al.* Conserved pan-cancer microenvironment subtypes predict response to immunotherapy. *Cancer Cell* **39**, 845-865 e847 (2021).
12. Jerby-Arnon, L. *et al.* A Cancer Cell Program Promotes T Cell Exclusion and Resistance to Checkpoint Blockade. *Cell* **175**, 984-997 e924 (2018).

13. Carreira, S. *et al.* Mitf regulation of Dia1 controls melanoma proliferation and invasiveness. *Genes Dev* **20**, 3426-3439 (2006).
14. Perez-Guijarro, E. *et al.* Multimodel preclinical platform predicts clinical response of melanoma to immunotherapy. *Nat Med* **26**, 781-791 (2020).
15. Joshi, S.S. *et al.* Characterization of a new, inducible transgenic mouse model with GFP expression in melanocytes and their precursors. *Gene Expr Patterns* **27**, 76-84 (2018).
16. Mort, R.L., Jackson, I.J. & Patton, E.E. The melanocyte lineage in development and disease. *Development* **142**, 1387 (2015).
17. McMenamin, P.G. *et al.* Melanoblasts Populate the Mouse Choroid Earlier in Development Than Previously Described. *Invest Ophthalmol Vis Sci* **61**, 33 (2020).
18. Schock, E.N. & LaBonne, C. Sorting Sox: Diverse Roles for Sox Transcription Factors During Neural Crest and Craniofacial Development. *Front Physiol* **11**, 606889 (2020).
19. Latremoliere, A. *et al.* Neuronal-Specific TUBB3 Is Not Required for Normal Neuronal Function but Is Essential for Timely Axon Regeneration. *Cell Rep* **24**, 1865-1879 e1869 (2018).
20. Boshuizen, J. *et al.* Reversal of pre-existing NGFR-driven tumor and immune therapy resistance. *Nat Commun* **11**, 3946 (2020).
21. Nataf, V., Lecoin, L., Eichmann, A. & Le Douarin, N.M. Endothelin-B receptor is expressed by neural crest cells in the avian embryo. *Proc Natl Acad Sci U S A* **93**, 9645-9650 (1996).
22. Aubin-Houzelstein, G. *et al.* Melanoblasts' proper location and timed differentiation depend on Notch/RBP-J signaling in postnatal hair follicles. *J Invest Dermatol* **128**, 2686-2695 (2008).
23. Osawa, M. & Fisher, D.E. Notch and melanocytes: diverse outcomes from a single signal. *J Invest Dermatol* **128**, 2571-2574 (2008).
24. Yang, Y., Xin, X., Fu, X. & Xu, D. Expression pattern of human SERPINE2 in a variety of human tumors. *Oncol Lett* **15**, 4523-4530 (2018).
25. Setty, M. *et al.* Characterization of cell fate probabilities in single-cell data with Palantir. *Nat Biotechnol* **37**, 451-460 (2019).
26. Jessen, K.R. & Mirsky, R. Schwann Cell Precursors; Multipotent Glial Cells in Embryonic Nerves. *Front Mol Neurosci* **12**, 69 (2019).
27. Cancer Genome Atlas, N. Genomic Classification of Cutaneous Melanoma. *Cell* **161**, 1681-1696 (2015).
28. Jiang, P. *et al.* Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat Med* **24**, 1550-1558 (2018).
29. Capparelli, C. *et al.* Targeting SOX10-deficient cells to reduce the dormant-invasive phenotype state in melanoma. *Nat Commun* **13**, 1381 (2022).
30. Zhu, C., Wei, Y. & Wei, X. AXL receptor tyrosine kinase as a promising anti-cancer approach: functions, molecular mechanisms and clinical applications. *Mol Cancer* **18**, 153 (2019).
31. Wang, K. *et al.* Deconvolving Clinically Relevant Cellular Immune Cross-talk from Bulk Gene Expression Using CODEFACS and LIRICS Stratifies Patients with Melanoma to Anti-PD-1 Therapy. *Cancer Discov* **12**, 1088-1105 (2022).
32. Van Allen, E.M. *et al.* Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science* **350**, 207-211 (2015).

33. Gide, T.N. *et al.* Clinical and Molecular Heterogeneity in Patients with Innate Resistance to Anti-PD-1 +/- Anti-CTLA-4 Immunotherapy in Metastatic Melanoma Reveals Distinct Therapeutic Targets. *Cancers (Basel)* **13** (2021).
34. Liu, D. *et al.* Integrative molecular and clinical modeling of clinical outcomes to PD1 blockade in patients with metastatic melanoma. *Nat Med* **25**, 1916-1927 (2019).
35. Baron, M. *et al.* The Stress-Like Cancer Cell State Is a Consistent Component of Tumorigenesis. *Cell Syst* **11**, 536-546 e537 (2020).
36. Falletta, P. *et al.* Translation reprogramming is an evolutionarily conserved driver of phenotypic plasticity and therapeutic resistance in melanoma. *Genes Dev* **31**, 18-33 (2017).
37. Garcia-Jimenez, C. & Goding, C.R. Starvation and Pseudo-Starvation as Drivers of Cancer Metastasis through Translation Reprogramming. *Cell Metab* **29**, 254-267 (2019).
38. Fidler, I.J. Selection of successive tumour lines for metastasis. *Nat New Biol* **242**, 148-149 (1973).
39. Mehrabadi, F.R. *et al.* Profiles of expressed mutations in single cells reveal subclonal expansion patterns and therapeutic impact of intratumor heterogeneity. *bioRxiv*, 2021.2003.2026.437185 (2021).
40. Tabula Muris, C. *et al.* Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. *Nature* **562**, 367-372 (2018).
41. Li, X. *et al.* Disseminated Melanoma Cells Transdifferentiate into Endothelial Cells in Intravascular Niches at Metastatic Sites. *Cell Rep* **31**, 107765 (2020).
42. Schiebinger, G. *et al.* Optimal-Transport Analysis of Single-Cell Gene Expression Identifies Developmental Trajectories in Reprogramming. *Cell* **176**, 1517 (2019).
43. Zaidi, M.R. & Merlino, G. The two faces of interferon-gamma in cancer. *Clin Cancer Res* **17**, 6118-6124 (2011).
44. Grasso, C.S. *et al.* Conserved Interferon-gamma Signaling Drives Clinical Response to Immune Checkpoint Blockade Therapy in Melanoma. *Cancer Cell* **38**, 500-515 e503 (2020).
45. Sun, S.C. The non-canonical NF-kappaB pathway in immunity and inflammation. *Nat Rev Immunol* **17**, 545-558 (2017).
46. Kim, Y.J. *et al.* Melanoma dedifferentiation induced by IFN-gamma epigenetic remodeling in response to anti-PD-1 therapy. *J Clin Invest* **131** (2021).
47. Kinker, G.S. *et al.* Pan-cancer single-cell RNA-seq identifies recurring programs of cellular heterogeneity. *Nat Genet* **52**, 1208-1218 (2020).
48. Hersey, P. & Zhang, X.D. Adaptation to ER stress as a driver of malignancy and resistance to therapy in human melanoma. *Pigment Cell Melanoma Res* **21**, 358-367 (2008).
49. Huang, S. On the intrinsic inevitability of cancer: from foetal to fatal attraction. *Semin Cancer Biol* **21**, 183-199 (2011).
50. Couturier, C.P. *et al.* Single-cell RNA-seq reveals that glioblastoma recapitulates a normal neurodevelopmental hierarchy. *Nat Commun* **11**, 3406 (2020).
51. Tirosh, I. *et al.* Single-cell RNA-seq supports a developmental hierarchy in human oligodendroglioma. *Nature* **539**, 309-313 (2016).
52. Bedoya-Reina, O.C. *et al.* Single-nuclei transcriptomes from human adrenal gland reveal distinct cellular identities of low and high-risk neuroblastoma tumors. *Nat Commun* **12**, 5309 (2021).

53. Laughney, A.M. *et al.* Regenerative lineages and immune-mediated pruning in lung cancer metastasis. *Nat Med* **26**, 259-269 (2020).
54. Young, M.D. *et al.* Single-cell transcriptomes from human kidneys reveal the cellular identity of renal tumors. *Science* **361**, 594-599 (2018).
55. Karras, P. *et al.* A cellular hierarchy in melanoma uncouples growth and metastasis. *Nature* **610**, 190-198 (2022).
56. Aiello, N.M. & Stanger, B.Z. Echoes of the embryo: using the developmental biology toolkit to study cancer. *Dis Model Mech* **9**, 105-114 (2016).
57. Huser, L. *et al.* HER3-Receptor-Mediated STAT3 Activation Plays a Central Role in Adaptive Resistance toward Vemurafenib in Melanoma. *Cancers (Basel)* **12** (2020).

Figure 1

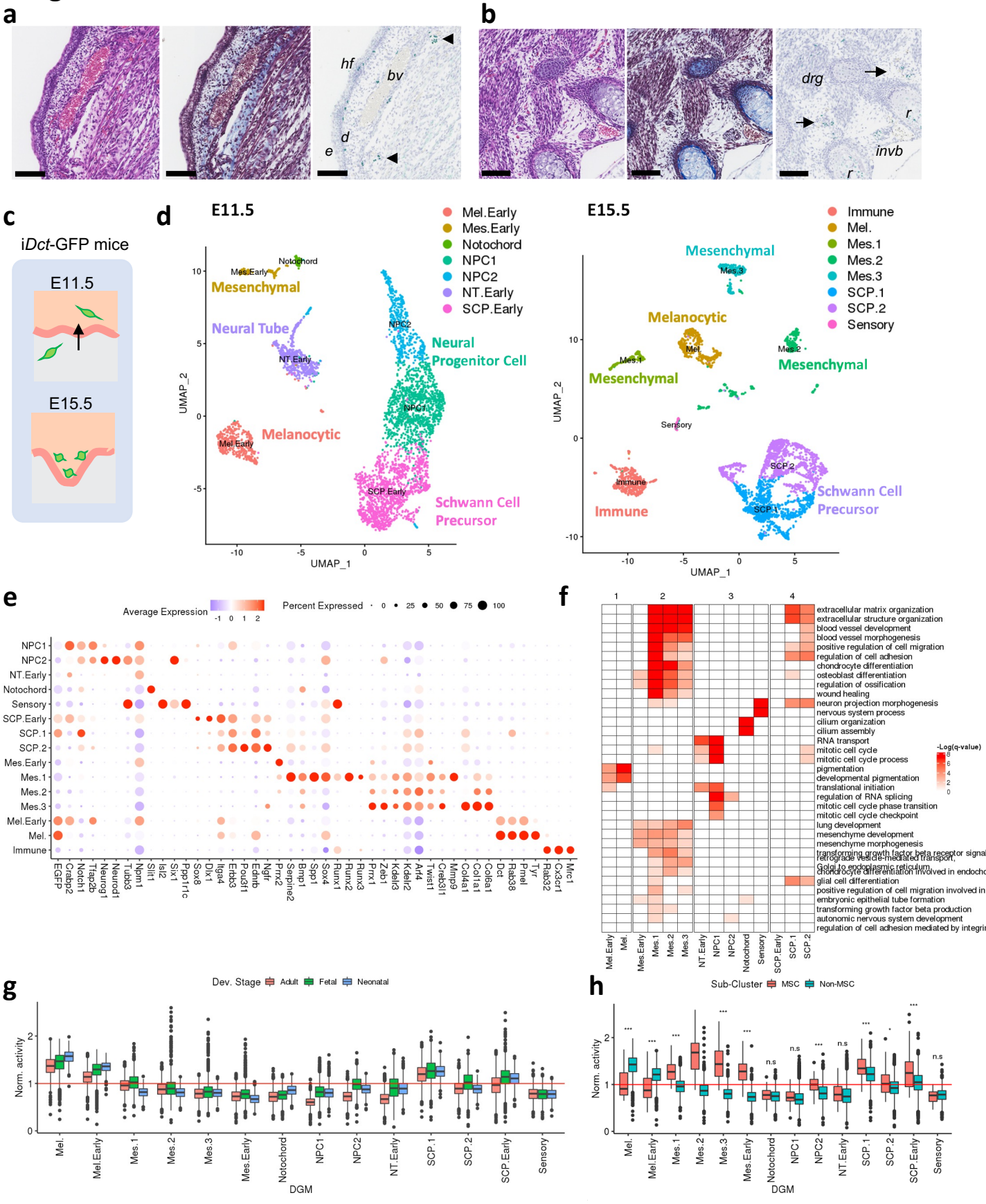
a
hf
bv
d
e
b
drg
r
invb
r
c
iDct-GFP mice
E11.5
E15.5
d
E11.5
Mel.Early
Mes.Early
Notochord
NPC1
NPC2
NT.Early
SCP.Early
Mesenchymal
Neural Tube
Neural Progenitor Cell
Melanocytic
Schwann Cell Precursor
UMAP_1
UMAP_2
E15.5
Immune
Mel.
Mes.1
Mes.2
Mes.3
SCP.1
SCP.2
Sensory
Mesenchymal
Melanocytic
Mesenchymal
Mesenchymal
Immune
Schwann Cell Precursor
e
Average Expression
Percent Expressed
NPC1
NPC2
NT.Early
Notochord
Sensory
SCP.Early
SCP.1
SCP.2
Mes.Early
Mes.1
Mes.2
Mes.3
Mel.Early
Mel.
Immune
EGFP
Crabp2
Notch1
Ttap2b
Neurog1
Neurod1
Tubb3
Npm1
Slit1
Isl2
Six1
Ppp1r1c
Sox8
Dlx1
Itga4
Erbb3
Pou3f1
Ednrb
Ngfr
Prrx2
Serpine2
Bmp1
Spp1
Sox4
Runx1
Runx2
Runx3
Prrx1
Zeb1
Kdelr3
Kdelr2
Arf4
Twist1
Creb3l1
Mmp9
Col4a1
Col1a1
Col6a1
Dct
Rab38
Pmel
Tyr
Rab32
Cx3cr1
Mrc1
f
extracellular matrix organization
extracellular structure organization
blood vessel development
blood vessel morphogenesis
positive regulation of cell migration
regulation of cell adhesion
chondrocyte differentiation
osteoblast differentiation
regulation of ossification
wound healing
neuron projection morphogenesis
nervous system process
cilium organization
cilium assembly
RNA transport
mitotic cell cycle
mitotic cell cycle process
pigmentation
developmental pigmentation
translational initiation
regulation of RNA splicing
mitotic cell cycle phase transition
mitotic cell cycle checkpoint
lung development
mesenchyme development
mesenchyme morphogenesis
transforming growth factor beta receptor signal
retrograde vesicle-mediated transport, Golgi to endoplasmic reticulum
chondrocyte differentiation involved in endocho
glial cell differentiation
positive regulation of cell migration involved in
embryonic epithelial tube formation
transforming growth factor beta production
autonomic nervous system development
regulation of cell adhesion mediated by integrin
-Log(q-value)
Mel.Early
Mel.
Mes.Early
Mes.1
Mes.2
Mes.3
NT.Early
NPC1
NPC2
Notochord
Sensory
SCP.Early
SCP.1
SCP.2
g
Dev. Stage
Adult
Fetal
Neonatal
Norm. activity
DGM
h
Sub-Cluster
MSC
Non-MSC
Norm. activity
DGM

Figure 2

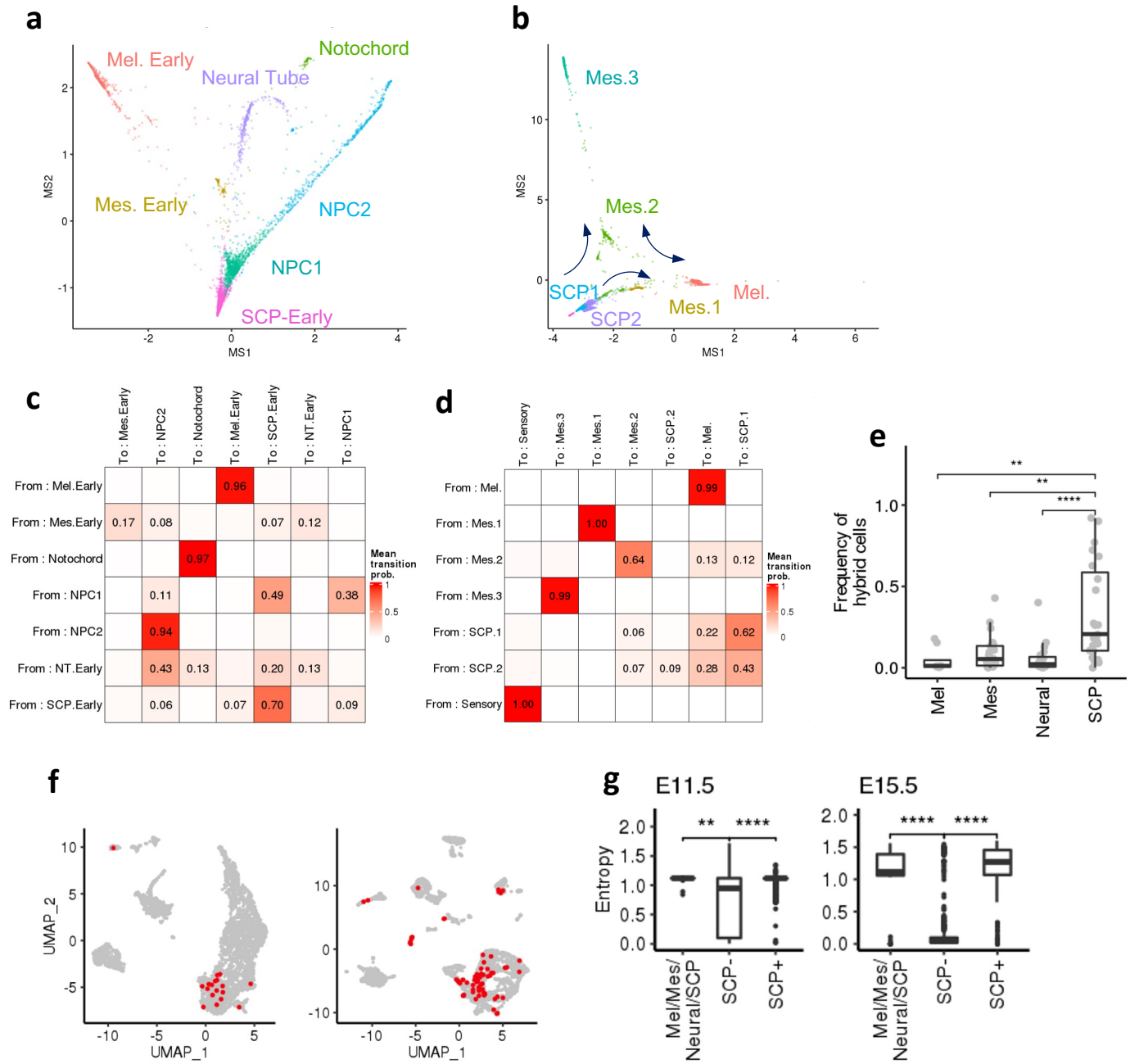
a
Mel. Early
Notochord
Neural Tube
Mes. Early
NPC2
NPC1
SCP-Early
MS1
MS2
b
Mes.3
Mes.2
SCP1
SCP2
Mes.1
Mel.
MS1
MS2
c
To : Mes.Early
To : NPC2
To : Notochord
To : Mel.Early
To : SCP.Early
To : NT.Early
To : NPC1
From : Mel.Early
From : Mes.Early
From : Notochord
From : NPC1
From : NPC2
From : NT.Early
From : SCP.Early
Mean transition prob.
d
To : Sensory
To : Mes.3
To : Mes.1
To : Mes.2
To : SCP.2
To : Mel.
To : SCP.1
From : Mel.
From : Mes.1
From : Mes.2
From : Mes.3
From : SCP.1
From : SCP.2
From : Sensory
Mean transition prob.
e
Frequency of hybrid cells
Mel
Mes
Neural
SCP
f
UMAP_1
UMAP_2
g
E11.5
E15.5
Entropy
Mel/Mes/ Neural/SCP
SCP-
SCP+

Figure 3

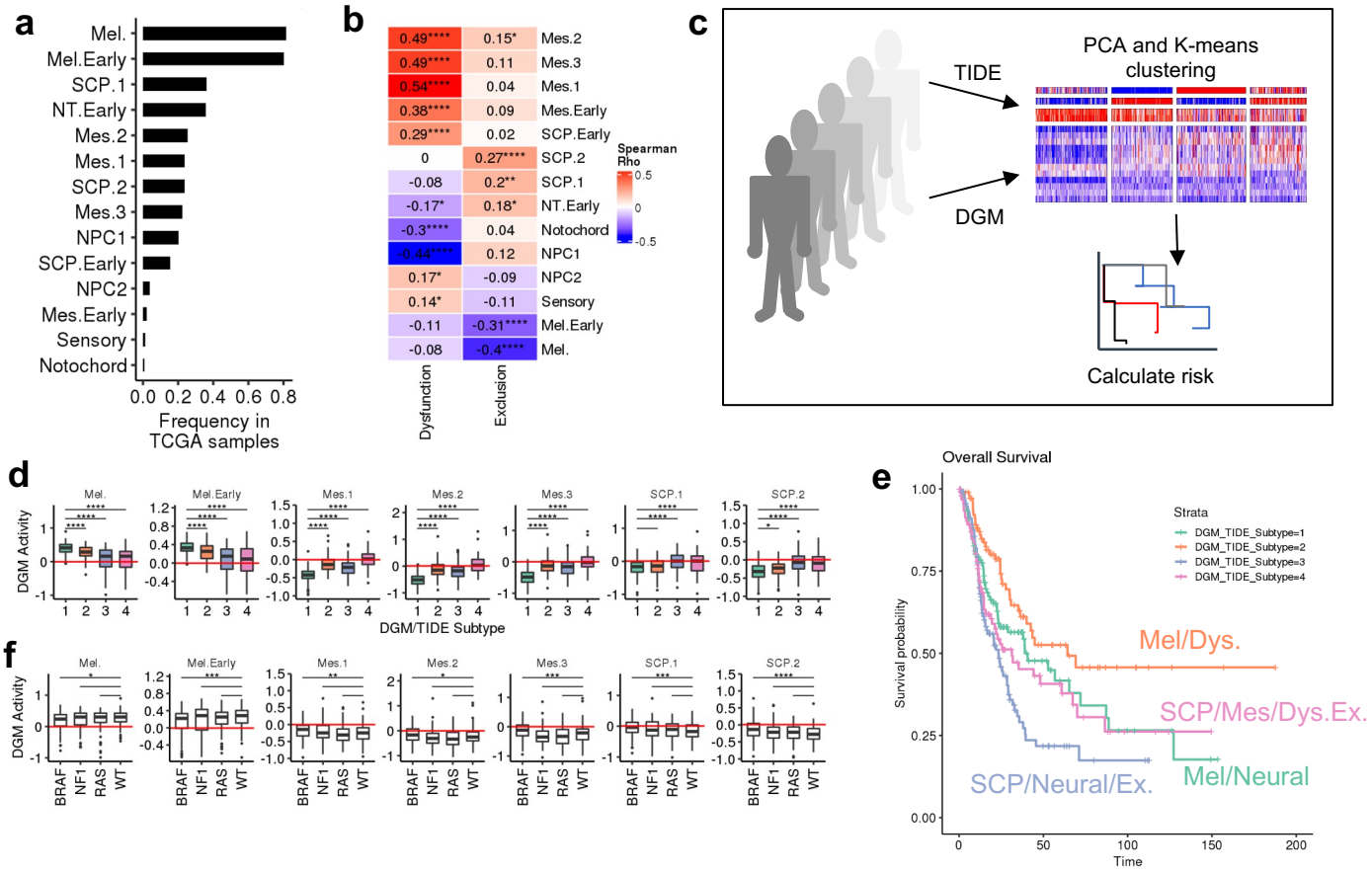
a
Mel.
Mel.Early
SCP.1
NT.Early
Mes.2
Mes.1
SCP.2
Mes.3
NPC1
SCP.Early
NPC2
Mes.Early
Sensory
Notochord
0.0 0.2 0.4 0.6 0.8
Frequency in TCGA samples
b
Dysfunction
Exclusion
Spearman Rho
c
TIDE
DGM
PCA and K-means clustering
Calculate risk
d
DGM Activity
DGM/TIDE Subtype
e
Overall Survival
Strata
DGM_TIDE_Subtype=1
DGM_TIDE_Subtype=2
DGM_TIDE_Subtype=3
DGM_TIDE_Subtype=4
Survival probability
Time
Mel/Dys.
SCP/Mes/Dys.Ex.
Mel/Neural
SCP/Neural/Ex.
f
BRAF
NF1
RAS
WT

Figure 4

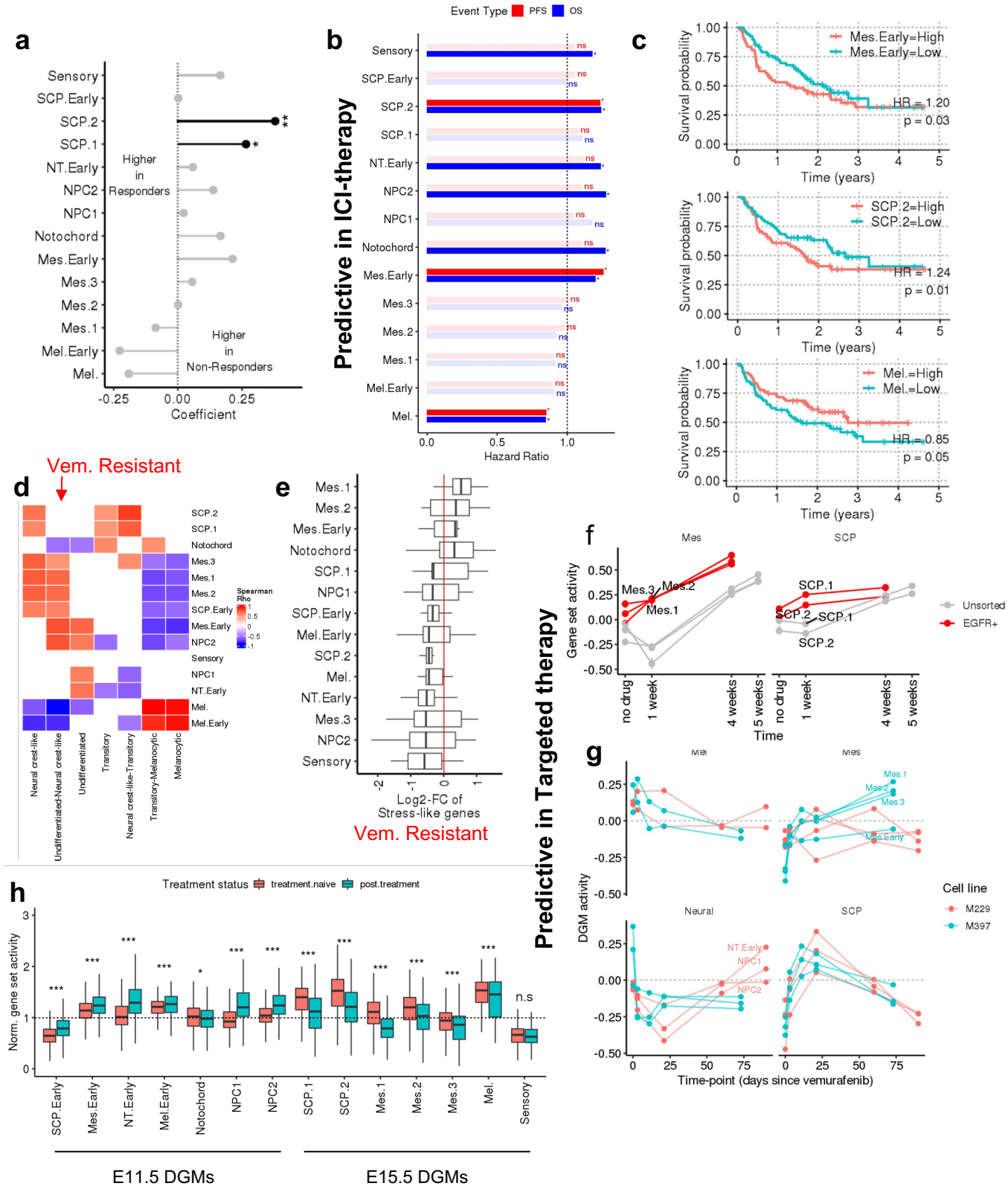
a
Sensory
SCP.Early
SCP.2
SCP.1
NT.Early
NPC2
NPC1
Notochord
Mes.Early
Mes.3
Mes.2
Mes.1
Mel.Early
Mel.
Higher in Responders
Higher in Non-Responders
-0.25
0.00
0.25
Coefficient
b
Event Type PFS OS
Predictive in ICI-therapy
ns
0.0
0.5
1.0
Hazard Ratio
c
Mes.Early=High
Mes.Early=Low
HR = 1.20
p = 0.03
SCP.2=High
SCP.2=Low
HR = 1.24
p = 0.01
Mel.=High
Mel.=Low
HR = 0.85
p = 0.05
Survival probability
Time (years)
d
Vem. Resistant
Spearman Rho
Neural crest-like
Undifferentiated-Neural crest-like
Undifferentiated
Transitory
Neural crest-like-Transitory
Transitory-Melanocytic
Melanocytic
e
Log2-FC of Stress-like genes
Vem. Resistant
Predictive in Targeted therapy
f
Mes
SCP
Gene set activity
Unsorted
EGFR+
no drug
1 week
4 weeks
5 weeks
Time
g
Mel
Mes
Neural
SCP
DGM activity
Cell line
M229
M397
Time-point (days since vemurafenib)
h
Treatment status treatment.naive post.treatment
Norm. gene set activity
n.s
E11.5 DGMs
E15.5 DGMs

Figure 5

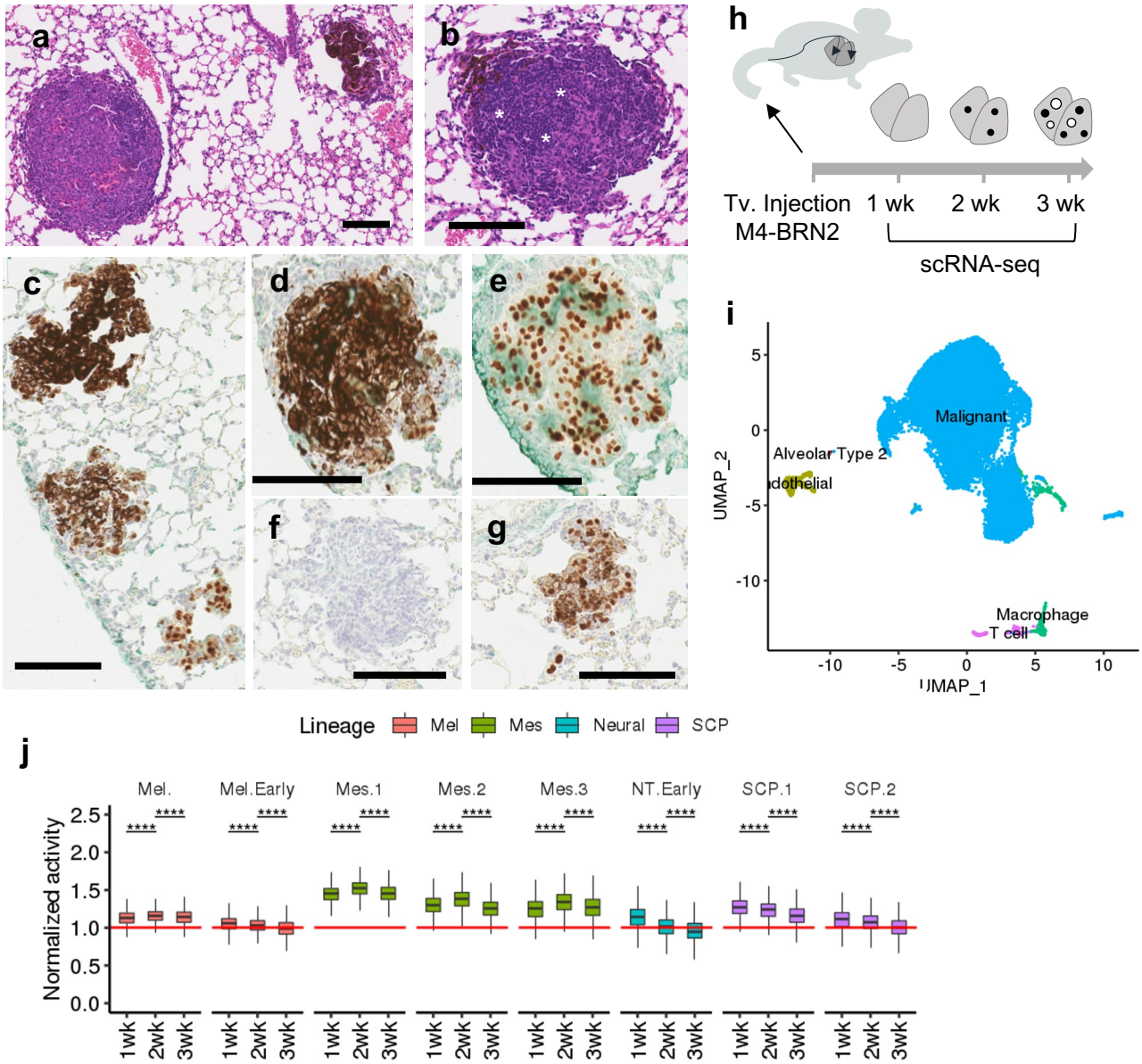
a
b
c
d
e
f
g
h
Tv. Injection
M4-BRN2
1 wk
2 wk
3 wk
scRNA-seq
i
Malignant
Alveolar Type 2
Macrophage
T cell
UMAP_2
UMAP_1
Lineage
Mel
Mes
Neural
SCP
j
Mel.
Mel.Early
Mes.1
Mes.2
Mes.3
NT.Early
SCP.1
SCP.2
Normalized activity
1wk
2wk
3wk

Figure 6

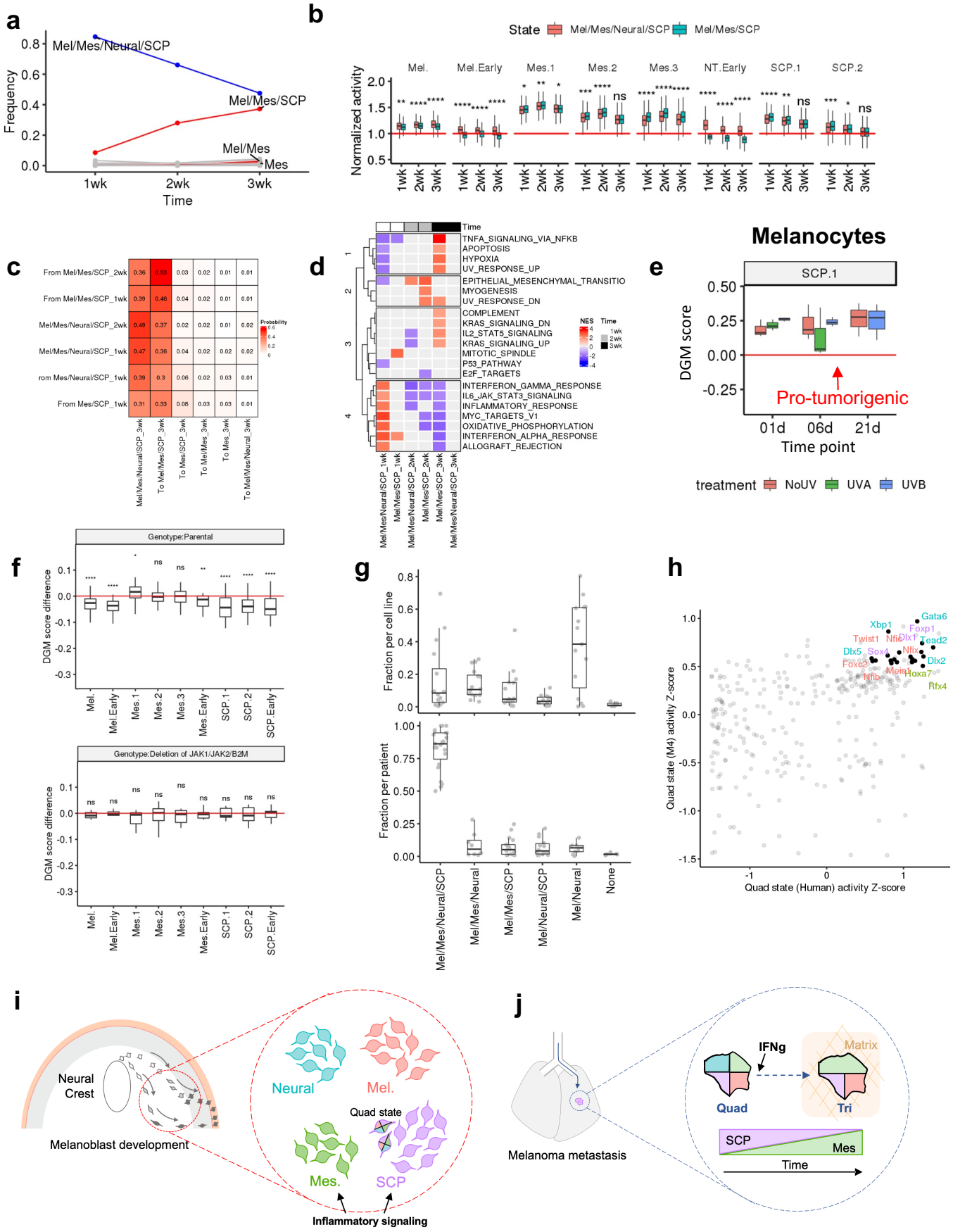
a
Mel/Mes/Neural/SCP
Mel/Mes/SCP
Mel/Mes
Mes
Frequency
Time
1wk
2wk
3wk
b
State
Mel/Mes/Neural/SCP
Mel/Mes/SCP
Mel.
Mel.Early
Mes.1
Mes.2
Mes.3
NT.Early
SCP.1
SCP.2
Normalized activity
c
From Mel/Mes/SCP_2wk
From Mel/Mes/SCP_1wk
Mel/Mes/Neural/SCP_2wk
Mel/Mes/Neural/SCP_1wk
rom Mes/Neural/SCP_1wk
From Mes/SCP_1wk
Probability
d
Time
TNFA_SIGNALING_VIA_NFKB
APOPTOSIS
HYPOXIA
UV_RESPONSE_UP
EPITHELIAL_MESENCHYMAL_TRANSITIO
MYOGENESIS
UV_RESPONSE_DN
COMPLEMENT
KRAS_SIGNALING_DN
IL2_STAT5_SIGNALING
KRAS_SIGNALING_UP
MITOTIC_SPINDLE
P53_PATHWAY
E2F_TARGETS
INTERFERON_GAMMA_RESPONSE
IL6_JAK_STAT3_SIGNALING
INFLAMMATORY_RESPONSE
MYC_TARGETS_V1
OXIDATIVE_PHOSPHORYLATION
INTERFERON_ALPHA_RESPONSE
ALLOGRAFT_REJECTION
NES
e
Melanocytes
SCP.1
DGM score
Pro-tumorigenic
01d
06d
21d
Time point
treatment
NoUV
UVA
UVB
f
Genotype:Parental
Genotype:Deletion of JAK1/JAK2/B2M
DGM score difference
g
Fraction per cell line
Fraction per patient
Mel/Mes/Neural/SCP
Mel/Mes/Neural
Mel/Mes/SCP
Mel/Neural/SCP
Mel/Neural
None
h
Quad state (M4) activity Z-score
Quad state (Human) activity Z-score
Gata6
Xbp1
Foxp1
Twist1
Dlx1
Tead2
Dlx5
Sox4
Nfix
Foxc2
Dlx2
Meis1
Nfib
Hoxa7
Rfx4
i
Neural Crest
Melanoblast development
Neural
Mel.
Quad state
Mes.
SCP
Inflammatory signaling
j
Melanoma metastasis
IFNg
Matrix
Quad
Tri
SCP
Mes
Time

Table 1

Pathway	pval	padj	NES	Cluster
TNFA/NFKB SIGNALING	0.001233046	0.01541307	2.365452708	SCP.1
TNFA/NFKB SIGNALING	0.001574803	0.008650519	2.350980028	Mes.1
INFLAMMATORY RESPONSE	0.001730104	0.008650519	2.144800115	Mes.1
IFNg RESPONSE	0.003361345	0.012928248	2.057876541	Mes.1
IL2/STAT5 SIGNALING	0.001605136	0.008650519	1.963255602	Mes.1
IFNg RESPONSE	0.002989537	0.017415535	1.920944316	Mes.3
INFLAMMATORY RESPONSE	0.00887574	0.028153153	1.809053276	Mes.3
INFLAMMATORY RESPONSE	0.014588859	0.065445026	1.734830081	SCP.1
IL2/STAT5 SIGNALING	0.017857143	0.089285714	1.661156276	Mes.2
IFNa RESPONSE	0.033123028	0.078864353	1.62836093	Mes.3
IL2/STAT5 SIGNALING	0.031767956	0.078864353	1.586483417	Mes.3
IL2/STAT5 SIGNALING	0.031553398	0.105177994	1.572688409	SCP.1
TNFA/NFKB SIGNALING	0.039641944	0.137634409	1.551010203	Mes.2
IL6/JAK/STAT3 SIGNALING	0.065989848	0.150045363	1.504197763	Mes.1
INFLAMMATORY RESPONSE	0.058676654	0.12224303	-1.563175511	NT.Early
TNFA/NFKB SIGNALING	0.042091837	0.140306122	-1.582729511	Mes.Early
TNFA/NFKB SIGNALING	0.037688442	0.206922498	-1.587610361	Notochord
IL2/STAT5 SIGNALING	0.038461538	0.206922498	-1.594923921	Notochord
IL2/STAT5 SIGNALING	0.025510204	0.091107872	-1.627645051	Mes.Early
IFNa RESPONSE	0.040944882	0.081889764	-1.631557133	NPC1
IL6/JAK/STAT3 SIGNALING	0.01910828	0.079617834	-1.662310183	Mes.Early
IL2/STAT5 SIGNALING	0.006319115	0.022568269	-1.741894431	NPC2
INFLAMMATORY RESPONSE	0.015965167	0.038012302	-1.768547716	NPC1
IL2/STAT5 SIGNALING	0.002554278	0.025542784	-1.903905735	Mel.Early
TNFA/NFKB SIGNALING	0.00921659	0.035448423	-1.943208983	Sensory
IL2/STAT5 SIGNALING	0.001240695	0.006729475	-2.049190043	NPC1
IL2/STAT5 SIGNALING	0.001132503	0.008116883	-2.055336607	NT.Early
TNFA/NFKB SIGNALING	0.0015625	0.008305648	-2.132381043	NPC2
TNFA/NFKB SIGNALING	0.001124859	0.008116883	-2.404165262	NT.Early
TNFA/NFKB SIGNALING	0.001236094	0.006729475	-2.463277931	NPC1

↑ Inflammation

↓ Inflammation