Single cell analysis of human site-specific melanocyte differentiation and the decoding of developmental programs in melanoma.

Rachel L. Belote[1*], Daniel Le[2*], Ashley Maynard[2], Ursula E. Lang[3], Adriane Sinclair[4], Vicente Planells-Palop[5], Laurence Baskin[4], Aaron D. Tward[5], Spyros Darmanis[2†] and Robert L. Judson-Torres[1,6,†,‡]

[(1)] Huntsman Cancer Institute, University of Utah, Salt Lake City, UT, USA
[(2)] Chan Zuckerberg Biohub, San Francisco, CA, USA
[(3)] Department of Dermatology and Department of Pathology, University of California, San Francisco, CA, USA
[(4)] Department of Urology and Division of Pediatric Urology, University of California, San Francisco, CA, USA
[(5)] Department of Otolaryngology–Head and Neck Surgery, University of California, San Francisco, CA, USA
[(6)] Department of Dermatology, University of Utah, Salt Lake City, UT, USA

Author list footnotes:
*These authors contributed equally
† Correspondence: spyros.darmanis@czbiohub.org, judsontorreslab@gmail.com
‡ Lead contact

SUMMARY:
Epidermal melanocytes are present throughout the skin, one of the largest organs with distinct anatomical, morphological and functional characteristics. Clear differences in melanocytic disease manifestation and phenotype exist across different anatomical locations. Here, we investigate human melanocyte heterogeneity during development, homeostasis, and disease progression using single cell RNA sequencing of freshly isolated human fetal, neonatal, and adult skin from a demographically diverse cohort. Comparative analysis across developmental stages and between anatomical sites revealed distinct subclasses of melanocytes from lineages that diverge early in human development. Using differentiation programs delineated from healthy melanocytes, we identified melanoma gene expression signatures that are developmental in origin and are re-acquired during disease, signatures that are unique to melanoma progression, as well as signatures that offer prognostic value. This dataset provides a valuable resource for further investigations of the melanocytic lineage in health and disease.

KEYWORDS:
melanocyte, single cell, RNA sequencing, pigment, development, melanoma, dedifferentiation, acral, volar, cutaneous

INTRODUCTION:

Epidermal melanocytes are among the most influential and distinguishing cell types in human history and evolution. Melanocytes establish human skin, hair and eye color - physical characteristics that have shaped political, economic and socio-cultural systems and injustices. Global genome-wide association studies have provided insights into pigment regulation and variation among individuals (Adhikari et al., 2019; Crawford et al., 2017); however, pigmentation also varies dramatically among anatomical sites within a single individual (e.g., between volar (palm and sole) and non-volar sites). Yet, little is known about melanocyte-intrinsic factors driving these differences (Nakamura et al., 2015; Okamoto et al., 2014). Melanocytes exist in a variety of other anatomical locations including the cochlea, eye, heart and mucosa with some populations lacking pigment (Colombo et al., 2011). Since the widely accepted melanocyte-specific markers (SOX10, MITF, DCT, MLANA, PMEL, TYR, TYRP1) are

involved in regulating and producing pigment, it is difficult to conclusively identify and investigate melanocyte populations that are not pigmented. Characterizing the transcriptional landscape of hypo-pigmented human melanocytes, such as those found in volar skin, would provide a valuable resource for elucidating melanocyte functions beyond pigment production.

Melanocytes can undergo oncogenic transformation and give rise to melanoma, the deadliest common skin cancer, with distinct phenotypic and genomic characteristics correlated with primary tumor location (Hayward et al., 2017; Rabbie et al., 2019). Whether these differences in presentation are the consequence of distinct cells of origin is unknown. Like many cancers, it is well established that melanoma progression can reflect a dedifferentiation trajectory (Malta et al., 2018) and it is generally appreciated that a greater degree of dedifferentiation is associated with a worse prognosis (Edge and Compton, 2010). However, studies of melanocyte development and melanocytic disease progression have focused on model organisms that harbor primarily non-epidermal melanocytes; while, in humans, the predominant population is functionally-distinct resident epidermal melanocytes (Adameyko et al. 2009; Marie et al. 2020; Mort et al. 2015). Consequently, melanocyte fate specification mechanisms identified from model organisms may not be conserved in human. Likewise, studies of *in vitro* human embryonic stem cell differentiation may not be reflective of *in vivo* human melanocyte development (Mica et al. 2013; Mort et al. 2015). Therefore, *in vivo* characterization of human melanocyte heterogeneity, inclusive of developmental trajectories, would provide much needed insight into the process of dedifferentiation that occurs during melanoma progression.

Here we used a single cell enrichment and RNA-sequencing pipeline to generate the first cell atlas of human epidermal melanocytes derived directly from skin, capturing transcriptional diversity across anatomical site, development, gender, and skin tone. Using donor-matched volar and non-volar skin, we discovered a subclass of melanocytes that diverge during gestation. In addition, we identified human melanocyte differentiation transcriptional programs that are distinct from previously published gene signatures generated from model systems. Finally, we use these programs to define patterns of dedifferentiation in two *in vivo* human melanoma datasets that are predictive of prognosis.

RESULTS:

Single cell transcriptomics of normal human melanocytes across developmental stages and anatomical locations.

We performed single cell RNA sequencing on post-surgery healthy skin discards from donors aged 9.5 fetal weeks (f.w.) to 81 years and from multiple anatomical locations (leg, arm, foreskin, palm and sole) (Table 1, Figure 1A). For each specimen, the epidermis was enzymatically removed from the dermis and further dissociated into a single cell suspension. Since melanocytes comprise a small fraction of the total epidermal cell mass, FACS was used to enrich for melanocytes (alpha 6 integrin-, CKIT+, CD11C-) (Figures 1A and S1A). Sorted cells were processed using the Smartseq2 single-cell RNA-seq protocol (Picelli et al., 2013), which yielded 9,688 annotated cells: melanocytes, keratinocytes, and three immune cell populations (Figure S1, S2 and STAR Methods). Individual cells were annotated as cycling (G2 or M phase) or non-cycling based on expression of established marker genes (Hsiao et al., 2019) (Figure S3A,B, see STAR methods). As expected, cycling keratinocytes were present across all developmental stages (fetal, neonatal, and adult) (Figure S3D); whereas, the proportion of cycling melanocytes was anti-correlated with age (Figure S3E).

Non-cycling melanocytes clustered that segregated by developmental stage and anatomical location (Figure 1B, Figure S3F-H). Adult and fetal populations, which contain multiple anatomical locations, exhibited further segregation based on volar and non-volar sites (Figure 1B). Within the fetal melanocyte population, an additional cluster was identified as melanocyte stem cells (discussed in

subsequent sections, Figure S2J). Separate from all adult and fetal populations, neonatal melanocytes occupied a distinct cluster composed of a single anatomical location, foreskin (Figure 1B).

Transcriptional profiling reveals molecular signature that distinguishes melanocyte subclasses.

The anatomical location of skin influences melanocyte survival and function but it remains unclear how site-specific specialization arises during melanocyte maturation (Yamaguchi et al., 2004). One prominent site-specific skin phenotype is reduced pigmentation of volar skin compared to non-volar cutaneous skin. Thus, we queried our dataset for transcriptional programs that could distinguish volar vs non-volar cutaneous melanocytes. Based on 6 sets of donor-matched volar and non-volar cutaneous specimens that spanned 10 f.w. to 77 years (two adults and four fetal), included both genders, varied ethnicities and skin pigmentation levels (Figure 1C, Table 1), differential gene expression analysis (Mann-Whitney U test, Benjamini-Hochberg FDR < 5%) revealed 2,059 transcripts with site-specific expression (Figure 1D, Table S1). Among the most differentially expressed genes, NTRK2 (neurotrophic receptor tyrosine kinase 2) and HPGD (15-hydroxyprostaglandin dehydrogenase) presented a striking level of inverse expression between volar and non-volar cutaneous melanocytes across all donor-matched skin, regardless of developmental stage, age, gender or ethnicity (Figure 1E). Thus, we hypothesized that NTRK2+/HPGD- and HPGD+/NTRK2- were molecular signatures for two distinct subclasses of human epidermal melanocytes: volar-enriched melanocytes (v-mel) and non-volar cutaneous melanocytes (c-mel).

Using the full donor cohort (n=22 donors, 6,904 melanocytes), we assessed the expression of NTRK2 and HPGD for all melanocytes. Consistent with our hypothesis, the fraction of v-mels (NTRK2+/HPGD-) and c-mels (HPGD-/NTRK+) was site-dependent in fetal and adult skin. Quantification of the number of v-mels vs c-mels from each skin sample showed an enrichment of v-mels in volar skin (mean: 87.3% palm, 83.7% sole) whereas c-mels were enriched in non-volar cutaneous skin (mean: 88.6% arm, 87.0% leg) (Figure 1F). Neonatal foreskin was composed primarily of c-mels (mean: 89.3%). Immunofluorescence of HPGD and quantification of HPGD positive melanocytes in volar and non-volar skin validated anatomical differential expression (Figure S4A,B).

Site-specific pigment divergence during melanocyte differentiation.

Hypopigmentation of palms and soles is present in neonates and continues through adulthood, indicating that site-specific pigmentation occurs during development (Yamaguchi et al., 2004). Using FACS back-scatter (BSC) measurements indexed to each cell as a proxy for pigment content (Choi et al., 2012), we queried the differential pigmentation between donor-matched volar and non-volar melanocytes. At 10 and 12 f.w., there was no detectable differences in pigmentation between volar- and non-volar cutaneous-derived melanocytes (Figure 2A). In contrast, differential pigmentation between the two subclasses was evident at 18 f.w., which is consistent with previous reports (Nakamura et al., 2015; Cramer & Fesyuk, 2012), and indicates that the bifurcation of melanocyte pigmentation occurs between 12 and 18 f.w. (Figure 2A). In agreement with this observation, Fontana-Masson staining of fetal skin showed an increase in melanin at 18.5 f.w. in non-volar cutaneous skin with no evidence of pigmentation in donor-matched volar skin (Figure 2B).

To better understand the molecular underpinnings of the bifurcation in pigmentation between 12 and 18 f.w., we analyzed age-dependent expression of known pigment genes (Baxter et al., 2019) between donor-matched volar and non-volar cutaneous melanocytes (Table S2). Not surprisingly, we found an overall increase in the relative expression of pigment-associated genes in non-volar cutaneous melanocytes compared to volar melanocytes after 12 f.w. (Figure 2C). Interestingly, we observed little to no increase in expression of the canonical melanocyte differentiation genes (SOX10, PAX3, MITF, DCT, TYRP1, TYR, PMEL) (Figure 2D). Instead, non-volar cutaneous melanocytes more highly expressed secondary pigmentation genes (Table S2), such as OCA2, EMX2, and MFSD12, which exhibited strong

site-specific expression patterns (Figure 2E and S4C). Consistent with our findings, allelic variants of these genes are known to play a role in human skin color variation among individuals (Adhikari et al., 2019; Crawford et al., 2017). Our data reveal that differential expression of canonical markers of melanocyte pigmentation do not correlate with differential pigmentation in human epidermal melanocytes from different donor-matched anatomical sites. Instead, changes in secondary pigmentation gene expression coincides with divergence in phenotypes between sites (Figure 2F).

Identification of human melanocyte developmental programs

To identify gene expression signatures that are associated with different developmental stages *in vivo*, we partitioned non-volar cutaneous melanocytes into four groups based on developmental stage: adult (ADT), neonatal (NEO), fetal (FET) and melanocyte stem cells (MSC) (Figure 3A). Each developmental stage-specific group exhibited distinct gene signatures (Figure S5A-D). Briefly, both MSC and FET were enriched for known developmental genes (SOX11, LYPD1) (Dessaud et al. 2006; Kuhlbrodt et al. 1998) and genes involved in extracellular matrix establishment/remodeling (COL1A2, PXDN) (Colon and Bhave 2016). The ADT group expressed genes involved in innate immunity, inflammation and regulating apoptosis/cell stress in other cell types and tissues (HLAs, APOD, CLU, LGALS3) (Dassati et al., 2014; Johannes et al., 2018; Zhang et al., 2014). Interestingly, the NEO group exhibited high expression of a subset of genes from both the FET and ADT stages, consistent with neonatal melanocytes being an intermediate developmental state (Figure S5D, Table S3).

To identify genes that best distinguish each developmental stage, a regularized logistic regression was trained using the single cell transcriptomes from the four aforementioned developmental groups (Figure 3B). The resultant Development Melanocyte (DevMel) model demonstrated excellent classification accuracy, with f1-scores ranging from 0.93-1.00 (Figure 3B). Elastic net regularization yielded parsimonious model variables (i.e. genes) that constituted developmental stage-specific expression programs: prg[MSC], prg[FET], prg[NEO] and prg[ADT] (Figure 3C and Table S4).

Melanocyte development has primarily been studied using non-human model organisms (Mort et al., 2015) or employing *in vitro* differentiation of human pluripotent stem cells (Mica et al. 2013). We sought to assess the agreement between: i) mouse melanocyte development (Joshi et al., 2019; Marie et al., 2020), ii) an *in vitro* embryonic stem cell (ESC) differentiation model (Mica et al., 2013) and iii) the *in vivo* developmental stage-specific profiles described above. Indicative of discovering novel melanocyte developmental markers, there was sparse overlap between DevMel programs and published gene sets derived from the aforementioned model systems (Figure 3D). To further examine the expression of each signature within the individual cells of our dataset, we generated a mean program expression score (Figure S5E-G). No developmental stage defined in this study was enriched for the *in vitro*-derived ESC or neural crest cell (NC) gene signatures, consistent with these populations representing stages of development preceding those captured in our cohort (Figure S5F). Unexpectedly, the *in vitro* differentiation MB and mouse MB gene signatures similarly presented only minor enrichments, equally distributed across MSC, FET and NEO groups (Figures 3E and S5G). An inverse, and equally surprising, pattern presented for the *in vitro* derived mature melanocyte gene signature – all *in vivo* developmental groups scored highly, with FET and NEO scoring the highest (Figures 3E and Figure S5F). These observations indicate that each human developmental stage program identified in this study is encompassed within the single *in vitro* differentiated state defined as mature melanocytes (Figures 3E and Figure 3E), whereas previously defined MB programs are not represented by the cells captured in our cohort. In striking contrast, the MSC group was enriched for the multipotent progenitor neural crest-like mouse CD34+ melanocyte stem cell program (Figures 3E and Figure S5G). Taken together, these analyses demonstrate that through profiling human cells *in vivo*, our DevMel model offers a new perspective and refinement of the transcriptional programs underpinning melanocyte differentiation. Further analyses of these programs might provide novel insight into human specific functions of

melanocytes, genes that regulate human pigmentation, or the cells of origin and trajectories of melanoma progression.

Identification of reacquired-developmental and melanoma-specific programs during tumorigenesis

Melanoma progression often coincides with the loss of melanocyte differentiation markers and upregulation of genes associated with earlier stages of embryonic development (Hoek et al., 2008; Landsberg et al., 2012; Richard et al., 2016; Tsoi et al., 2018). This process is broadly described as dedifferentiation. We reasoned tumors are composed of cells at various stages of dedifferentiation, resulting in diverse transcriptional landscapes that can potentially influence disease progression and treatment. To unravel such tumor heterogeneity, through the lens of development-associated expression, we classified published single-cell malignant melanoma samples (Tirosh et al. 2016; Jerbey-Arnon et al. 2018) using our DevMel model. Individual malignant melanoma cells were labeled MAL^{ADT}, MAL^{NEO}, MAL^{FET} and MAL^{MSC} based on transcriptional similarity to corresponding normal healthy melanocyte developmental stage programs (Figure 4A). We observed inter- and intra-tumor heterogeneity in the representation of each melanoma group (Figure S6A), indicating tumors are composed of a mix of dedifferentiated states.

To better define the course of dedifferentiation during melanoma progression, we looked for differential expression patterns across each of the four MAL groups that were consistent with different forms of cellular reprogramming: (1) down-regulation of differentiation programs, (2) a retrograde unfolding of the differentiation cascade (sequential dedifferentiation), (3) direct reprogramming to a more pluripotent stage (direct dedifferentiation), or (4) the acquisition of a melanoma-specific program from another lineage (Figure 4B). Of the 511 total genes, inclusive of DevMel model variables and MAL group top differentially expressed genes (Figure 4C and STAR Methods), 45% exhibited expression patterns consistent with sequential dedifferentiation, in which the relative expression across healthy melanocyte developmental groups was conserved among MAL groups (Figure 4C, D and Table S5). We found that 3.1% of genes exhibited a direct dedifferentiation pattern, which indicates that expression of these genes may be a prerequisite for disease progression and metastasis (Figure 4C, E and Table S5). Supporting this interpretation, this small set of genes includes known markers of aggressive melanoma such as AXL, and HMGA2 (Raskin et al., 2013; Tirosh et al., 2016). We also identified genes expressed in healthy melanocyte groups that were down regulated in all melanoma groups (Figure 4C, F and Table S5), thus characterizing aspects of normal melanocyte expression that are either non-essential or potentially inhibitory to melanoma progression and/or metastasis.

Next, we identified 1,272 genes highly expressed in melanoma that were absent from the melanocyte developmental stages captured in our cohort (Figure 4G and Table S5). Among the top differentially expressed genes was the well-known melanoma-associated antigen PRAME, further supporting its use as a candidate melanoma molecular diagnostic. In addition, MTRNR2L family transcripts were highly expressed in all melanoma groups. These transcripts encode short peptides with anti-apoptotic activity (Guo et al., 2003), which may be important for melanoma progression.

Therapeutic resistance programs are expressed in healthy melanocytes

Recent reports have identified a population of slow cycling persisting cells responsible for therapeutic resistance in melanoma (Webster et al. 2015; 2020). Webster et al. discovered that WNT5A, a non-canonical Wnt ligand, uses wild type TP53 to drive the slow cycling state in melanoma cells. Consistent with these studies, MAL^{MSC}, had the highest expression of TP53 and WNT5A and was enriched for genes previously associated with a slow cycling, stem like state (Figure 4H). Interestingly, only 1 of the 14 tumors analyzed in this study was a treatment-naïve tumor (mel-81) but it too contained a small MAL^{MSC} population (approximately 11% of tumor cells), confirming previous reports of the presence of a stem-like dedifferentiation state prior to therapy (Figure S6A). MSC healthy melanocytes were also enriched in

the WNT5A gene signature indicating that the transcriptional program associated with therapy-resistant, stress-induced, slow cycling tumor cells is a feature of normal fetal hair follicle melanocyte stem cells.

Distinct MITF-regulated programs classify melanoma dedifferentiation states.

A number of studies have identified two transcriptional programs associated with the differentiation status of melanoma cell-lines and tumors. These two programs have been defined as "proliferative" and "invasive" (Widmer et al., 2012) and characterized as MITF high versus AXL high, respectively (Tirosh et al., 2016). Both classifications center on the average expression level of MITF itself as well as a MITF regulated transcriptional program. Increasing expression of the MITF program is thought to drive a progressively differentiated state in a hypothesis referred to as the rheostat model (Strub et al., 2011). In agreement with this model, the more differentiated MAL^{ADT} and MAL^{NEO} melanoma groups exhibited higher MITF program expression compared to the less differentiated MAL^{FET} and MAL^{MSC} groups (Figure 4I, J). However, despite the similarity in MAL^{ADT} and MAL^{NEO} overall expression levels of MITF programs identified in melanoma cells (Tirosh et al., 2016 and Widmer et al., 2012), we identified developmental stage-specific gene expression of MITF programs previously identified in healthy melanocytes (Hoek et al, 2008). Surprisingly, the expression level of individual genes in the healthy melanocyte MITF program was reciprocal between MAL^{ADT} and MAL^{NEO} groups (Figure 4K). Based on this trend, we identified three MITF sub-programs (MITFsp) with distinct expression patterns during melanoma dedifferentiation: (1) MITFsp1 specific to MAL^{ADT}, (2) MITFsp2 shared by both MAL^{ADT} and MAL^{NEO}, and (3) MITFsp3 specific to MAL^{NEO}. These observations demonstrate that the high MITF program expression is generally associated with more differentiated melanoma states; however, the degree of differentiation is not necessarily proportional to unified MITF program expression. Rather, switching between distinct MITF-regulated sub-programs explains the development-associated differences between more differentiated melanoma cells.

Classifying bulk tumors by developmental stage program predicts survival

To determine whether knowledge of the precise developmental stage programs readopted by melanomas offers prognostic value, we used CIBERSORT (Newman et al., 2015) to estimate the fraction of melanoma cells similar to ADT, NEO, FET, MSC for all skin cutaneous melanoma (SKCM) tumor samples from The Cancer Genome Atlas (TCGA) (Akbani et al., 2015) (see STAR methods). Briefly, CIBERSORT employs support vector regression to estimate the proportions of labeled cell subpopulations in bulk RNA-seq samples. Here, we used the single-cell transcriptomes of healthy normal ADT, NEO, FET, and MSC melanocytes as reference labels in the deconvolution, resulting in relative estimates of label fraction for each bulk SKCM tumor. Similar to the single cell melanoma dataset (Figure S6A), there was inter-tumor heterogeneity in the fractional representation of the four developmental stages (Figure 5A). Hierarchical clustering of SKCM label distributions grouped tumor samples according to the observed predominant developmental group label: $SKCM^{ADT}$, $SKCM^{NEO}$, $SKCM^{FET}$, $SKCM^{MSC}$ (Table S6). The development-specific SKCM groups were not enriched with known clinicopathological features, such as tissue of origin, driver mutation, and pigmentation level (Figure 5A, B and Figure S6B-D), suggesting that neither genetic driver nor tumor site orchestrates the developmental stage program of the tumor.

We compared our SKCM categorization to previously published classifications. Consistent with representing the most dedifferentiated state, $SKCM^{MSC}$ was composed of tumor samples that were in the least differentiated categories of each study, namely the AXL-high/MITF-low (Tirosh et al., 2015) and invasive (Widmer et al., 2012) tumors (Figure 5A), and the MITF-low (Akbani et al. 2015) and NC-like / undifferentiated (Tsoi et al. 2018) categories (Figure 5C and Figure S6B,D). In contrast, $SKCM^{FET}$, $SKCM^{NEO}$, and $SKCM^{ADT}$ were comprised of relatively even distributions of the remaining previously reported transcriptional categories. This observation suggests that, beyond MSC-like tumors, our

SKCM^{FET}, SKCM^{NEO}, and SKCM^{ADT} groups offer a novel partitioning of the remaining SKCM sample cohort based on development stage program.

Using our developmentally-defined subclasses of melanoma tumors, we evaluated the correlation of bulk tumor differentiation status with patient outcome. Similar to previous reports (Jerby-Arnon et al., 2018), we confirmed that overall survival does not vary linearly with differentiation status (Figure 5D). While SKCM^{ADT} exhibited best median overall survival (SKCM^{ADT} = 11.0 yr vs rest = 5.3 yr), the more dedifferentiated groups (SKCM^{FET}, SKCM^{MSC}) did not did not have the worst outcomes. Strikingly, SKCM^{NEO} exhibited the shortest median overall survival (SKCM^{NEO} = 4.2 yr vs rest = 8.2 yr). This observation indicates that progressive dedifferentiation does not result in incrementally worse prognosis; rather, the most aggressive SKCM melanomas re-acquired a neonatal-like developmental program. To better understand this unexpected finding, we characterized transcriptional programs associated with clinical response to therapeutics. Consistent with our observation, SKCM^{NEO} tumors expressed high levels of transcripts associated with immune resistance and a dearth of both immune infiltration signatures and FDA-approved therapeutic targets. In contrast, SKCM^{ADT} and SKCM^{MSC} tumors were depleted of immune resistance genes and enriched for immune infiltrating signatures. The SKCM^{MSC} tumor group was unique in its increased expression of FDA-approved therapeutic targets (Figure 5E). Of particular interest are highly expressed ion channels: a family of molecular targets for which small molecule inhibition in melanoma cell lines has been shown to reduce proliferation by up to 90% (D'Arcangelo et al., 2019) (Figure 5E, inset and Table S7). Taken together, these data demonstrate that while some amount of dedifferentiation is associated with worse prognosis, overall survival, immune evasion and immune resistance are not linearly correlated with dedifferentiation.

DISCUSSION:

We have provided, as a resource, a human epidermal melanocyte atlas. To our knowledge, this is the first human cell atlas entry that encompasses human development, gender and diverse ethnicities and includes multiple matched anatomical locations. This is a foundational resource for understanding the development, function and diversity of human melanocytes. In addition, it is valuable for understanding human pigmentation and associated conditions and diseases, such as vitiligo and melanoma. Our observations should motivate future studies to sample melanocytes from even earlier developmental stages and from more anatomical locations, including adult follicles, dermis, mucosal membranes, inner ear and eye. This resource is an important step toward a comprehensive human cell atlas, and the analyses described here have revealed fundamental discoveries of anatomic site-specific melanocyte subclasses, mechanisms of human melanocyte development and malignant transformation trajectories.

From the analysis of donor-matched volar and non-volar cutaneous skin we identified two distinct subclasses of epidermal melanocytes that diverge early in human development and are retained in adults. Specifically, we discovered a subclass of human melanocytes, characterized by NTRK2 expression and low pigmentation that bifurcates early in development and populates specific anatomical locations of the adult skin with differential pigmentation. Elegant epidermal transplant studies have attributed differences in the epidermal phenotype in fully developed volar and non-volar skin to site specific mesenchymal–epithelial interactions (Yamaguchi et al., 2001, 2004). Our discovery of site-dependent enrichment, but non-exclusivity, of melanocyte subclasses suggests that site specific signaling provides more permissive conditions for one melanocyte subclass over another, resulting in a net change in pigmentation. The methods used here provide the groundwork for deciphering melanocyte biology through intra-donor differences in melanocyte subclass distribution. By comparing within individuals, there is reduced noise from donor to donor variability permitting discovery of programs common across donors. Through these analyses, we identified programs that define non-pigmented melanocytes, which are a valuable resource for furthering characterizing melanocyte

functions in addition to pigmentation. This class of non-pigmented melanocytes awaits further detailed characterization and could uncover new functions for melanocytes.

"Melanoblast" is a broad term used to describe committed immature melanocytes along a wide range of developmental time points. It describes cells from initial commitment to melanocyte fate, which occurs prior to epidermal infiltration, along two distinct lineage specification pathways, through hair follicle morphogenesis. Here we have refined the "melanoblast" identity in the developing human epidermis and hair follicle. In humans, hair follicle formation is reported to start around 10 f.w. with mature hair follicles appearing around 20 f.w. dependent on anatomical location and study (Gleason et al., 2008; Holbrook et al., 1989). Mouse hair follicle morphogenesis occurs from E14.5 and is completed postnatally by P8 as a fully mature hair-bearing follicle in anagen phase (Nowak et al., 2008). Thus, it is not surprising that the mouse MB program, derived from E15.5 and E17.5 mice, is most highly expressed in the MSC population. Comparison of the MB and differentiated melanocyte programs derived from the *in vitro* ESC system indicated that the *in vitro* MB stage likely precedes 9.5 f.w. in humans; while, the *in vitro* differentiated melanocyte stage most closely resembles 9.5 f.w. to neonatal epidermal melanocytes. The relatively high level of expression of the *in vitro* differentiated melanocyte program across MSC, FET, NEO and ADT melanocytes, compared to the analogous MB program, suggests that the final timepoint in the differentiation protocol contains an asynchronous population of cells representing the full span of developmental stages in our cohort. That being said, although we report which model system programs are most similar to those characterized here, the overall poor level of overlap is notable. The lack of observed conservation could have several sources: species- or condition- specific differences, differences between scRNAseq and bulk mRNAseq, and/or differences in timescale of hair follicle development between species.

The concept of dedifferentiation during melanoma progression is broadly defined as a loss of characteristics associated with the mature melanocyte. On the transcriptomic level, dedifferentiation can occur via a variety of trajectories: some ordered like differentiation itself and others chaotic, such as selection of stochastically-expressed programs (Parchem et al., 2014). Our analyses suggest that sequential dedifferentiation, which recapitulates the ordered cascade of differentiation in reverse, is prevalent in melanoma progression. This discovery is consistent with recent reports that show therapeutic resistance development coinciding with sequential dedifferentiation (Rambow et al., 2018; Tsoi et al., 2018). Our dataset permits objective defining of degrees of dedifferentiation and assessment of the relationship between each stage and tumor characteristics. We found that individual tumors are comprised of a heterogeneous mix of cells at different stages in the dedifferentiation cascade (Figure 6A). It has been theorized that a greater degree of dedifferentiation confers a worse outcome. However, our analyses identified a "sweet spot" (or perhaps more appropriately, a "sour patch") in the dedifferentiation cascade that is linked to worst overall survival: the intermediate developmental state associated with neonatal melanocytes. It is interesting to note that due to tissue availability and ease of culture, the neonatal melanocyte transcriptome is usually considered the baseline "normal differentiated program" for comparison to melanoma transcriptomes. This technical nuance can explain why this program has been previously underappreciated and highlights the importance of conducting *in vivo* human cell atlas projects for interpreting the progression of human cancers.

Furthermore, we enumerated McSC- and melanoma-specific genes directly acquired in all stages of dedifferentiation (Figure 6B), suggesting that these genes may undergo positive selection during early metastatic dissemination. Along with the widely-accepted diagnostic melanoma biomarker PRAME (Lezcano et al., 2018) and the master regulator of invasion AXL (Tirosh et al., 2016), we identified novel melanoma-associated genes. Further investigation into the mechanistic roles of this gene set could reveal drivers of melanoma metastasis. Taken together, we hope this resource will provide a foundation for discovery of dedifferentiation stage-specific diagnostics and therapeutics.

ACKNOWLEDGMENTS:
We thank the University of California, San Francisco Biospecimen Resource Program for their help with tissue acquisition, and Life Science Editors for critical editing of the manuscript. We thank the UCSF Program for Breakthrough Biomedical Research Sandler Fellows Program (to R.L.JT.) for financial support. We would like to thank Norma Neff and Michelle Tan for all the help with library quality control and sequencing.

AUTHOR CONTRIBUTIONS:
Conceptualization, R.L.B, D.L., A.D.T, S.D., and R.L.JT.; Methodology, R.L.B and A.M.; Validation, R.L.B. and D.L.; Formal Analysis, R.L.B. and D.L.; Investigation, R.L.B, D.L., and A.M; Resources, U.E.L., A.S., V.PP, L.B., and A.D.T.; Data Curation: D.L., and A.M.; Writing – Original Draft, R.L.B and R.L.JT.; Writing – Review & Editing, D.L., A.M., U.E.L. and S.D.; Visualization, R.L.B. and D.L.; Supervision, S.D. and R.L.JT.; Funding Acquisition, S.D. and R.L.JT.

FIGURE LEGENDS:

Figure 1: Melanocyte transcriptomic profiles differ based on development and anatomical location
A) Single cell analysis pipeline. **B**) UMAP projection of melanocytes from 22 donors labeled by developmental stage and anatomic location. **C**) Schematic illustrating cohort containing donor-matched non-volar and volar skin. n = 6 total donors. **D**) Volcano plot of genes enriched in non-volar vs volar melanocytes. Top differentially expressed genes labeled. **E)** Expression level of NTRK2 and HPGD in donor-matched volar and non-volar melanocytes. Box plots: Interquartile range with median, standard deviation and outliers (grey circles). **F**) Fraction of v-mels and c-mels across anatomical locations from 22 donors. Volar skin is enriched for v-mels (NTRK2+/HPGD-). Whereas, non-volar skin, including foreskin, is enriched for c-mels (HPGD+ /NTRK2-). * p-value = 0.1, ** p-value <0.02, ***p-value < 0.0002.

Figure 2: Divergent pigment developmental trajectories in volar and non-volar melanocytes
A) Raw (top) and average normalized (bottom) BSC values of volar and cutaneous melanocytes prior to 18 weeks (pre-bifurcation) and after 18weeks (post-bifurcation). ns, p-value =0.1; * p-value = $4x10^{-184}$. **B**) Fontana Masson staining for melanin in fetal and adult non-volar and volar skin. **C**) Increased pigment content coincides with upregulation of the pigment transcriptional program in cutaneous melanocytes at 18 wks. Normalized mean expression of 170 pigment associated genes (thin lines) in volar (blue) and cutaneous (red) melanocytes. Thick lines: average expression of all pigment associated genes. **D-E**) Ratio of non-volar:volar expression of (D) canonical melanocyte differentiation and (E) secondary pigment genes. **F**) Model depicting two anatomical site-dependent divergent development trajectories where bifurcation in pigment content and pigment transcription program occurs between 12 and 18 f.w. Canonical genes associated with melanocyte development and pigment production are present early in both melanocyte lineages (early genes). The appearance of OCA2 expression coincides with appearance of melanin in cutaneous fetal melanocytes at 18 wks with differential expression persisting into adulthood. Site-specific MFSD12 and EMX2 expression occurs later in development.

Figure 3: Transcriptional profiling of human melanocyte differentiation

A) Illustration depicting the three developmental stages used to identify the transcriptional programs involved in melanocyte differentiation in c-mel enriched anatomical locations. **B**) Left, schematic of the logistical regression model used to generate (i) and validate (ii) unique transcription profiles for each developmental stage of melanocytes (programs, prg). Right, classifier performance on hold out validation set. **C**) Heatmap of the relative expression (row z score) of genes in each DevMel profile from (B). **D**) Number and percent of genes (color scale) from previously published gene sets that are present in the positive correlated component of each DevMel profile. ECS, embryonic stem cell; NC, neural crest; MB, melanoblast; Mel, differentiated melanocyte; CD34+ and CD43-: two hair follicle MSC states. **E**) Data summary of the mean program expression from model systems of melanocyte development (see Figure S5E-G) in MSC, FET, NEO, and ADT groups from (A). The genes with positive coefficients from the DevMel programs in (D) were enriched in the respective melanocyte developmental groups from (A). *In vitro* differentiated and mouse melanoblast (MB) programs are minorly expressed in MSC, FET (i*n vivo* human melanoblasts), and NEO. *In vitro* differentiated human melanocytes (Mel) primarily encompass fetal and neonatal *in vivo* human programs. The melanocyte stem cell program (CD34+) is relatively conserved in human and mouse.

Figure 4: Decoding melanoma dedifferentiation using normal melanocyte DevMel transcriptional profiles

A) Schematic showing categorization of individual melanoma cells by melanocyte developmental stages using the LOGIT DevMel classifier (Figure 3). Every melanoma cell (MAL) was categorized by the predominantly expressed developmental stage program. **B**) Theoretical categories of cancer-associated transcriptional reprogramming: sequential dedifferentiation, a reverse stepwise unfolding of development; direct dedifferentiation, direct reacquisition of programs from early developmental stages; melanoma specific, acquisition of programs not associated with the stages melanocyte development identified here. **C**) Workflow to generate genes used to identify genes associated with melanoma dedifferention (top). Percent of genes across MAL groups that exhibit patterns consistent with categories described in (A) or melanocyte specific DevMel transcriptional programs from Figure 3 (bottom). **D-G**) Examples of each category are visualized as heatmaps of the relative expression (row z score). Light purple indicates healthy human melanocytes from each indicated developmental stage. Dark purple indicates melanoma cell expressing the defining program of each indicated developmental stage. See Table S5 for complete list. **H**) Heatmap showing the relative expression levels (row z score) of WNT5A high, TP53 high slow cycling cell associated genes. **I-J)** Violin plots showing single cell expression scores in each MAL group of previously published genes sets used to classify two different melanoma (I) transcriptomic states or (J) phenotypes. **K**) Expression levels of genes correlated with healthy melanocyte associated MITF program (Hoek et al., 2008) across MAL groups. Three differentially expressed MITF network sub-programs (sp) are indicated.

Figure 5: Classification of TGCA data by normal melanocyte developmental stage has prognostic value

A) Hierarchical clustering of TGCA tumors based on fractional composition of normal melanocyte developmental stages assigned using CIBERSORT (top) with clinicopathological features and previously described categorization of each tumor indicated (bottom panels). **B**) Distribution of each clinicopathological feature and **C**) transcriptional categorization by SKCM group ($SKCM^{ADT}$, $SKCM^{NEO}$, $SKCM^{FET}$, $SKCM^{MSC}$). **D**) Kaplan Meier curves for each SKCM group from (A). Enrichment for cells similar to ADT is associated with increased survival, whereas enrichment for NEO is associated with worse survival. **E**) Relative expression (column z score) of immune infiltration program, immune evasion program and FDA-approved therapeutic targets in SKCM groups (left). Enrichment of ion channel druggable targets (red) among DEGs comparing $SKCM^{MSC}$ tumor to rest (Fisher exact test p-value = 2.31×10^{-6}). Y axis ordered by log2 fold change (right).

Figure 6: Heterogeneous reacquisition of defined developmental programs in melanoma dedifferentiation.

A) Individual melanoma tumors are comprised of a heterogeneous mix of malignant cells expressing defined melanocyte developmental programs. The fraction of cells expressing each program within the tumor is predictive of overall survival and correlates to signatures of immune infiltration, evasion and potential therapeutic options. **B**) Each melanoma cell can occupy a different degree of dedifferentiation, associated with differences in TP53 and WNT5A expression, specific MITF sub-programs, and antigen presentation. Genes common to all cells include melanoma specific genes, such as PRAME, or direct reprogramming to an early developmental stage (direct dedifferentiation).

TABLES:

Table 1: Donor demographics for skin samples used for single cell RNAseq

SAMPLE	AGE[+]	SEX	ANATOMICAL LOCATION	SKIN PIGMENTATION LEVEL [++]	RACE / ETHNICITY
9.5WK02	9.5 f.w.	unknown	leg	n/a	unknown
			arm	n/a	
***10WK03**	10 f.w.	unknown	leg	n/a	unknown
			arm	n/a	
			palm	n/a	
16WK04	16 f.w.	unknown	palm	n/a	unknown
***12WK01**	12 f.w.	M	leg	n/a	unknown
			arm	n/a	
			palm	n/a	
			sole	n/a	
***12WK05**	12 f.w.	unknown	leg	n/a	unknown
			arm	n/a	
			sole	n/a	
***18WK06**	18 f.w.	unknown	leg	n/a	unknown
			palm	n/a	
			sole	n/a	
FS030	neo	M	foreskin	LM	unknown
FS043	neo	M	foreskin	LM	unknown
A1021	24 yr	F	leg	M	Asian
***A1038**	35 yr	M	leg	LM	White
			sole	L	
A1012	37 yr	F	arm	M	Asian
A1022	42 yr	M	arm	M	Hispanic-Latinx
A1015	52 yr	F	arm	LM	Asian
A1025	56 yr	M	leg	L	White
A1016	58 yr	M	arm	LM	White
A1020	60 yr	F	arm	LM	Asian
A1033	61 yr	M	leg	M	Hispanic-Latinx
A1011	65 yr	M	arm	L	White
A1026	66 yr	M	leg	L	White
A1014	68 yr	F	leg	L	White
***A1046**	77 yr	F	leg	M	Hispanic-Latinx
			sole	L	
A1017	81 yr	M	leg	LM	Hispanic-Latinx

[*] included in multi-site patent matched analysis

[+] f.w., fetal weeks; yr, years

[++] pigment score based on appearance of skin prior to epidermal dissociation: L, light; LM, light-medium; M, medium

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Subject details

All skin was collected from surgical discards with informed consent and approval from the UCSF Institutional Review Board. All ages, races/ethnicities and sexes were included in the eligibility criteria for this study. Adult tissue was obtained from surgical remnants of heathy skin taken for reconstructive surgery or from amputations with heathy skin. Neonatal foreskins were obtained after routine circumcision. Anonymous fetal specimens were obtained from elective terminations and fetal age (stated as fetal weeks) was estimated by heel-toe length (Drey et al., 2005). When possible, fetal gender was determined by visual inspection using a dissecting microscope. All samples were collected in cold CO_2 Independent Media(Gibco–Thermo Fisher Scientific) or Medium 154 (Gibco) with 1× Antibiotic-Antimycotic (Gibco) at 4°C until dissociation.

Human skin sample preparation

Tissue dissociation was started the same day as sample acquisition. For adult and neonatal skin, the epidermis was enzymatically dissociated from the dermis with a dispase, neutral protease, grade II (Roche–Sigma-Aldrich), incubation for 14 hours at 4˚C. Epidermal sheets were manually separated from dermis, finely minced, and incubated with 0.5% trypsin (Gibco) for 3 minutes at 37˚C. After manual trituration, trypsin was deactivated using ice cold soybean trypsin inhibitor (Gibco), then diluted 2:3 in ice cold Hanks' balanced salt solution, no Mg^{2+}, no Ca^{2+} (Gibco). The dissociated cell suspension was centrifuged at 500g, 4˚C , for 4 minutes, resuspended in FACS buffer (0.1% bovine serum albumin (Sigma) and 25mM Hepes (Gibco) in Dulbecco's phosphate-buffered saline (DPBS) (Gibco)) and strained with a 70µM filter to achieve a single cell suspension. For fetal tissue, the developing epidermis was manually removed from the dermis following a 20 - 30 minute incubation with 10mM EDTA (Invitrogen), DPBS at 37˚C. The resulting epidermal layer was incubated with 0.5% trypsin (Gibco) for 1 min at 37˚C and manually triturated. Trypsin was deactivated using ice cold soybean trypsin inhibitor (Gibco), then diluted 2:3 in ice cold Hanks' balanced salt solution (Gibco). The dissociated cell suspension was centrifuged at 500g, 4˚C , for 4 minutes, resuspended in FACS buffer, and strained with a 70µM filter to achieve a single cell suspension.

Data availability

Jupyter notebooks with detailed analysis scripts are available here:
https://github.com/czbiohub/human_melanocytes

GEO: GSE151091
Bioproject: PRJNA625154
SAMN14593853 : A1015LM
SAMN14593854 : 12WKM01
SAMN14593855 : FS043_LM
SAMN14593856 : A1011L
SAMN14593857 : A1020LM
SAMN14593858 : A1033M
SAMN14593859 : A1022M
SAMN14593860 : A1014L
SAMN14593861 : A1038LM
SAMN14593862 : A1026L
SAMN14593863 : A1021M

SAMN14593864 : FS030_LM
SAMN14593865 : A1016LM
SAMN14593866 : A1025L
SAMN14593867 : A1017LM
SAMN14593868 : A1012M
SAMN14593869 : 18WKM06
SAMN14593870 : 12WK05
SAMN14593871 : 10WK03
SAMN14593872 : 16WKM04
SAMN14593873 : 9.5WK02
SAMN14593874 : A1046M

METHOD DETAILS

FACS analysis and single cell sorting:

Single cell suspensions were counted, diluted to $1x10^6$ cells/100ul with ice cold FACS buffer containing dye conjugated antibodies (anti-CKIT (104D2), 15ng/100µl (CD11705, Thermo Fisher Scientific), anti-alpha6 integrin (GoH3), 15ng/100µl (12-0495-82, Thermo Fisher Scientific) and CD11c, 1:20 dilution (46-0116-41, Thermo Fisher Scientific)) and incubated on ice for 25 minutes. Cells were washed one time with 10x volume of FACS buffer, centrifuged for 2 minutes at 500g, resuspended in 30ng/mL Dapi (D3571, Molecular Probes), FACS buffer. Resuspended cells were strained through at 35 µm nylon mesh filter and kept on ice until sorted.

Single cells were sorted into 384-well plates using the "Ultra purity" setting on a SH800S (Sony) sorter. For a typical sort, a tube containing 0.3-1ml the pre-stained cell suspension was vortexed gently and loaded onto the FACS machine. A small number of cells were flowed at low pressure to check cell concentration and amount of debris. Then the pressure was adjusted, flow was paused, the first destination plate was unsealed and loaded. Single cells were sorted into plates by gating to exclude dead/dying cells (DAPI+) and doublets. The majority of the plate contained melanocytes (CD11c-/CKIT+) with several columns of basal (CD11c-/CKIT-/ITGA6+) and suprabasal keratinocytes (CD11c-/CKIT-/ITGA6). Immediately, after sorting, plates were sealed with a pre-labeled aluminum seal, centrifuged at 4°C and flash frozen on dry ice, before storage at -80 for later use.

Immunofluorescence

Skin samples were fixed in 4% paraformaldehyde (Electron Microscopy Sciences) at 4 ˚C overnight, washed with cold DPBS, followed by 30% sucrose infiltration prior to OCT embedding. Fixed frozen skin sections were incubated in blocking buffer: 2.5% donkey serum, 2.5% goat serum (Jackson ImmunoResearch Laboratories), 1% bovine serum albumin (Sigma-Aldrich), and 0.1% Triton X-100 (Sigma-Aldrich) for 1–2 hours at room temperature. The following primary antibodies were used at the indicated concentration in blocking buffer overnight at 4°C: mouse monoclonal anti-TYRP1 1:200 (TA99, ab3312, Abcam), mouse monoclonal anti-CKIT 1:100 (CD11700, Invitrogen–Thermo Fisher Scientific), rabbit polyclonal anti-HPGD 1:100 (HPA005679, Sigma-Aldrich). Secondary antibodies against mouse IgG, or rabbit IgG conjugated to DyLight 488 or 594 (Thermo Fisher Scientific) were used at a 1:1,000 dilution for 1–2 h at room temperature followed by Dapi, 1:1000 (Molecular Probes) for 1 minute. Sections were mounted in VECTASHIELD Vibrance (Vector Laboratories) prior to imaging.

Fluorescence imaging

Images were acquired using Nikon NIS-Elements multi-platform acquisition software on an fully automated Nikon Ti-E inverted microscope with an Apo TIRF, 60x, 1.49 NA, oil objective (Nikon) and a Clara CCD camera (Andor).

Percent v-mel and c-mel

For each single cell, non-zero gene expression level for HPGD or NTRK2 was labeled "+", otherwise the "-" label was given. Fraction of cells was calculated as the number of NTRK2+/HPGD- cells or HPGD+/NTRK2- cells divided by the number of NTRK2+/HPGD- and HPGD+/NTRK2-. To determine the percent of HPGD positive melanocytes in tissue sections, melanocytes (TYPR1+ cells) were manually counted. Fraction of cells was determined by the number of HPGD+ TYRP1+ cells divided by the total number of TYRP1+ cells from each fixed frozen section. Immunofluorescence image analysis was performed in Fiji (http://fiji.sc/).

Fontana Masson staining

Melanin staining was performed on fixed frozen sections, from patient matched volar and non-volar cutaneous skin, using the Fontana-Masson Stain Kit (ab150669, Abcam) following the manufacturers protocol.

Single-cell processing and analysis

Single cell reads were mapped to the human reference hg38 containing ERCC sequences using STAR aligner (10.1093/bioinformatics/bts635). HTSeq (10.1093/bioinformatics/btu638) was used to create gene count tables. These count tables were compiled and processed using Scanpy (10.1186/s13059-017-1382-0). Low-quality cells were filtered based on the following criteria: number of genes < 1,250 OR number of reads < 50,000. Each gene in the transcriptome exhibited read counts in at least 3 cells. Cells exhibiting > 2-fold higher number of genes than average were labeled as putative doublets and removed. Iterative Louvain clustering yielded cell type-specific clusters, which were annotated using published marker genes based on inter-cluster differential expression analysis (two-sided Mann Whitney U test, Benjamini-Hochberg FDR < 5%). Briefly, Louvain clustering was performed on the k-nearest neighbor graph in principle component space of scaled highly variable genes. Cells were visualized using 2-dimensional UMAP embeddings. Cell cycle status was inferred by the mean ranked expression of marker genes, referred to as the cell cycle program score (10.1101/gr.247759.118). Cells below the 95th-percentile of the cell cycle program score were labeled G2/M; conversely, cell equal to or greater than 95th-percentile of cell cycle program score were labeled G1/S.

Backscatter analysis

Normalized FACS backscatter (BSC) was computed as the ratio of non-volar cutaneous cell BSC over volar cell BSC. Similarly, normalized mean ranked expression was computed as a ratio of non-volar cutaneous cell expression over that of volar cells. Only genes with mean ranked expression greater than the 10th-percentile were used.

Single cell developmental stage melanonocyte (DevMel) logistic regression model

Input data was composed to single cell transcriptomes from the following 4 groups: MSC, FET, NEO and ADT. The input examples were randomly sampled and the number of examples was balanced among all labels. The combination of normal and melanoma transcriptomes was used to scale and center the data. The input data was split into testing and training partitions at a ratio 33:67. We implemented elasticnet regularization with an l1 ratio = 0.8. Single cell transcriptomes were evaluated by the model to yield a developmental stage label.

Classification of genes in melanoma dedifferentiation categories

Logistic regression variables and top-100 differentially expressed genes for each melanoma cell population grouped by DevMel label were used in subsequent dedifferentiation pathway analysis. DevMel group mean ranked expression was compared between normal and melanoma datasets to determine pathway based on the following criteria:

Direct dedifferentiation: All cancer DevMel group >= mean normal non-ADT DevMel group, 4-fold normal ADT DevMel group < mean normal non-ADT DevMel group

Sequential dedifferentiation:Max cancer DevMel group == Max normal DevMel group

Melanoma-specific: For each, cancer DevMel group > corresponding normal DevMel group, All cancer DevMel groups > 40^{th}-percentile of expression, All normal DevMel groups < 10^{th}-percentile of expression

Normal-specific: For each, cancer DevMel group < corresponding normal DevMel group, All cancer DevMel groups < 10^{th}-percentile of expression, All normal DevMel groups > 15^{th}-percentile of expression

Down regulated: All cancer DevMel group <= mean normal non-ADT DevMel group, Normal ADT DevMel group > 1.5-fold mean normal non-ADT DevMel group

Not-readopted: All cancer DevMel group <= normal ADT DevMel group, 1.5-fold normal ADT DevMel group < mean normal non-ADT DevMel group

Bulk tumor deconvolution

CIBERSORT (Newman et al. 2015) was used to deconvolve bulk RNA-seq from the SKCM-TCGA cohort. As input, CIBERSORT requires cell type-labeled transcriptomes to estimate the proportion of each cell type in a bulk RNA-seq sample. Here, trimmed both single cell and bulk RNA-seq transcriptomes to include only genes that are shared in both datasets. Adopting a k-fold cross-validation approach, we prepared 10 sets of single cell input transcriptomes from normal melanocytes across 4 developmental stages: MSC, FET, NEO and ADT (balanced cell counts across all labels). Each input transcriptome set was used to devolve the SKCM-TCGA bulk RNA-seq samples, yielding 10 estimates of cell proportion. For each individual sample in the SKCM-TCGA dataset, the label means were used as the final estimate of label proportion. Hierarchical clustering was used to group SKCM-TCGA samples based on similar label proportions. One-sided Fisher Exact test was used to determine significant enrichment between two gene lists. The lifelines python package (10.5281/zenodo.3833188) was used to create Kaplan-Meier survival plots and perform logrank tests using curated SKCM-TCGA metadata (10.1016/j.cell.2018.02.052).

SUPPLEMENTAL FIGURE LEGENDS:

Figure S1: **Single cell RNA sequencing quality control.**

A) Indexed FACS gate assignments for cells from each cell type identified. **B**) Number of reads and **C**) number of genes per cell for all 14,370 sequenced cells. Dashed line: quality control threshold, cells with < 50,000 reads and < 500 genes were excluded from further analysis. **D**) Genes expressed in more than 3 cells (dashed line) were including for subsequent analysis.

Figure S2: Identification of cell types.

A) Lovain cluster identification of 6 cell types from freshly isolation human skin. UMAP projection of the 9,688 cells that passed quality control. **B**) Heat map showing the relative expression of top differentially

expressed genes for each cluster in (A). **C-J**) UMAP projection with cell type specific expression score overlay.

Figure S3: Characterization of cell cycle state and patent age distribution
A) UMAP projection of cycling cell program score used to determine which cell were designated as **B**) cycling (blue, in G2 & M phase) vs non-cycling (red). **C**) Fraction of cycling and non-cycling cells for each cell type identified in Figure S2. **D-E**) Fraction of cycling and non-cycling cells by age for **D**) keratinocytes and **E**) melanocytes. **F-H**) UMAP projection of melanocytes with donor age overlay for **F**) adult, **G**) neonatal, and **H**) fetal cells.

FigureS4: Single cell expression of NTRK2 and HPGD
A) Immunofluorescence staining of adult volar and non-volar skin cyro-sections for melanocytes (>, epidermal CKIT+ cells, magenta) and HPGD (green). Dashed line, epidermal-dermal junction. **B**) Percent HPGD positive melanocytes per donor volar and non-volar skin. Adult skin: A1046, n=78 total cells; A1038, n = 39 total cells; A1018, n= 48 total cells; A1026, n= 15 total cells. Fetal skin: 9WK07, n= 41 total cells; 16WK04, n = 10 total cells. **C**) Expression level of OCA2, EMX2 and MFSD12 for all donor-matched melanocytes in Figure 2. Box plots: Interquartile range with median, standard deviation and outliers (grey circles).

Figure S5: Characterization of melanocyte developmental groups
A) Volcano plot showing the top ten differentially expressed genes between MSC (yellow) and fetal (teal) melanocyte populations. **B**) MSCs express known melanocyte stem cell markers and have high expression of mesenchymal markers. **C**) Volcano plot showing the top differentially expressed genes between fetal (teal) and adult (magenta) melanocytes. **D**) Heatmap visualization of the relative expression (row z score) of differentially expressed genes from MSC in (A) and fetal, adult in (C) showing enrichment of both fetal and adult genes in neonatal melanocytes. Gene in red are discussed in the main text. **E-G**) Violin plots showing MSC, FET, NEO, and ADT single cell program expression scores of **E**) the positively correlated genes in the DevMel profiles from Figure 3C, **F**) *in vitro* melanocyte differentiation programs, and **G**) mouse melanocyte developmental programs.

Figure S6: Tumor composition, categories, and enrichment analyses from figure 4 and figure 5.
A) Individual tumors are a heterogeneous mix of malignant cells in different dedifferentiation states. Fraction of MAL^{ADT}, MAL^{NEO}, MAL^{FET} and MAL^{MSC} cells in each of the 14 tumors analyzed from Tirosh et al. and Jerby-Arnon et al. in Figure 4. **B**) There is little to no difference in the enrichment of pigment level, mutation category, or tissue origin between SKCM groups in Figure 5. Whereas, $SKCM^{MSC}$ is enriched in Undifferentiated and NC-like (Tsoi et al.), and MITF-low (TCGA) tumors. Negative log Fisher exact test adjusted p-values. **C**) Distribution of clinicopathological features and **D**) TCGA and Tsoi et al. categorizations for each SKCM group in Figure 5.

SUPPLEMENTAL TABLES:
Supplemental Table 1: Related to Figure1D. Top differentially expressed genes between volar and non-volar cutaneous melanocytes.

Supplemental Table 2: Pigment genes used in Figure2C.

Supplemental Table 3: Differentially expressed genes for MCS vs FET and FET vs ADT in Figure S5A,C

Supplemental Table 4: DevMel profile genes from Figure 3.

Supplemental Table 5: The genes used for the dedifferentiation analysis in Figure 4.

Supplemental Table 6: SKCM group categorization of TCGA tumors in Figure 5

Supplemental Table 7: FDA approved drugs and immune signatures from Figure 5.

REFERENCES:

Adhikari, K., Mendoza-Revilla, J., Sohail, A., Fuentes-Guajardo, M., Lampert, J., Chacón-Duque, J.C., Hurtado, M., Villegas, V., Granja, V., Acuña-Alonzo, V., et al. (2019). A GWAS in Latin Americans highlights the convergent evolution of lighter skin pigmentation in Eurasia. Nat. Commun. *10*, 1–16.

Akbani, R., Akdemir, K.C., Aksoy, B.A., Albert, M., Ally, A., Amin, S.B., Arachchi, H., Arora, A., Auman, J.T., Ayala, B., et al. (2015). Genomic Classification of Cutaneous Melanoma. Cell *161*, 1681–1696.

Baxter, L.L., Watkins-Chow, D.E., Pavan, W.J., and Loftus, S.K. (2019). A curated gene list for expanding the horizons of pigmentation biology. Pigment Cell Melanoma Res. *32*, 348–358.

Cajot, J.F., Schleuning, W.D., Medcalf, R.L., Bamat, J., Testuz, J., Liebermann, L., and Sordat, B. (1989). Mouse L cells expressing human prourokinase-type plasminogen activator: Effects on extracellular matrix degradation and invasion. J. Cell Biol. *109*, 915–925.

Choi, H.R., Park, S.H., Choi, J.W., Kim, D.S., and Park, K.C. (2012). A simple assay method for melanosome transfer. Ann. Dermatol. *24*, 90–93.

Colombo, S., Berlin, I., Delmas, V., and Larue, L. (2011). Classical and Nonclassical Melanocytes in Vertebrates. In Melanins and Melanosomes, (Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA), pp. 21–61.

Colon, S., and Bhave, G. (2016). Proprotein convertase processing enhances peroxidasin activity to reinforce collagen IV. J. Biol. Chem. *291*, 24009–24016.

Cramer, S.F., and Fesyuk, A. (2012). On the development of neurocutaneous units - Implications for the histogenesis of congenital, acquired, and dysplastic nevi. Am. J. Dermatopathol. *34*, 60–81.

Crawford, N.G., Kelly, D.E., Hansen, M.E.B., Beltrame, M.H., Fan, S., Bowman, S.L., Jewett, E., Ranciaro, A., Thompson, S., Lo, Y., et al. (2017). Loci associated with skin pigmentation identified in African populations. Science (80-.). *358*.

D'Arcangelo, D., Scatozza, F., Giampietri, C., Marchetti, P., Facchiano, F., and Facchiano, A. (2019). Ion channel expression in human melanoma samples: In silico identification and experimental validation of molecular targets. Cancers (Basel). *11*.

Dassati, S., Waldner, A., and Schweigreiter, R. (2014). Apolipoprotein D takes center stage in the stress response of the aging and degenerative brain. Neurobiol. Aging *35*, 1632–1642.

Dessaud, E., Salaün, D., Gayet, O., Chabbert, M., and DeLapeyrière, O. (2006). Identification of lynx2, a novel member of the ly-6/neurotoxin superfamily, expressed in neuronal subpopulations during mouse development. Mol. Cell. Neurosci. *31*, 232–242.

Drey, E.A., Kang, M.S., McFarland, W., and Darney, P.D. (2005). Improving the accuracy of fetal foot length to confirm gestational duration. Obstet. Gynecol. *105*, 773–778.

Edge, S.B., and Compton, C.C. (2010). The american joint committee on cancer: The 7th edition of the AJCC cancer staging manual and the future of TNM. Ann. Surg. Oncol. *17*, 1471–1474.

Frontini, M., Kukalev, A., Leo, E., Ng, Y.M., Cervantes, M., Cheng, C.W., Holic, R., Dormann, D., Tse, E.,

Pommier, Y., et al. (2012). The CDK Subunit CKS2 Counteracts CKS1 to Control Cyclin A/CDK2 Activity in Maintaining Replicative Fidelity and Neurodevelopment. Dev. Cell *23*, 356–370.

Gleason, B.C., Crum, C.P., and Murphy, G.F. (2008). Expression patterns of MITF during human cutaneous embryogenesis: Evidence for bulge epithelial expression and persistence of dermal melanoblasts. J. Cutan. Pathol. *35*, 615–622.

Guo, B., Zhai, D., Cabezas, E., Welsh, K., Nouraini, S., Satterthwait, A.C., and Reed, J.C. (2003). Humanin peptide suppresses apoptosis by interfering with Bax activation. Nature *423*, 456–461.

Hayward, N.K., Wilmott, J.S., Waddell, N., Johansson, P.A., Field, M.A., Nones, K., Patch, A.M., Kakavand, H., Alexandrov, L.B., Burke, H., et al. (2017). Whole-genome landscapes of major melanoma subtypes. Nature *545*, 175–180.

Hoek, K.S., Eichhoff, O.M., Schlegel, N.C., Döbbeling, U., Kobert, N., Schaerer, L., Hemmi, S., and Dummer, R. (2008). In vivo switching of human melanoma cells between proliferative and invasive states. Cancer Res. *68*, 650–656.

Holbrook, K.A., Underwood, R.A., Vogel, A.M., Gown, A.M., and Kimball, H. (1989). The appearance, density and distribution of melanocytes in human embryonic and fetal skin revealed by the anti-melanoma monoclonal antibody, HMB-45. Anat. Embryol. (Berl). *180*, 443–455.

Hsiao, C.J., Tung, P., Blischak, J.D., Burnett, J.E., Barr, K.A., Dey, K.K., Stephens, M., and Gilad, Y. (2019). Characterizing and inferring quantitative cell cycle phase in single-cell RNA-seq data analysis. BioRxiv 526848.

Jerby-Arnon, L., Shah, P., Cuoco, M.S., Rodman, C., Su, M.J., Melms, J.C., Leeson, R., Kanodia, A., Mei, S., Lin, J.R., et al. (2018). A Cancer Cell Program Promotes T Cell Exclusion and Resistance to Checkpoint Blockade. Cell *175*, 984-997.e24.

Johannes, L., Jacob, R., and Leffler, H. (2018). Galectins at a glance. J. Cell Sci. *131*.

Joshi, S.S., Tandukar, B., Pan, L., Huang, J.M., Livak, F., Smith, B.J., Hodges, T., Mahurkar, A.A., and Hornyak, T.J. (2019). CD34 defines melanocyte stem cell subpopulations with distinct regenerative properties. PLOS Genet. *15*, e1008034.

Kuhlbrodt, K., Herbarth, B., Sock, E., Enderich, J., Hermans-Borgmeyer, I., and Wegner, M. (1998). Cooperative function of POU proteins and SOX proteins in glial cells. J. Biol. Chem. *273*, 16050–16057.

Landsberg, J., Kohlmeyer, J., Renn, M., Bald, T., Rogava, M., Cron, M., Fatho, M., Lennerz, V., Wölfel, T., Hölzel, M., et al. (2012). Melanomas resist T-cell therapy through inflammation-induced reversible dedifferentiation. Nature *490*, 412–416.

Lezcano, C., Jungbluth, A.A., Nehal, K.S., Hollmann, T.J., and Busam, K.J. (2018). PRAME Expression in Melanocytic Tumors. Am. J. Surg. Pathol. *42*, 1456–1465.

Malta, T.M., Sokolov, A., Gentles, A.J., Burzykowski, T., Poisson, L., Weinstein, J.N., Kamińska, B., Huelsken, J., Omberg, L., Gevaert, O., et al. (2018). Machine Learning Identifies Stemness Features Associated with Oncogenic Dedifferentiation. Cell *173*, 338-354.e15.

Marie, K.L., Sassano, A., Yang, H.H., Michalowski, A.M., Michael, H.T., Guo, T., Tsai, Y.C., Weissman, A.M., Lee, M.P., Jenkins, L.M., et al. (2020). Melanoblast transcriptome analysis reveals pathways promoting melanoma metastasis. Nat. Commun. *11*, 1–18.

Mica, Y., Lee, G., Chambers, S.M., Tomishima, M.J., and Studer, L. (2013). Modeling Neural Crest Induction, Melanocyte Specification, and Disease-Related Pigmentation Defects in hESCs and Patient-

Specific iPSCs. Cell Rep. *3*, 1140–1152.

Mort, R.L., Jackson, I.J., and Elizabeth Patton, E. (2015). The melanocyte lineage in development and disease. Dev. *142*, 620–632.

Nakamura, M., Fukunaga-Kalabis, M., Yamaguchi, Y., Furuhashi, T., Nishida, E., Kato, H., Mizuno, T., Sugiura, M., and Morita, A. (2015). Site-specific migration of human fetal melanocytes in volar skin. J. Dermatol. Sci. *78*, 143–148.

Newman, A.M., Liu, C.L., Green, M.R., Gentles, A.J., Feng, W., Xu, Y., Hoang, C.D., Diehn, M., and Alizadeh, A.A. (2015). Robust enumeration of cell subsets from tissue expression profiles. Nat. Methods *12*, 453–457.

Nowak, J.A., Polak, L., Pasolli, H.A., and Fuchs, E. (2008). Hair follicle stem cells are specified and function in early skin morphogenesis. Cell Stem Cell *3*, 33–43.

Okamoto, N., Aoto, T., Uhara, H., Yamazaki, S., Akutsu, H., Umezawa, A., Nakauchi, H., Miyachi, Y., Saida, T., and Nishimura, E.K. (2014). A melanocyte--melanoma precursor niche in sweat glands of volar skin. Pigment Cell Melanoma Res. *27*, 1039–1050.

Parchem, R.J., Ye, J., Judson, R.L., Larussa, M.F., Krishnakumar, R., Blelloch, A., Oldham, M.C., and Blelloch, R. (2014). Two miRNA clusters reveal alternative paths in late-stage reprogramming. Cell Stem Cell *14*, 617–631.

Picelli, S., Björklund, Å.K., Faridani, O.R., Sagasser, S., Winberg, G., and Sandberg, R. (2013). Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nat. Methods *10*, 1096–1100.

Rabbie, R., Ferguson, P., Molina-Aguilar, C., Adams, D.J., and Robles-Espinoza, C.D. (2019). Melanoma subtypes: genomic profiles, prognostic molecular markers and therapeutic possibilities. J. Pathol. *247*, 539–551.

Rambow, F., Rogiers, A., Marin-Bejar, O., Aibar, S., Femel, J., Dewaele, M., Karras, P., Brown, D., Chang, Y.H., Debiec-Rychter, M., et al. (2018). Toward Minimal Residual Disease-Directed Therapy in Melanoma. Cell *174*, 843-855.e19.

Raskin, L., Fullen, D.R., Giordano, T.J., Thomas, D.G., Frohm, M.L., Cha, K.B., Ahn, J., Mukherjee, B., Johnson, T.M., and Gruber, S.B. (2013). Transcriptome profiling identifies HMGA2 as a biomarker of melanoma progression and prognosis. J. Invest. Dermatol. *133*, 2585–2592.

Richard, G., Dalle, S., Monet, M., Ligier, M., Boespflug, A., Pommier, R.M., Fouchardière, A., Perier-Muzet, M., Depaepe, L., Barnault, R., et al. (2016). ZEB 1-mediated melanoma cell plasticity enhances resistance to MAPK inhibitors . EMBO Mol. Med. *8*, 1143–1161.

Strub, T., Giuliano, S., Ye, T., Bonet, C., Keime, C., Kobi, D., Le Gras, S., Cormont, M., Ballotti, R., Bertolotto, C., et al. (2011). Essential role of microphthalmia transcription factor for DNA replication, mitosis and genomic stability in melanoma. Oncogene *30*, 2319–2332.

Tirosh, I., Izar, B., Prakadan, S.M., Wadsworth, M.H., Treacy, D., Trombetta, J.J., Rotem, A., Rodman, C., Lian, C., Murphy, G., et al. (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science (80-.). *352*, 189–196.

Tsoi, J., Robert, L., Paraiso, K., Galvan, C., Sheu, K.M., Lay, J., Wong, D.J.L., Atefi, M., Shirazi, R., Wang, X., et al. (2018). Multi-stage Differentiation Defines Melanoma Subtypes with Differential Vulnerability to Drug-Induced Iron-Dependent Oxidative Stress. Cancer Cell *33*, 890-904.e5.

Webster, M.R., Xu, M., Kinzler, K.A., Kaur, A., Appleton, J., O'Connell, M.P., Marchbank, K., Valiga, A.,

Dang, V.M., Perego, M., et al. (2015). Wnt5A promotes an adaptive, senescent-like stress response, while continuing to drive invasion in melanoma cells. Pigment Cell Melanoma Res. *28*, 184–195.

Webster, M.R., Fane, M.E., Alicea, G.M., Basu, S., Kossenkov, A. V., Marino, G.E., Douglass, S.M., Kaur, A., Ecker, B.L., Gnanapradeepan, K., et al. (2020). Paradoxical Role for Wild-Type p53 in Driving Therapy Resistance in Melanoma. Mol. Cell *77*, 633-644.e5.

Widmer, D.S., Cheng, P.F., Eichhoff, O.M., Belloni, B.C., Zipser, M.C., Schlegel, N.C., Javelaud, D., Mauviel, A., Dummer, R., and Hoek, K.S. (2012). Systematic classification of melanoma cells by phenotype-specific gene expression mapping. Pigment Cell Melanoma Res. *25*, 343–353.

Yamaguchi, Y., Kubo, T., Tarutani, M., Sano, S., Asada, H., Kakibuchi, M., Hosokawa, K., Itami, S., and Yoshikawa, K. (2001). Epithelial-mesenchymal interactions in wounds: treatment of palmoplantar wounds by nonpalmoplantar pure epidermal sheet grafts. Arch. Dermatol. *137*, 621–628.

Yamaguchi, Y., Itami, S., Watabe, H., Yasumoto, K.I., Abdel-Malek, Z.A., Kubo, T., Rouzaud, F., Tanemura, A., Yoshikawa, K., and Hearing, V.J. (2004). Mesenchymal-epithelial interactions in the skin: Increased expression of dickkopf1 by palmoplantar fibroblasts inhibits melanocyte growth and differentiation. J. Cell Biol. *165*, 275–285.

Zhang, F., Kumano, M., Beraldi, E., Fazli, L., Du, C., Moore, S., Sorensen, P., Zoubeidi, A., and Gleave, M.E. (2014). Clusterin facilitates stress-induced lipidation of LC3 and autophagosome biogenesis to enhance cancer cell survival. Nat. Commun. *5*, 1–13.

Figure 1

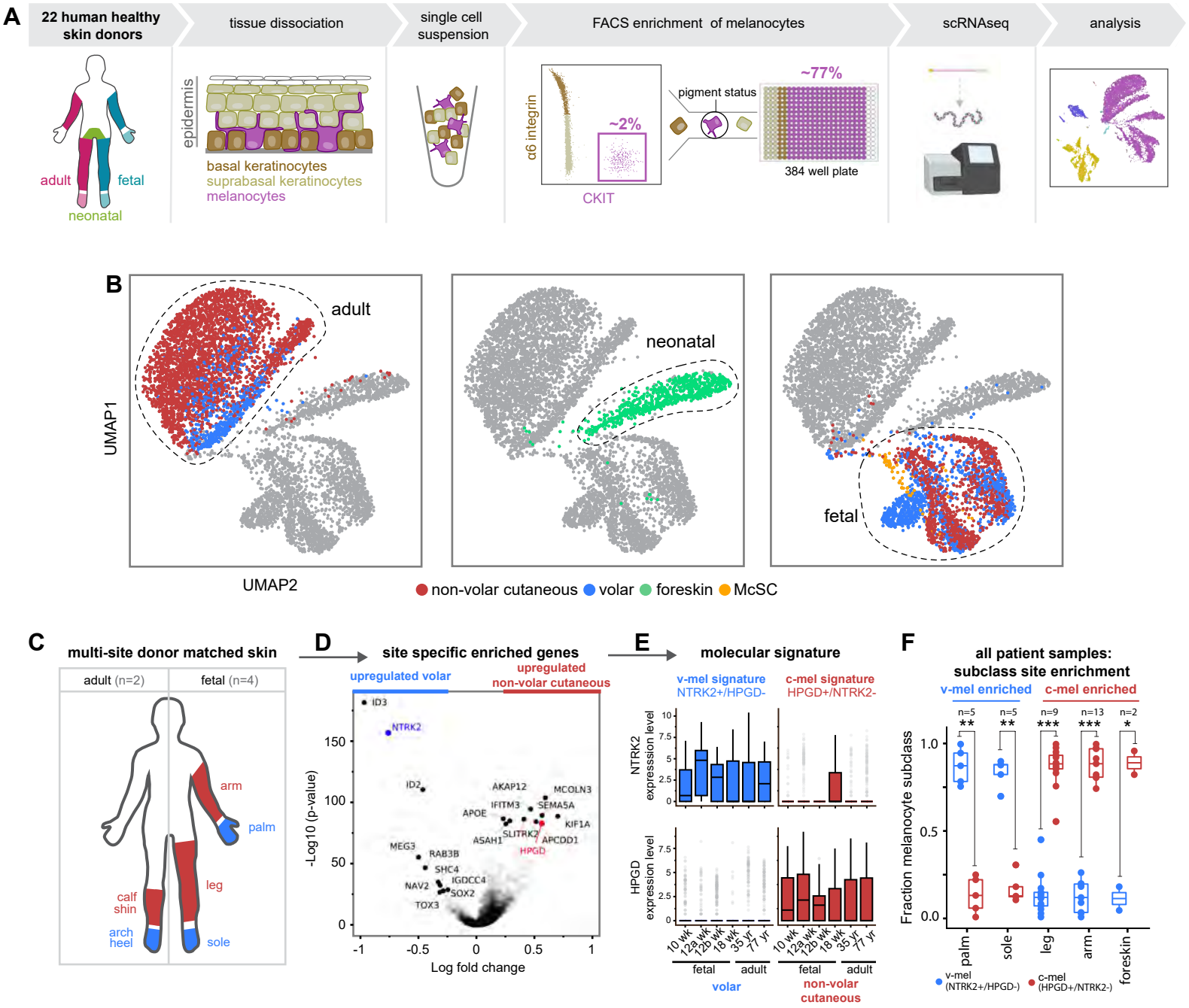


Figure 2

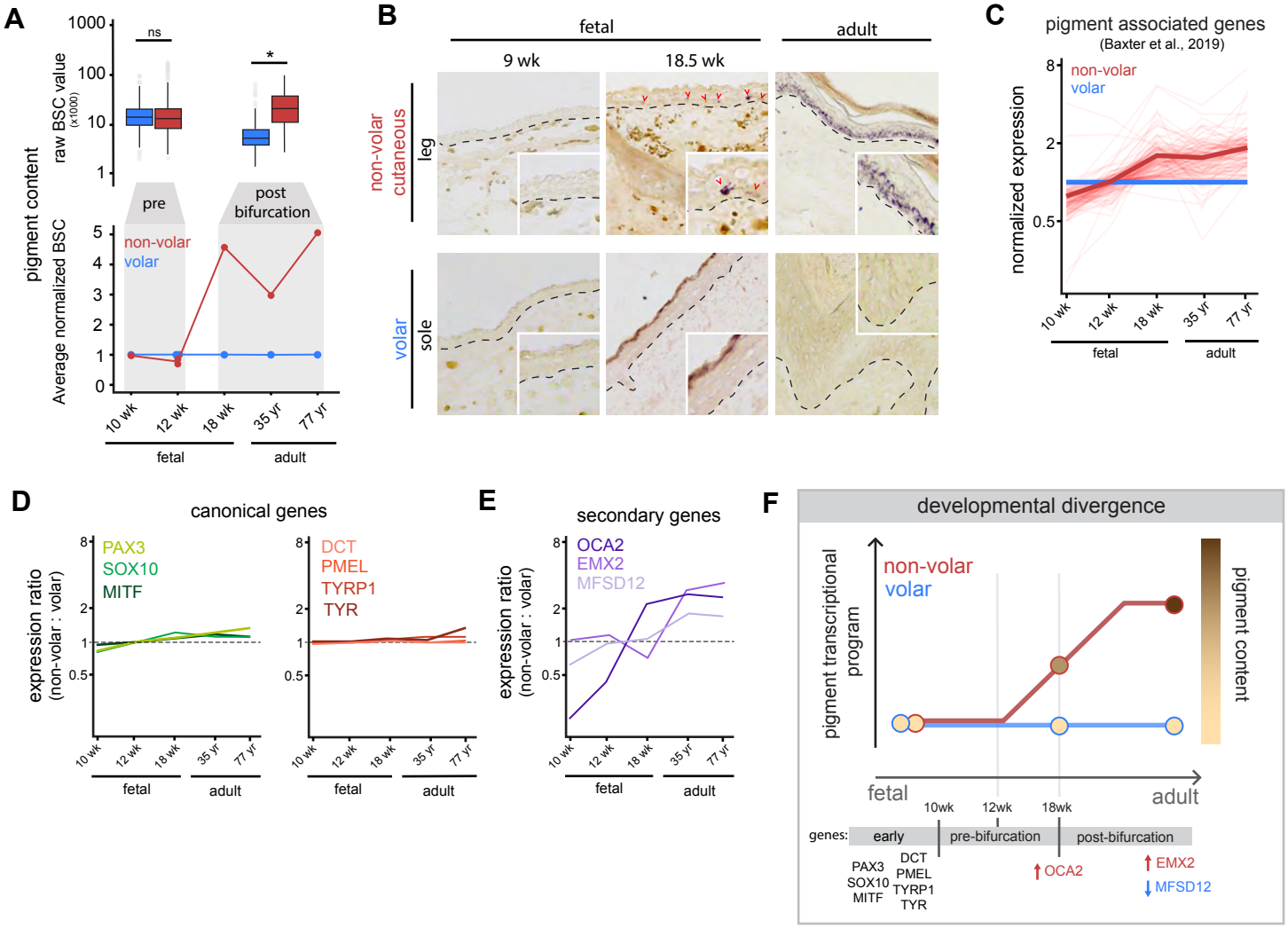
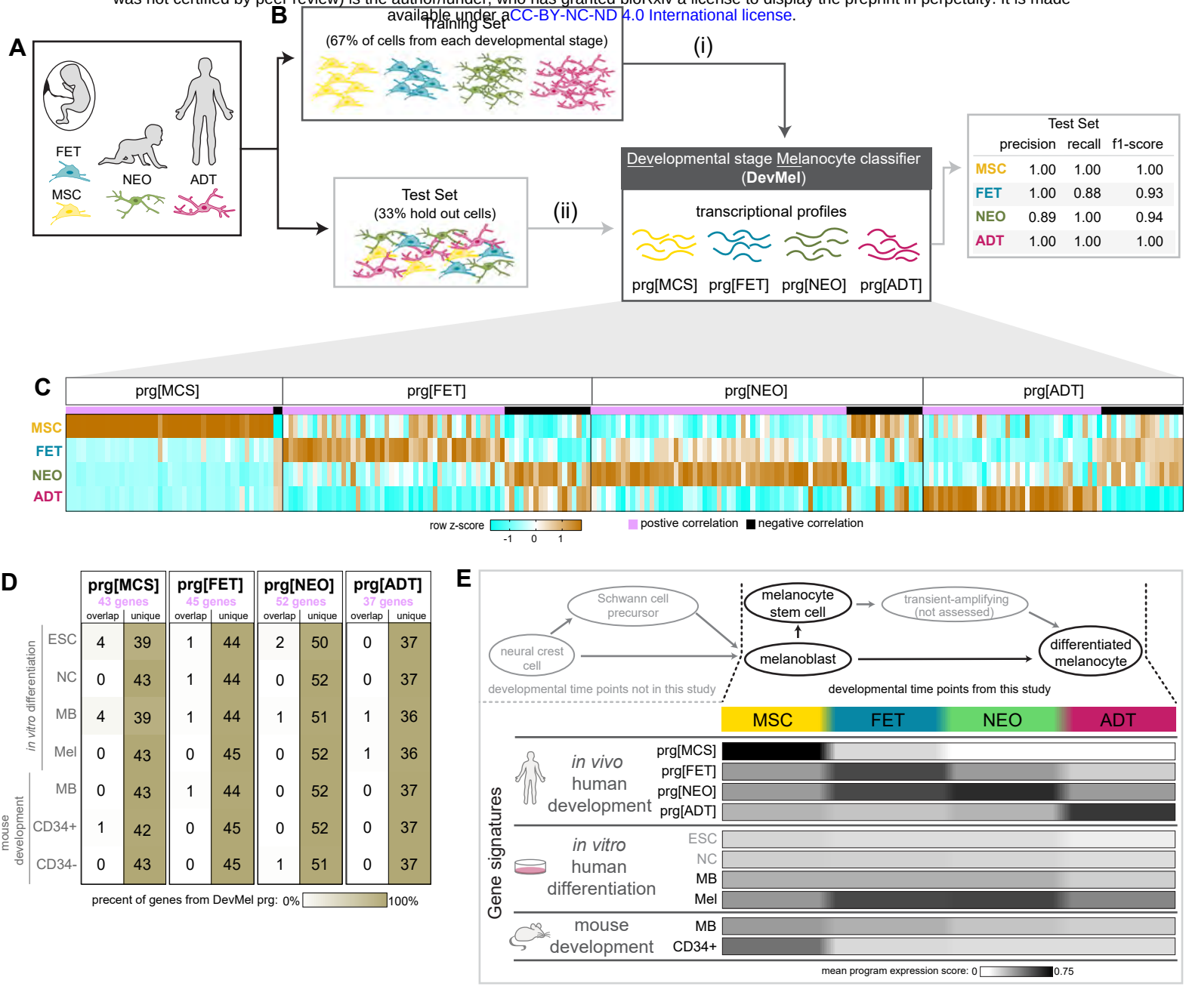


Figure 3



D

		prg[MCS] (43 genes)		prg[FET] (45 genes)		prg[NEO] (52 genes)		prg[ADT] (37 genes)	
		overlap	unique	overlap	unique	overlap	unique	overlap	unique
in vitro differentiation	ESC	4	39	1	44	2	50	0	37
	NC	0	43	1	44	0	52	0	37
	MB	4	39	1	44	1	51	1	36
	Mel	0	43	0	45	0	52	1	36
mouse development	MB	0	43	1	44	0	52	0	37
	CD34+	1	42	0	45	0	52	0	37
	CD34-	0	43	0	45	1	51	0	37

precent of genes from DevMel prg: 0% 100%

E

Figure 4

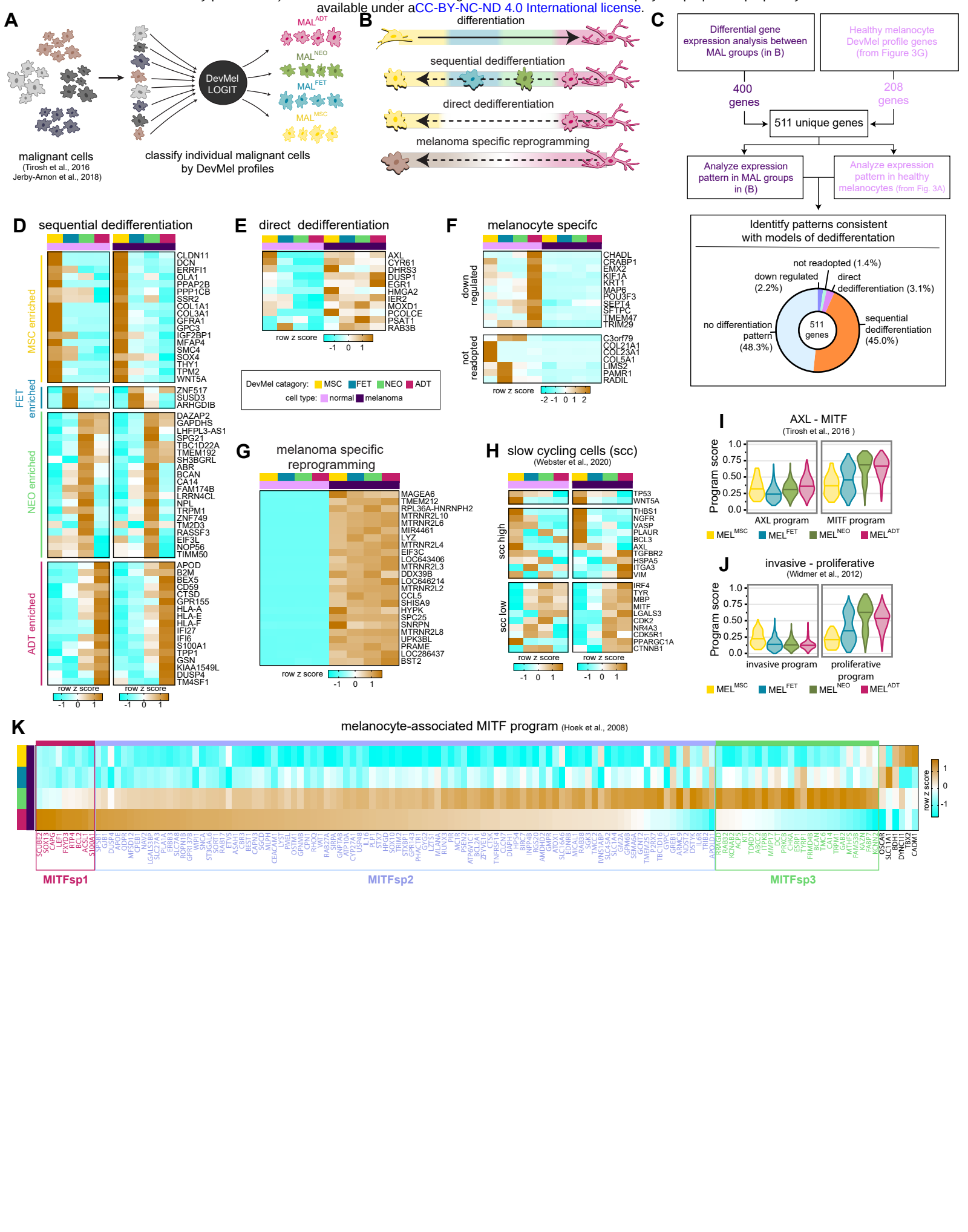


Figure 5

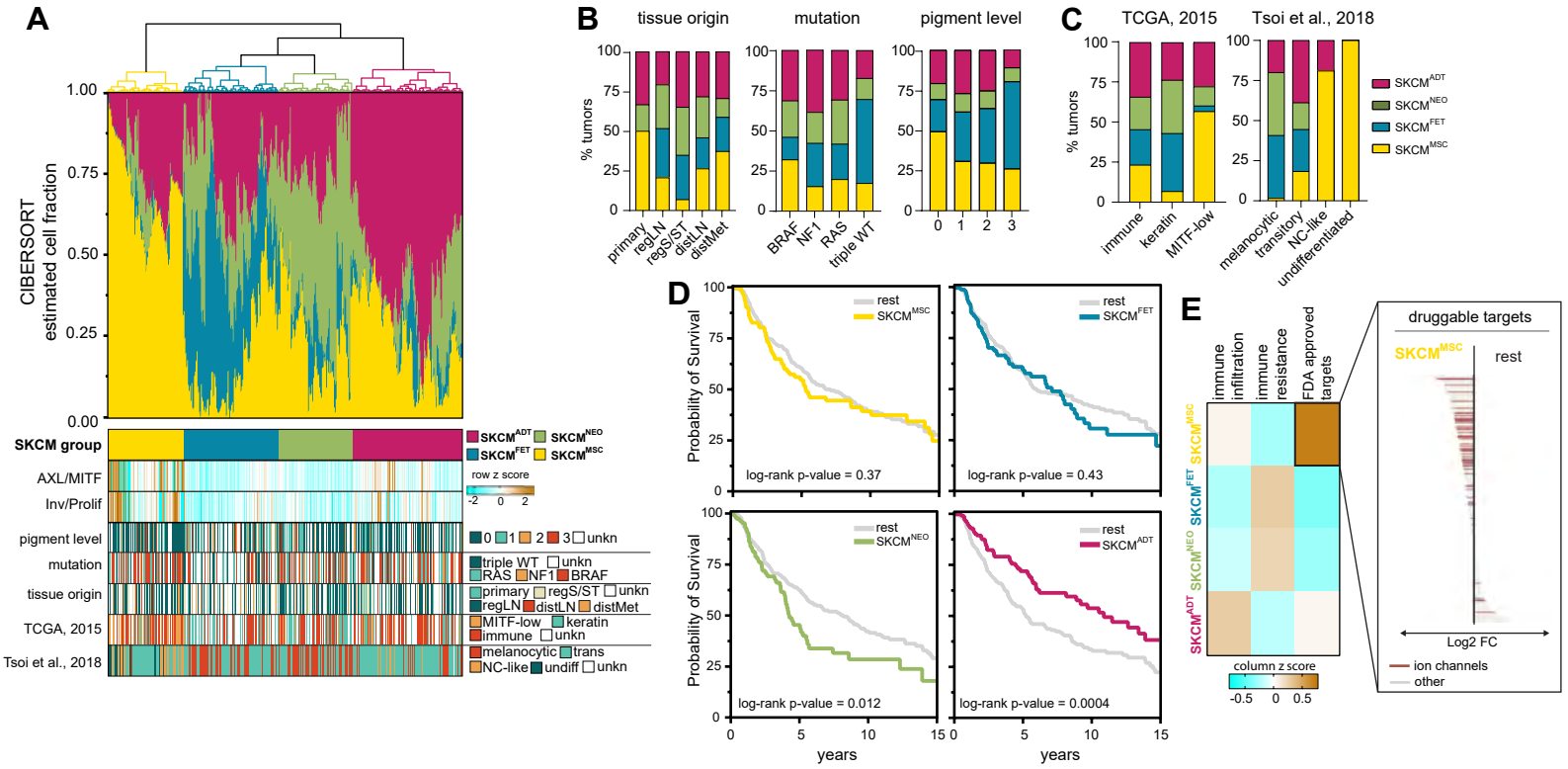


Figure 6

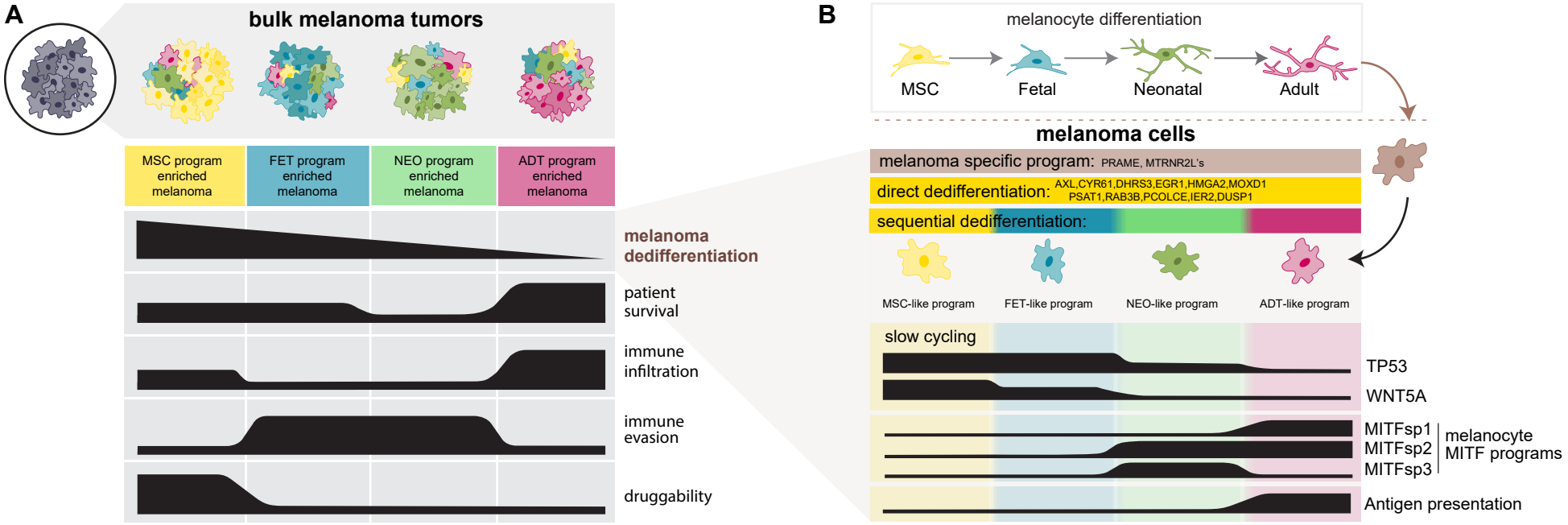

A
bulk melanoma tumors
MSC program enriched melanoma
FET program enriched melanoma
NEO program enriched melanoma
ADT program enriched melanoma
melanoma dedifferentiation
patient survival
immune infiltration
immune evasion
druggability
B
melanocyte differentiation
MSC
Fetal
Neonatal
Adult
melanoma cells
melanoma specific program: PRAME, MTRNR2L's
direct dedifferentiation: AXL,CYR61,DHRS3,EGR1,HMGA2,MOXD1 PSAT1,RAB3B,PCOLCE,IER2,DUSP1
sequential dedifferentiation:
MSC-like program
FET-like program
NEO-like program
ADT-like program
slow cycling
TP53
WNT5A
MITFsp1
MITFsp2
MITFsp3
melanocyte MITF programs
Antigen presentation

Figure S1

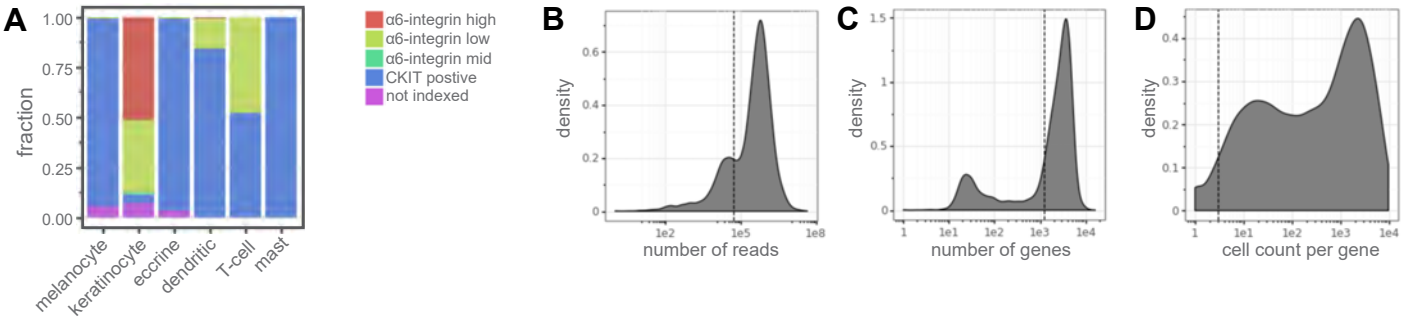
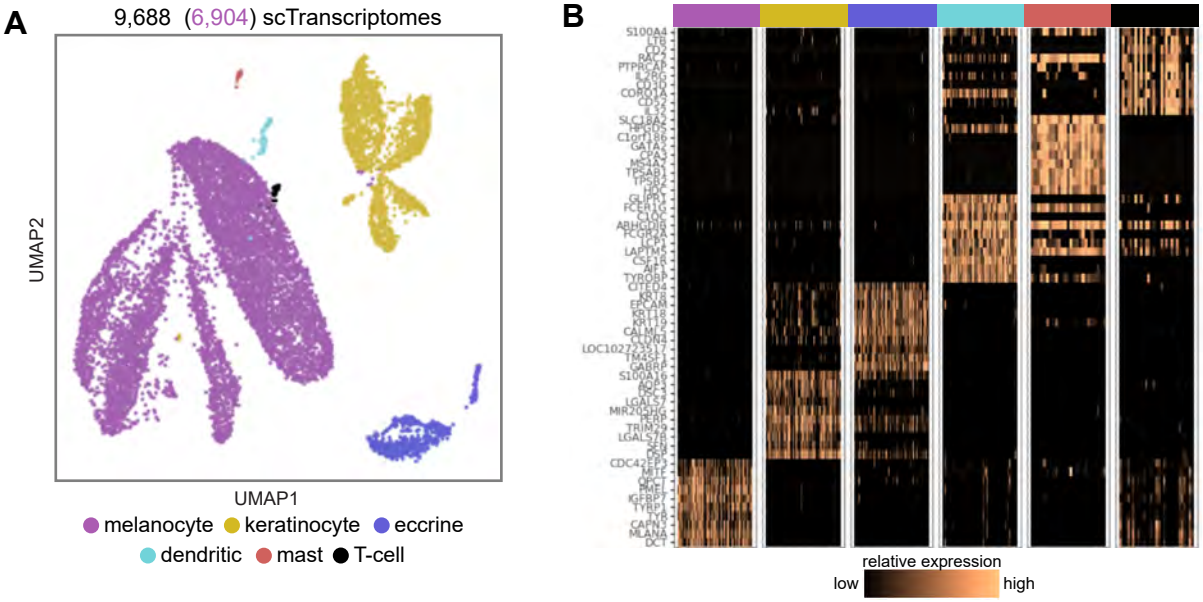
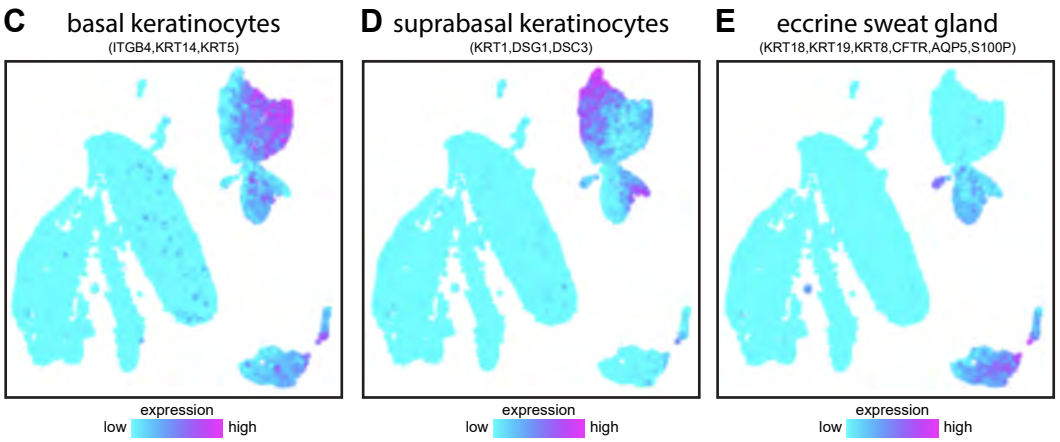
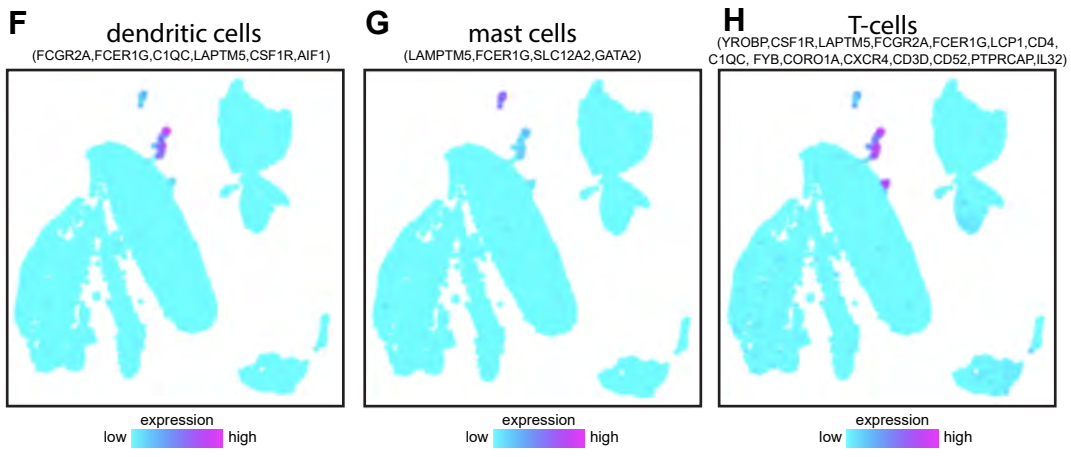


Figure S2







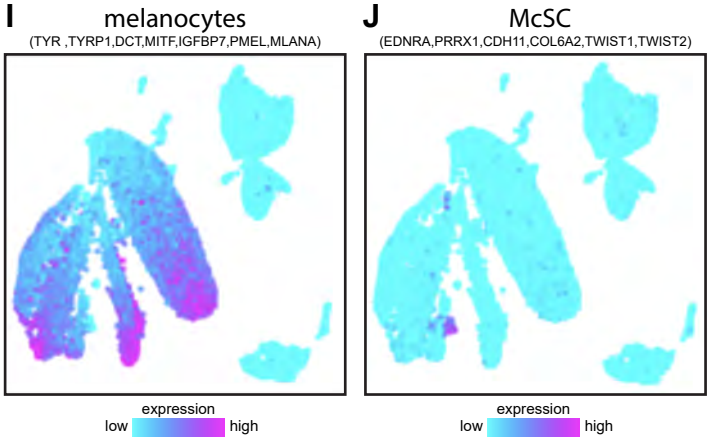


Figure S3

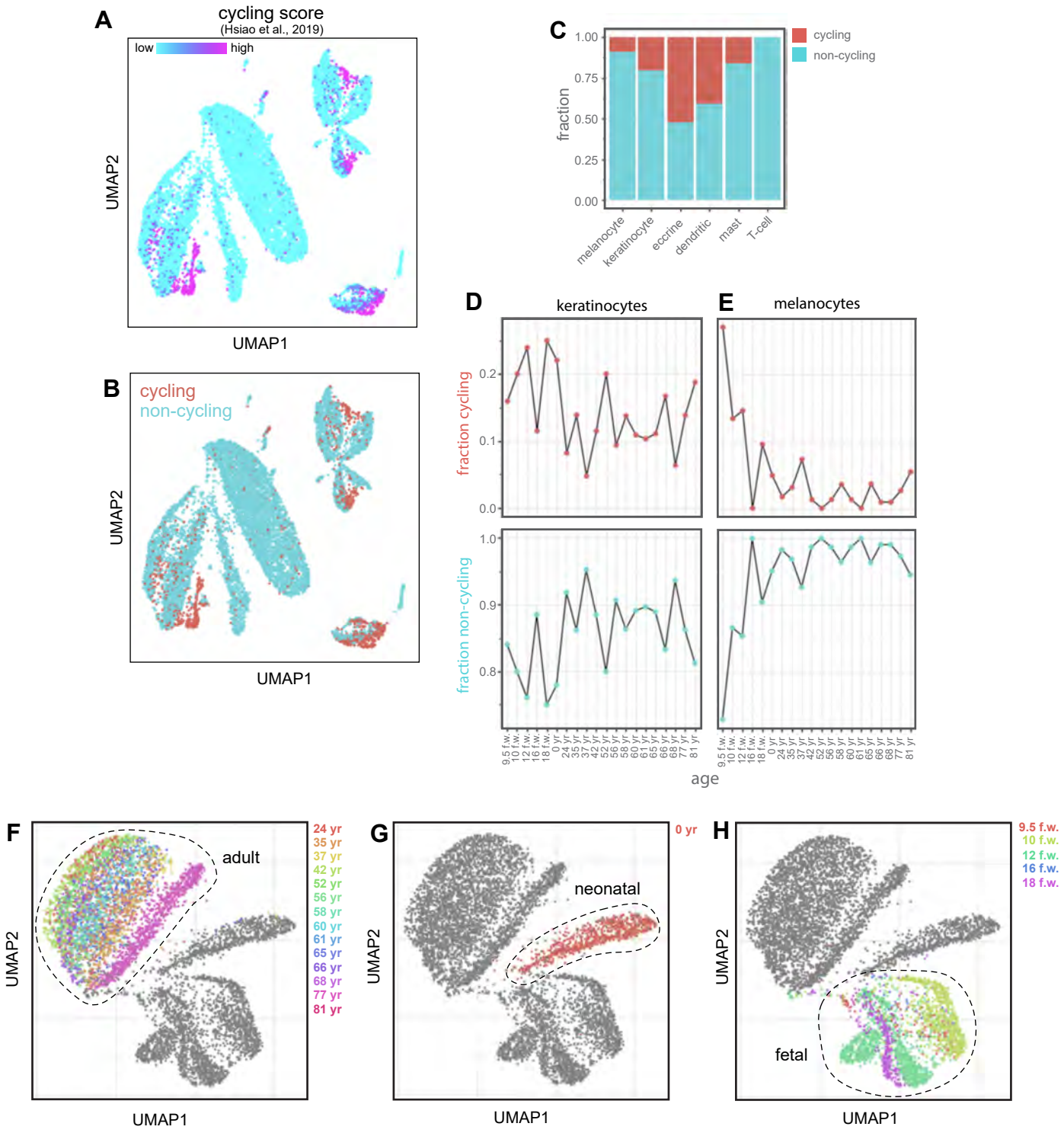


Figure S4

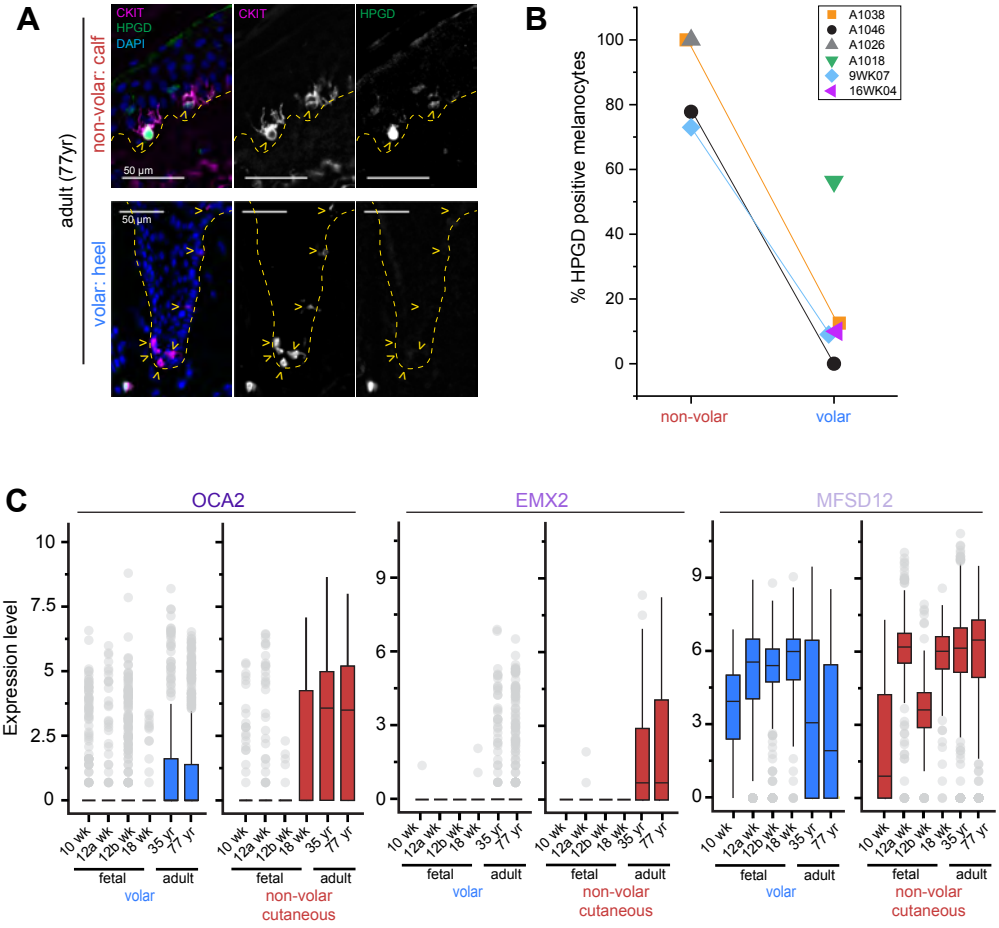


Figure S5

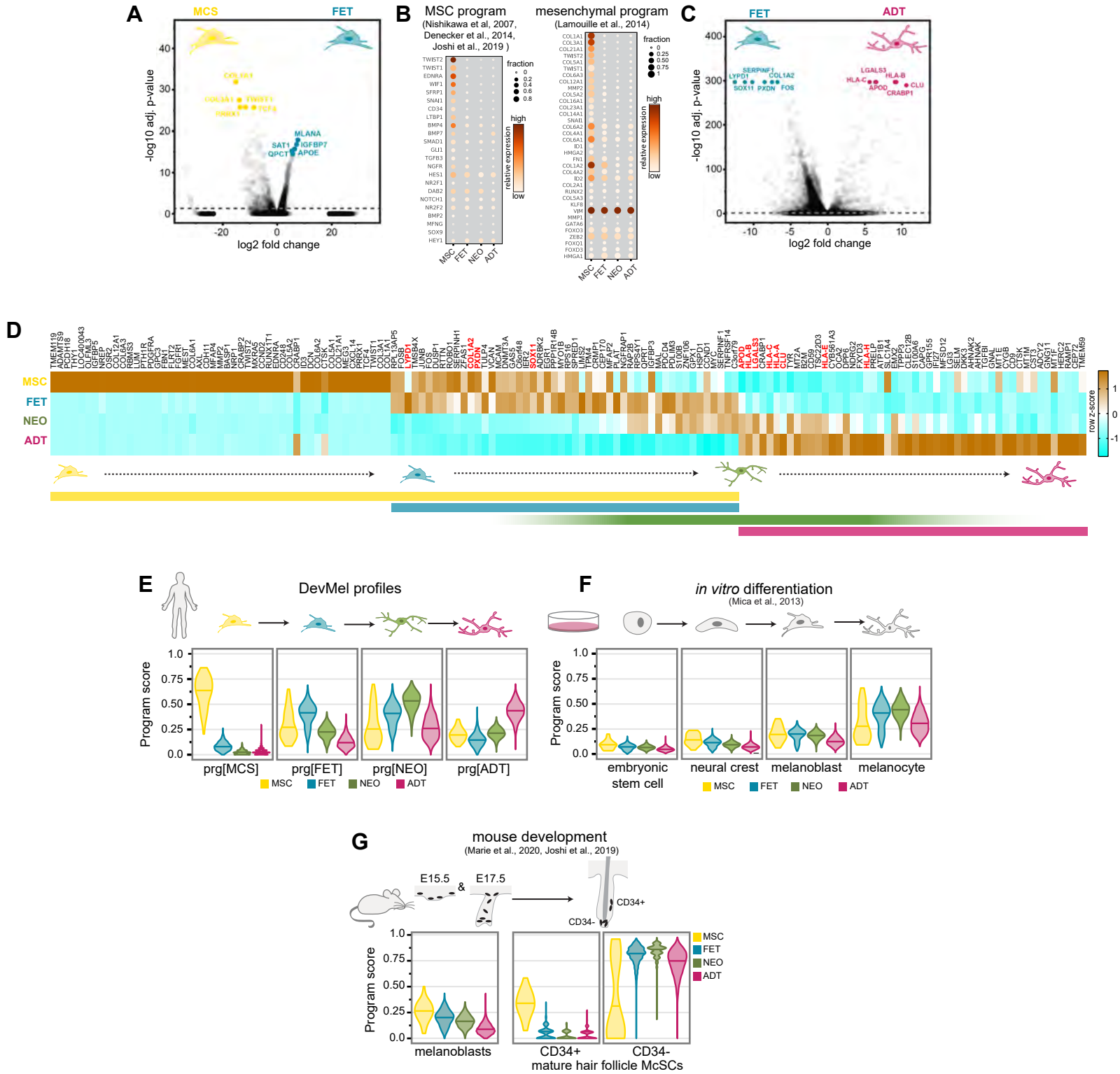


Figure S6