RESEARCH ARTICLE SUMMARY

MELANOMA INITIATION

A zebrafish melanoma model reveals emergence of neural crest identity during melanoma initiation

Charles K. Kaufman, Christian Mosimann, Zi Peng Fan, Song Yang, Andrew J. Thomas, Julien Ablain, Justin L. Tan, Rachel D. Fogley, Ellen van Rooijen, Elliott J. Hagedorn, Christie Ciarlo, Richard M. White, Dominick A. Matos, Ann-Christin Puller, Cristina Santoriello, Eric C. Liao, Richard A. Young, Leonard I. Zon*

INTRODUCTION: The "cancerized field" concept posits that cells in a given tissue sharing an oncogenic mutation are cancer-prone, yet only discreet clones within the field initiate tumors. Studying the process of cancer initiation has remained challenging because of (i) the rarity of these events, (ii) the difficulty of visiualizing initiating clones in living organisms, and (iii) the transient nature of a newly transformed clone emerging before it expands to form an early tumor. A more complete understanding of the molecular processes that regulate cancer initiation could provide important prognostic information about which precancerous lesions are most prone to becoming cancer and also implicate druggable molecular pathways that, when inhibited, may prevent the cancer from ever starting.

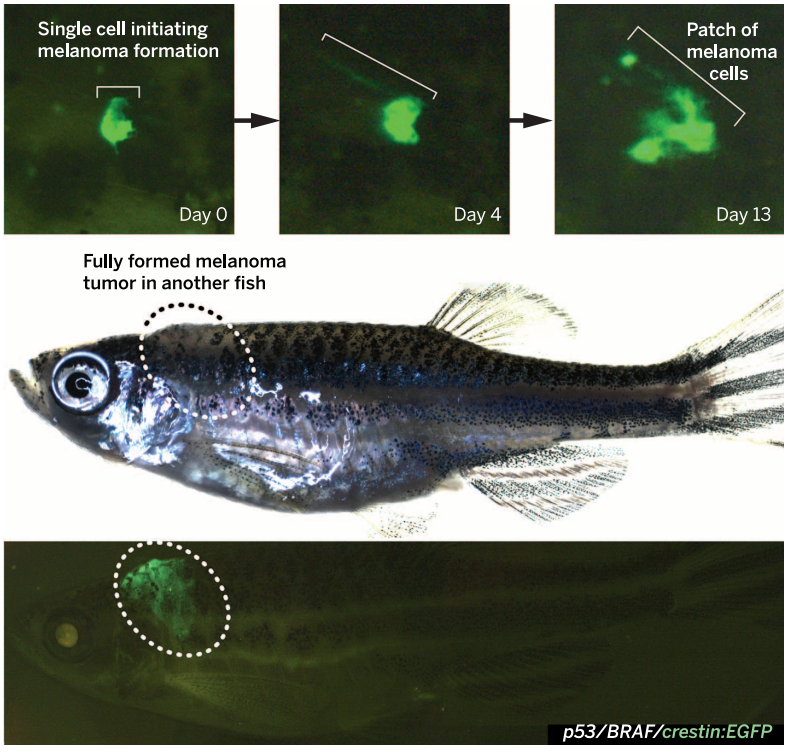
RATIONALE: The majority of benign nevi carry oncogenic *BRAF*^{V600E} mutations and can be considered a cancerized field of melanocytes, but they only rarely convert to melanoma. In an effort to define events that initiate cancer, we used a melanoma model in the zebrafish in which the human *BRAF*^{V600E} oncogene is driven by the melanocyte-specific *mitfa* promoter. When bred into a *p53* mutant background, these fish develop melanoma tumors over the course of many months. The zebrafish *crestin* gene is expressed embryonically in neural crest progenitors (NCPs) and is specifically reexpressed only in melanoma tumors, making it an ideal candidate for tracking melanoma from initiation onward.

RESULTS: We developed a *crestin:EGFP* reporter that recapitulates the embryonic neural crest expression pattern of *crestin* and its expression in melanoma tumors. We show through live imaging of transgenic zebrafish *crestin* reporters that within a cancerized field (*BRAF*^{V600E}-mutant; *p53*-deficient), a single melanocyte reactivates the NCP state, and this establishes that a fate change occurs at melanoma initiation in this model. Early *crestin*$^{+}$ patches of cells expand and are transplantable in a manner consistent with their possessing tumorigenic activity, and they exhibit a gene expression pattern consistent with the NCP identity readout by the *crestin* reporter. The *crestin* element is regulated by NCP transcription factors, including *sox10*. Forced *sox10* overexpression in melanocytes accelerated melanoma formation, whereas CRISPR/Cas9 targeting of *sox10* delayed melanoma onset. We show activation of super-enhancers at NCP genes in both zebrafish and human melanomas, identifying an epigenetic mechanism for control of this NCP signature leading to melanoma.

ON OUR WEB SITE

Read the full article at http://dx.doi.org/10.1126/science.aad2197

CONCLUSION: This work using our zebrafish melanoma model and in vivo reporter of NCP identity allows us to see cancer from its birth as a single cell and shows the importance of NCP-state reemergence as a key event in melanoma initiation from a field of cancer-prone melanocytes. Thus, in addition to the typical fixed genetic alterations in oncogenes and tumor supressors that are required for cancer development, the reemergence of progenitor identity may be an additional rate-limiting step in the formation of melanoma. Preventing NCP reemergence in a field of cancer-prone melanocytes may thus prove therapeutically useful, and the association of NCP genes with super-enhancer regulatory elements implicates the associated druggable epigenetic machinery in this process. ■



Neural crest reporter expression in melanoma. The *crestin:EGFP* transgene is specifically expressed in melanoma in *BRAF*^{V600E}/*p53* mutant melanoma-prone zebrafish. (Top) A single cell expressing *crestin:EGFP* expands into a small patch of cells over the course of 2 weeks, capturing the initiation of melanoma formation (bracket). (Bottom) A fully formed melanoma specifically expresses *crestin:EGFP*, whereas the rest of the fish remains *EGFP*-negative.

The list of author affiliations is available in the full article online.
*Corresponding author. E-mail: zon@enders.tch.harvard.edu
Cite this article as C. K. Kaufman *et al.*, *Science* **351**, aad2197 (2016). DOI: 10.1126/science.aad2197

RESEARCH ARTICLE

MELANOMA INITIATION

A zebrafish melanoma model reveals emergence of neural crest identity during melanoma initiation

Charles K. Kaufman,[1,2,3,4] Christian Mosimann,[5] Zi Peng Fan,[6,7] Song Yang,[1,2] Andrew J. Thomas,[1] Julien Ablain,[1,2,4] Justin L. Tan,[1] Rachel D. Fogley,[1] Ellen van Rooijen,[1,2,4] Elliott J. Hagedorn,[1,2,4] Christie Ciarlo,[1,4] Richard M. White,[8] Dominick A. Matos,[9] Ann-Christin Puller,[10] Cristina Santoriello,[1,11] Eric C. Liao,[2,4,12] Richard A. Young,[6,13] Leonard I. Zon[1,2,3,4,11]*

The "cancerized field" concept posits that cancer-prone cells in a given tissue share an oncogenic mutation, but only discreet clones within the field initiate tumors. Most benign nevi carry oncogenic *BRAF*^{V600E} mutations but rarely become melanoma. The zebrafish *crestin* gene is expressed embryonically in neural crest progenitors (NCPs) and specifically reexpressed in melanoma. Live imaging of transgenic zebrafish *crestin* reporters shows that within a cancerized field (*BRAF*^{V600E}-mutant; *p53*-deficient), a single melanocyte reactivates the NCP state, revealing a fate change at melanoma initiation in this model. NCP transcription factors, including *sox10*, regulate *crestin* expression. Forced *sox10* overexpression in melanocytes accelerated melanoma formation, which is consistent with activation of NCP genes and super-enhancers leading to melanoma. Our work highlights NCP state reemergence as a key event in melanoma initiation.

Understanding the earliest events in cancer formation remains an incompletely fulfilled goal in biology, with important implications for human health. In cancer initiation, an activated oncogene or inactivated tumor suppressor can trigger tumor formation. However, it is unclear as to why only sporadic cells with these genetic alterations complete the conversion to a malignant state when they are present in a large group of cancer-prone cells, sometimes described as a "cancerized field" (*1*). Better characterizing initiating events would help identify targets for early therapeutic interventions and also provide prognostic information about which precancerous lesions are most likely to progress to cancer.

Melanoma is a cancer of transformed melanocytes, which are pigment-producing cells derived from the embryonic neural crest lineage. It is frequently driven by *BRAF* or *RAS* mutations (~80% of cases) (*2*, *3*). Melanoma is treatable and curable when it is localized and resected completely but remains largely incurable once it has spread, even when treated with new kinase- and immune checkpoint–targeted therapies (*4*). Our laboratory previously developed an animal model of a *BRAF*^{V600E}-driven cancer by placing the human *BRAF*^{V600E} gene under the control of the melanocyte-specific *mitfa*-promoter in transgenic zebrafish (*5*). When crossed into a *p53* mutant loss-of-function background, these zebrafish (referred to here as *p53/BRAF*) invariably develop nevi and, after several months, invasive melanoma (*5*). Despite creating this extensive "cancerized field" in which all melanocytes harbor both oncogenic *BRAF*^{V600E} and *p53* loss throughout their life span, these *p53/BRAF* melanoma–prone zebrafish all develop one to three melanoma tumors after several months of age, indicating that other molecular alterations are important for tumor initiation.

crestin transgenics mark neural crest

To investigate the dynamics and mechanism of sporadic melanoma formation, we visualized and characterized melanoma lesion initiation. The functionally uncharacterized zebrafish *crestin* gene marks the neural crest during embryonic development, becomes undetectable by ~72 hours after fertilization (*6*, *7*), and is specifically reexpressed in melanoma tumors in adult zebrafish (*8*). We reasoned that a *crestin*-based reporter transgene would allow us to track embryonic neural crest cells as well as melanoma tumors in vivo, potentially from their earliest onset. We amplified by means of polyermase chain reaction (PCR) a 4.5-kb upstream region common to multiple *crestin* insertions in the zebrafish genome and cloned this element upstream of an *enhanced green fluorescent protein* (*EGFP*) reporter (Fig. 1A, *crestin:EGFP*). In stable transgenic zebrafish embryos, this construct reproduced *crestin* mRNA expression through EGFP fluorescence (Fig. 1, B and C, and fig. S1A), and time-lapse videos demonstrated the dorsal emergence and wide migration of these *crestin*-expressing putative neural crest progenitor cells (movies S1 and S2). Neural crest expression was reproducible in multiple independent lines and with additional reporter genes (*creERT2* and *mCherry*) (Fig. 1, D to G, and fig. S5, A to C). As with endogenous *crestin* expression, transgenic *crestin:EGFP* expression was not detectable after 3 days after fertilization and did not come back on in wild-type juvenile or adult zebrafish.

To confirm that the *crestin* transgenes target neural crest progenitors, we also generated transgenics for *crestin:creERT2* to genetically mark *crestin*-expressing embryonic cells using a Cre/*lox*–dependent *EGFP*-to-*mCherry* switching line ("*ubi:switch*") (*9*) and genetically labeled neural crest-derived cells, including melanocytes/pigment cells (Fig. 1, D and E, red cells), jaw cartilage (Fig. 1F), and lateral line glia (Fig. 1G). Because the *crestin* gene is specific to zebrafish, we wanted to ensure that *crestin* reporter embryonic expression is consistent with another conserved early neural crest marker, the transcription factor *sox10*. Confocal analysis of double-transgenic *Tg(crestin_1kb:EGFP)* and *Tg(sox10:mCh)* (*10*) zebrafish embryos showed a high degree of overlap in reporter gene expression (Fig. 1H), with any differences matching published in situ hybridization (ISH) data (*11*). Thus, our *crestin* transgenic lines recapitulate *crestin* expression and specifically mark the embryonic neural crest stem/progenitor cell population.

crestin transgenics visualize melanoma initiation

We next determined whether *crestin:EGFP* is reexpressed in melanoma tumors, as noted previously by ISH (*8*). We found *crestin:EGFP* is expressed in tumors arising on triple transgenic *p53/BRAF/crestin:EGFP* adult zebrafish but is absent in the remainder of the animal, highlighting its specificity to the tumor (Fig. 2A). We next followed developing zebrafish in order to observe the onset of *crestin:EGFP*$^{+}$ expression. Before *EGFP*-expressing patches of cells formed raised melanoma lesions on a given fish (Fig. 2B), we were able to detect single isolated *EGFP*$^{+}$ cells in *p53/BRAF/crestin:EGFP* zebrafish (Fig. 2C). We could track their persistence and enlargement (fig. S2, A and B). Small patches of *EGFP*$^{+}$ cells, containing <50 cells, are readily tractable as they enlarge (fig. S2C). Analysis of single scales with

[1]Stem Cell Program and Division of Hematology/Oncology, Children's Hospital Boston, Howard Hughes Medical Institute, Boston, MA 02115, USA. [2]Harvard Stem Cell Institute, Boston, MA 02115, USA. [3]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA 02215, USA. [4]Harvard Medical School, Boston, MA 02115, USA. [5]Institute of Molecular Life Sciences, University of Zürich, 8057 Zürich, Switzerland. [6]Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142, USA. [7]Computational and Systems Biology Program, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [8]Memorial Sloan Kettering Cancer Center, Weill Cornell Medical College, New York, NY 10075, USA. [9]Massachusetts General Hospital Cancer Center, Harvard Medical School, Charlestown, MA 02129, USA. [10]Research Institute Children's Cancer Center Hamburg and Department of Pediatric Hematology and Oncology, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany. [11]Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA 02138, USA. [12]Center for Regenerative Medicine, Massachusetts General Hospital, Boston, MA 02114, USA. [13]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.
*Corresponding author. E-mail: zon@enders.tch.harvard.edu

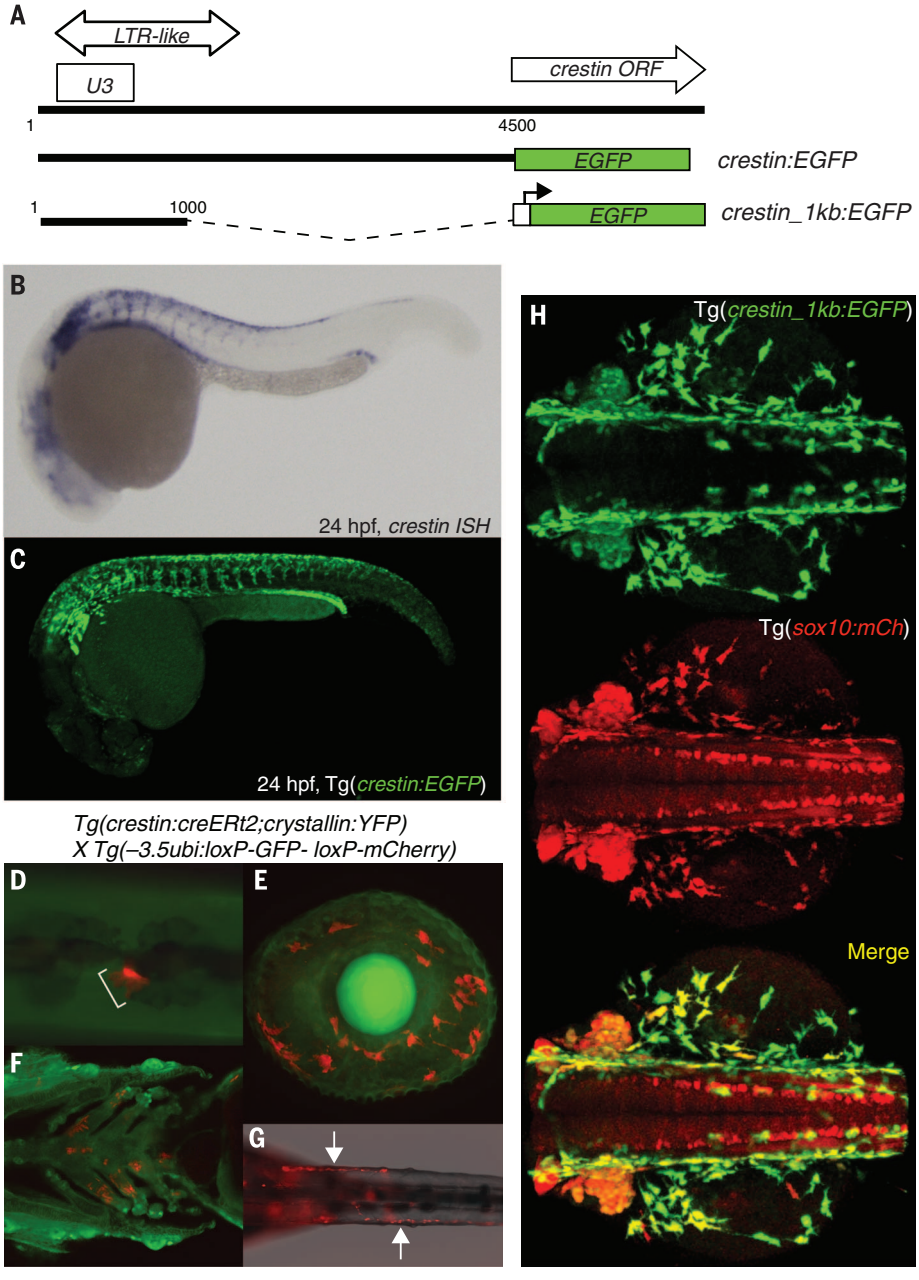


Fig. 1. The *crestin* promoter/enhancer drives neural crest-specific gene expression. (**A**) Prototypical *crestin* retrotransposon locus with predicted ORF, LTR-like, and U3-like promoter regions. Shown are locations of 4.5- and 1-kb segments used for *crestin:EGFP* constructs (white box/promoter arrow indicate β-globin gene minimal promoter). (**B**) Endogenous expression pattern of *crestin* transcript by means of ISH (purple staining) at 24 hours after fertilization marks developing and migrating neural crest cells. (**C**) This expression pattern (green) is recapitulated by a stable *Tg(crestin:EGFP)* embryo at 24 hours after fertilization. (**D** to **G**) Genetic lineage tracing of cells that express *crestin* [*Tg(crestin:creERt2;crystallin:YFP)* X *Tg(–3.5ubi:loxP-GFP-loxP-mCherry)*] marks multiple neural crest lineages (red cells), including melanocytes (bracket) on (D) the dorsum and (E) the eye (72 hours after fertilization), (F) jaw cartilage (ventral view, 5 days after fertilization), and (G) glial cells of the lateral line (arrows, dorsal view posterior to the yolk, 72 hours after fertilization). (**H**) *Tg(crestin:EGFP)* expression overlaps substantially with a *sox10:mCh* transgene (confocal image, dorsal view over yolk, 24 hours after fertilization).

discrete *crestin:EGFP*$^+$ patches demonstrated that transgene expression detectable with fluorescence microscopy overlaps with *crestin* mRNA detected with ISH (Fig. 2D). Together, these observations reveal that after pan-neural crest expression confined to the embryo, our *crestin* reporter expresses specifically and reproducibly in melanoma tumors, thus providing an in vivo genetic label for melanoma cells that is earlier than with previous detection methods (*5*, *12*).

We next addressed the dynamics of reemerging *crestin* expression in cohorts of *p53/BRAF/crestin:EGFP* zebrafish. At the population level, *crestin:EGFP*$^+$ patches of cells (fig. S1, B and C) were visible before the appearance of grossly raised melanoma lesions (Fig. 2E, fig. S1D, and movie S3). The *crestin:EGFP* expression is undetectable in the *p53/BRAF* fish from 3 to >21 days after fertilization, which is again consistent with previous in situ analyses for endogenous *crestin*. We tracked individual small patches of *crestin:EGFP*$^+$ cells over time as they progressed into fully formed raised melanoma lesions (Fig. 2E) and found that all melanomas tracked in this manner initiated from *crestin:EGFP*$^+$ patches of cells (30 out of 30). Thus, if a patch is seen in the *p53/BRAF* background, it will become an overt melanoma. These data demonstrate that reemergence of *crestin:EGFP* expression, and a neural crest progenitor state, correlates with melanoma initiation in an in vivo model of de novo melanoma formation.

To establish that pretumor patches of *crestin:EGFP*$^+$ cells are tumorigenic and can autonomously expand locally after transplant, we performed scale auto-transplants on *p53/BRAF/crestin:EGFP* zebrafish (*13*). After transplant, patches of *crestin:EGFP*$^+$ cells survive and expand at the new site. The *EGFP*$^+$ cells persist and further expand when later removing the transplanted scale, suggesting that the cells have invaded the hypodermis (Fig. 2G, representative example). We achieved similar results with isolated scales placed in tissue culture but on a shorter time scale (figs. S3A, and S4, A and B) and with allotransplants to sublethally irradiated recipient zebrafish (fig. S3B) (*8*). Thus, early patches of *crestin:EGFP*–expressing cells are transplantable in a manner suggesting that they are already tumorigenic.

Transcriptional regulators of *crestin* expression

As the *crestin* element proved to be a highly specific and distinct tool for monitoring neural crest and melanoma development, we aimed to identify (i) a minimal element within the 4.5-kb *crestin* promoter/enhancer that could drive this expression pattern and (ii) key transcriptional regulators within the element. Sequence analysis of the *crestin* locus, which is replicated throughout the zebrafish genome >40 times, is similar to another retroelement called *bhikari* that is expressed in early mesendoderm (fig. S6A) (*6*, *7*, *14*). Both a 1-kb segment from the putative retroelement promoter region and a smaller 296–base pair (bp) subregion fully reproduced the neural crest- and melanoma-specfic expression pattern of the full 4.5-kb *crestin* element (Fig. 1H and fig. S5, A to G), with slightly weaker expression for the 296-bp element. Hence, key neural crest regulatory elements are contained in this 296-bp of DNA, although additional contributory binding sites may also be functional in the context of the larger *crestin* element.

Database searches identified multiple predicted transcription factor binding sites for important neural crest developmental regulators within the 296-bp segment, including two *sox10*, one *pax3*,

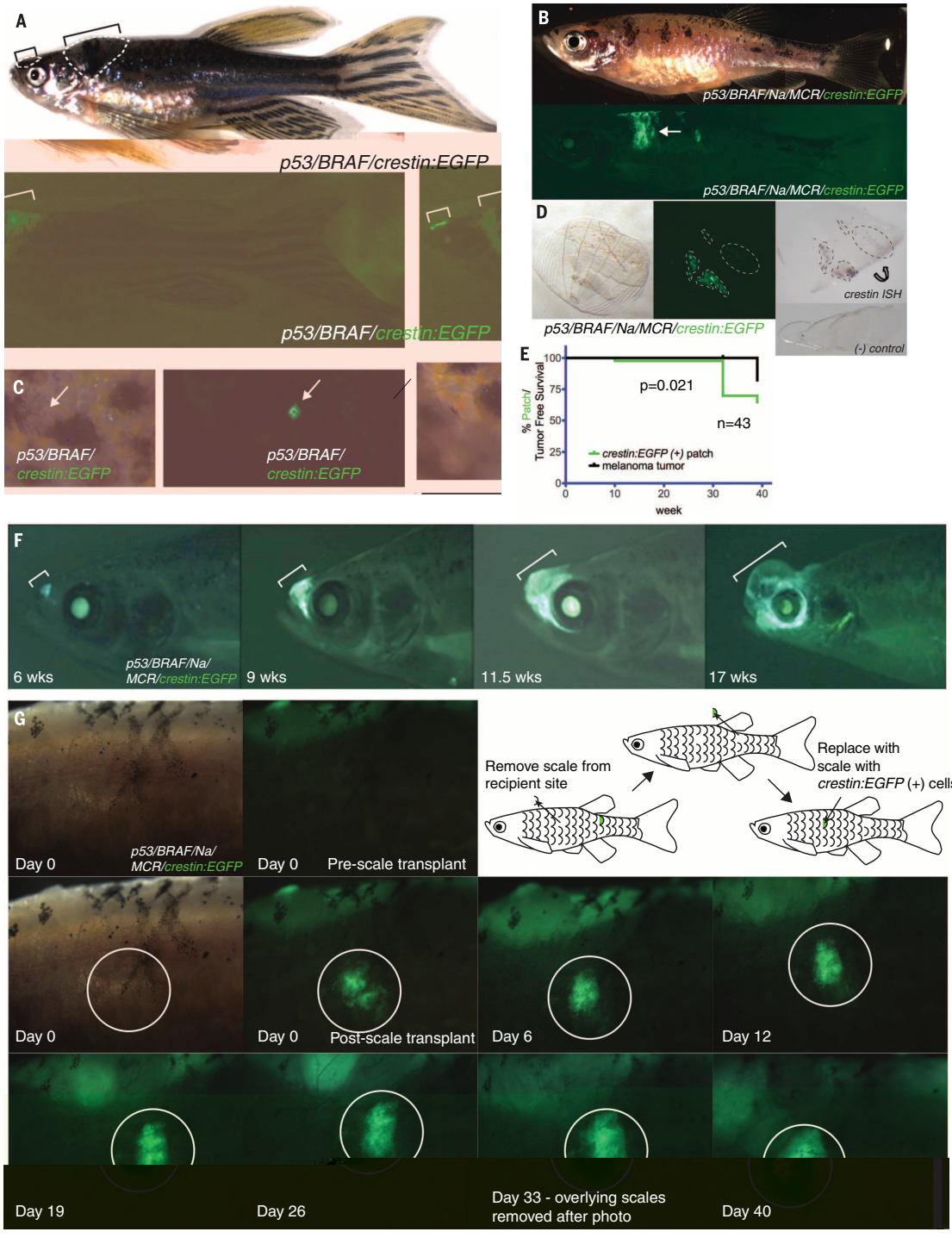


Fig. 2. *Tg(crestin:EGFP)* specifically marks melanoma tumors and precursor lesions. (**A**) Spontaneously arising tumors (outlined) in *p53/BRAF/crestin:EGFP* zebrafish express *EGFP* (brackets), whereas the remainder of the animal is negative. (**B**) *crestin:EGFP* expression is also visible in precursor, nonraised lesions. (**C**) Example of a single *crestin:EGFP*$^+$ cell in *p53/BRAF* background. (**D**) Scales expressing *crestin:EGFP* from precursor, nonraised regions [(B), bottom, arrow] were plucked, photographed [(D), left and middle], and subjected to ISH for *crestin* transcript [(D), right]. There is a concordance of *EGFP* (green) and *crestin* transcript (purple, dotted outlines, scales curl during ISH procedure, indicated by the curved arrow, observed in 5 of 5 scales). (Bottom right) *crestin:EGFP*$^-$ scales are negative for *crestin* ISH staining (observed in 7 of 7 tested scales). (**E**) Cohorts of *p53/BRAF/crestin:EGFP* zebrafish were tracked over time for the appearance of *crestin:EGFP*$^+$ patches and tumors, with *crestin:EGFP*$^+$ cells/patches (green line) identifiable before raised melanoma tumors (black line). (**F**) Example of an *EGFP*$^+$ preclinical patch tracked over time (6, 9, 11.5, and 17 weeks) as it expands into a clinically apparent melanoma tumor. (**G**) Scale autotransplant and expansion of *crestin:EGFP*$^+$ patch of cells. At day 0, the recipient site is free of *crestin:EGFP*$^+$ cells (pre–scale transplant), but immediately after transplant of a single scale (post–scale transplant), the patch of *EGFP*$^+$ cells is apparent (white circle). This patch expands outward, and even upon removal of the original transplanted scale after the day 33 photograph, *EGFP*$^+$ cells remain in place and continue to expand. The magnification and size of white circle is the same in each image.

one *E-box* (*myc* or *mitf*-binding site), and one *tfap2* site (Fig. 3A and table S1) (*15*). To determine which sites are functionally required for *crestin* transgene expression, we individually mutated the core consensus for each site (Fig. 3A and fig. S6B) and tested expression at 24 hours after fertilization in F0 embryos injected with the different *crestin_296bp:EGFP* constructs at the one-cell stage, examining >80 successfully injected F0 embryos per construct. Whereas mutation of the predicted *pax3* site left the expression pattern largely unchanged from that of the wild type (~55% of embryos with neural crest predominant expression in both), mutation of either *sox10* site drastically reduced neural crest expression (≤20%), as did mutation of the *tfap2* site or the *E-box* site (<10% and none, respectively) (Fig. 3A and fig. S6, C and D). These functional transcription factor binding sites provide an explanation for the neural crest specificity of *crestin* transgene expression, which integrates regulatory signals of multiple neural crest transcription factors, including *sox10*.

Neural crest signature in melanoma initiation

To test whether early precursor melanoma lesions express other melanoma and neural crest progenitor markers in addition to *crestin*, we isolated individual scales from *p53/BRAF/Na/MiniCoopR/crestin:EGFP* zebrafish with early *crestin:EGFP*$^+$ patches and compared them with adjacent individual scales without *crestin:EGFP* expression (Fig. 3, B and C) and performed Affymetrix microarrays (*12*). Such scales appear well-matched in their cell make-up, particularly in regards to melanocytes, as shown through *mitf:mCh* co-expression (marking melanocytes), and *crestin:EGFP* (fig. S7). *Crestin:EGFP*$^+$ scale-enriched genes include neural crest– (such as *crestin*, *mitf*, and *dlx2a)* and melanoma-expressed [such as *mia* and *mt* (metallothionein)] genes (table S2) (*16*–*18*). We confirmed enrichment by means of quantitative reverse transcription PCR (RT-PCR) on independent *crestin:EGFP*$^+$ and *crestin:EGFP*$^-$ scales (including *crestin*, *dlx2a*, and *mia*) (Fig. 3D). We also found *sox10* expression enriched in the *crestin:EGFP*$^+$ samples (Fig. 3D). *Sox10* is a known marker and key regulator of neural crest identity (Fig. 1H) (*19*) and, from our data, of *crestin* expression (Fig. 3A); it is also sufficient to direct reprogramming of human fibroblasts to induced neural crest cells (*20*) and can be highly express in melanoma, where it is involved in growth control (*21*–*23*). We used gene set enrichment analysis (GSEA) to query a rank-ordered list of the *crestin:EGFP*$^+$ scale enriched genes for an association with all neural crest–expressed genes in the Zebrafish Information Network (ZFIN) database, and we found a significant correlation [false discovery rate (FDR) Q = 0.019, and familywise error rate (FWER) P = 0.019] (Fig. 3E, left, and table S6). Similarly, we used GSEA to compare a rank-ordered list of genes enriched in embryonic stem (ES)–derived human neural crest cells (*24*) with genes enriched ≥2-fold in *crestin:EGFP*$^+$ scales, and we found a positive correlation, detectable even across species (FDR Q = 0.089, FWER P = 0.089) (Fig. 3E, right, and table S6). These data collectively support the concept that key aspects of NCP state reemerge at the time of melanoma initiation, as read out by the *crestin:EGFP* reporter.

Neural crest progenitor identity and melanomagenesis

On the basis of our analysis of *crestin* expression, which provides an in vivo readout of NCP identity at the time of melanoma initiation, we reasoned that favoring entry into or inappropriately maintaining the NCP state in a cancerized field of melanocytes would accelerate the onset of melanoma formation (Fig. 3F). The neural crest master transcription factor *sox10* has been shown to increase *crestin* mRNA in embryos when overexpressed (*25*), and we found similar results for our *crestin:EGFP* reporter (fig. S8). Misexpression of *SOX10* in postnatal fibroblasts also generates multipotent neural crest cells in culture (*20*). We therefore overexpressed *sox10* in melanocytes using the transgenic *MiniCoopR* system (*12*) and found that *sox10* overexpression in melanocytes accelerated melanoma onset significantly versus controls (Fig. 3G). To examine the consequence of *sox10* inactivation, we used a melanocyte-specific CRISPR/Cas9 system to target *sox10* in the *p53/BRAF/Na* background (fig. S9A). As compared with controls in which *p53* is redundantly targeted (already mutated in our system) using an analogous vector, we found a significant slowing of median tumor onset in the *sox10* CRISPR/Cas9 setting [133 days (*p53*) versus 180 days (*sox10*), P < 0.0001] (fig. S9B). When melanomas developed in the *sox10*-targeted background, the *sox10* target genomic locus exhibited a propensity for mutations that preserve predicted *sox10* function (for example, point mutations or in-frame deletions, ~60% of sequenced genomes) as opposed to inactivating mutations (for example, frame-shifts, ~40% of sequenced genomes), suggesting a selective pressure for retention of *sox10* function (fig. S9, C and D). These gain- and loss-of-function results together strongly support our hypothesis that reemergence of NCP state is an important event in melanoma tumor initiation.

Neural crest super-enhancers and melanoma

In order to understand how the expression of neural crest genes such as *sox10* may be regulated in zebrafish and human melanoma, we used a combination of chromatin immunoprecipitation sequencing (ChIP-seq) and assay for transposase-accessible chromatin using sequencing (ATAC-seq). Chromatin regions with high levels of H3K27Ac histone marks have been referred to as super-enhancers (SEs), or stretch-enhancers (*26*, *27*), and have been identified as key transcriptional regulatory elements that modulate cell type–specific and cancer-related gene expression (*26*–*29*). We used ChIP-seq to identify H3K27Ac-enriched regions in a zebrafish *crestin:EGFP*$^+$ melanoma cell line (zcrest1) that we derived from *p53/BRAF/Na/MiniCoopR/crestin:EGFP* zebrafish and noted substantial regions of H3K27Ac enrichment at *crestin* loci, identified as SEs (representative locus shown in Fig. 4A, red bar indicates SE). We also identified Sox10 binding by means of ChIP-seq across the *crestin* locus (Fig. 4A, bottom track), which is consistent with our promoter analysis linking *sox10* to *crestin* transcriptional regulation (Fig. 3A). We examined the *sox10* locus in the zcrest 1 zebrafish melanoma cell line and also identified H3K27Ac SE marks (Fig. 4B). These SEs were similarly found at *sox10* and *crestin* via H3K27Ac ChIP-seq performed on a freshly isolated *primary* zebrafish melanoma tumor (fig. S10, A and B, red bars), supporting our findings on the cell lines as being representative of the in vivo landscape. ATAC-seq identified open and accessible chromatin corresponding to the SEs at *crestin* and *sox10* (Fig. 4, A and B) and other SE-associated loci in two zebrafish melanoma cell lines (fig. S10, E and F) (*30*). These data suggest a molecular basis for the epigenetic state readout by *crestin* of NCP identity in initiating melanoma cells.

To compare our fish studies with human melanoma, for which SE analysis is limited, we examined the Cancer Cell Line Encyclopedia (CCLE) database (*31*) and found that most human melanoma lines (51 of 60) express *SOX10* according to Affymetrix microarray data (fig. S11A). As with zebrafish melanomas, ChIP-seq showed enriched H3K27Ac marks near the *SOX10* locus in six *SOX10*-expressing human melanoma lines tested but not in a rare *SOX10*-negative human melanoma cell line (LOXIMVI) (Fig. 4, C and E, and fig. S11A). The *SOX10* SE's were ranked 3 and 6 out of 842 SEs from ~15,000 total enhancers in the A375 line (fig. S11B and table S3). Clustering based on the SE landscape yielded two distinct groups of *SOX10*-expressing lines that correlated with the presence or absence of expression of melanocyte differentiation markers, *TYR* and *DCT* (fig. S11, C and D). H3K4me1, a histone modification typically at active enhancers, was also enriched at *SOX10* in the representative A375 melanoma line (Fig. 4C). Remarkably, these *SOX10* SE peaks were also found in published H3K27Ac data from human ES-derived neural crest cells (hNCCs) (Fig. 4, C and E) (*32*). Examining multiple normal and cancer cell types (66 and 18 types, respectively), the enrichment of H3K27Ac signal at *SOX10* was evident and specific to melanoma cells, hNCCs, and brain tissue, which contains *SOX10*-expressing oligodendrocytes (Fig. 4E) (*26*). Beyond *SOX10*, a similar SE epigenetic signature was shared for the neural crest transcription factor *DLX2* among melanomas across species and was enriched in melanomas and hNCCs (Fig. 4, D and E, and fig. S10, C and E). *DLX2* expression is enriched in sorted *crestin*$^+$ embryonic neural crest cells (table S4), in *crestin*$^+$ precursor melanoma patches (Fig. 3D), and in the less differentiated, *TYR/DCT*$^-$ melanomas relative to cultured normal human melanocytes (fig. S11, C, D, and F; and table S5). SEs were

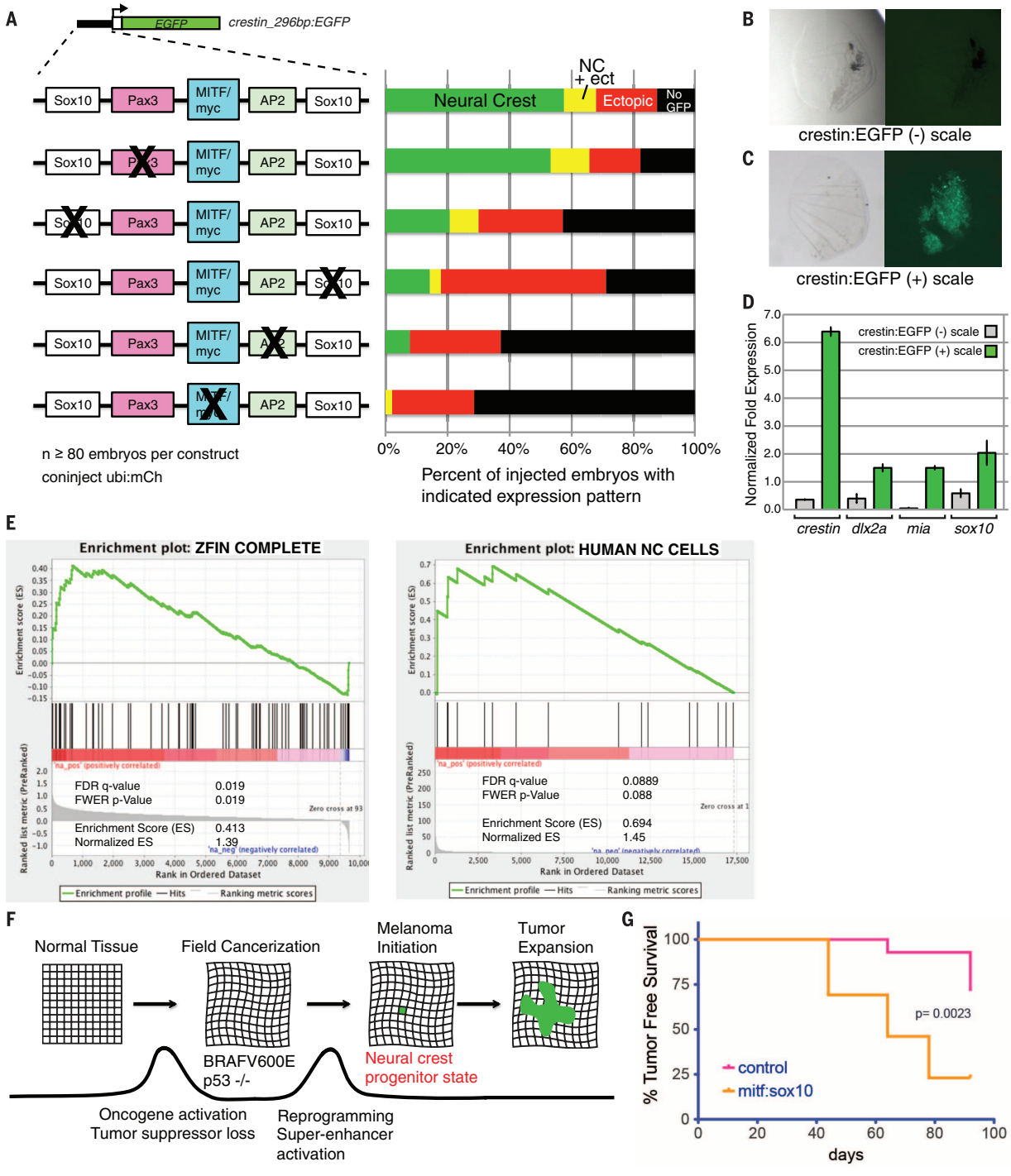


Fig. 3. Reemergence of neural crest progenitor identity in melanoma initiation. (**A**) Mutation of key neural crest transcription factor binding sites in the 296-bp *crestin* element, including *sox10*, *tfap2*, and an E-box for *myc* or *mitf*, substantially reduces neural crest *EGFP* expression at 24 hours after fertilization, whereas mutation of the predicted *pax3* site does not alter expression. Coinjection of a ubiquitous *ubi:mCh* transgene confirmed successful injection for the >80 independently injected F0 embryos analyzed for each construct. Scales from *p53/BRAF/Na/MiniCoopR/crestin:EGFP* adult zebrafish (**B**) with and (**C**) without $EGFP^+$ cells were collected, and total RNA was isolated for microarray analysis. (**D**) Quantitative RT-PCR of $crestin:EGFP^+$ versus $crestin:EGFP^-$ scales reveals enrichment of neural crest (*crestin*, *dlx2a*, *sox10*) and melanoma marker expression (*crestin*, *mia*, *sox10*). (**E**) GSEA analysis shows a positive association between $crestin:EGFP^+$ patch-enriched genes and neural crest–expressed genes in zebrafish (left) and in human ES–derived neural crest cells (right). (**F**) Model for the importance of reemergence of NCP state through SE activation as an essential step in melanoma initiation. The acquisition of genetic lesions in normal tissue leads to oncogene activation ($BRAF^{V600E}$) and tumor suppressor loss ($p53^{-/-}$) and represents an initial barrier that generates a cancerized field from which rare clones (green) overcome the additional barrier of achieving a NCP state to initiate melanoma formation and then tumor expansion. Favoring reemergence of the neural crest progenitor state would then increase melanoma formation, and strengthening this barrier to inhibit adoption of the $crestin^+$ NCP state would block melanoma initiation. (**G**) Misexpression of the NCP transcription factor *sox10* accelerates melanoma onset as compared with controls in *p53/BRAF/Na* zebrafish rescued with the *MiniCoopR* construct.

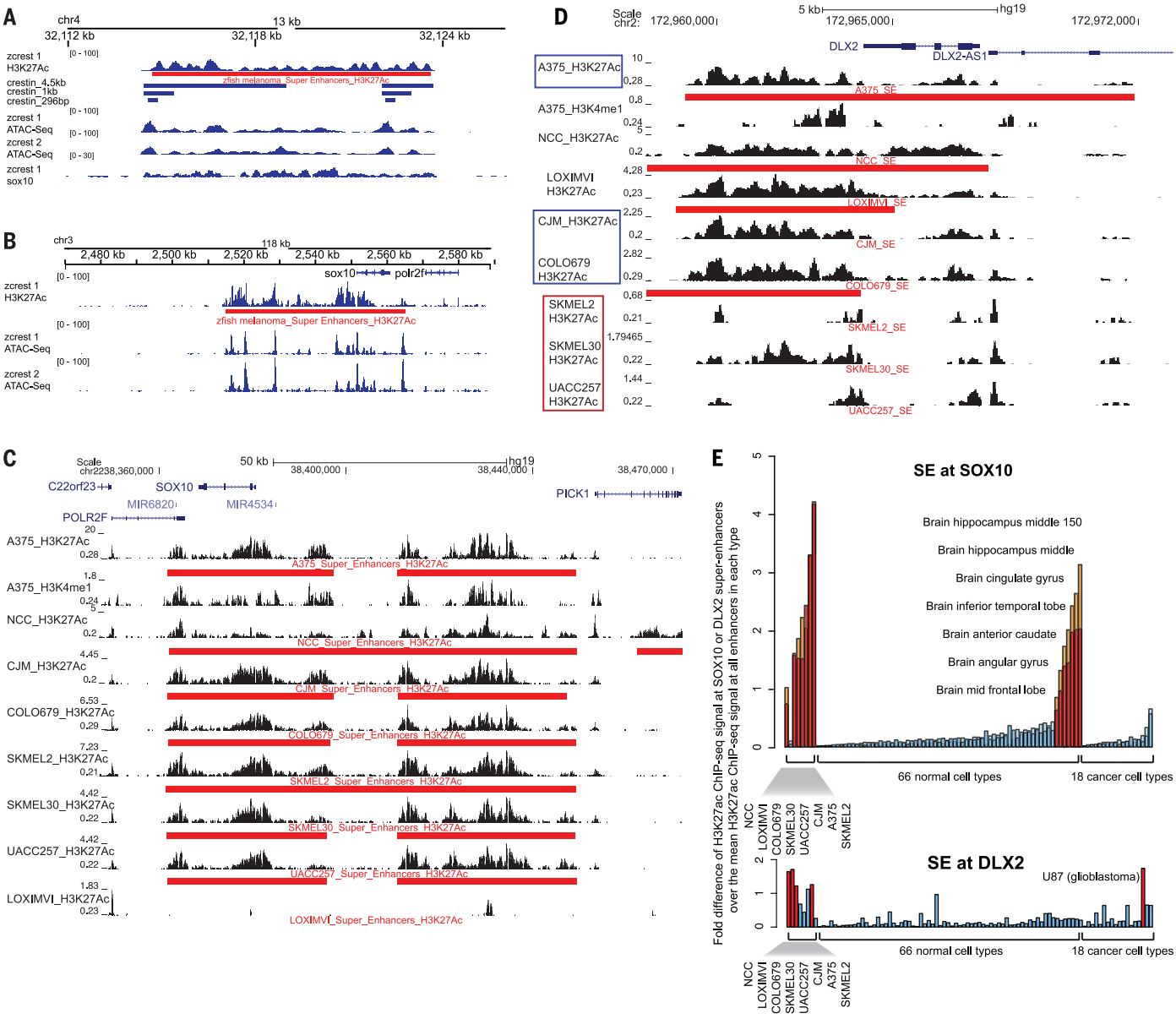


Fig. 4. A SE signature in zebrafish and human melanoma. (**A**) ChIP-seq for the H3K27Ac histone mark (top row) in a *crestin:EGFP*$^{+}$ zebrafish melanoma cell line (zcrest 1) reveals enriched peaks, identified as a SE (red bar), at a representative *crestin* locus. Sequences of *crestin_4.5kb*, *crestin_1kb*, and *crestin_296bp* shown with blue horizontal bars. ATAC-seq on two zebrafish melanoma lines (zcrest 1 and zcrest 2) identifies open chromatin coincident with the H3K27Ac marks at *crestin* loci. ChIP-seq for Sox10 shows enrichment across the *crestin* locus (bottom row) in zcrest 1 cells. (**B**) ChIP-seq for the H3K27Ac histone mark (top row) on the zcrest 1 line identifies robust enrichment and a SE at *sox10* (red bar). ATAC-seq identifies corresponding regions of open chromatin (rows 2 and 3). (**C**) ChIP-seq for the H3K27Ac mark on multiple *SOX10*-expressing melanoma lines (A375, CJM, COLO679, SKMEL2, SKMEL30, and UACC257) and a rare *SOX10*-negative melanoma line (LOXIMVI). Robust peaks corresponding to SEs (red bars) are identified in all lines, except the *SOX10*-negative LOXIMVI line. Published H3K27Ac ChIP-seq data from hNCCs reveal a similar SE pattern. ChIP-seq for H3K4Me1, an enhancer mark, on a representative melanoma line, A375 (row 2), identifies regions corresponding to the H3K27Ac marks. (**D**) H3K27Ac signal is robust at the *DLX2* locus in melanoma cell lines not expressing the melanocyte differentiation genes *TYR* and *DCT* (blue box), in hNCCs, and in the *SOX10*-negative LOXIMVI melanoma line. Human genomic track images were generated at http://genome.ucsc.edu. (**E**) High relative H3K27Ac signal at *SOX10* (top) and *DLX2* (bottom) identifies SEs (presence = red/orange bar, absence = blue bars) and is largely enriched in melanomas and hNCCs compared with 66 normal and 18 cancer cell types.

also found at *TFAP2* family members, whose functional binding site was identified in the 296-bp *crestin* element (Fig. 3A), in A375 human melanoma (*TFAP2C*) (fig. S12A), and in the zebrafish melanoma cell line and primary melanoma (*tfap2a*) (fig. S10, D and F). Such SEs are also found near other *crestin*$^{+}$ cell-enriched genes *MIA* and *MT2A* (metallothionein genes) in A375 melanoma cells (fig. S12, B and C).

Discussion

This work shows that melanoma precursor cells reinitiate an embryonic neural crest signature and activate a melanoma gene program. Although the expression of stem cell factors or embryonic genes has been noted previously in advanced malignancies including melanoma (*8*, *23*), it remained uncertain whether this was simply due to aberrant misexpression of these genes or whether these genes were present from tumor outset (Fig. 3F). Our data support a model in which the stem/progenitor

cell gene programs are an integral part of cancer initiation and not reacquired later (Fig. 3F). Analogously, this theme was also recently described in basal cell carcinoma (BCC), a cancer arising from a different cell lineage, in which careful analysis of gene expression from tissue collected during early BCC formation in a mouse model of the disease exhibited an embryonic hair follicle progenitor signature, indicating that this may be a general feature of cancer initiation (*33*). There also is a conceptual similarity between the reprogramming-induced pluripotent stem process and the tumorigenic reemergence of NCP state seen in our system as melanocytes with the required, but insufficient, *p53/BRAF* genetic mutations stochastically reenter a NCP state (as read out by *crestin* expression) early at the initiation of melanoma. As the initiating stages of more cancers are analyzed, stem cell/progenitor phenotype reacquisition may be a generally observed phenomenon in most cancers (*33*).

Our work establishes a method to live-image cancer development when a tumor starts, potentially as a single cell, providing a singular view of the initiating events. NCP identity arises, and is likely a necessary step, early in tumor development. In our *p53/BRAF/crestin:EGFP* model, all $EGFP^{+}$ patches of *crestin:EGFP* cells that we track go on to enlarge and form tumors; given the rarity of capturing single cells in this background, we cannot rule out that clones may rarely fail to progress to patches. The *crestin* gene has the distinct characteristic of being expressed in neural crest cells in embryos, but not being expressed in adult tissues except when melanoma arises. Super-enhancers shared between melanoma and neural crest are specifically activated. The neural crest signature encompasses a combinatorial code including *SOX10*, *DLX2*, and the *TFAP2* family. It is unlikely that any one of these genes is sufficient for the reprogramming event, but it is the combination of multiple transcription factors that participate. In contrast to *SOX10*, *DLX2* expression is characteristic of genes that are expressed by neural crest and melanoma, but not expressed in melanocytes. As such, this gene set may have diagnostic importance in the initiating cancer cell. A reevaluation of published expression data for nevi, primary melanoma, and metastatic melanoma does reveal increased *SOX10* expression in malignant melanomas (*34*). Nevi show a wide range and consistently lower amount of *SOX10* expression than that of melanoma, raising the possibility that the histologically defined category of nevus is capturing a range of melanoma-initiating capacity (those nevi expressing higher *SOX10* levels may have initiated or may be more prone to initiating melanoma).

Several major questions remain about an initiating cancer cell that may now be more accessible with our live visualization tool. The niche environment must participate in the process of initiation, perhaps through the activation of neural crest signaling pathways akin to the development of the normal neural crest or by stress pathways related to irradiation or oxidative damage. Further genetic mutations, as opposed to isolated epigenetic changes, may be required for tumor initiation from the cancerized field; however, the absence of identifiable functionally relevant exomic mutations in a study of 53 zebrafish melanoma tumors to which we contributed would tend to favor that the key genetic drivers ($BRAF^{V600E}$ and mutant *p53*) are already present in our model (*35*). Work on this question will provide information on how cancer initiates. The reprogramming event appears to occur in one melanocyte or progenitor in a cancerized field, and defining why the process initiates in that single cell rather than an adjacent cell will provide an understanding of protective mechanisms in cancer formation.

Materials and methods

Cloning of crestin *promoter/enhancer*

The transcript for *crestin*, originally described as AF195881, was used with Basic Local Alignment Search Tool (BLAST) on the zebrafish genome and, as has been previously noted, identified many partial or complete highly similar (>90% identical) sequences spread throughout the genome (*6*, *7*). Multiple insertions are present on chromosome 4, and a ~4.5-kb sequence located upstream of the predicted *crestin* open reading frame (ORF) was noted to be present in multiple instances. Reasoning that this segment may contain the relevant regulatory elements of *crestin*, primers were designed to amplify via PCR upstream sequences of the *crestin* locus (LOC796814) on chromosome 4 in the TU background (primers 299 and 302; primers supplemental table). The 1-kb fragment and 296-bp fragment were isolated by using PCR primers 302/340 and 517/516, respectively. Fragments were cloned into pENTR5′ (Life Technologies, Grand Island, NY) per manufacturers instructions. Expression vectors were derived by using Multisite Gateway technology per manufacturer's specifications and the Tol2 Kit (*36*) [*EGFP*, vector 383; *mCh*, vector 386; SV40 poly A, vector 302; *EGFP* with mouse minimal ß-globin promoter (*37*); destination vectors, vector 394 alone and modified with addition of *crystallin:YFP* marker; *zmitfa* middle entry vector, gift of Craig Ceol; MiniCoopR from (*15*)].

Production of transgenic zebrafish and lineage tracing

One-cell-stage embryos from the AB strain grown under standard, Institutional Animal Care and Use Committee (IACUC)–approved conditions were injected with the given DNA construct at 25 ng/μl with Tol2 mRNA at 20 ng/μl (*36*). Embryos were screened at 24 hours after fertilization for neural crest expression of *EGFP* or *mCh*, or in the case of *Tg(crestin:CreERt2;crystallin:YFP)*, screened for yellow fluorescent protein (YFP)–positive lenses at 4 days after fertilization. These were grown to adulthood and outcrossed so as to identify founders that gave germline transmission. For each DNA construct, ≥2 independent lines were generated to confirm the expression pattern. For lineage tracing of *crestin*-expressing cells, a stable transgenic line of *Tg(crestin:creERt2;crystalline:YFP)* was crossed to the *Tg(-3.5ubi:loxP-GFP-loxP-mCherry)* line (*ubi:Switch*) (*9*); embryos were collected and treated with 10 μM 4-hydroxytamoxifen (4-OHT) at 50% epiboly and 24 hours after fertilization while grown at 28°C in E3 medium, as per standard protocol.

Embryo and adult imaging

Transmitted light and fluorescence images of adult and nonconfocal images of embryos were collected on a Zeiss Discovery V.8 Stereoscope with an Axiocam HRc. Static confocal images were collected on a Nikon C2si Laser Scanning Confocal using 25× objective on embryos mounted in 1% low melt agarose. Maximum intensity projections of Z stacks or three-dimensional reconstructions are presented here. Movies of developing embryos were collected on a Nikon Eclipse Ti Spinning Disk Confocal with a 10× objective, with tiled images collected every 6 to 7 min. Images were processed by using Photoshop, ImageJ, or Imaris. Multiple tiled images of adult zebrafish were stitched together by using the automated Photomerge function in Photoshop or manually aligned.

Melanoma model and MiniCoopR *system*

Experiments were performed as outlined in (*15*). Briefly, *p53/BRAF/Na* embryos were injected with equal amounts of *MiniCoopR* alone or *MiniCoopR; mitf:sox10* and selected for melanocyte rescue at 48 hours. Equal numbers of melanocyte-rescued embryos were grown to adulthood (n = 14 for control and n = 13 for mitf:sox10) and scored for the emergence of raised melanoma lesions as per (*15*). Survival curves and statistics were generated in Prism.

Quantitative PCR, microarray, and in situ hybridization

Adult *p53/BRAF/Na/MiniCoopR/crestin:EGFP* fish were anesthetized with Tricaine, viewed under the fluorescent dissecting scope, and precursor patches of *crestin:*$EGFP^{+}$ cells were identified and associated single scales removed. These fish contained stable (germ-line transmitted) alleles of *crestin:EGFP* and control *MiniCoopR* (no test gene). Neighboring *crestin:EGFP*–negative scales from the same zebrafish were selected for controls. Scales were immediately placed in Trizol, and total RNA was purified following the manufacturer's instructions, with the additional use of Sigma GenElute LPA carrier. RNA was analyzed on a BioAnalyzer, and high-quality samples were chosen for microarray libraries generated by using the Ovation Pico WTA System V2 and Encore Biotin labeling system for hybridization on the Zebrafish Gene 1.0 ST Affymetrix Array or used for cDNA preparation using SuperScript III. RMA-normalized Affymetrix Array results were sorted for maximum fold (log2) increase in gene expression in GFP-positive samples versus negative samples. Quantitative PCR reactions were run on triplicate biological samples with triplicate technical replicates and normalized to β-actin expression, with representative results presented.

For in situ hybridizations, we used the methods of (*38*) and the *crestin* probe from (*7*).

ChIP-seq and SE analysis

Human melanoma cells were grown to confluence in Dulbecco's Modified Eagle medium (DMEM) + 10% fetal calf serum (FCS), and ~ 1×10^8 cells were formaldehyde cross-linked and collected. ChIP was performed by using the methods of (*30*, *39*) with antibodies against H3K27Ac (ab4729) (Abcam), H3K4Me1 (ab8895) (Abcam), and SOX10 [Santa Cruz Biotechnology (Dallas, TX), sc-17342x]. Libraries were prepared by using the NEBNext Multiplex Oligos for Illumina kit (NEB) and run on an Illumina HiSeq 2000. Data analysis, including enhancer and SE calling, was performed as described in (*26*). Genomic track images were generated by using the IGV package (*40*) and the University of California, Santa Cruz Browser (*41*).

All human ChIP-seq data sets were aligned to build version NCBI37/HG19 of the human genome using Bowtie (version 0.12.9) (*42*) with the following parameters: –n2, –e70, –m2, –k2, –best. We used the MACS version 1.4.1 (Model based analysis of ChIP-seq) (*43*) peak finding algorithm to identify regions of ChIP-seq enrichment over background. A P value threshold of enrichment of 1×10^{-9} was used for all data sets. Wiggle files for gene tracks were created by using MACS with options –w –S –space = 50 to count reads in 50-bp bins. They were normalized to the total number (in millions) of mapped reads producing the final tracks in units of reads per million mapped reads per base pair (rpm/bp).

Identifying SEs

The identification of SEs has previously been described in detail (*26*). Briefly, H3K27Ac peaks were used to identify constituent enhancers. These were stitched if within 12.5 kb, and peaks fully contained within ±2kb from a TSS were excluded from stitching. H3K27Ac signal (less input control) was used to rank enhancers by their enrichment. Super-enhancers were assigned to active genes by using the ROSE software package (younglab.wi.mit.edu/super_enhancer_code.html).

ATAC-seq

Zebrafish melanoma cell lines were grown to 80% confluence, trypsinized, and counted, and 50,000 cells were lysed and subjected to "tagmentation" reaction and library construction as described in (*30*). Libraries were run on an Illumina HiSeq 2000. All zebrafish ATAC-seq data sets were aligned to build version Zv9 of the zebrafish genome by using Bowtie2 (version 2.2.1) (*42*) with the following parameters: –end-to-end, –N0, –L20. We used the MACS2 version 2.1.0 (*43*) peak-finding algorithm to identify regions of ATAC-seq peaks, with the following parameter–nomodel–shift –100–extsize 200. A Q value threshold of enrichment of 0.05 was used for all data sets.

Scale transplants

Adult *p53/BRAF/Na/MiniCoopR/crestin:EGFP* were anesthetized, viewed under the fluorescent dissecting scope; precursor patches of *crestin:EGFP*$^+$ cells were identified (not from raised melanoma lesions); and single associated scales were removed and placed in 50-µl drops of E3 buffer on a petri dish lid. Anesthetized recipient zebrafish (in the case of allotransplants) or the same zebrafish (in the case of autotransplants) were gently placed on a wet sponge and a recipient site selected (free of *crestin:EGFP*–expressing cells). One scale was removed at the donor site, and the previously selected donor scale from the drop of E3 was placed on the zebrafish and slid posterior to anterior into place by using surrounding scales to hold it in place. Recipients were quickly placed in fresh zebrafish water and monitored for recovery from anesthesia. Transplants were monitored frequently with some loss of transplanted scales (~20%) occurring quickly within 1 day because of simple dislodgement. Once in place for ~4 days, scales were firmly incorporated and could be monitored over time and photographed at the same magnification under the dissecting scope after mild Tricaine anesthesia of the recipient zebrafish. In single-scale autotransplants in a representative cohort of 10 fish, three scales were lost in the first week (we consider a technical failure with the scales falling out), five showed expansion of the *crestin:EGFP* cells, and three showed no change or loss of *crestin:EGFP* cells. In single-scale allotransplants onto sublethally irradiated *casper* recipient fish, eight *EGFP*$^+$ scales were transplanted onto different recipients with one scale lost in the first week, five scales showing expansion of the *crestin:EGFP* cells for ≥2.5 weeks, and two showing stable appearance for ≥ 2 weeks. Six of six *EGFP*$^-$ scales showed loss of pigmented cells in this time frame.

Promoter analysis

Transcription factor binding sites were predicted by using JASPAR (*14*). We used the Q5 Site-Directed Mutagenesis Kit from NEB to introduce mutations to destroy the chosen transcription factor binding site (primers supplemental table). To analyze expression in vivo, equal volumes of a mixture of the *crestin_296bp:EGFP* construct variant at 20 ng/µl mixed with 5 ng/µl of *ubi:mCh* (*9*, *36*) (for an injection control) were injected with 20 ng/µl of Tol2 mRNA into single-cell AB embryos. At 24 hours after fertilization, embryos were fixed in paraformaldehyde (PFA), washed and stored in phosphate-buffered saline (PBS), and scored for *mCh* expression to identify successfully injected embryos and for *EGFP* to bin based on the predominant expression pattern. More than 80 transgenic F0 embryos for each construct were scored.

Investigating SOX10 and DLX2 SEs in healthy normal cells and melanoma cells

To investigate whether the two SOX10-asscociated SEs identified in A375 cells are also present in other cell types, they were compared with the SEs identified in other melanoma cell types, neural crest cells, and 84 additional cell types from normal or cancer cells described in (*26*).

First, the SOX10-asscociated SEs were operationally called "present" in a cell type if the SEs identified in the cell-type overlap with the SOX10 SEs in A375 cells by at least 1 bp. Second, the H3K27Ac signal density of the SOX10-asscociated SEs identified in A375 cells was compared with the mean H3K27Ac signal density of all enhancer clusters in each cell type. For each cell type, the average H3K27Ac ChIP-seq read density was calculated in rpm/bp for the two SOX10 SEs identified in A375 cells as well as all enhancers clusters identified by using the ROSE software package in the cell type (younglab.wi.mit.edu/super_enhancer_code.html). The fold difference of H3K27Ac ChIP-seq signal at SOX10 SEs over the mean H3K27Ac ChIP-seq signal at all enhancers in each cell type was plotted in Fig. 4E. The same process was undertaken for the SEs at the *DLX2* locus.

Pairwise comparison of SEs between different melanoma cell lines and neural crest cells

The set of SE regions in each cell type (SKMEL2, SKMEL30, UACC257, LOXIMVI, A375, CJM, COLO679, and neural crest cells) were merged together if overlapping by 1 bp, resulting in a total of 3407 merged SE regions. The neural crest cell data were previously published in (*32*). The average H3K27Ac ChIP-seq read density was calculated in rpm/bp for each of the merged regions. The pair-wise comparisons by Pearson correlation were performed on all data sets by using the average read density at the merged regions. The average linkage hierarchical clustering of the Pearson correlation was shown in the heatmap (fig. S11C).

Zebrafish melanoma cell lines and in vitro scale imaging

A single *EGFP*$^+$ melanoma tumor arising in a *p53/BRAF/Na* fish injected with *MiniCoopR* plasmid and *crestin:EGFP* plasmid (zcrest 1 line) or a *p53/BRAF/Na/crestin:EGFP* fish injected with MiniCoopR plasmid (zcrest 2 line) was removed after killing the adult fish. Briefly, tumors were dissociated with a razor blade and trypsin, filtered, and plated on a fibronectin-coated well and grown in rich media supplemented with FBS and zebrafish embryo extract as described (*44*). After several passages, the zcrest 1 line was sorted for *EGFP*$^+$ cells, which were continued as the line. For the zcrest 2 line, after several passages, most if not all cells remaining were *EGFP*$^+$. Both lines continue to be *EGFP*$^+$ and have been grown for >50 passages on plastic in standard DMEM + 10% FBS with 1X GlutaMAX supplement and penicillin/streptomycin antibiotics.

Scales from *p53/BRAF/Na/crestin:EGFP/MiniCoopR* and *p53/BRAF/crestin:EGFP* fish with and without *EGFP*$^+$ patches of cells were placed in zebrafish melanoma growth medium in fibronectin coated wells in 384-well format, flat-bottomed plates and imaged daily on a Yokogawa

CV7000 confocal imager with brightfield and z-stack image projections collected.

GSEA analysis

Using our microarrays comparing *crestin:EGFP*$^{+}$ versus *crestin:EGFP*$^{-}$ scale gene expression, zebrafish genes were rank-ordered (10,705 genes) from high to low for enrichment (crestin scales rank.rnk). We generated a gene list of all neural-crest–expressed genes in the ZFIN database (317 genes, zfincomp.gmt) (*45*) and used the preranked GSEA analysis tool. Using a list of genes with ≥2-fold enrichment in *crestin:EGFP*$^{+}$ scales by means of microarray (Crestin Scales.gmt), we queried a list of 17,575 genes rank-ordered for their enrichment in human neural crest cells versus parental ES cells from (*32*) (Rada Ranked.rnk), also using the preranked GSEA analysis tool (*46*).

RNA-seq analysis

Multiple human melanoma cell lines and adult human epidermal melanocytes (purchased from Life Technologies) were grown to near confluence, and total RNA was isolated by using the standard Trizol protocol. Illumina libraries were prepared by using Ribo-Zero Magnetic Gold Kit (epicenter) and NEBNext Ultra RNA Library Prep Kit (NEB) and run on a HiSeq 2500, reads aligned by using Tophat 2.0, and fragments per kilobase of exon per million fragments mapped values determined by using Cufflinks. For sorted *crestin*$^{+}$ cells, transgenic *crestin_1kb:EGFP* adults were mated, and embryos collected and grown to the 15-somite stage. These were homogenized, filtered, and sorted by using fluorescence-activated cell sorting into PBS, collecting ~5500 *EGFP*$^{+}$ cells and 100K *EGFP*$^{-}$ cells. Total RNA was again collected by using Trizol and GenElute LPA carrier per manufacturer instructions. Libraries were prepared by using Ribogone kit (Clontech) and the SMARTer Universal Low RNA Kit (Clontech) and sequenced on the Illumina HiSeq 2500, with post-analysis performed as above with the zebrafish genome.

CRISPR/Cas9 experiment

Cas9 mRNA was produced by means of in vitro transcription from a pCS2 *Cas9* vector (*47*) by using mMESSAGE mMACHINE SP6 kit (Invitrogen). Guide RNAs (gRNAs) were generated by following established methods (*48*). The *Sox10* target sequence was GGCCGCGCGCAGGAAACTGG. Six hundred picograms of *Cas9* mRNA and 25 pg of gRNA were injected into embryos of the AB strain. After microinjection, embryos were raised in E3 medium at 28.5°C. The T7E1 assay was performed as reported (*49*). Briefly, genomic DNA was extracted from 2-day-old embryos by using the hotSHOT method (*50*). A fragment of 434 bp was amplified from genomic DNA by using the following primers: GAAGTCCGACGAGGAAGAT and CTTGACTGAGTAAATAGTGCGT. The PCR amplicons were then purified on a 1% agarose gel. Two hundred nanograms of purified DNA were denatured at 95°C for 5 min and slowly reannealed before digestion with 10 units of T7E1 enzyme (NEB) for 1 hour at 37°C. The digestion product was finally run on a 2.5% agarose gel.

CRISPR/Cas9 tumor-free survival curves

In order to inactivate *sox10* specifically in the melanocytes of our zebrafish melanoma model, the *MiniCoopR* vector was engineered to express Cas9 under the control of the melanocyte-specific *mitfa* promoter and a gRNA efficiently mutating *sox10*, described above, off a *U6* promoter. A gRNA against *p53* was used as a negative control (*51*). The two vectors were injected into one-cell stage, Tg (*mitfa:BRAF*^{V600E}), *p53*$^{-/-}$, *mitfa*$^{-/-}$ embryos, and tumor formation was monitored.

To sequence genomic DNA from tumors, tumor tissue was dissected carefully, digested in buffer with proteinase K (*52*), and after inactivation of proteinase K, PCR was performed by using the primers described above as for the T7E1 reaction. PCR fragments were cloned by using TopoTA cloning per manufacturer instructions, and colony PCR was performed on resulting individual clones and submitted for Sanger sequencing for fig. S9C. For next-generation sequencing in fig. S9D, nested PCR (primer sequences TGAACGGGTACGACTGGACGCT and TGTTGTAGCAGTGCGTTTA, yielding a 238-bp amplicon) was performed on the initially amplified genomic locus so as to bring the amplicon ends closer to the CRISPR target sequence in order to allow for coverage by using a MiSeq-based 150-bp paired-end Illumina run with pooled and barcoded samples at the MGH DNA Core. Both Sanger and compiled next-generation sequences were aligned to the wild-type locus by using Lasergene Seqman in order to identify changes at the CRISPR target sequence. Wild-type *sox10* reads, although potentially from nontargeted *sox10* loci in melanocytes and not necessarily from other tissue types in the sample, were excluded from the calculations of fractions of allele types. If included, these would only increase the fraction of active *sox10* alleles and would further favor our interpretation of the results. When determining the fractions of alleles in fig. S9D, each tumor was weighted equally so as to avoid skewing from more reads from a given tumor.

Injection of sox10 mRNA

The *sox10* cDNA from zebrafish was cloned into pENTR/D-TOPO and transferred into pCSDest (*53*) by using Gateway cloning, all per manufacturer instructions. mRNA was generated by using SP6 mMessage mMachine Kit (Ambion) per manufacturer, and 1 nl mRNA mix was injected into single-cell *p53/BRAF/crestin:EGFP* embryos at multiple concentrations, with 20 pg being the highest tolerated dose without substantial toxicity. Embryos were imaged on a Zeiss Discovery V.8 Stereoscope and scored for EGFP expression.

REFERENCES AND NOTES

1. D. P. Slaughter, H. W. Southwick, W. Smejkal, Field cancerization in oral stratified squamous epithelium; clinical implications of multicentric origin. *Cancer* **6**, 963–968 (1953). doi: 10.1002/1097-0142(195309)6:5<963::AID-CNCR2820060515>3.0.CO;2-Q; pmid: 13094644
2. R. L. Mort, I. J. Jackson, E. E. Patton, The melanocyte lineage in development and disease. *Development* **142**, 620–632 (2015). doi: 10.1242/dev.106567; pmid: 25670789
3. The Cancer Genome Atlas Research Network, Comprehensive molecular profiling of lung adenocarcinoma. *Nature* **511**, 543–550 (2014). pmid: 25079552
4. J. A. Lo, D. E. Fisher, The melanoma revolution: From UV carcinogenesis to a new era in therapeutics. *Science* **346**, 945–949 (2014). doi: 10.1126/science.1253735; pmid: 25414302
5. E. E. Patton *et al*., BRAF mutations are sufficient to promote nevi formation and cooperate with p53 in the genesis of melanoma. *Curr. Biol.* **15**, 249–254 (2005). doi: 10.1016/j.cub.2005.01.031; pmid: 15694309
6. R. Luo, M. An, B. L. Arduini, P. D. Henion, Specific pan-neural crest expression of zebrafish Crestin throughout embryonic development. *Dev. Dyn.* **220**, 169–174 (2001). pmid: 11169850
7. A. L. Rubinstein, D. Lee, R. Luo, P. D. Henion, M. E. Halpern, Genes dependent on zebrafish *cyclops* function identified by AFLP differential gene expression screen. *Genesis* **26**, 86–97 (2000). doi: 10.1002/(SICI)1526-968X(200001)26:1<86::AID-GENE11>3.0.CO;2-Q; pmid: 10660676
8. R. M. White *et al*., DHODH modulates transcriptional elongation in the neural crest and melanoma. *Nature* **471**, 518–522 (2011). doi: 10.1038/nature09882; pmid: 21430780
9. C. Mosimann *et al*., Ubiquitous transgene expression and Cre-based recombination driven by the *ubiquitin* promoter in zebrafish. *Development* **138**, 169–177 (2011). doi: 10.1242/dev.059345; pmid: 21138979
10. Y. Kong *et al*., Neural crest development and craniofacial morphogenesis is coordinated by nitric oxide and histone acetylation. *Chem. Biol.* **21**, 488–501 (2014). pmid: 24684905
11. M. Grzmil *et al*., The *INT6* cancer gene and MEK signaling pathways converge during zebrafish development. *PLOS ONE* **2**, e959 (2007). doi: 10.1371/journal.pone.0000959; pmid: 17895999
12. C. J. Ceol *et al*., The histone methyltransferase SETDB1 is recurrently amplified in melanoma and accelerates its onset. *Nature* **471**, 513–517 (2011). pmid: 21430779
13. W. H. Hildemann, Tissue transplantation immunity in goldfish. *Immunology* **1**, 46–54 (1958). pmid: 13513141
14. A. M. Vogel, T. Gerster, Promoter activity of the zebrafish *bhikhari* retroelement requires an intact activin signaling pathway. *Mech. Dev.* **85**, 133–146 (1999). doi: 10.1016/S0925-4773(99)00104-5; pmid: 10415354
15. A. Sandelin, W. Alkema, P. Engström, W. W. Wasserman, B. Lenhard, JASPAR: An open-access database for eukaryotic transcription factor binding profiles. *Nucleic Acids Res.* **32**, D91–D94 (2004). doi: 10.1093/nar/gkh012; pmid: 14681366
16. M. Uhlén *et al*., A human protein atlas for normal and cancer tissues based on antibody proteomics. *Mol. Cell. Proteomics* **4**, 1920–1932 (2005). doi: 10.1074/mcp.M500279-MCP200; pmid: 16127175
17. A. K. Bosserhoff, Melanoma inhibitory activity (MIA): An important molecule in melanoma development and progression. *Pigment Cell Res.* **18**, 411–416 (2005). pmid: 16280006
18. G. Weinlich *et al*., Metallothionein—Overexpression as a highly significant prognostic factor in melanoma: A prospective study on 1270 patients. *Br. J. Cancer* **94**, 835–841 (2006). doi: 10.1038/sj.bjc.6603028; pmid: 16508630
19. R. N. Kelsh, Sorting out *Sox10* functions in neural crest development. *BioEssays* **28**, 788–798 (2006). doi: 10.1002/bies.20445; pmid: 16927299
20. Y. J. Kim *et al*., Generation of multipotent induced neural crest by direct reprogramming of human postnatal fibroblasts with a single transcription factor. *Cell Stem Cell* **15**, 497–506 (2014). doi: 10.1016/j.stem.2014.07.013; pmid: 25158936
21. J. C. Cronin *et al*., SOX10 ablation arrests cell cycle, induces senescence, and suppresses melanomagenesis. *Cancer Res.* **73**, 5709–5718 (2013). pmid: 23913827
22. S. A. Graf, C. Busch, A. K. Bosserhoff, R. Besch, C. Berking, SOX10 promotes melanoma cell invasion by regulating melanoma inhibitory activity (MIA). *J. Investig. Dermatol.* **134**, 2212–2220 (2014). doi: 10.1038/jid.2014.128
23. O. Shakhova *et al*., Sox10 promotes the formation and maintenance of giant congenital naevi and melanoma. *Nat. Cell Biol.* **14**, 882–890 (2012). pmid: 22772081
24. A. Rada-Iglesias *et al*., Epigenomic annotation of enhancers predicts transcriptional regulators of human neural crest. *Cell Stem Cell* **11**, 633–648 (2012). doi: 10.1016/j.stem.2012.07.006; pmid: 22981823
25. E. Olesnicky, L. Hernandez-Lagunas, K. B. Artinger, *prdm1a* Regulates *sox10* and *islet1* in the development of neural crest

and Rohon-Beard sensory neurons. *Genesis* **48**, 656–666 (2010). doi: 10.1002/dvg.20673; pmid: 20836130
26. D. Hnisz *et al*., Super-enhancers in the control of cell identity and disease. *Cell* **155**, 934–947 (2013). doi: 10.1016/j.cell.2013.09.053; pmid: 24119843
27. S. C. J. Parker *et al*., NISC Comparative Sequencing Program, National Institutes of Health Intramural Sequencing Center Comparative Sequencing Program Authors, NISC Comparative Sequencing Program Authors, Chromatin stretch enhancer states drive cell-specific gene regulation and harbor human disease risk variants. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 17921–17926 (2013). pmid: 24127591
28. W. A. Whyte *et al*., Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319 (2013). doi: 10.1016/j.cell.2013.03.035; pmid: 23582322
29. J. Lovén *et al*., Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* **153**, 320–334 (2013). doi: 10.1016/j.cell.2013.03.036; pmid: 23582323
30. J. D. Buenrostro, P. G. Giresi, L. C. Zaba, H. Y. Chang, W. J. Greenleaf, Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–1218 (2013). doi: 10.1038/nmeth.2688; pmid: 24097267
31. J. Barretina *et al*., The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603–607 (2012). doi: 10.1038/nature11003; pmid: 22460905
32. A. Rada-Iglesias *et al*., A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* **470**, 279–283 (2011). doi: 10.1038/nature09692; pmid: 21160473
33. K. K. Youssef *et al*., Adult interfollicular tumour-initiating cells are reprogrammed into an embryonic hair follicle progenitor-like fate during basal cell carcinoma initiation. *Nat. Cell Biol.* **14**, 1282–1294 (2012). doi: 10.1038/ncb2628; pmid: 23178882
34. L. Rönnstrand, B. Phung, Enhanced SOX10 and KIT expression in cutaneous melanoma. *Med. Oncol.* **30**, 648 (2013). doi: 10.1007/s12032-013-0648-y; pmid: 23801280
35. J. Yen *et al*., The genetic heterogeneity and mutational burden of engineered melanomas in zebrafish models. *Genome Biol.* **14**, R113 (2013). doi: 10.1186/gb-2013-14-10-r113; pmid: 24148783
36. K. M. Kwan *et al*., The Tol2kit: A multisite gateway-based construction kit for *Tol2* transposon transgenesis constructs. *Dev. Dyn.* **236**, 3088–3099 (2007). doi: 10.1002/dvdy.21343; pmid: 17937395
37. O. J. Tamplin, B. J. Cox, J. Rossant, Integrated microarray and ChIP analysis identifies multiple Foxa2 dependent target genes in the notochord. *Dev. Biol.* **360**, 415–425 (2011). doi: 10.1016/j.ydbio.2011.10.002; pmid: 22008794
38. C. Thisse, B. Thisse, High-resolution in situ hybridization to whole-mount zebrafish embryos. *Nat. Protoc.* **3**, 59–69 (2008). doi: 10.1038/nprot.2007.514; pmid: 18193022
39. T. I. Lee, S. E. Johnstone, R. A. Young, Chromatin immunoprecipitation and microarray-based analysis of protein location. *Nat. Protoc.* **1**, 729–748 (2006). doi: 10.1038/nprot.2006.98; pmid: 17406303
40. J. T. Robinson *et al*., Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011). doi: 10.1038/nbt.1754; pmid: 21221095
41. W. J. Kent *et al*., The human genome browser at UCSC. *Genome Res.* **12**, 996–1006 (2002). doi: 10.1101/gr.229102; pmid: 12045153
42. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012). doi: 10.1038/nmeth.1923; pmid: 22388286
43. Y. Zhang *et al*., Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008). doi: 10.1186/gb-2008-9-9-r137; pmid: 18798982
44. S. Heilmann *et al*., A quantitative system for studying metastasis using transparent zebrafish. *Cancer Res.* **75**, 4272–4282 (2015). doi: 10.1158/0008-5472.CAN-14-3319; pmid: 26282170
45. D. G. Howe *et al*., ZFIN, the Zebrafish Model Organism Database: Increased support for mutants and transgenics. *Nucleic Acids Res.* **41** (D1), D854–D860 (2013). doi: 10.1093/nar/gks938; pmid: 23074187
46. A. Subramanian *et al*., Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15545–15550 (2005). doi: 10.1073/pnas.0506580102; pmid: 16199517
47. L.-E. Jao, S. R. Wente, W. Chen, Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 13904–13909 (2013). doi: 10.1073/pnas.1308335110; pmid: 23918387
48. J. A. Gagnon *et al*., Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. *PLOS ONE* **9**, e98186 (2014). doi: 10.1371/journal.pone.0098186; pmid: 24873830
49. H. J. Kim, H. J. Lee, H. Kim, S. W. Cho, J.-S. Kim, Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly. *Genome Res.* **19**, 1279–1288 (2009). doi: 10.1101/gr.089417.108; pmid: 19470664
50. G. E. Truett *et al*., Preparation of PCR-quality mouse genomic DNA with hot sodium hydroxide and tris (HotSHOT). *BioTechniques* **29**, 52–54 (2000).
51. J. Ablain, E. M. Durand, S. Yang, Y. Zhou, L. I. Zon, A CRISPR/Cas9 vector system for tissue-specific gene disruption in zebrafish. *Dev. Cell* **32**, 756–764 (2015). doi: 10.1016/j.devcel.2015.01.032; pmid: 25752963
52. E. Wienholds *et al*., Efficient target-selected mutagenesis in zebrafish. *Genome Res.* **13**, 2700–2707 (2003). doi: 10.1101/gr.1725103; pmid: 14613981
53. J. A. Villefranc, J. Amigo, N. D. Lawson, Gateway compatible vectors for analysis of gene function in the zebrafish. *Dev. Dyn.* **236**, 3077–3087 (2007). doi: 10.1002/dvdy.21354; pmid: 17948311

ACKNOWLEDGMENTS

C.K.K. and research reported here is supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health under award K08 AR061071. C.M. received support from a European Molecular Biology Organization long-term fellowship, a Human Frontier Science Program long-term fellowship, and a Swiss National Science Foundation advanced postdoctoral fellowship. L.I.Z. is supported by R01 CA103846, the Ellison Foundation, the V Foundation, and the Melanoma Research Alliance and is a Howard Hughes Medical Institute Investigator. R.A.Y. is supported by NIH HG002668. Z.P.F. was supported by NIH HG002668. E.J.H. is a Howard Hughes Medical Institute Fellow of the Helen Hay Whitney Foundation. L.I.Z. is a founder and stock holder of Fate, Inc. and Scholar Rock. R.A.Y. is a founder of Syros Pharmaceuticals. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We thank the Massachusetts General Hospital DNA Core for providing technical assistance with deep sequencing. We thank J. Ablain for critical review of the manuscript, C. Ceol for the gift of the zebrafish *mitfa* middle entry clone, and G. Musso for helpful discussions on microarray data and GSEA. All microarray, Chip-seq, and RNA-seq are deposited at Gene Expression Omnibus, ID GSE75356.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/351/6272/aad2197/suppl/DC1
Figs. S1 to S12
Tables S1 to S6
Movies S1 to S3

11 August 2015; accepted 22 December 2015
10.1126/science.aad2197

A zebrafish melanoma model reveals emergence of neural crest identity during melanoma initiation

Charles K. Kaufman *et al.*

Science **351**, (2016);
DOI: 10.1126/science.aad2197