Anatomic position determines oncogenic specificity in melanoma

Joshua M. Weiss[1,2,3], Miranda V. Hunter[2], Mohita Tagore[2], Yilun Ma[1,2,3], Sandra Misale[4], Theresa Simon-Vermot[2], Nathaniel R. Campbell[1,2,5,6], Felicity Newell[7], James S. Wilmott[8], Peter A. Johansson[7], John F. Thompson[8,9,10], Georgina V. Long[8,9,11], John V. Pearson[7], Graham J. Mann[8,12,13], Richard A. Scolyer[8,9,10,14], Nicola Waddell[7,15], Emily D. Montal[2], Ting-Hsiang Huang[2], Philip Jonsson[16], Mark T.A. Donoghue[17], Christopher C. Harris[17], Barry S. Taylor[17,18,19], Charlotte E. Ariyan[20], David B. Solit[17,19,21], Jedd D. Wolchok[17,21,22,23], Taha Merghoub[23], Neal Rosen[4], Nicholas K. Hayward[7], Richard M. White[2,24*]

[1] Weill Cornell/Rockefeller/Sloan Kettering Tri-Institutional MD-PhD Program, New York, NY, 10065, USA
[2] Department of Cancer Biology and Genetics, Memorial Sloan Kettering Cancer Center, New York, NY, 10065, USA
[3] Cell and Developmental Biology Program, Weill Cornell Graduate School of Medical Sciences, New York, NY, 10065, USA
[4] Molecular Pharmacology Program, Memorial Sloan Kettering Cancer Center, New York, NY, 10065, USA
[5] Computational and Systems Biology, Memorial Sloan Kettering Cancer Center, New York, NY, 10065, USA
[6] Physiology, Biophysics & Systems Biology Graduate Program, Weill Cornell Graduate School of Medical Sciences, New York, NY, 10065, USA
[7] QIMR Berghofer Medical Research Institute, Brisbane, QLD, 4006, Australia
[8] Melanoma Institute Australia, The University of Sydney, Sydney, NSW, 2006, Australia
[9] Faculty of Medicine and Health, The University of Sydney, Sydney, NSW, 2050, Australia
[10] Royal Prince Alfred Hospital, Sydney, NSW, 2050, Australia
[11] Royal North Shore Hospital, Sydney, NSW, 2065, Australia
[12] John Curtin School of Medical Research, Australian National University, ACT, 2601, Australia
[13] Centre for Cancer Research, Westmead Institute for Medical Research, The University of Sydney, Westmead, Sydney, NSW, 2528, Australia
[14] New South Wales Health Pathology, Sydney, NSW, 2099, Australia
[15] School of Medicine, The University of Queensland, Brisbane, QLD, 4072, Australia
[16] Celsius Therapeutics, Cambridge, MA, 02139, USA
[17] Human Oncology & Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY, 10065, USA
[18] Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY, 10065, USA
[19] Marie-Josée and Henry R. Kravis Center for Molecular Oncology, Memorial Sloan Kettering Cancer Center, New York, NY, 10065, USA
[20] Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, NY, 10065, USA
[21] Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, 10065, USA
[22] Parker Institute for Cancer Immunotherapy, San Francisco, California, 94129, USA
[23] Weill Cornell Medicine, New York, NY, 10065, USA
[24] Lead Contact
*Correspondence: whiter@mskcc.org

Summary

Oncogenic alterations to DNA are not transforming in all cellular contexts[1,2]. This may be due to pre-existing transcriptional programs in the cell of origin. Here, we define anatomic position as a major determinant of why cells respond to specific oncogenes. Cutaneous melanoma arises throughout the body, whereas the acral subtype has a unique tropism for the limbs, specifically the hands and feet[3]. We sequenced the DNA of cutaneous and acral melanomas from a large cohort of human patients and found a specific enrichment for BRAF mutations in cutaneous melanoma but CRKL amplifications in acral melanoma. We modeled these changes in transgenic zebrafish models and found that CRKL-driven tumors predominantly formed in the fins of the fish. The fins are the evolutionary precursors to tetrapod limbs, indicating that melanocytes in these acral locations may be uniquely susceptible to CRKL. RNA profiling of these fin/limb melanocytes, compared to body melanocytes, revealed a positional identity gene program typified by posterior HOX13 genes. This positional gene program synergized with CRKL to drive tumors at acral sites. Abrogation of this CRKL-driven program eliminated the anatomic specificity of acral melanoma. These data suggest that the anatomic position of the cell of origin endows it with a unique transcriptional state that makes it susceptible to only certain oncogenic insults.

Introduction

During development, cells express lineage-specific gene programs as well as position-specific gene programs. This coordinated transcriptional state ensures that cells pattern appropriately and fulfill the unique role they have at a given anatomic location[4]. Transcriptional programs unique to a particular anatomic site are referred to as positional identity[5]. Whether the transcriptional programs that mediate positional identity determines the response to oncogenes is unknown. Across many cancer types, the anatomic location of the tumor is associated with distinct molecular and clinical subtypes[6-9]. This is exemplified most strikingly by acral melanoma, a melanoma subtype defined by its position on the palms and soles of the hands and feet[3,8]. Compared to other forms of cutaneous melanoma, the acral subtype has a unique genetic profile, a lower response rate to both targeted and immunotherapy, and a worse overall survival[6,8,10-12]. Using data from human patients and transgenic zebrafish models, we identify genetic drivers unique to this anatomically restricted type of melanoma and find that the positional identity gene program in the cell-of-origin determines the competence to respond to those oncogenes.

Results

Acral driver genes lead to anatomically distinct tumors in transgenic zebrafish

To discern the genetic differences between acral and cutaneous melanoma, we performed targeted DNA sequencing of 100 acral and 839 cutaneous melanoma patients using the MSK-IMPACT platform[13-16] , a focused sequencing panel of 468 clinically relevant genes (Fig. 1a) . We calculated the log2-fold ratio of genetic alterations in acral versus cutaneous melanoma to generate an “acral enrichment score” (Fig. 1a). Consistent with previous reports, we found BRAF mutations as the most common coding mutation in cutaneous melanoma (Extended Data Fig. 1a). In contrast, we found that acral melanoma has a significantly higher frequency of copy number alterations (CNAs), a significantly lower frequency of activating BRAF mutations, and a lower overall mutational burden (Fig. 1a-b, Extended Data Fig. 1a-b)[6,17,18]. To better define which of the frequently amplified genes were likely acral-specific drivers, we also performed RNA-sequencing (RNA-seq) on an independent set of 61 acral and 53 cutaneous melanoma patient samples. Amplification of CRKL and GAB2 were amongst the top acral enriched genes identified by a combined analysis of DNA and RNA-sequencing (Fig. 1b-c, Extended Data Fig. 1c). These were of particular interest to us since CRKL and GAB2 are both oncogenic signal amplifier proteins that together form a complex with receptor tyrosine kinases (RTKs) to amplify downstream MAPK and PI3K signaling pathways through the recruitment of signaling mediators, such as SOS1 and SHP2 (Fig. 1d)[19]. They have previously been implicated in driving other cancer

types[20-23], but limited investigation has been performed in melanoma[24,25]. Other frequently altered genes included NF1 and TERT (Fig. 1a-b, Extended Data Fig. 1a-b), which significantly co-occur with alterations in CRKL and GAB2 across cancer (Extended Data Fig. 1e). This led to the hypothesis that alterations in CRKL, GAB2, NF1, and TERT may synergize to specifically drive acral compared to cutaneous melanoma (Extended Data Fig. 1d). As an example, we identified an acral melanoma patient at Memorial Sloan Kettering Cancer Center (MSKCC) who had amplification of CRKL and GAB2, deletion of NF1, and an activating promoter mutation in TERT in both a primary tumor and metastasis (Extended Data Fig. 1f-h).

Since there are currently no widely available animal models of acral melanoma, we used a rapid transgenic zebrafish system to model these potential drivers *in vivo*. Using the melanocyte-specific mitfa promoter, we created transgenic animals expressing putative acral melanoma drivers (CRKL, GAB2, TERT and NF1 loss) and then compared this to a previously developed cutaneous melanoma model (BRAFV600E;p53) or to wildtype melanocytes (Fig 1d., Extended Data Fig. 2a, Extended Data Fig. 3). We developed a rigorous set of criteria to detect tumors in both acral and cutaneous melanoma models by imaging (Supplemental Table 1) and confirmed the validity of these criteria with histology. Melanocytes in all models were fluorescently labeled with GFP (Fig. 1d, Extended Data Fig. 2a-b), and darkly pigmented cells with dendritic morphology on H&E, positive immunohistochemistry (IHC) staining for *mitfa*-GFP, RNA expression of *tyrp1a*, and immunofluorescence (IF) for *sox10* indicate that tumors generated by both models are of the melanocyte lineage (Extended Data Fig. 2b-d, Extended Data Fig. 3d). While wildtype melanocytes never formed melanomas, we found that 68% of the animals expressing the acral genes resulted in melanoma and 100% of the animals expressing the cutaneous genes resulted in melanoma by 1 year (Fig. 2f). We observed heterogeneous staining for both pERK and pAKT in the acral model (Extended Data Fig. 2b), consistent with previous studies demonstrating the ability for CRKL and GAB2 to activate both the MAPK and PI3K pathways[19]. We found a similar heterogeneous pattern of pERK staining from the in-transit metastasis of the representative acral melanoma patient identified at MSKCC, supporting that our model is representative of the human disease (Extended Data Fig. 2b).

Extensive research across many fields ranging from paleontology to genetics has demonstrated an evolutionary link between the morphogenesis of fish fins and tetrapod limbs, which include the acral sites of the hands and feet[26-29]. Although fins and limbs are structurally distinct, the genes involved in fin and limb development are well conserved[27,29]. This led us to ask whether the acral versus cutaneous driver genes would yield differences in anatomic distribution of the tumors. We monitored the fish as they developed tumors over the course of a year and calculated the relative proportion of tumors arising on the head, body, and fins. To visually aid our ability to determine the effect of genetic drivers on the anatomic distribution of tumors, we utilized the ternary plot, which compresses 3-dimensional categorical data into 2-dimensional space (Fig. 1e, Extended Data Fig. 2e)[30]. This revealed that the acral model developed a significantly higher proportion of fin tumors than the cutaneous model (53% versus 30%, $p < 0.0001$) (Fig. 1e, Extended Data Fig. 2e). The fin specificity of melanocytes in the acral model manifested early in development. At 3 days post-fertilization, melanocytes in the acral melanoma model demonstrated a greater total melanocyte area in the tailfin than in WT melanocytes (Fig. 1g). By 6 weeks post-fertilization the acral melanoma model developed fin hyperpigmentation and a significantly greater expansion of the melanocyte population in the fins than in the body (Extended Data Fig. 2f).

CRKL drives melanoma to acral sites

To more specifically determine which of the acral driver genes were sufficient for fin positioning, we expressed each one individually (Fig. 2a). CRKL was the only genetic driver sufficient to form tumors without any additional genetic alterations (Fig. 2b-c). Strikingly, 62% of CRKL-alone

tumors arose in the fins (Fig. 2d-e), indicating that fin melanocytes were more efficiently transformed by CRKL than melanocytes at other locations. Tumor initiation in CRKL tumors was slower compared to all 4 genes together, suggesting that GAB2 and TERT overexpression with NF1 loss accelerated disease progression (Fig. 2c). We then investigated whether any of the 4 putative genetic drivers were necessary for tumor progression by removing CRKL, GAB2, TERT, or NF1 loss from the full 4-gene acral model (Fig. 2f). While tumors still formed after withdrawal of GAB2, TERT, or NF1 loss, tumors never formed after withdrawal of CRKL (Fig. 2g-h), indicating that CRKL is both necessary and sufficient for tumor formation in the acral melanoma model. Withdrawal of GAB2, TERT, or NF1 loss from the acral model still demonstrated fin specificity (Fig. 2i), which suggests that CRKL is solely responsible for the preponderance of acral tumors. Together, these data demonstrate that CRKL is a key driver of acral melanoma and that fin melanocytes are more vulnerable to CRKL than melanocytes at other locations.

Positional identity gene programs determine the response to CRKL

These data suggest that fin melanocytes, analogous to human hand/feet melanocytes, are uniquely sensitive to CRKL. We hypothesized that this susceptibility was due to intrinsic differences in the gene program of these melanocytes, and that this gene program was due to anatomic positioning. To address this, we performed RNA-sequencing of body versus fin melanocytes. We did this in both the acral melanoma model as well as in wild-type melanocytes. GFP+ melanocytes and GFP- microenvironment cells were isolated from adult fish using a combination of dissection and FACS (Fig. 3a). As expected, the GFP+ population (compared to the GFP- population) showed a marked enrichment for melanocyte markers, such as *mitfa*, *pmela*, *tyr*, *tyrp1a*, *dct*, *sox10*, and *mlpha*, as well as the relevant transgenes (GFP, CRKL, etc), confirming that the successful isolation of melanocytes from both body skin and fins (Extended Data Fig. 4). We performed unsupervised hierarchical clustering on all of the samples. Surprisingly, this showed that the samples clustered first by cell lineage (melanocyte or not), second by anatomic location (fin vs body), and third by genotype (CRKL vs WT) (Fig. 3b). Along the same lines Principal Component Analysis (PCA) showed clustering by cell lineage on PC1 and by anatomic position on PC2 (Fig. 3c). This data suggests that anatomic location contributes a greater amount of transcriptional variation than the genotype and can shape the response to oncogenes, such as CRKL.

We performed pathway analysis comparing fin versus body melanocytes, which found appendage development as one of the top enriched pathways in fin melanocytes (Fig. 3d-e). This indicated that the differences in the transcriptional programs of fin versus body melanocytes were largely associated with anatomic position gene programs known as positional identity. More specifically, comparing fin melanocytes to body melanocytes, we found a significant upregulation of limb-specific HOX genes in the fin melanocytes (*hoxc12a*, *hoxa13a*, *hoxa13b*, *hoxb13a*, *hoxc13a*, *hoxd13a*) whereas the body melanocytes upregulated axial-specific HOX genes (*hoxb5a*, *hoxb6a*, *hoxb7a*, *hoxb8a*, *hoxb9a*) (Fig. 3f-g). Posterior HOX13 genes are master regulators of limb development[27], whereas the more anterior HOXB5-9 regulate many segments of axial body development[31]. Unsupervised clustering based on expression of all detected HOX genes clustered samples first by anatomic location, demonstrating that positional identities of fin versus body melanocytes are associated with a HOX code (Extended Data Fig. 5).

Human acral versus cutaneous melanoma has a positional identity gene program

We next wished to understand if these anatomic position gene programs were conserved in human acral versus cutaneous melanomas. We performed RNA-seq of human acral versus cutaneous melanoma and analyzed this by GSEA pathway analysis (Fig. 4a-b). Strikingly, appendage development and other positional identity-related pathways were amongst the most enriched pathways in human acral melanoma (Fig. 4b-c). When comparing just metastatic

samples by melanoma subtype, the appendage development pathway was also enriched (although due to small sample size, underpowered), indicating that limb positional identity may be maintained even when acral melanoma spreads to distant sites (Fig. 4c). Similar to what we observed in zebrafish, human acral melanoma had significantly greater expression of HOXA13, HOXB13, HOXD13 as well as other key homeobox transcription factors that regulate limb development, such as TBX4, DLX4, and HAND2 (Fig. 4d-e)[32-34]. PCA of all samples based on just HOX gene expression shows clustering by melanoma subtype (Extended Data Fig. 6). While the majority of melanoma samples expressed HOX genes in a pattern consistent with a melanoma HOX code, 14 out of 62 acral melanomas and 16 out of 53 cutaneous melanomas did not follow this pattern, suggesting that some melanomas may lose or acquire new positional identities over the course of tumorigenesis (Extended Data Fig. 6). Because many genes other than HOX genes are differentially expressed between acral and cutaneous melanoma, we also performed HOMER motif enrichment analysis to determine what transcription factors regulate the broader set of differentially expressed genes. A motif for HOXB13 was one of the top motifs indicating that HOX genes may regulate many of the transcriptional differences between the two subtypes (Fig. 4f).

Limb-specific HOX13 genes regulates IGF signaling in acral melanocytes

This data suggested that oncogenic drivers such as CRKL might synergize with transcriptional targets of limb-specific HOX13 genes. To identify the genes regulated by HOX13 in the limbs, we analyzed HOX13 ChIP-seq data from developing mouse limb bud tissue[35] (Fig. 5a). We used Cistrome-GO[36] to perform pathway analysis for genes with promoters or enhancers occupied by HOX13 and found that IGF/insulin signaling, a key regulator of limb development and regeneration, was amongst the topmost dysregulated pathways[37-40] (Fig. 5b-c, Extended Data Fig. 7a-b). We observed significant binding of HOXA13 and HOXD13 at multiple sites along the promoters of both IGF1 and IGF2 indicating that HOX13 genes regulate IGF (Extended Data Fig. 7c-d). We also found significant H3K27 acetylation peaks at HOX13 binding sites along the promoters of IGF1 and IGF2, indicating that HOX13 likely is associated with IGF1 and IGF2 expression (Extended Data Fig. 7c-d). Consistent with these data, we found that one of the top enriched pathways in fin versus body melanocytes was regulation of multicellular organism growth, where fin melanocytes have significantly elevated expression of *igf1*, *igf2a*, and *igf2b* compared to the body melanocytes (Fig 5d-e., Extended Data Fig. 7e). These data raised the possibility that growth of fin melanocytes may be cell-autonomously driven by activation of insulin/IGF signaling downstream of HOX13. In normal physiology, CRKL is recruited to RTKs upon ligand activation and amplifies downstream signaling[19], and our data suggested that IGF may be one such ligand. To test this, we performed a phospho-RTK array of 49 different RTKs in a human NF1-null melanoma cell line with or without overexpression of CRKL (Fig. 5f). CRKL overexpression specifically increased the phosphorylation of the insulin and IGF1 receptors (INSR/IGF1R) (Fig. 5f), consistent with increased activation of these RTKs in the setting of CRKL overexpression.

Abrogation of IGF signaling eliminates anatomic positioning of acral melanoma

This data raised the possibility that IGF signaling, downstream of the intrinsic HOX13 program, might synergize with CRKL and explain why that oncogene is enriched at acral sites. To test this functionally *in vivo*, we overexpressed a previously validated dominant negative IRS2 (dnIRS2) to block insulin/IGF signaling in the CRKL-driven acral melanoma model[41,42] (Fig. 5g). We found that the dnIRS2 significantly slowed tumor initiation (Extended Data Fig. 8a), and most importantly led to a loss of fin specificity in the acral melanoma model (Fig. 5h). We also tested whether pharmacologic inhibition of this pathway would block the fin phenotype of the acral melanoma model. Treatment of the acral melanoma model with INSR/IGF1R antagonists (BMS-754807 and NVP-AEW541)[42-44] led to a significant decrease in the total melanocyte area in the tailfin (Fig. 5i, Extended Data Fig. 8b). To determine whether INSR/IGF1R was operating through activation of

the MAPK or PI3K pathway, we treated the acral melanoma model with a PI3K or a dual RAF/MEK inhibitor (LY294002 and CH5126766)[45,46] and found that only the PI3K inhibitor, but not the RAF/MEK inhibitor, significantly decreased total melanocyte area in the tailfin (Fig. 5j, Extended Data Fig. 8c). Together, these data demonstrate that CRKL-mediated transformation depends on a HOX13/IGF-associated limb positional gene program (Extended Data Fig. 9).

Discussion

What are critical factors necessary for a cell to become a cancer? Many studies have documented phenotypically normal cells with classic oncogenic driver mutations[1,2], suggesting that tumorigenesis depends on not just on mutations but also on a particular cell state. Here we introduce positional identity, a transcriptional program linked to a particular anatomic location[5], as a key determinant of the transforming potential of oncogenes in cancer. Through a combination of DNA and RNA-sequencing of human patients, we identified CRKL amplification as specifically enriched in acral melanoma, a subtype defined by its anatomic position on the hands and feet[3]. Due to the evolutionary relationship between zebrafish fins and human hands and feet[27,28], we were able to establish the first *in vivo* model of acral melanoma. We found that fin melanocytes were more susceptible to CRKL transformation than melanocytes at other anatomic sites. Mechanistically we found that fin melanocytes have a limb positional identity defined by HOX13 genes, which drive higher levels of IGF. We identified the same positional identity programs in human melanoma. Because CRKL potentiates IGF signaling, CRKL amplification synergizes with a limb positional identity to drive melanoma most efficiently at acral anatomic sites (Extended Data Fig. 9).

Positional identity can address a longstanding question, which is why do cancers arising at different locations often possess distinct genetic and clinical characteristics? Examples include not just other skin cancers like basal and squamous cell carcinoma[2], but also colon cancer[7], gastric cancer[47], cholangiocarcinoma[48], and glioma[49]. While the HOX13-IGF mechanism we propose may be specific to acral melanoma, the concept that oncogenes interact with location-specific programs in the cell of origin may apply to a variety of cancer types. Because oncogenes like CRKL are dependent on positional identity programs to drive tumorigenesis (Extended Data Fig. 9), positional identity programs themselves open up new opportunities in the treatment of cancer. Future studies to comprehensively characterize the positional dependencies of oncogenes will be a valuable resource in elucidating the origins of cancer and to treat cancer based on its positional identity.

Acknowledgements
J.M.W. was supported by the Kirschstein-NRSA predoctoral fellowship (F30) from the NIH under award number F30CA236442 and by a predoctoral fellowship (T32) by the Cell and Developmental Biology Program at Weill Cornell Graduate School supported by the the NIH under award number T32GM008539. N.R.C was supported by the Kirschstein-NRSA predoctoral fellowship (F30) from the NIH under award number F30CA220954. J.M.W., Y.M., and N.R.C were supported by a Medical Scientist Training Program grant from the NIH under award number T32GM007739. M.T. was supported by the Marie-Josée Kravis Women in Science Endeavor Postdoctoral Fellowship. E.D.M. was supported by the Predoctoral to Postdoctoral Fellow Transition Award (F99/K00) from the NIH under award number K00CA223016. R.A.S., G.V.L., J.F.T., G.J.M. and N.K.H. are supported by an Australian National Health and Medical Research Council Program Grant. R.A.S., N.K.H. and G.V.L. are supported by the Australian National Health and Medical Research Council Fellowship program. The authors acknowledge support from colleagues at their respective institutions and support from the Cameron Family and the Ainsworth Foundation is also gratefully acknowledged. This research was funded in part through the NIH/NCI Cancer Center Support Grant P30 CA008748, Ludwig Collaborative and Swim Across America Laboratory, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA; Parker Institute for Cancer Immunotherapy, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA; Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA; Weill Cornell Medicine, New York, NY 10065, USA. This work was supported by awards from the Melanoma Research Alliance, The Debra and Leon Black Family Foundation, NIH Research Program Grants R01CA229215 and R01CA238317, NIH Director's New Innovator Award DP2CA186572, The Pershing Square Sohn Foundation, The Mark Foundation, The Alan and Sandra Gerry Metastasis Research Initiative at the Memorial Sloan Kettering Cancer Center, The Harry J. Lloyd Foundation, Consano and the Starr Cancer Consortium (all to R.M.W.). Thank you to Lambros T. Koufariotis for technical assistance with the alignment of human RNA-seq data.

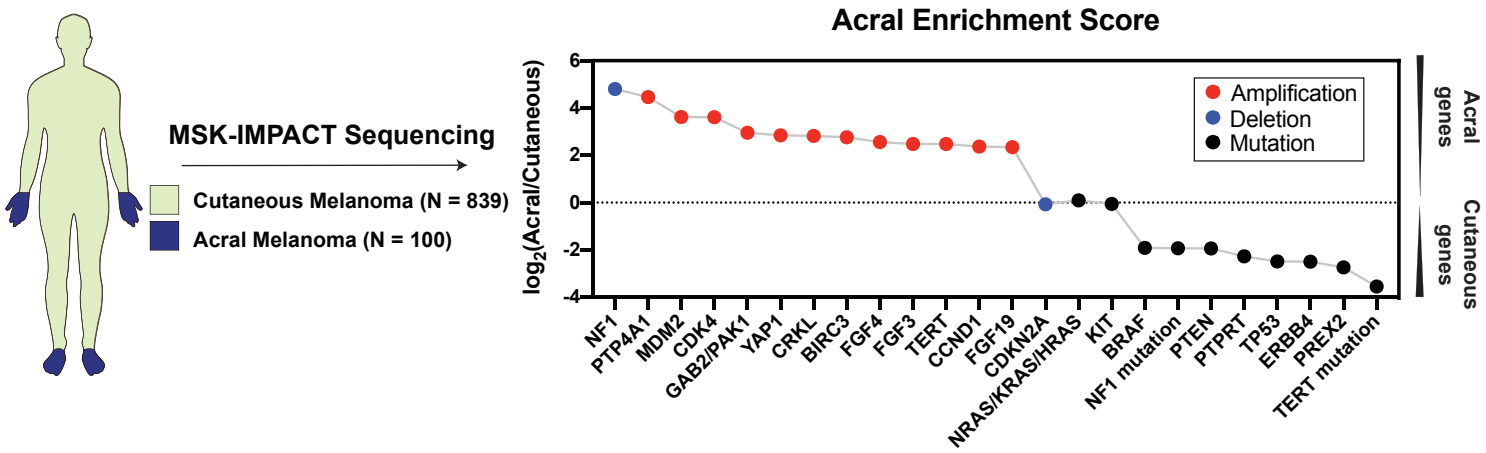
Author Contributions
Conceptualization, J.M.W. and R.M.W; Methodology, J.M.W. and T.S.V; Software, J.M.W., M.V.H., and N.R.C.; Formal Analysis, J.M.W, M.V.H., M.T., N.R.C., P.J., C.C.H., M.D., B.S.T., D.B.S.; Investigation, J.M.W., Y.M., S.M., T.S.V., E.D.M., T.H.H., Resources, R.M.W.; Data Curation, F.N., J.S.W., P.A.J., J.F.T., G.V.L., J.V.P, G.J.M., R.A.S., N.W., P.J., C.C.H., M.D., B.S.T, C.E.A., D.B.S, J.D.W., T.M., and N.K.H.; Writing – Original Draft, J.M.W. and R.M.W.; Writing – Review & Editing, J.M.W., M.V.H., S.M., C.E.A., and R.M.W.; Visualization, J.M.W; Supervision, C.E.A., D.B.S., J.D.W., T.M., N.R., and R.M.W.; Project Administration, J.M.W., C.E.A., D.B.S., J.D.W., T.M., N.R., and R.M.W.; Funding Acquisition, R.M.W.

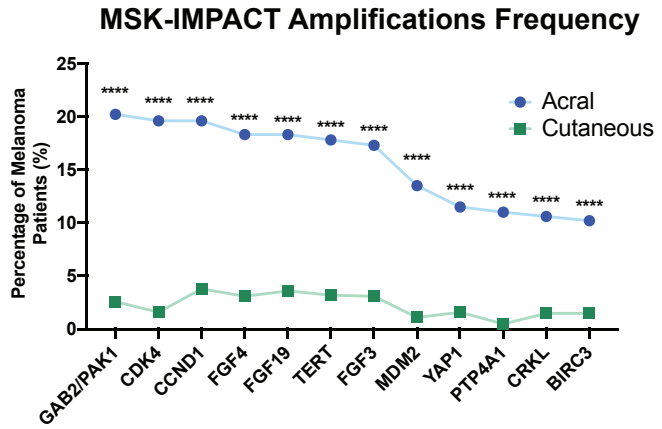
Declaration of Interests
S.M. consulted for Boehringer-Ingelheim. J.F.T. has received honoraria for advisory board participation from BMS Australia, MSD Australia, GSK and Provectus Inc, and travel support from GSK and Provectus Inc. R.A.S. has received fees for professional services from Qbiotics, Novartis, MSD Sharp & Dohme, NeraCare, AMGEN Inc., Bristol-Myers Squibb, Myriad Genetics, GlaxoSmithKline. J.V.P. and N.W. are equity holders and Board members of genomiQa PTY LTD. P.J. is currently employed by Celsius Therapeutics. B.S.T. reports receiving Honoria and research funding from Genentech and Illumina and advisory board activities for Boehringer Ingelheim and Loxo Oncology, a wholly owned subsidiary of Eli Lilly. B.S.T. is currently employed by Loxo Oncology. D.B.S. has consulted with/received honoraria from Pfizer, Loxo Oncology, Lilly Oncology, Vivideon Therapeutics, Q.E.D. Therapeutics, and Illumina. J.D.W. is a consultant for Amgen; Apricity; Arsenal; Ascentage Pharma; Astellas; Boehringer Ingelheim; Bristol Myers Squibb; Eli Lilly; F Star; Georgiamune; Imvaq; Kyowa Hakko Kirin; Merck; Neon Therapeutics;

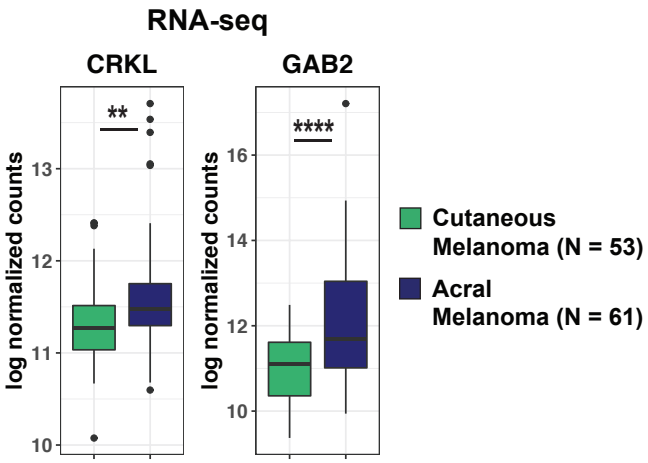
Polynoma; Psioxus, Recepta; Trieza; Truvax; Sellas. J.D.W. has grant and research support from Bristol Meyers Squibb and Sephora. J.D.W. has equity in Tizona Pharmaceuticals; Imvaq; Beigene; Linneaus, Apricity, Arsenal IO; Georgiamune. T.M. is a consultant for Leap Therapeutics, Immunos Therapeutics and Pfizer, and co-founder of Imvaq therapeutics. T.M. has equity in Imvaq therapeutics. T.M. reports grants from Bristol Myers Squibb, Surface Oncology, Kyn Therapeutics, Infinity Pharmaceuticals, Peregrine Pharmeceuticals, Adaptive Biotechnologies, Leap Therapeutics and Aprea. T.M. is inventor on patent applications related to work on oncolytic viral therapy, alphavirus-based vaccines, neo-antigen modeling, CD40, GITR, OX40, PD-1 and CTLA-4. N.R. is on the SAB and receives research funding from Chugai, on the SAB and owns equity in Beigene, and Fortress. N.R. is also on the SAB of Daiichi-Sankyo, Astra-Zeneca-MedImmune, and F-Prime, and is a past SAB member of Millenium-Takeda, Kadmon, Kura, and Araxes. N.R. is a consultant to Novartis, Boehringer Ingelheim, Tarveda, and Foresight and consulted in the last three years with Eli Lilly, Merrimack, Kura Oncology, Araxes, and Kadman. N.R. owns equity in ZaiLab, Kura Oncology, Araxes, and Kadman. N.R. also collaborates with Plexxikon. R.M.W. is a paid consultant to N-of-One Therapeutics, a subsidiary of Qiagen. R.M.W. is on the Scientific Advisory Board of Consano but receives no income for this. R.M.W. receives royalty payments for the use of the casper line from Carolina Biologicals.

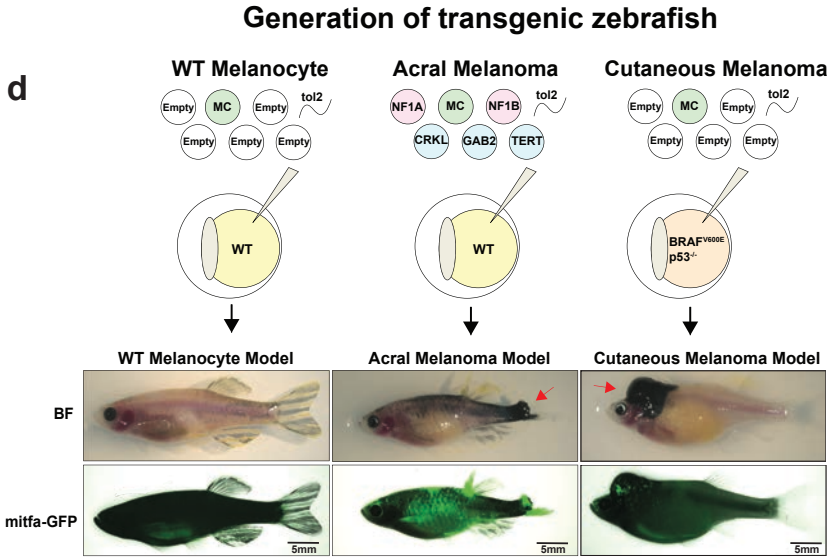
Fig. 1: Acral versus cutaneous melanoma driver genes lead to anatomically distinct tumors in transgenic zebrafish.
(a-b) MSK-IMPACT targeted sequencing of 100 acral and 839 cutaneous melanoma patients. (a) Acral enrichment score was calculated by dividing the frequency a gene is altered in acral melanoma by the frequency it is altered in cutaneous melanoma. Positive scores indicate genes enriched in acral melanoma and negative scores indicate genes enriched in cutaneous melanoma. (b) Frequency of gene amplifications compared by melanoma subtype using Fisher's exact test. (c) RNA sequencing on a separate cohort of 61 acral and 53 cutaneous melanoma patients shows significantly higher expression of CRKL and GAB2 in acral versus cutaneous melanoma. P-values calculated using DESeq2. (d) Transgenic zebrafish models. The WT melanocyte model expresses no oncogenic drivers. The acral melanoma model has overexpression of CRKL, GAB2, TERT and knockout of *nf1a* and *nf1b.* The cutaneous melanoma model expresses $BRAF^{V600E}$ in a $p53^{-/-}$ genetic background. (e) Ternary diagram portraying the percentage of tumors arising in the head, body, or fins between the acral and cutaneous melanoma models. Anatomic distribution was compared using a Chi-squared test. See Extended Data Figure 2e for histogram representation. (f) Tumor-free survival of WT melanocytes (n = 94), acral (n = 230), cutaneous (n = 182) transgenic zebrafish models. P-values generated by log-rank Mantel-Cox test. (g) WT melanocyte vs acral melanoma model compared for melanocyte area in tailfin at 3 days post-fertilization. Data represents pooling of four biological replicates. P-values generated by a student's t-test. Error bars = SEM. * p-value < 0.05, ** p-value < 0.01, *** p-value < 0.001, **** p-value < 0.0001.

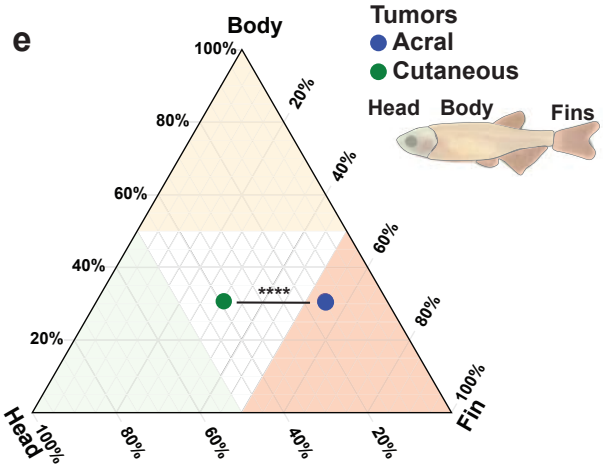
Fig. 1: Acral versus cutaneous melanoma driver genes lead to anatomically distinct tumors in transgenic zebrafish.

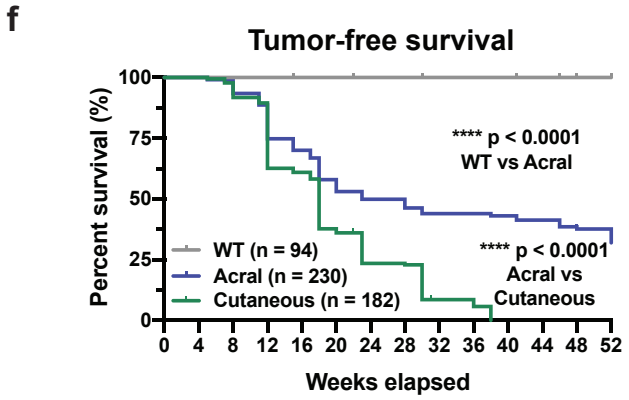












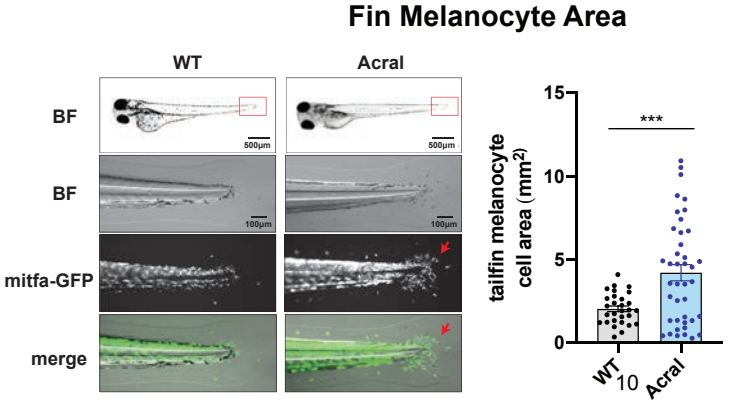
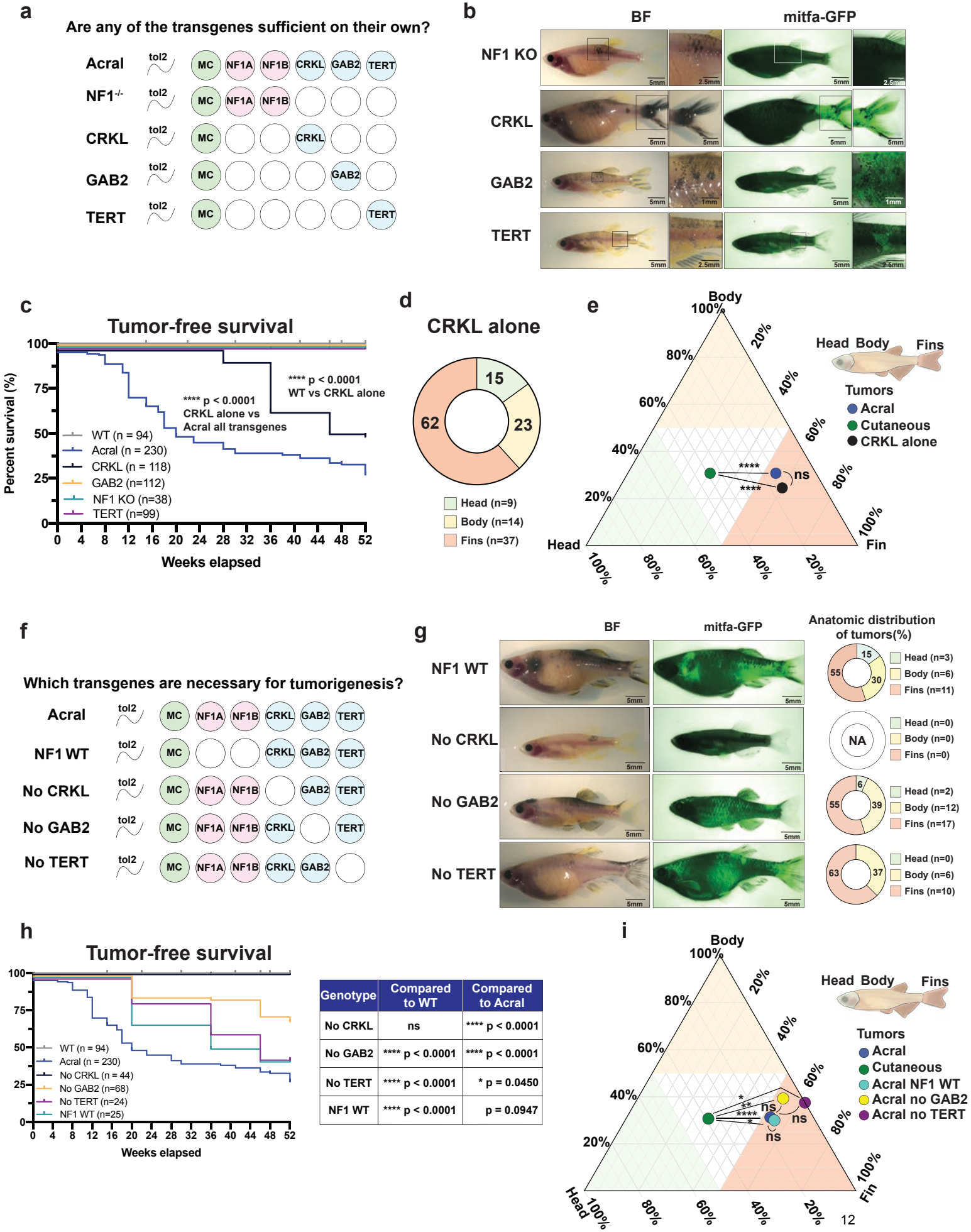


Fig. 2: CRKL drives melanoma to acral sites.
(a) The 4 genes (CRKL, GAB2, TERT, NF1 KO) used to drive the acral melanoma model were introduced separately to assess sufficiency in tumorigenesis. (b) Representative images of zebrafish for each genotype. (c) Tumor-free survival of indicated genotypes. Number of fish for each genotype indicated in figure key. P-values generated by log-rank Mantel-Cox test. (d) Pie chart showing the percentage of CRKL-alone melanomas arising at each anatomic location. (e) Ternary diagram portraying the percentage of tumors forming in the head, body, or fins of indicated genotypes. A Chi-squared test was performed to compare the anatomic distribution between the different transgenic models. (f) Each one of the 4 acral drive genes was removed to assess which driver genes were necessary for tumorigenesis. (g) Representative images of zebrafish for each genotype. Right panel shows pie charts with the anatomic distribution of tumors. Number of tumors are indicated. (h) Tumor-free survival of indicated genotypes. Number of fish for each genotype indicated in figure key. P-values generated by log-rank Mantel-Cox test. (i) Ternary diagram comparing the indicated genotypes. Chi-squared test used to compare anatomic distribution between genotypes. * p-value < 0.05, ** p-value < 0.01, *** p-value < 0.001, **** p-value < 0.0001.

Fig. 2: CRKL drives melanoma to acral sites.



Genotype	Compared to WT	Compared to Acral
No CRKL	ns	**** p < 0.0001
No GAB2	**** p < 0.0001	**** p < 0.0001
No TERT	**** p < 0.0001	* p = 0.0450
NF1 WT	**** p < 0.0001	p = 0.0947

Fig. 3: Positional identity gene programs determine the response to CRKL.
(a) Schematic illustrating the fin versus body RNA-seq experiment. Body skin and fins were isolated from WT melanocyte and acral melanoma fish by dissection and then FACS sorted for GFP+ melanocytes and GFP- microenvironmental cells. (b) Unsupervised clustering using the 500 most variable genes across all samples, which clustered by cell lineage, anatomic location, and then genotype. (c) Principal component analysis (PCA) for all samples showing principal components 1 (PC1) and 2 (PC2). The percent of transcriptional variation captured by each principal component is indicated. (d) GSEA pathway analysis comparing fin vs body melanocytes from the WT melanocyte model listing the top enriched pathways in fin melanocytes. (e) GSEA barcode plot showing enrichment of genes in the GO: Appendage Development pathway. NES and FDR are indicated. (f-g) Volcano plots comparing fin melanocytes vs body melanocytes from the (f) WT melanocyte model and (g) acral melanoma model. Genes with FDR-adjusted p-value < 0.05 indicated in blue.

Fig. 3: Positional identity gene programs determine the response to CRKL.

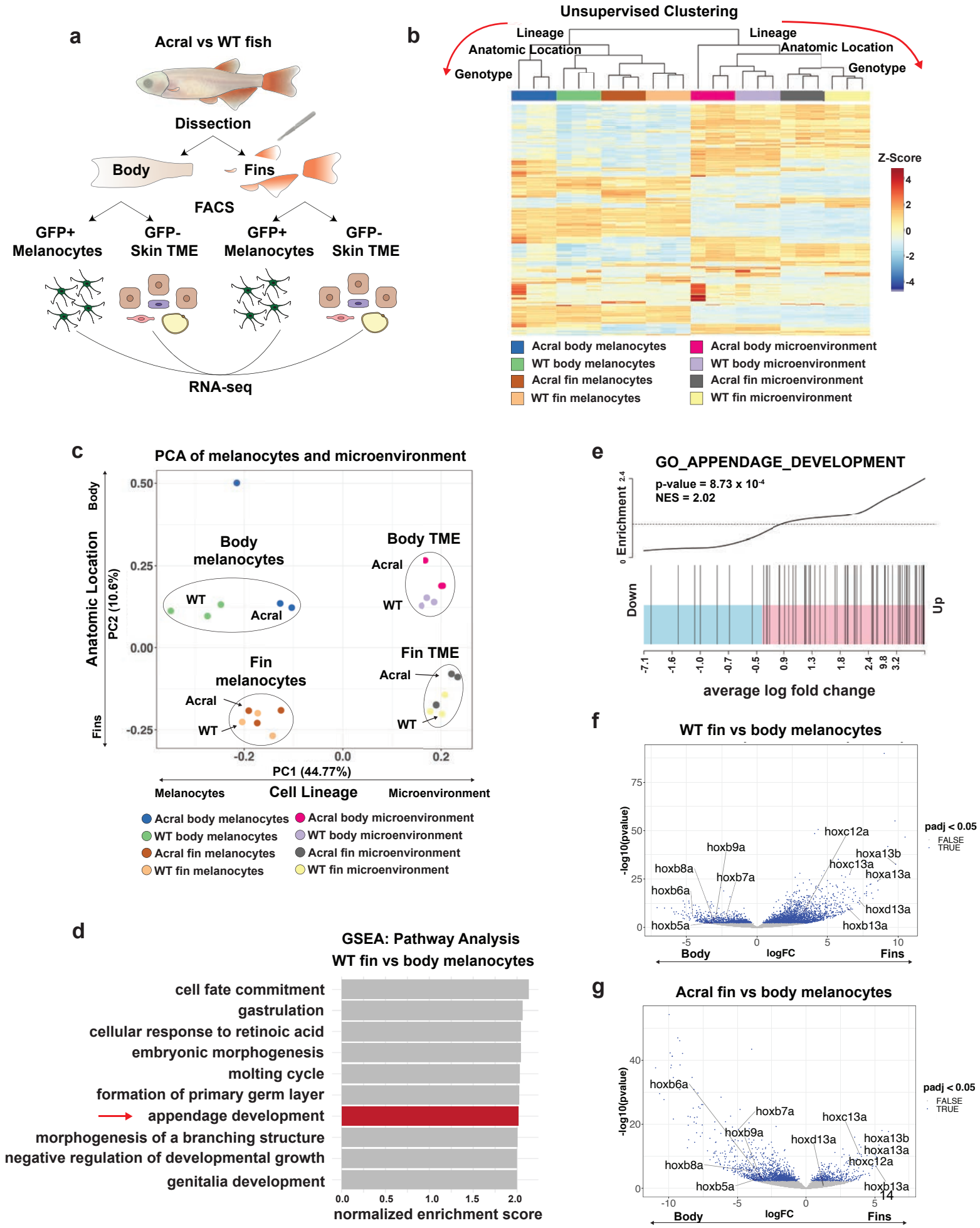


Fig. 4: Human acral versus cutaneous melanoma has a positional identity gene program. (a) RNA-seq performed on 61 acral and 53 cutaneous melanoma human patient samples. (b) GSEA pathway analysis for top enriched pathways in human acral melanoma and is controlled for disease stage. (c) GSEA barcode plot showing enrichment of genes in the GO: Appendage Development pathway. NES and FDR are indicated. (d) Volcano plot comparing showing differentially expressed genes between all acral and cutaneous melanoma samples. Genes with FDR-adjusted p-value < 0.05 indicated in blue. (e) Boxplots showing differences in gene expression between all acral melanoma and cutaneous melanoma samples. P-values adjusted for FDR = 0.05 indicated on the plot. Error bars = SEM. (f) Genes differentially expressed between all acral and cutaneous melanoma samples underwent HOMER known motif enrichment analysis for predicted transcription factor binding motifs. * p-value < 0.05, ** p-value < 0.01, *** p-value < 0.001, **** p-value < 0.0001.

Fig. 4: Human acral versus cutaneous melanoma has a positional identity gene program.

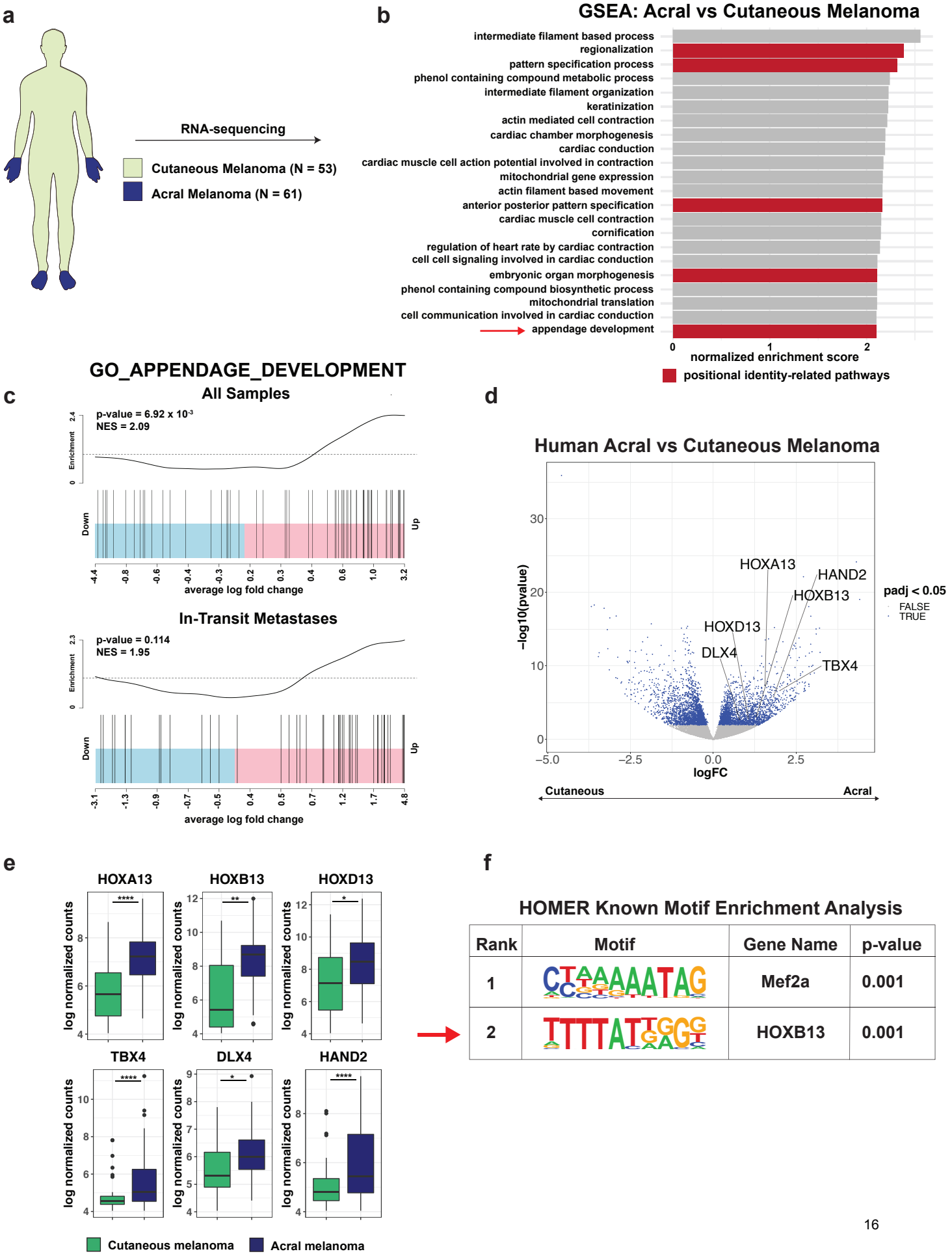
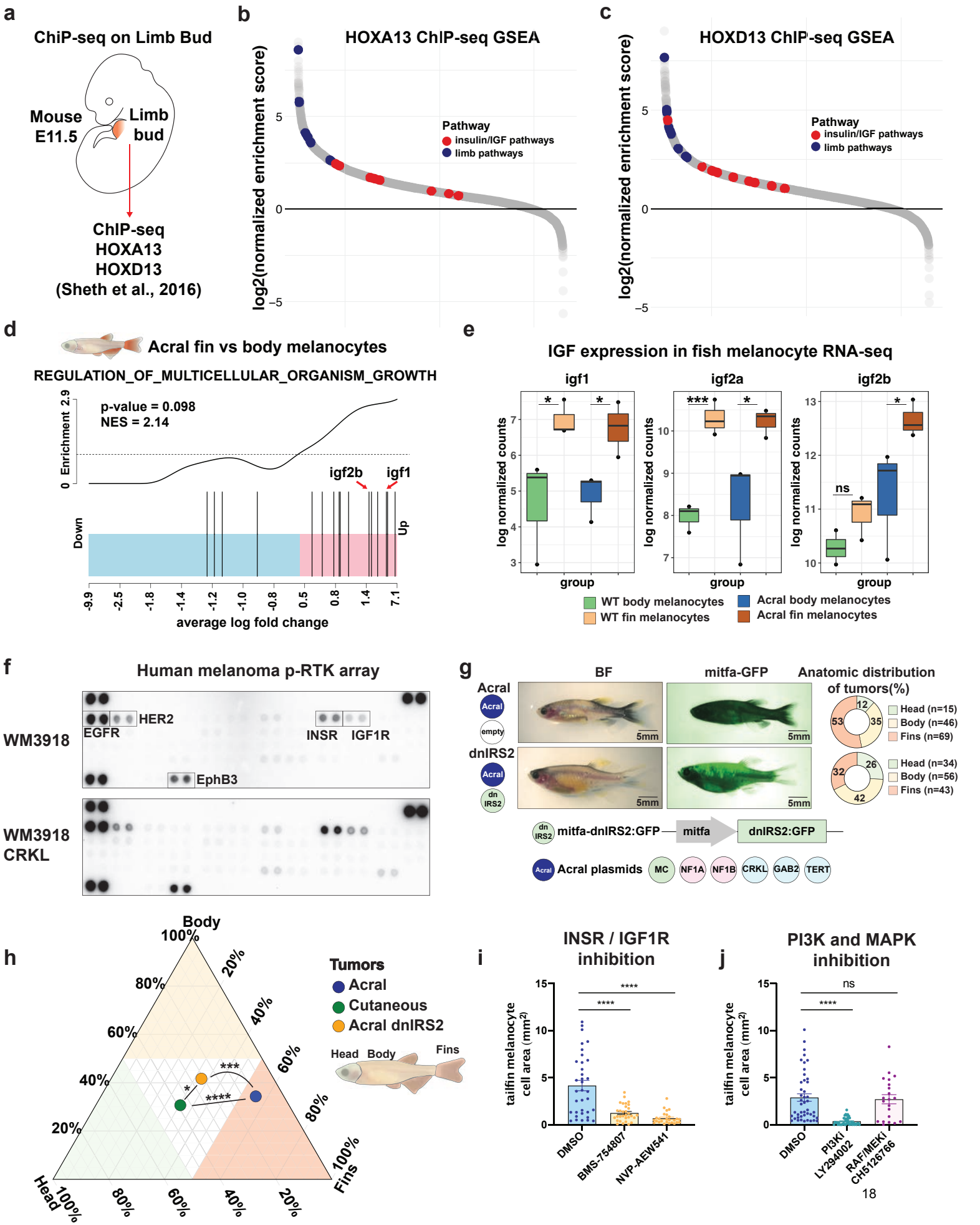


Fig. 5: HOX13 synergizes with CRKL to drive acral melanoma through insulin/IGF signaling.
(a) HOXA13 and HOXD13 ChIP-seq performed on developing limb buds from E11.5 mouse embryos[35]. (b-c) Waterfall plot showing gene pathways regulated by HOXA13 and HOXD13 binding, highlighting limb and insulin/IGF-related pathways. (d) GSEA barcode plot comparing Acral model fin vs body melanocytes showing enrichment of genes in the GO: Regulation of Multicellular Organisms Growth pathway. NES and FDR are indicated. (e) Boxplot showing log normalized counts of zebrafish *igf1*, i*gf2a*, and *igf2b* expression in all melanocyte samples. P-values are adjusted for FDR = 0.05. (f) Phospho-RTK array performed on WM3918 human melanoma cell line with or without overexpression of CRKL. Array tests phosphorylation status of 49 RTKs. (g) Schematic of dnIRS2 transgene used to block insulin/IGF signaling in zebrafish melanocytes. Representative images of acral melanoma model with or without overexpression of dnIRS2-GFP. Pie charts demonstrate the anatomic distribution of each genotype. (h) Ternary plot showing the anatomic distribution of tumors with indicated genotypes. P-values generated by Chi-squared test. (i-j) Acral melanoma model imaged for melanocyte tailfin area at 3-days post-fertilization after pharmacologic treatment. P-value generated with student's t-test. Error bars = SEM. (i) Insulin/IGF1 receptor antagonists BMS-754807 at 7.5μM and NVP-AEW541 at 60μM. Data is pooled from four independent biological replicates. (j) PI3K inhibitor LY294002 at 15μM and RAF/MEK inhibitor CH5126766 at 1μM. Data is pooled from three independent biological replicates. * p-value < 0.05, ** p-value < 0.01, *** p-value < 0.001, **** p-value < 0.0001.

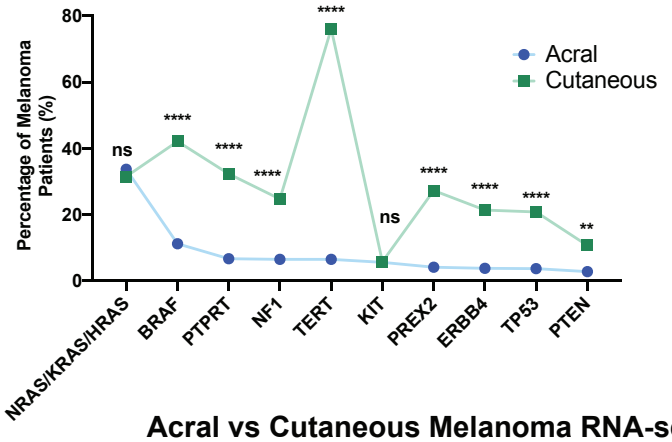
Fig. 5: HOX13 synergizes with CRKL to drive acral melanoma through insulin/IGF signaling.



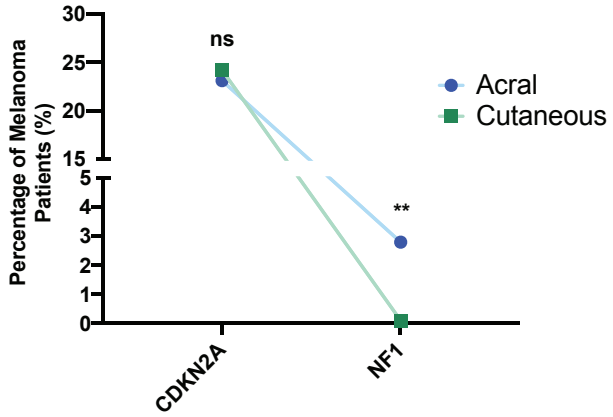
Extended Data Fig. 1: Identification of acral versus cutaneous melanoma genes. Related to Fig. 1. (a-b) MSK-IMPACT of acral and cutaneous melanoma. Fisher's exact test was used to compare the frequency of the most recurrently mutated and deleted genes by melanoma subtype. Both coding and promoter mutations were counted for TERT. (c) RNA-seq of 61 acral and 53 cutaneous melanoma patient samples. Boxplots compare indicated genes by subtype. Error bars = SEM. (d) Schematic detailing predicted synergistic interaction between putative driver genes in acral melanoma. (e) TCGA pan-cancer analysis shows significant co-occurrence of CRKL, GAB2, NF1, and TERT alterations. (f) Clinical course of an acral melanoma patient. (g) WGS copy number profile showing copy number changes in GAB2, CRKL, and NF1. (h) Putative drivers of patient's acral melanoma. * p-value < 0.05, ** p-value < 0.01, *** p-value < 0.001, **** p-value < 0.0001.

Extended Data Fig. 1: Identification of acral versus cutaneous melanoma genetic drivers.

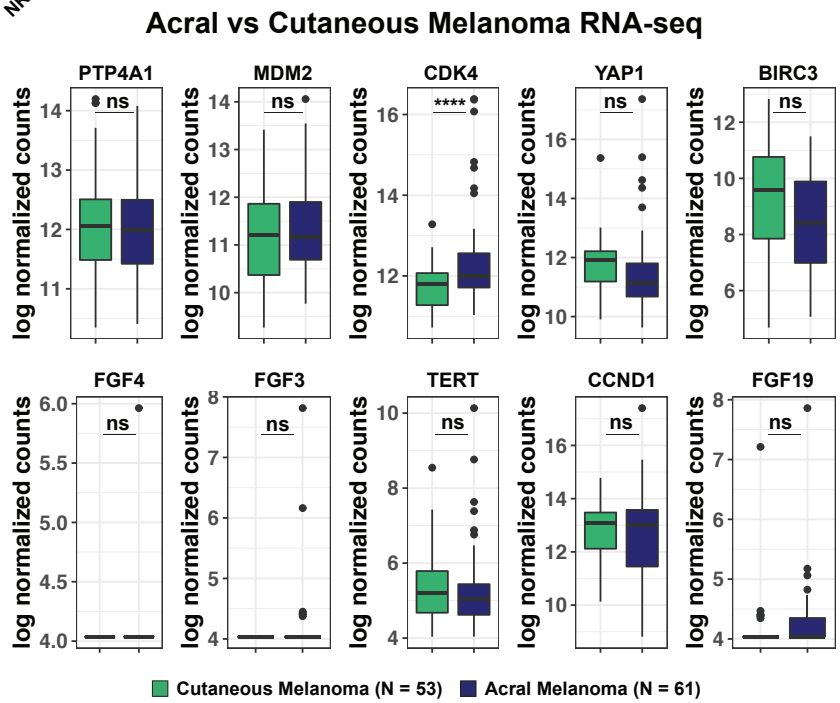
a



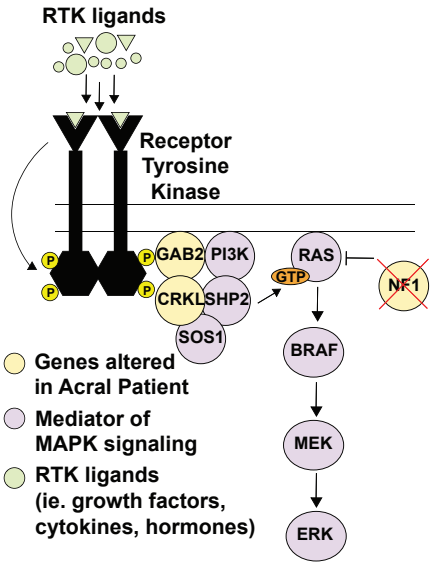
b



c



d

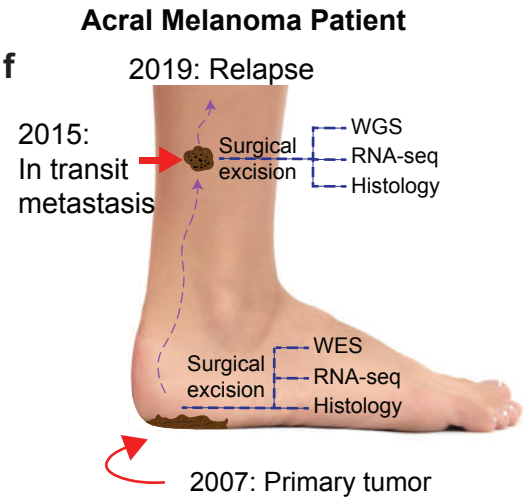


e

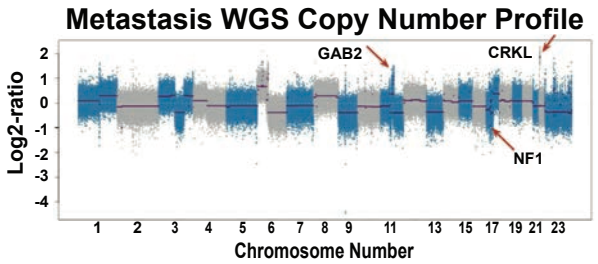
Co-Occurrence of CRKL, GAB2, NF1, and TERT in Cancer, TCGA

Gene A	Gene B	Neither	A Not B	B Not A	Both	Log_2 Odds Ratio	p-Value	q-Value	Tendency
CRKL	GAB2	23,159	412	838	57	1.935	<0.001	<0.001	Co-occurrence
CRKL	NF1	22,347	389	1,744	82	1.434	<0.001	<0.001	Co-occurrence
GAB2	NF1	21,858	793	1,713	102	0.715	<0.001	<0.001	Co-occurrence
CRKL	TERT	22,597	400	1,494	71	1.425	<0.001	<0.001	Co-occurrence
GAB2	TERT	22,194	779	1,377	116	1.263	<0.001	<0.001	Co-occurrence
NF1	TERT	21,447	1,638	1,364	203	0.962	<0.001	<0.001	Co-occurrence

f



g



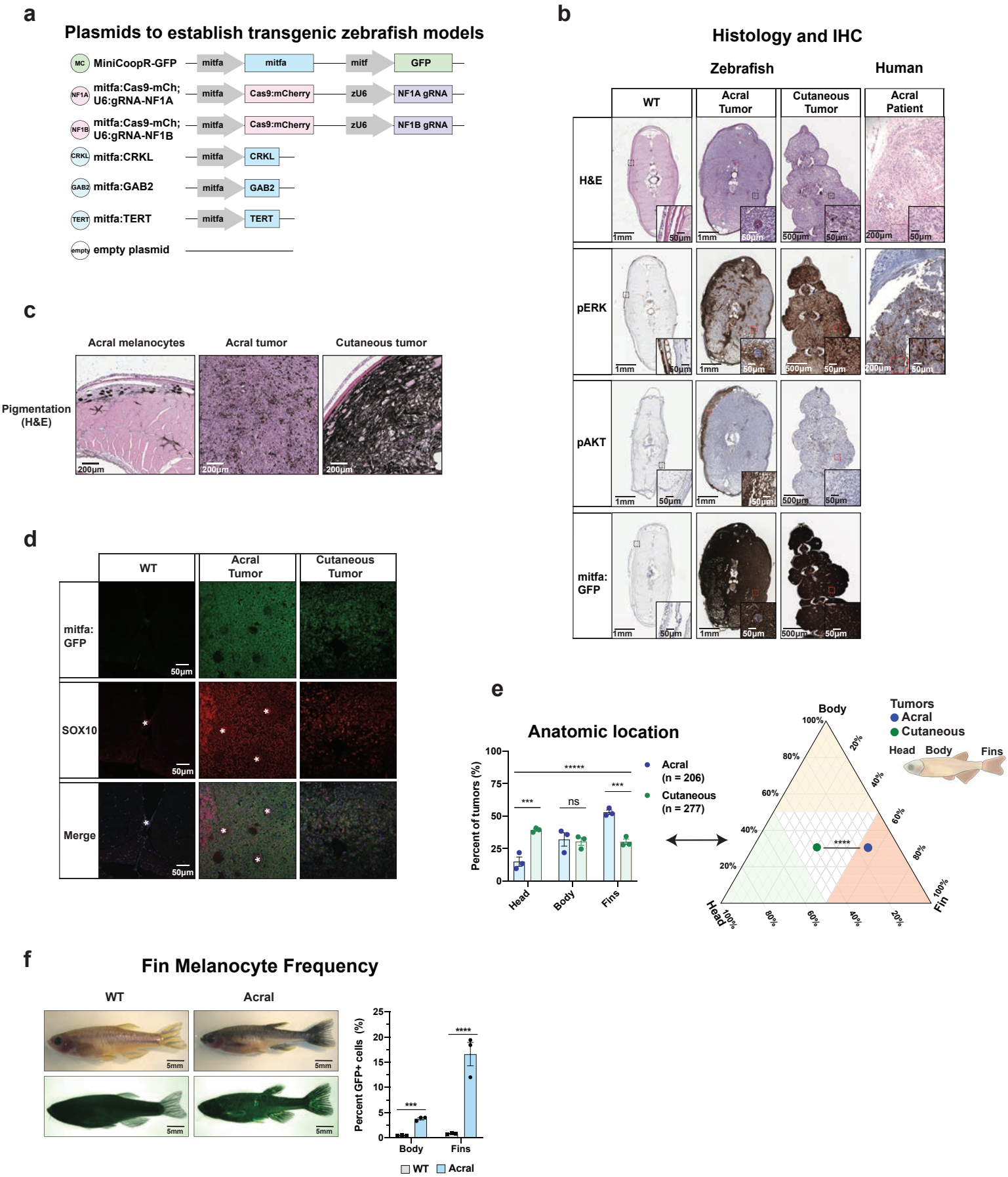
h

Putative Driver Mutations

Gene	Mutation Type	Primary	Metastasis
NF1	Deletion	+	+
CRKL	Amplification	+	+
GAB2	Amplification	+	+
TERT	Promoter→ Upregulation	+	+

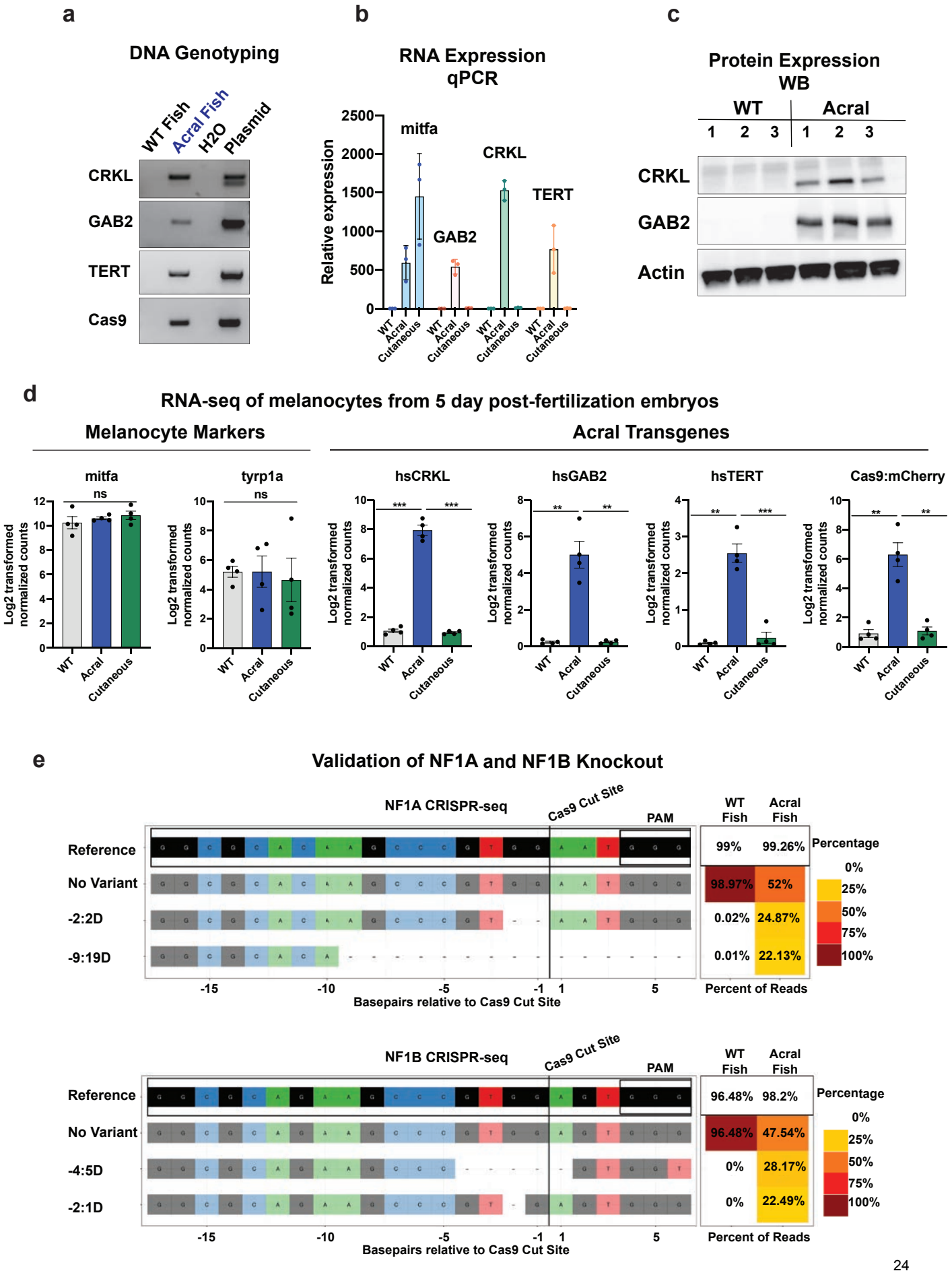
Extended Data Fig. 2: A transgenic zebrafish model of acral melanoma.
Related to Fig. 1. (A) Plasmids used to create the WT melanocyte model, acral melanoma model, and cutaneous melanoma models. (b) Histology and IHC on transverse sections of WT melanocyte model, acral melanoma model, cutaneous melanoma model, and a human patient's in-transit metastasis. Sections were stained with H&E and IHC for GFP, pERK, and pAKT. (c) H&E on transverse sections of acral and cutaneous melanoma models showing pigmentation. (d) Immunofluorescence on transverse sections of WT melanocyte model, tumor-bearing Acral melanoma model, and tumor-bearing Cutaneous melanoma model for GFP and SOX10. GFP labels all tumor cells and SOX10 is used as a melanocyte lineage marker. Asterisks indicate blood vessels with autofluorescence. (e) Histogram portraying the percentage of tumors forming in the head, body, or fins of the fish between the Acral fish and Cutaneous fish melanoma models. Data represents three biological replicates. To compare overall anatomic distribution between the two genotypes, a Chi-squared test was performed. To compare the frequency of tumors at each anatomic location, a student's t-test was performed. Error bars = SEM. Ternary plot represents the same data presented in the histogram. (f) Adult WT melanocyte and acral melanoma models were dissected to isolate body skin and fins and then analyzed via flow cytometry. The images show representative differences in pigment patterning between the two models. The histogram shows percentage of total cells that were GFP+ grouped by anatomic location and genotype. Data represents three biological replicates. P-values generated by student's t-test. Error bars = SEM. * p-value < 0.05, ** p-value < 0.01, *** p-value < 0.001, **** p-value < 0.0001.

Extended Data Fig. 2: A transgenic zebrafish model of acral melanoma.



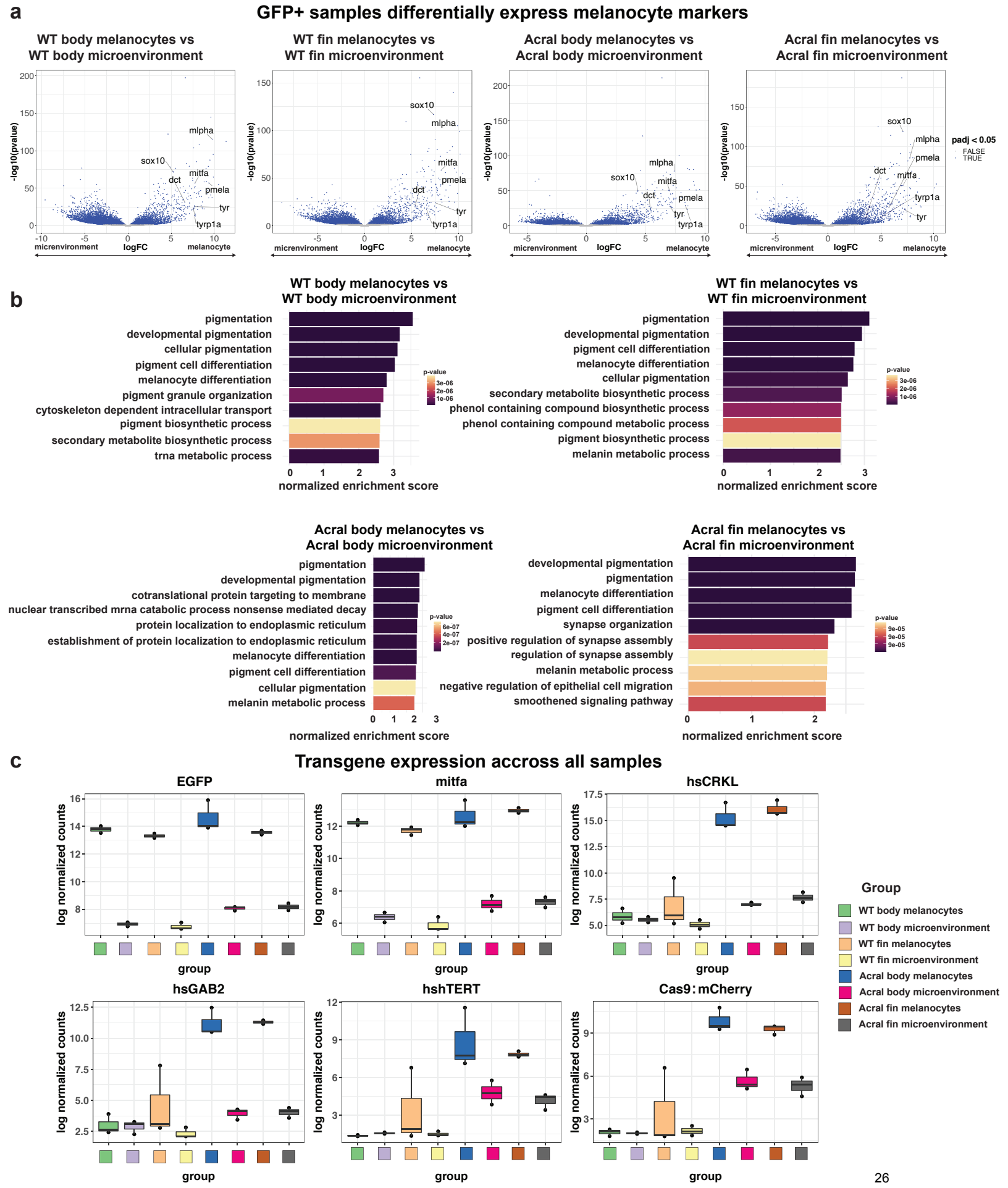
Extended Data Fig. 3: Acral melanoma model demonstrates overexpression and knockout of multiple driver genes.
Related to Fig. 1. (a) Genotyping PCR to confirm integration of plasmids into zebrafish genome. (b) qPCR to validate RNA expression of transgenes. (c) Western blot was performed for human CRKL and GAB2 to confirm transgene expression. Three biological replicates are shown. (d) RNA-seq on embryos at 5 days post-fertilization. Log normalized expression of melanocyte markers and acral transgenes are indicated. (e) CRISPR-seq performed on predicted cut locus of zebrafish *nf1a* and *nf1b* to sensitively detect Cas9-mediated editing. Reference genome as well as the two most commonly altered reads are shown. The right panel is a heatmap with the frequency of reference and edited sequences in WT melanocyte and acral melanoma models. Variants displayed for both *nf1a* and *nf1b* are frameshift mutations in exon 1 leading to a predicted loss of function. * p-value < 0.05, ** p-value < 0.01, *** p-value < 0.001, **** p-value < 0.0001.

Extended Data Fig. 3: Acral melanoma model demonstrates overexpression and knockout of multiple driver genes.



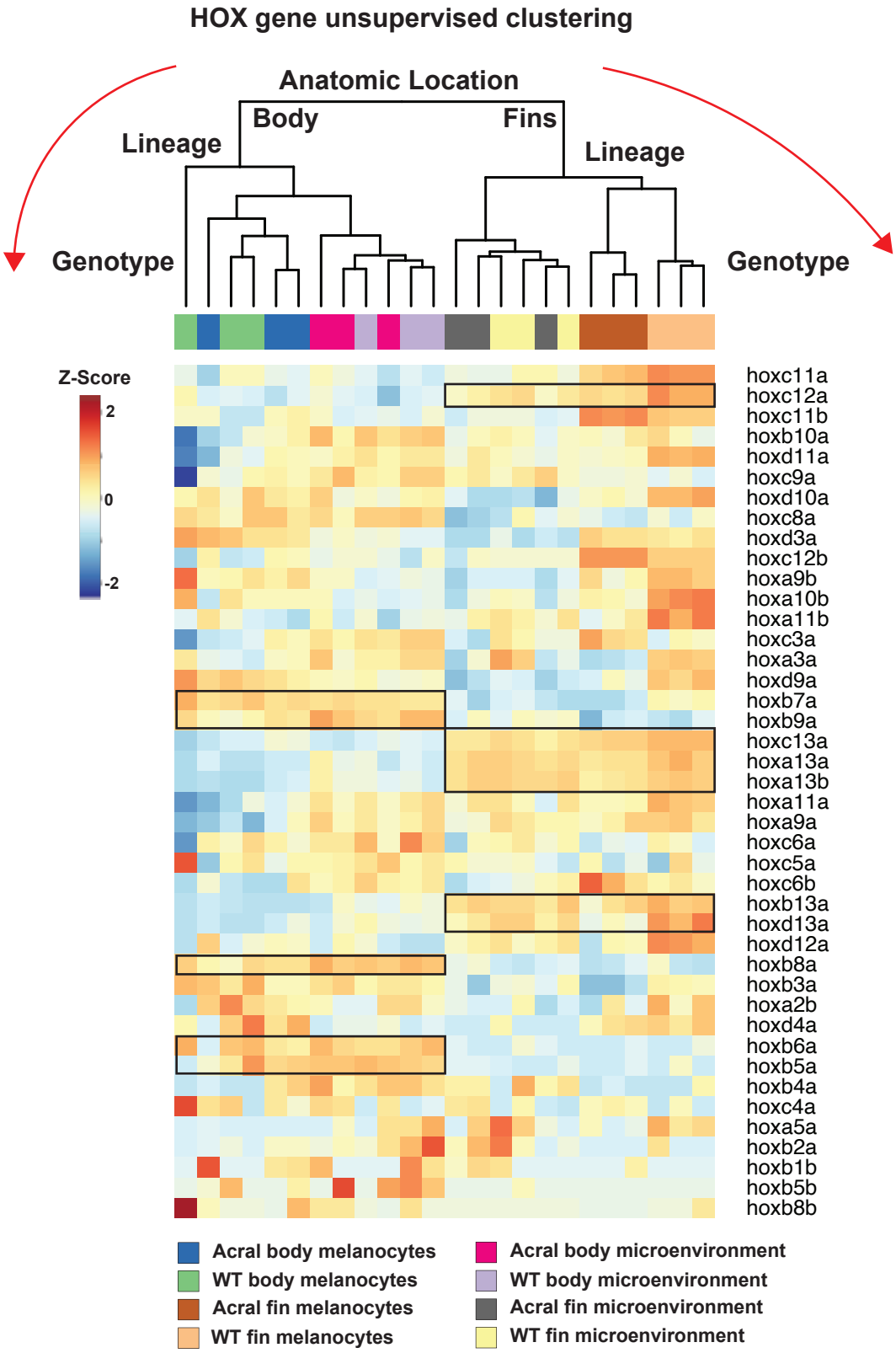
Extended Data Fig. 4: Validation of zebrafish melanocyte RNA-seq.
Related to Fig. 3. (a) Volcano plots comparing melanocytes to their microenvironment across locations and transgenic models. Melanocytes markers are labeled. Genes with FDR-adjusted p-value < 0.05 indicated in blue. (b) GSEA showing the list of the top pathways enriched in melanocytes compared to their microenvironment. Colors indicate p-value adjusted for FDR = 0.05. (c) Log normalized counts for the expression of transgenes across all samples. EGFP and *mitfa* expression are high in all melanocyte samples and CRKL, GAB2, TERT, and Cas9-mCherry expression is high only in acral melanoma model melanocytes. Error bars = SEM.

Extended Data Fig. 4: Validation of zebrafish melanocyte RNA-seq.



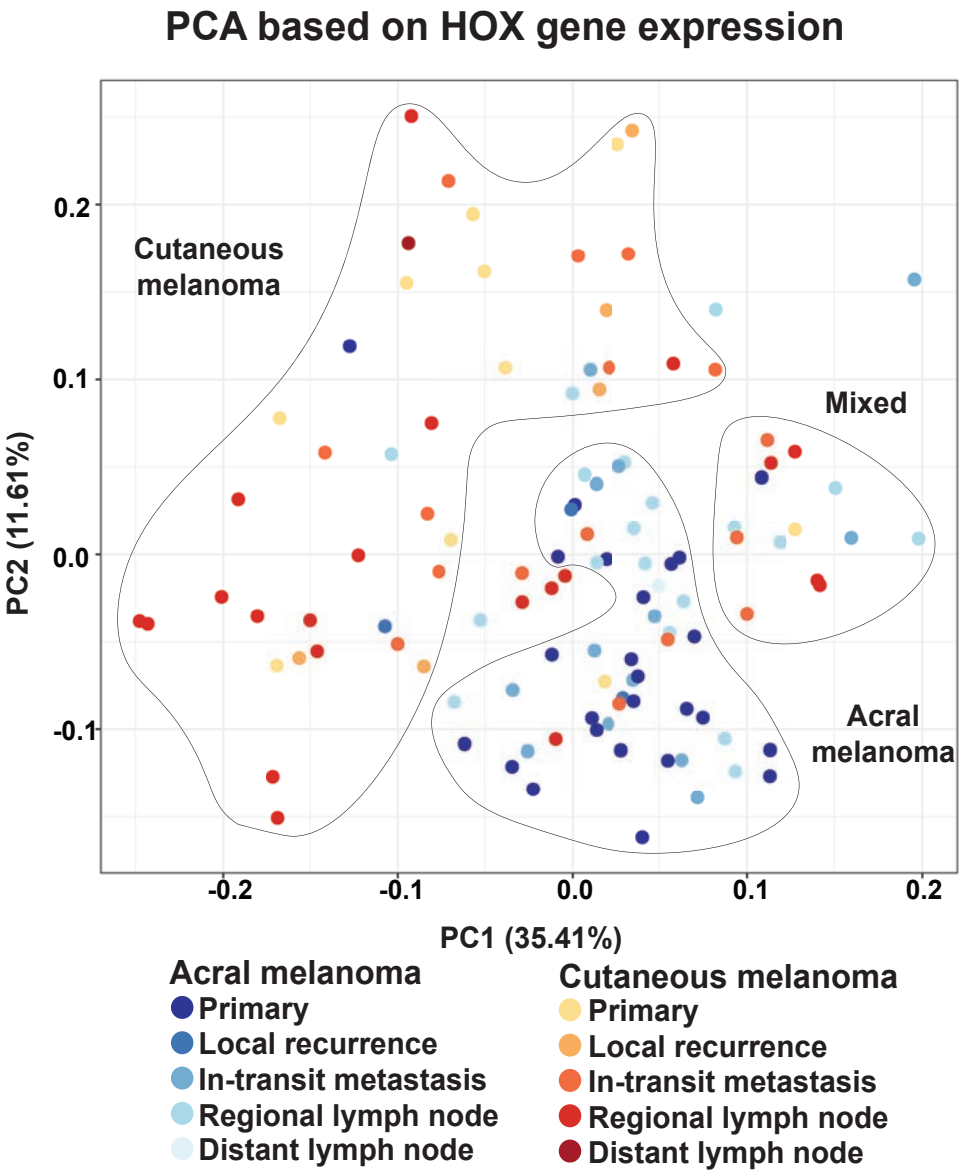
Extended Data Fig. 5: Zebrafish melanoma has a positional identity defined by *hox* genes.
Related to Fig. 3. Zebrafish *hox* genes detected by RNA-seq were used to perform unsupervised clustering and visualized with a heatmap. Samples clustered by anatomic location, then lineage, and then genotype in that order, indicating the *hox* genes predominantly associate with anatomic position. Particular hox genes that are differentially expressed across all body and fin samples are indicated with a black box.

Extended Data Fig. 5: Zebrafish melanoma has a positional identity defined by *hox* genes.



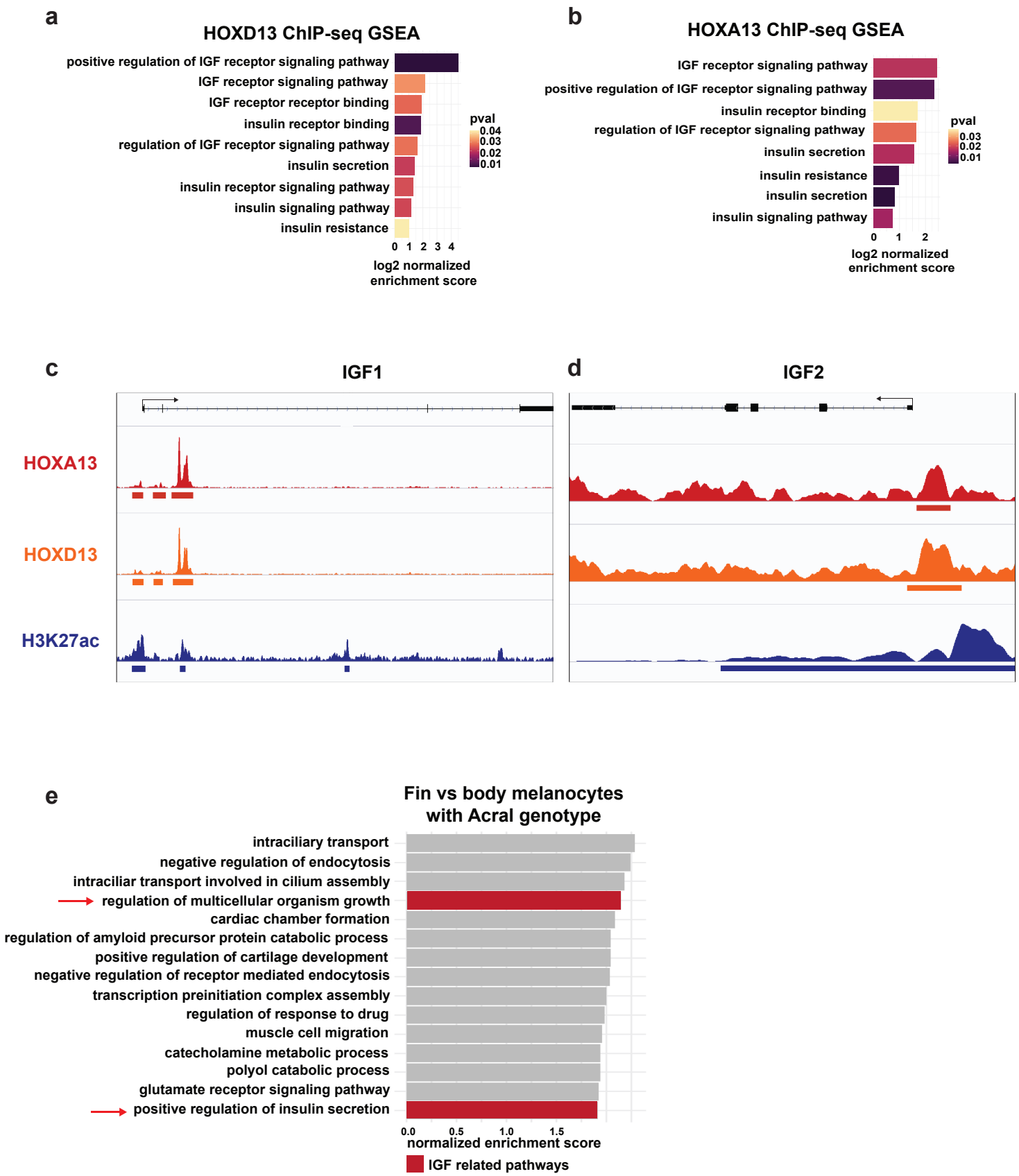
Extended Data Fig. 6: Human melanoma has a positional identity defined by HOX genes. Related to Fig. 4. PCA of all samples based on the expression of just HOX genes. Color indicates the combination of melanoma subtype and disease stage.

Extended Data Fig. 6: Human melanoma has a positional identity defined by HOX genes.



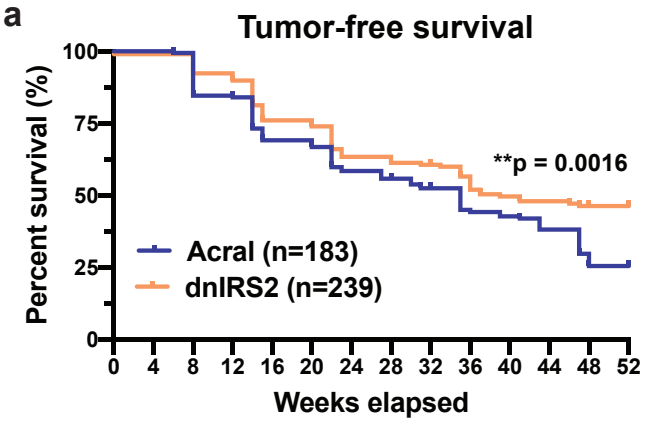
Extended Data Fig. 7: HOX13 regulates insulin/IGF signaling.
Related to Fig. 5. (a-d) Analysis of HOXA13, HOXD13, H3K27ac ChIP-seq data analyzed from Sheth et al., 2016[35] performed on developing limb buds from E11.5 mouse embryos. (a-b) Histogram showing the significantly enriched insulin/IGF pathways regulated by HOXA13 and HOXD13. NES and p-value are indicated. (c-d) Integrated genome browser tracks for HOXA13 and HOXD13 ChIP-seq binding and H3K27 acetylation near transcription start site of IGF1 and IGF2. (e) GSEA pathway analysis for the top pathways enriched in fin melanocytes versus body melanocytes from the zebrafish acral melanoma model. IGF-related pathways indicated in red.

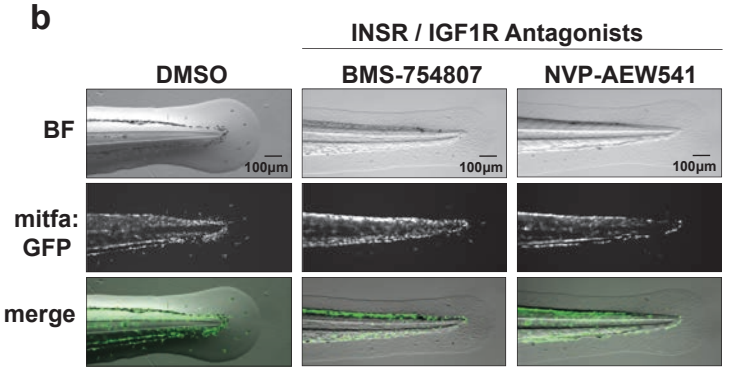
Extended Data Fig. 7: HOX13 regulates insulin/IGF signaling.

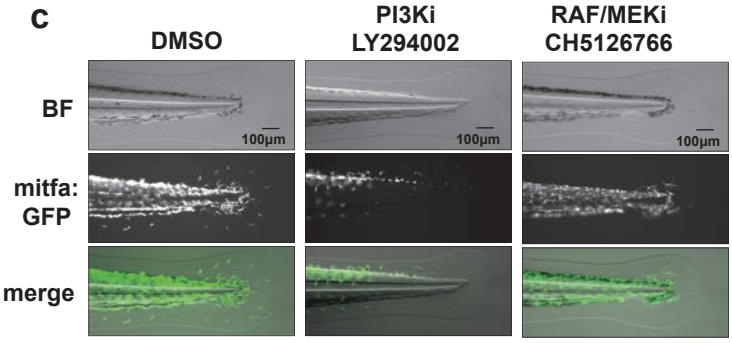


Extended Data Fig. 8: HOX13 synergizes with CRKL to drive acral melanoma through insulin/IGF signaling.
Related to Fig. 5. (a) Tumor-free survival comparing acral to acral dnIRS2 genotypes. Number of fish for each genotype indicated in figure key. P-values generated by log-rank Mantel-Cox test. (b-c) Representative images that correspond to histograms in Fig. 5i-j. Acral melanoma model imaged for melanocyte tailfin area at 3-days post-fertilization after indicated treatment with (b) INSR/IGF1R antagonists BMS-754807 at 7.5μM, NVP-AEW541 at 60μM, (c) PI3K inhibitor LY294002 at 15μM, and RAF/MEK inhibitor CH5126766 at 1μM. * p-value < 0.05, ** p-value < 0.01, *** p-value < 0.001, **** p-value < 0.0001.

Extended Data Fig. 8: HOX13 synergizes with CRKL to drive acral melanoma through insulin/IGF signaling.



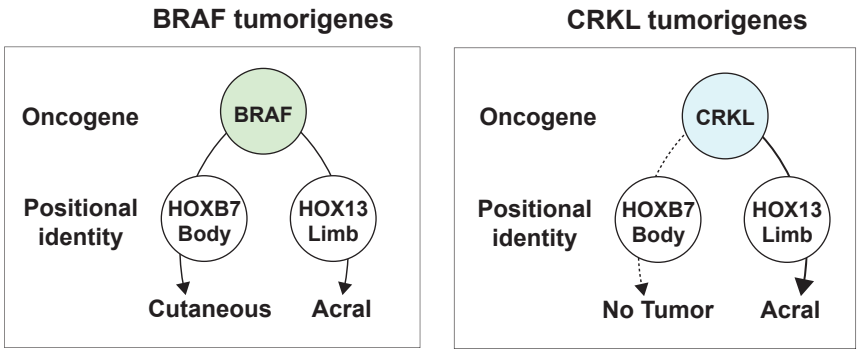




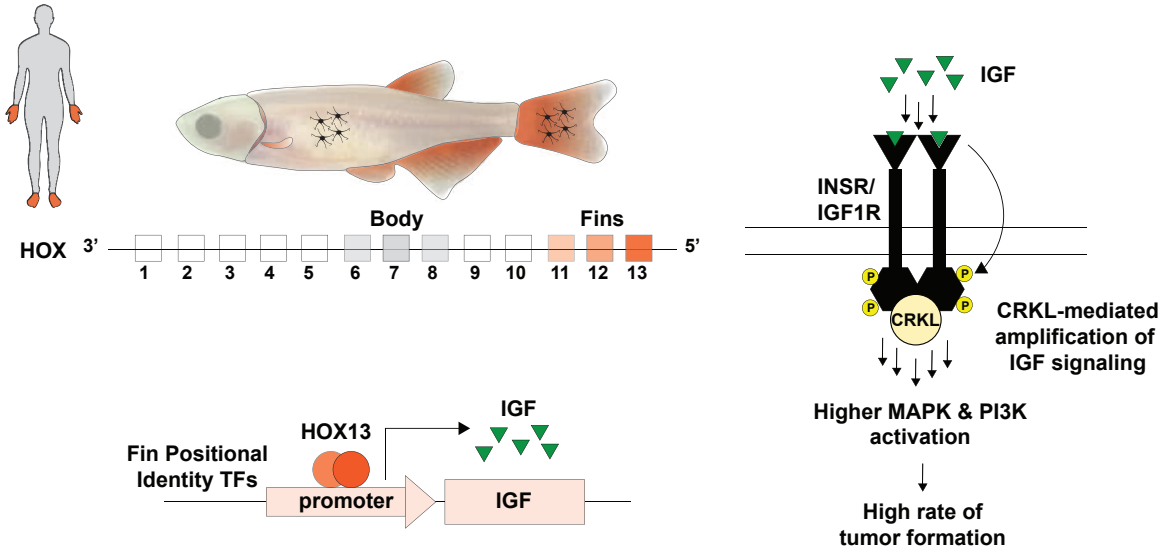
Extended Data Fig. 9: Anatomic position determines oncogenic specificity in melanoma. Related to Figure 5. (a) While BRAF mediates tumorigenesis irrespective of anatomic location, CRKL most efficiently mediates tumorigenesis in the limbs, specifically the hands and feet. (b) Melanocytes at different anatomic locations have different positional identities determined by HOX genes. While body melanocytes have higher expression of more anterior HOX genes like HOXB6-8, fin/limb melanocytes have higher expression of posterior HOX13 genes. HOX13 drives higher expression of IGF and increases vulnerability of fin melanocytes to CRKL-mediated transformation.

Extended Data Fig. 9: Anatomic position determines the oncogenic specificity in melanoma.

a



b



Supplemental Table 1: CAMP criteria for diagnosis in zebrafish models of melanoma
Related to Methods.

Supplemental Table 2: Primer Sequences
Related to Methods.

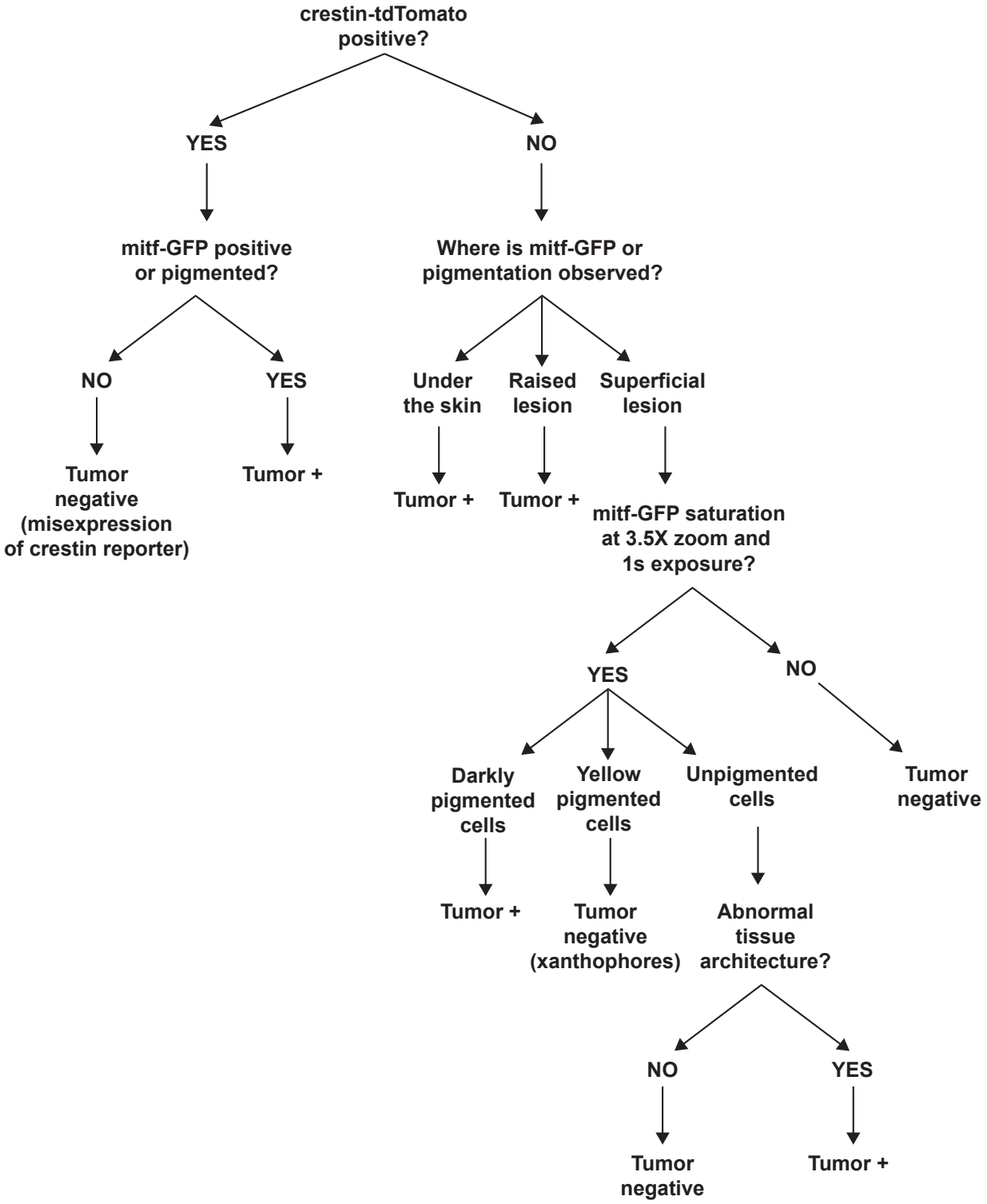
Supplemental Table 3: Human Acral vs Cutaneous RNA-seq
Related to Fig. 1 and Fig. 4. Differential expression and pathway analysis provided.

Supplemental Table 4: Zebrafish Fin vs Body RNA-seq
Related to Fig. 3 and Fig. 5. Differential expression and pathway analysis provided.

Supplemental Table 5: HOX13 ChIP-seq pathway analysis
Related to Fig. 5. Pathway analysis results provided.

Supplemental Table 1: CAMP criteria for diagnosis in zebrafish models of melanoma
Related to Fig.1. CAMP criteria: Crestin positivity, Anatomic localization, *Mitfa* expression, Pigmentation. Schematic shows the decision tree used in calling lesions a melanoma vs not a melanoma in both Acral and Cutaneous melanoma models. These criteria have been validated by performing histology on fish from both acral and cutaneous melanoma models.

Supplemental Table 1: CAMP Criteria for Diagnosis in Zebrafish Models of Melanoma
(**C**restin positivity, **A**natomic localization, **M**ITF expression, **P**igmentation)



Primer	**Sequence (5'-3')**
Acral genotype F1	GTTGAACGCAAGTTTGTACA
hsCRKL genotype R	TGATGTAGTGGGAGACCCGC
hsGAB2 genotype R	AGTATTCCAGAACATCTGGG
hsTERT genotype R	AGGCAGGACACCTGGCGGAA
Cas9-mCherry	CATGTGCACCTTGAAGCGCA
zebrafish nf1a CRISPR-seq F	TCGGGATCGCAAAAGTGATT
zebrafish nf1a CRISPR-seq R	CACCAAGCTCACATCTTCAA
zebrafish nf1b CRISPR-seq F	CACCATCTTCATCATCCTCCT
zebrafish nf1b CRISPR-seq F	ACTACTCTCTGTCCCGTGTC
zebrafish mitfa RT-PCR F	GCCCTATGGCCCTTCTCAC
zebrafish mitfa RT-PCR R	CATCCATGAACCCAAGAATGTCA
hsCRKL RT-PCR F&R	Sino Biological cat #HP101145
hsGAB2 RT-PCR F	GCGGCGACGTGGTGT
hsGAB2 RT-PCR R	CTTCCAGGCATAGCGCCTC
hsTERT RT-PCR F	GGAGCAAGTTGCAAAGCATTG
hsTERT RT-PCR R	TCCCACGACGTAGTCCATGTT
zebrafish beta-actin RT-PCR F	GCCAACAGAGAGAAGATGACAC
zebrafish beta-actin RT-PCR R	CAGAGAGAGCACAGCCTGG

Table S2. Related to Methods: primer sequence.

Methods

Resource Availability

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by Richard M. White (whiter@mskcc.org).

Materials Availability

- Plasmids generated in this study will be deposited to Addgene [*mitfa*:hsCRKL, *mitfa*:hsGAB2, *mitfa*:hsTERT, *mitfa*:Cas9-mCherry;zU6:*nf1a*:gRNA, *mitfa*:Cas9-mCherry;zU6:*nf1b*:gRNA, *mitfa*:dnIRS2-GFP, catalog number and unique identifiers pending] and will also be available by request.
- Zebrafish lines generated in this study will be made available through Zebrafish International Resource Center (ZIRC) or by request.
- Cell lines generated in this study are available by request.

Data and Code Availability

- All RNA-seq data generated and utilized in this study will be made publicly available.
- Human RNAseq data that support the findings of this study have been deposited in the European Genome-phenome Archive (EGA) and are available under study accession EGAS00001001552 (https://ega-archive.org/studies/EGAS00001001552). Patient sample information and differential expression tables are made available in the Supporting Information (Supplemental Table 3). Data is also available through Mendeley Data: http://dx.doi.org/10.17632/5f5mf5gw26.1.
- The raw data from the zebrafish fin versus body RNA-seq experiment will be made available via the NCBI GEO repository under identifier code GSE158538, with bulk RNA-seq counts and differential expression tables in the Supporting Information (Supplemental Table 4). Data is also available through Mendeley Data: http://dx.doi.org/10.17632/5f5mf5gw26.1. Samples are annotated as follows:
    - MBG -> WT body melanocyte
    - MBN -> WT body skin microenvironment
    - MFG -> WT fin melanocyte
    - MFN -> WT fin microenvironment
    - XBG -> Acral body melanocyte
    - XBN -> Acral body skin microenvironment
    - XFG -> Acral fin melanocyte
    - XFN -> Acral fin microenvironment

- Original source data for ChIP-seq analysis is from [35] and can be found at GEO GSE81358. The pathway analysis is made available in the Supporting Information (Supplemental Table 5). Data is also available through Mendeley Data: http://dx.doi.org/10.17632/5f5mf5gw26.1.
- Matlab script for tailfin image analysis is available upon request.

Experimental Model and Subject Details

Zebrafish

Zebrafish Husbandry

Fish stocks were kept under standard conditions at 28.5°*C* under 14:10 light:dark cycles, pH (7.4), and salinity-controlled conditions. Animals were fed standard zebrafish diet consisting of

brine shrimp followed by Zeigler pellets. The animal protocols described in this manuscript are approved from the Memorial Sloan Kettering Cancer Center (MSKCC) Institutional Animal Care and Use Committee (IACUC), protocol number 12-05-008. All anesthesia was performed using Tricaine-S (MS-222, Syndel USA, 712 Ferndale, WA) with a 4g/L, pH 7.0 stock. Both male and female zebrafish were utilized in equal proportions for all experiments utilizing adult fish. Sex determination in embryos is not possible at 3 days post fertilization (dpf). Embryos were collected from natural mating and incubated in E3 buffer (5mM NaCl, 0.17 mM KCl, 0.33 mM CaCl2, 0.33 mM MgSO4) at 28.5 °*C.*

Generating Zebrafish Transgenic Lines
One-cell-stage embryos of indicated genotype were injected with indicated plasmids at 5ng/uL per plasmid with tol2 mRNA at 20ng/uL. The total amount of plasmid used did not exceed 40ng. Embryos were grown to adulthood at 2 months post-fertilization (2mpf) and then screened for melanocyte rescue. To create stable lines, these fish were outcrossed with casper or BRAF/p53 caspers fish for two generations (F1 and F2) before in-crossing for F3 stable line generation. Monitoring for tumor-free survival began at 2mpf and ended at 1 year. For imaging of tailfin melanocytes, embryos at 3dpf were utilized. Fish utilized for FACS and RNA-seq were 6mpf.

Zebrafish Transgenic Lines
Zebrafish strains used in these studies included casper (mitfa$^{-/-}$,mpv17$^{-/-}$)[50] and casper p53$^{-/-}$ with *mitfa*:hBRAF^{V600E} ("BRAF/p53 casper"). The following transgenic lines were all generated using one-cell injection of tol2 mRNA at 20ng/uL and indicated plasmids. The WT melanocyte model was generated by injecting casper fish with 5ng of MiniCoopR-eGFP and 25ng of "empty" plasmid, which is a plasmid backbone with no transgene. This was used as filler to ensure all conditions were injected with the same total amount of DNA. The Cutaneous melanoma model was generated by injecting BRAF/p53 casper fish with 5ng of MiniCoopR-eGFP and 25ng of "empty" plasmid. The Acral melanoma model was generated by injecting casper fish with 5ng of each of the following plasmids: MiniCoopR-eGFP, *mitfa*:hsCRKL, *mitfa*:hsGAB2, *mitfa*:hsTERT *mitfa*:Cas9-mCherry;zU6-*nf1a*-gRNA, of mitfa:Cas9-mCherry;zU6-*nf1b*-gRNA. For transgenic models evaluating the role of a single transgene (CRKL, GAB2, and TERT), 5ng of MiniCoopR-eGFP and 25ng of the indicated transgene was used. The NF1 KO melanoma model was generated by injecting 5ng of MiniCoopR-eGFP and 12.5ng of *mitfa*:Cas9-mCherry;zU6-*NF1a*-gRNA and 12.5ng of mitfa:Cas9-mCherry;zU6-*NF1b*-gRNA. For transgenic models evaluating the removal of a given transgene from Acral melanoma model, 5ng was used for all indicated plasmids. For transgenic models evaluating the effect of dnIRS2 on the Acral melanoma model, 5ng was used for all indicated plasmids. *crestin*:tdTomato reporter was also utilized to aid in identifying tumors to monitor tumor-free survival[51,52]. Genotypes were regularly monitored by PCR.

<u>Humans</u>
MSK-IMPACT DNA Sequencing
Genomic analysis of human acral and cutaneous melanoma patients came from a broad institutional effort to sequence cancer patients that arrive at MSK and make available for the clinical and scientific community[14,16]. Patients analyzed were clearly indicated as having either acral or cutaneous melanoma. Melanoma patients without subtype designations or with other subtypes (i.e. uveal or mucosal) were not included. Patients of all sexes, ethnicities, and ages were included for analysis.

WES and WGS of an Individual Acral Melanoma Patient
The female patient was enrolled at Memorial Sloan-Kettering Cancer Center (MSKCC) and consented on a protocol approved by MSKCC and analyzed as part of a previous acral

melanoma study[17]. Samples were obtained in accordance with standard biopsy or surgical procedures.

RNA-sequencing of Human Melanoma Patients
For human RNA-seq sequencing analysis, fresh-frozen tissue samples were obtained from the biospecimen bank of Melanoma Institute Australia (MIA) and all samples were accrued prospectively with written informed patient consent. The protocol for the study was approved by the Sydney Local Health District Ethics Committee (Protocol No X15-0454 (prev X11-0289) & HREC/11/RPAH/444 and Protocol No X17-0312 (prev X11-0023) & HREC/11/RPAH/32) and cases were also approved by institutional ethics committees of Melanoma Institute of Australia and QIMR Berghofer Medical Research Institute (HREC approval P452 & P2274). Details of patient samples can be found in Supplemental Table 2.

Cell Lines and Virus Preparation
WM3918 melanoma cell line was grown with Dulbecco's Modified Eagle Medium (Gibco #11965) supplemented with 10% FBS (Seradigm), 1X penicillin/streptomycin/glutamine (Gibco #10378016). To make CRKL-overexpressing cell line, CRKL Retrovirus was produced by transfecting 293T packaging cells with CRKL-pWzl, VSVg and pCL-Ampho plasmids. WM3918 human melanoma cell line was transduced with CRKL retroviral particles plus polybrene and then selected with10 µg/mL blasticidin. CRKL-pWzl was a kind gift from professor William Hahn.

Method Details
Acral Enrichment Score Analysis of MSK-IMPACT DNA Sequencing
Only melanoma patients clearly indicated as having either acral or cutaneous melanoma were analyzed. We selected the 13 most frequently observed copy number alterations and 11 most frequently mutated genes in acral melanoma and compare to the frequency observed in cutaneous melanoma. We also included well described melanoma drivers, such as deletion of NF1[53] and mutations in PTEN[54]. To calculate the acral enrichment score we used the following equation log2(Acral frequency/Cutaneous frequency). Statistical differences in the frequency of copy number alterations and mutations were determined using a Fisher's exact test.

WES/WGS Analysis of an Acral Melanoma Patient
Tissue was selected by the pathologist to limit the amount of necrotic tissue and adjacent normal tissue was collected for DNA extraction of germline DNA. An H&E slide was made to confirm normal or malignant tissue and <50% necrosis, which was reviewed by the pathologist. The patient had both a primary tumor sample, which was formalin-fixed and paraffin embedded (FFPE) and an in-transit metastasis, which was fresh frozen by placing in a vial and submerging into a liquid nitrogen container. Paired tumor/normal whole-exome and whole-genome sequencing (WES and WGS, respectively) libraries were constructed and sequenced on the Illumina HiSeq using V3 reagents. WES and WGS data was aligned to the genome with BWA [55,56] and mutations called using MuTect[57]. Coverage-based copy-number analysis was performed using custom scripts for WGS and using the allele-specific method FACETS for the WES data[58].

RNA-seq Analysis of Human Melanoma Patients
Fresh-frozen tumor RNA was extracted using the AllPrep® DNA/RNA/miRNA Universal kit (Qiagen #80224) and were quantified using the Qubit® RNA HS Assay (Q32852, Life Technologies). The TruSeq RNA library prep kit (Illumina, San Diego, California, USA) was used to prepare libraries from RNA and these were sequenced with 100bp paired-end reads using Illumina HiSeq2000 or HiSeq2500 platforms. RNA-seq reads were trimmed for adapter sequences using Cutadapt (version 1.9) and aligned with STAR (version 2.5.2a)[59] to the

GRCh37 assembly using the gene, transcript, and exon features of Ensembl (release 70) gene model. Gene expression counts were estimated using RSEM (version 1.2.30)[60]. Differential expression was calculated with DESeq2[61] using the output of the quantMode and GeneCounts feature of STAR. The vst function was used to generate log2 transformed normalized counts. Differential expression for Figure 5D-F, including HOMER analysis used comparison of all acral vs cutaneous melanoma samples. Differential expression used for pathway analysis for Figure 5B-C used comparison of all acral vs cutaneous melanoma samples normalized by specimen type (i.e. primary tumor or lymph node metastasis). Pathway and Gene Ontology (GO) analysis were performed with GSEA using FGSEA-multilevel[62]. Known motif analysis was performed with the HOMER [63] function findMotifs.pl, using the human genome (GRCh37) and searching for motifs of lengths 8, 10, and 16 within ± 500bp of the TSS of differentially expressed genes. Motifs were annotated using JASPAR [64].

Co-Occurrence of CRKL, GAB2, NF1, and TERT in Cancer
The frequency of alterations occurring across multiple human cancer types was analyzed in cBioPortal[13,15] looking at the following TCGA studies: Firehose Legacy [65], PanCancer Atlas [66], Cell 2017[67-70], Nature 2014[71-73], Nature 2012[74-76], Cell 2015[53,77,78], Nature 2008[79], Nature 2015[80], Nature 2013[81,82], Cancer Cell 2014[83], NEJM 2013[84], Nature 2011[85], Cell 2013[86], Cell 2014[87]. The odds ratio for co-occurrence of alterations was calculated and used to generate a p-value adjusted for FDR = 0.05.

Plasmid Construction
The following plasmids were constructed using the Gateway Tol2kit.

mitfa:hsCRKL-394
mitfa:hsGAB2-394
mitfa:hsTERT-394
mitfa:Cas9-mCherry;zU6:*NF1a*:gRNA-394
mitfa:Cas9-mCherry;zU6:*NF1b*:gRNA-394
mitfa:dnIRS2-GFP-394

cDNAs in tol2-compatible pENRT223.1 plasmids were ordered through Horizon Discovery for the following genes.

hsCRKL (Clone ID: 100000145)
hsGAB2 (Clone ID: 56810)
hsTERT (Clone ID: 100061944)

gRNA sequences

nf1a: GGCGCACAAGCCCGTGGAAT
nf1b: GGCGCAGAAGCCCGTGGAGT

Zebrafish dnIRS2-GFP was a generous gift from [41] and then further cloned into pENTR/D-TOPO vector.

Zebrafish Genotyping by PCR
Tail clips from adult zebrafish were placed in microcentrifuge tubes or thermal cycler plates containing 50 ml of 50 mM NaOH. Samples were boiled at 95ºC for 30min, then cooled down with 5 ml of 1M Tris-HCL (pH=8.0). A 1:10 dilution of the supernatant was used in PCR. DNA was PCR amplified with Promega GoTag green mastermix (Promega #M7123). PCR amplicon sizes for CRKL, GAB2, TERT, and Cas9 are 221bp, 125bp, 296bp, and 207bp. All transgenes have the same forward primer. GTTGAACGCAAGTTTGTACA going off the *mitfa* promoter Primer sequences are listed below.

CRKL: TGATGTAGTGGGAGACCCGC
GAB2: AGTATTCCAGAACATCTGGG
TERT: AGGCAGGACACCTGGCGGAA

Cas9-mCherry: CATGTGCACCTTGAAGCGCA

CRISPR-seq
Genomic DNA of transgenic zebrafish was isolated via tail clip genotyping described above. DNA was PCR amplified with Phusion polymerase (NEB #M0531S), run on an agarose gel, and then gel purified using NucleoSpin® Gel & PCR Clean-up Midi (Takara # 740986.20). for deep sequencing using the CRISPR-seq platform [88]. Primers used are listed below. Sequencing data was aligned to the zebrafish genome (GRCz10) and analyzed with CrispRVariantsLite version 1.2 [89].

nf1a: TCGGGATCGCAAAAGTGATT
CACCAAGCTCACATCTTCAA
nf1b: CACCATCTTCATCATCCTCCT
ACTACTCTCTGTCCCGTGTC

RT-PCR Validation of Transgene Expression
Wildtype melanocyte skin, Acral melanoma model tumor, and Cutaneous melanoma model tumor tissue were dissected from zebrafish and RNA was isolated with the Zymo Quick RNA Miniprep kit (#R1050) using kit protocol. cDNA was synthesized using SuperScript III First Strand Synthesis System (Life Technologies #18080-400) following the kit protocol. RT-PCR was performed using iQ Sybr Green Supermix (Biorad #1708882). RT-PCR primer sequences are listed below:

mitfa: GCCCTATGGCCCTTCTCAC
CATCCATGAACCCAAGAATGTCA
hsGAB2: GCGGCGACGTGGTGT
CTTCCAGGCATAGCGCCTC
hsCRKL: Sino Biological qPCR primer pairs (HP101145)
hsTERT: GGAGCAAGTTGCAAAGCATTG
TCCCACGACGTAGTCCATGTT
beta-actin: GCCAACAGAGAGAAGATGACAC
CAGAGAGAGCACAGCCTGG

Western Blot
Zebrafish lysates were collected by sonication in RIPA buffer (Thermo #89901) with 1X Halt Protease and Phosphatase Inhibitor Cocktail (Thermo #78441) followed by centrifugation (14,000rpm for 10min at 4ºC) and collection of the supernatant. Protein concentration was quantified by Bradford (Sigma B6916-500mL) according to manufacturer's protocol. Samples were mixed with 6X reducing loading buffer (Boston BioProducts #BP-788 111R) and denatured at 95ºC for 10 minutes. Samples were run on a Mini PROTEAN TGX gel (BioRad) and transferred using Turbo Mini Nitrocellulose Transfer Pack (Bio-Rad, catalog #1704158). Membranes were blocked with 5% nonfat dry milk in TBST (1X TBS + 0.1% Tween 20) for 1 hour before incubation with primary antibody in PBS overnight at 4ºC. Membranes were washed with TBST and incubated with secondary antibody in 5% nonfat dry milk for 1 hour at room temperature. Membranes were washed with TBST and developed with ECL (Amersham, RPN2109) using an Amersham Imager 600 (GE) or chemiluminescence film.

WM3918 parental and WM3918-CRKL cell lines were serum starved for 24 hours and then stimulated with 20% FBS (Seradigm) for 10 minutes before collection. Cells were washed with cold PBS and lysed in RIPA buffer (Pierce #89901) plus phosphatase and protease inhibitors (Thermo Scientific #1861277, #1861278). Lysates were cleared by centrifugation at 14000rpm at 4°C and quantified using BCA method (Pierce #23224). Samples were prepared using LDS+Reducing agent Novex buffers (Invitrogen #NP0008, #NP0009). 10 to 20µg of lysates

were loaded and run on NuPageTM 4-12% Bis-Tris gels (ThermoFisher #NP0321BOX) followed by transfer to nitrocellulose membranes (Biorad #1620233). Membranes were incubated over night with the indicated antibodies, washed and incubated again for 45 minutes with anti-rabbit or anti-mouse secondary antibodies. Detection was performed using Immobilion Western (Millipore #WBKLS0500).

Primary antibodies are: CRKL (32H4) #3182S (Cell Signaling Technology) (used for fish lysates), CRKL (B-1) sc-365092 Lot#B0819 (Santa Cruz Biotechnology) (used for WM3918 cell lines), GAB2 (26B6) #3239 (Cell Signaling Technology), IGF-1R beta (D23H3) #9750S (Cell Signaling Technology), phospho-IGF-1R beta Y1135 (DA7A8) #3918S (Cell Signaling Technology) and β-Actin (13E5) #4970S (Cell Signaling Technology). All the primary antibodies were used at 1:1000 dilution except for β-Actin that was used at 1:4000 dilution.

Phospho-RTK array
Human Phospho-RTK arrays (R&D Systems, ARY001B) were used to detect activated RTKs according to manufacturer's instructions. WM3918 parental and WM3918-CRKL cell lines were serum starved for 24 hours and then stimulated with 20% FBS (Seradigm) for 10 minutes before collection. Cells were washed with cold PBS and lysed using the provided lysis buffer plus phosphatase and protease inhibitors. Lysates were cleared by centrifugation at 14000rpm at 4°C and quantified using BCA method (Pierce #23224). 200 µg lysates were incubated on membranes overnight. Membranes were subsequently washed and exposed to chemi-luminescent reagent (Millipore #WBKLS0500).

Zebrafish ImagingTumor-free Survival Curves
Zebrafish embryos were injected at the one-cell stage with indicated plasmids. *crestin*:tdTomato reporter was also utilized to aid in identifying tumors to monitor tumor-free survival[51,52]. All injected embryos were grown to adulthood and then screened for melanocyte rescue at 2mpf based on a combination of GFP fluorescence and pigmentation. Zebrafish were regularly monitored every 2-4 weeks for the development of new tumors. Tumors were called based on the CAMP criteria outlined in Table S1. Kaplan Meir curves were generated and analyzed in Prism version 8. Statistical differences in tumor-free survival was determined by log-rank Mantel-Cox test.

Analysis of Anatomic Distribution of Tumors
During monitoring of tumor-free survival, the anatomic location (head vs body vs fins) of the tumor is noted. If a fish has either >2 tumors or a tumor that encroaches on the border of multiple anatomic sites, all tumors and anatomic sites are counted. To compare the anatomic distribution of tumors, a Chi-squared test was performed using Prism software version 8. Ternary diagrams were generated using ternaryplot.com.

Histology of Zebrafish Samples
Zebrafish were sacrificed using ice-cold water. The head and tail were them from the body separated via dissection with a clean razor. The sample of interest was then placed in 4%PFA (Santa Cruz #30525-89-4) for 72 hours on a shaker at 4ºC and delivered to Histowiz who performed all H&E and IHC staining.

IF of Zebrafish Melanoma Samples
Unstained histology slides of zebrafish samples were deparaffinized using 2x10 minutes xylene, 4 minutes 100% ethanol, 1 minute 95% ethanol, 1 minute 70% ethanol, 1 minute 50% ethanol, and then rinsed with water. Antigen retrieval was achieved by placing slides in 10mM sodium citrate pH = 6.2 by heating in the microwave for approximately 2 minutes until the solution

begins to bubble. The slides then sat for 5 minutes and then were microwaved again for 40 seconds. The slides cooled down to room temperature for 30 minutes in the same buffer. The slides were blocked in a blocking solution consisting of 5% donkey serum (Millipore Sigma #S30-M), 1% BSA, 0.4% Triton X-100 (Fisher # PI85111) in PBS. 50uL of blocking buffer was added per section. The blocking solution incubated at room temperature for 1 hour. Primary antibody were diluted at 1:100 in blocking solution and incubated overnight at 4°C. Slides were washed 3 x 5 minutes in PBS and then applied with secondary antibody solution 1:250 in blocking solution. Hoescht (Fisher # H3570) was added to be 1:1000 dilution. The slides were washed again 3 x 5 minutes in PBS. Coverslips were mounted onto the slides using Vectashield Vibrance Antifade Mounting Medium (Vector Laboratories #H-1800-2). Slides were then imaged on a confocal microscope. Primary antibodies used were Sox10 (GeneTex #GTX128374) and GFP (Abcam #5450). Secondary antibodies used were Alexa-anti Rabbit 594 (Cell Signaling Technology #8889) and Alexa-anti Goat 488 (ThermoFisher #A32814).

Pharmacologic Treatment of Zebrafish Embryos
Zebrafish embryos were collected at 1dpf and placed in a 40-micron cell strainer (Thermo Fisher # 08-771-2) in a 6-well dish (Fisher #08-772-1B) in 6mL of E3 water. 20 embryos were used per well. Zebrafish embryos were treated with the indicated compounds: insulin/IGF1 receptor antagonists BMS-754807 (Sigma-Aldrich #BM0003-5MG) at 7.5μM and NVP-AEW541(Selleckchem # S1034) at 60μM, PI3K inhibitor LY294002 (Sigma-Aldrich #L9908-1MG) at 15μM and RAF/MEK inhibitor CH5126766 (Selleckchem # S7170) at 1μM. Compound stocks were at 1000X in DMSO. Treatment started at 1dpf and reapplied at 2dpf.

Measurement of Zebrafish Melanocyte Cell Area in Tailfin
The extent of melanocyte area in the tailfin was calculated as the area of the melanophore covering the tailfin mesenchyme. This was quantified by using MATLAB to perform background subtraction of autofluorescence and then used FIJI to threshold on GFP (from MiniCoopR-eGFP) intensity to highlight the pixels that represent melanocytes in each image. All zebrafish were imaged at 3dpf. If treated pharmacologically, treatment started at 1dpf at the indicated concentration for 48 hours and imaged at 3dpf.

Flow *Cytometry of Zebrafish Melanocytes*
Zebrafish Embryos
Approximately 250 5dpf embryos were euthanized via tricaine and transferred to an Eppendorf tube (Eppendorf # 022431021) and spun at 500G x 3 minutes at room temperature to remove E3 supernatant. 400uL of trypsin (Invitrogen # 25200-114) was added and incubated for 20 minutes at 28°C. Every 5 minutes an RNase-free disposable pellet pestle (Fisher # 12-141-364) was used to mash the embryos for 2-3 minutes into a single cell solution. After digestion was complete, 500uL of phenol-free Dulbecco's Modified Eagle Medium (ThermoFisher # 21063029) with 10% FBS (Seradigm) (DMEM10) was added to each sample. The samples were centrifuged 500G x 5 minutes at room temperature and supernatant was gently aspirated. Calcein red stock (Cayman Chemicals # 20632) at 1mM was diluted 1:4000 to desired concentration of 250nM in phenol-free DMEM10 and 500uL was added to embryo pellet and incubated for 30 minutes at 28°C to be used to identify viable cells. Samples were centrifuged again 500G x 5 minutes, supernatant was removed, and pellet was resuspended in phenol-free DMEM with 2% FBS (Seradigm) (DMEM2). Samples were filtered through 40-micron cell-strainer (Thermo Fisher # 08-771-2) two times and then placed in a flow cytometry tube (Fisher # 08-771-23). 1mL of DMEM2 and 1uL of 1000X DAPI (Sigma-Aldrich #D9542-10MG) was added and analyzed on the FACS sorter (BD FACSAria). No color and single-color controls were used to gate for GFP+/Calcein Red+ double positive cells to calculate the frequency of

viable melanocytes from the bulk cell suspension. The reporter frequency of viable melanocytes in Figure S3B were calculated on the FACS sorter.

Zebrafish Adults
Zebrafish were euthanized via placement in ice-cold water and then dissected to separate body skin and all fins. A clean razor is used to dice the sample into small pieces that can fit through a wide bore p1000 tip (ThermoFisher #2069G). Samples were placed into a 15mL Falcon tube (Fisher 14-959-49D) with 3mL of 1X PBS (Invitrogen 14190-250) and 187.5uL of 2.5mg/mL liberase (Millipore Sigma # 05401020001) and incubated at room temperature for 30 minutes on a shaker to gently keep tissue in suspension. At 15 minutes, a wide bore p1000 tip was used to pipette up and down gently for 3 minutes to dissociate the tissue. After the 30 minutes incubation in liberase at room temperature, 250uL of FBS (Seradigm) was added and then another 3 minutes of pipetting up and down using a wide bore p1000 tip was performed. Cells were then filtered through a 40-micron cell strainer (Thermo Fisher # 08-771-2) into a 50mL conical. Samples were spun at 500G x 5 minutes at 4 °*C* and the supernatant was carefully aspirated using a Pasteur pipette (Fisher #13-678-20D) with low vacuum suction. The pellet was then resuspended in 500uL of PBS with 5% FBS (Seradigm) to be used as flow buffer. Draw up cells in a p1000 tip with a regular bore size and then fit on 40um filter onto the tip and place in a flow cytometry tube (Fisher # 08-771-23). Add 0.5uL of 1000X DAPI (Sigma-Aldrich #D9542-10MG) and samples were placed on ice. Samples were then FACS sorted (BD FACSAria) for GFP-positive signal and gated based on a GFP-negative control. Each biological replicate represents the pooling of 2 males and 2 female adult zebrafish of the indicated genotype. The reporter frequency of viable melanocytes in Figure 2H were calculated on the FACS sorter.

Preparation of Zebrafish Sample for RNA-seq
Zebrafish Embryos
Samples were FACS sorted as described above into 200uL of Trizol (Invitrogen #15596026) in RNase-free LoBind Eppendorf tubes (Eppendorf # 022431021) and snap-frozen by placing on dry ice. Samples were shipped to Genewiz (South Plainfield, NJ), where the RNA isolation and RNA-sequencing was performed.

Zebrafish Adult Body Skin and Fins
Samples were FACS sorted as described above into 750uL of Trizol LS (Invitrogen # 10296010) in RNase-free LoBind Eppendorf tubes (Eppendorf # 022431021) and snap-frozen by placing on dry ice. RNA isolation was performed by following Trizol LS protocol. Glycogen (Millipore Sigma #10901393001 was used to co-precipitate the RNA. RNA quality and quantity was measured by Bioanalyzer (Agilent). RNA samples were shipped to Genewiz (South Plainfield, NJ), where RNA-sequencing was performed.

Analysis of Zebrafish RNA-seq
RNA-sequencing was performed using SMART-seq v4 Ultra Low Input RNA Kit (Clonetech). Libraries were constructed using Illumina Nexterna XT kit and were analyzed for concentration by Qubit and for size distribution by Agilent Bioanalyzer. Paired-end sequencing was performed on Illumina HiSeq 2500. After quality control with FASTQC (Babraham Bioinformatics) and trimming with TRIMMOMATIC [90] when necessary, reads were aligned to GRCz10 (Ensembl version 81) using STAR [59], with quality control via SeQC [91]. Differential expression was calculated with DESeq2 [61] using the output of the --quantMode GeneCounts feature of STAR. The rlog function was used to generate log2 transformed normalized counts Pathway and Gene Ontology (GO) analysis were performed with GSEA using FGSEA-multilevel[62]. Ortholog mapping between zebrafish and human was performed with DIOPT[92]. Only orthologs with a DIOPT score greater than 6 were used for GSEA and heatmap generation. In cases of more

than one zebrafish ortholog of a given human gene, the zebrafish gene with the highest average expression was selected.

ChIP-seq Pathway Analysis
ChIP-seq analysis was performed from raw reads of publicly available GEO datasets from Sheth et al, 2016 (GSE81358). Paired end reads from the following samples were processed: GSM2151011 (H3K27ac), GSM2151013 (HOXA13), GSM2151014 (HOXD13), GSM2151016 (H3K27ac input DNA), GSM2151017 (HOXA13 input DNA), GSM2151018 (HOXD13 input DNA). SRA files were converted to FASTQ format using the SRA toolkit. Raw reads were checked for sequence quality, adapter content, overrepresented sequences and Kmer content using FASTQC (Babraham Bioinformatics). Adapters and low quality sequences were filtered out using Trimmomatic[90] and filtered reads were mapped to the mm9 genome using Bowtie2[93]. Mapped reads were analyzed using Samtools[94] and the sorted bam files were processed using Deeptools[95] to generate input normalized bigwig files. MACS v1.4[96] was used to perform peak calling using input DNA for each antibody as control. Pathway analysis of enriched peaks was performed using Cistrome-GO [36].

Quantification and Statistical Analysis

Statistical comparisons were performed with the aid of GraphPad PRISM 8.4.3 and the statistical details including sample size can be found in the figure legends. A p value of >0.05 is not considered statistically significant. * indicates p<0.05, ** indicates p<0.01, *** indicates p<0.001 and **** indicates p<0.0001.

RNA-seq and ChIP-seq analysis was analyzed in R version 4.0.2 and R Studio version 1.2.5033.

ChIP-seq data was visualized in IGB (BioViz)[97].

Image analysis performed in Matlab version Update 4 (9.6.0.1150989) and FIJI version 2.0.0.

References

1 Tang, J. *et al.* The genomic landscapes of individual melanocytes from human skin. *Nature*, doi:10.1038/s41586-020-2785-8 (2020).

2 Fowler, J. C. *et al.* Selection of oncogenic mutant clones in normal human skin varies with body site. *Cancer Discovery*, doi:10.1158/2159-8290.Cd-20-1092 (2020).

3 Reed, R. *New concepts in surgical pathology of the skin*. 89-90 (Wiley, 1976).

4 Shah, N. & Sukumar, S. The Hox genes and their roles in oncogenesis. *Nat Rev Cancer* **10**, 361-371, doi:10.1038/nrc2826 (2010).

5 Wang, K. C., Helms, J. A. & Chang, H. Y. Regeneration, repair and remembering identity: the three Rs of Hox gene expression. *Trends Cell Biol* **19**, 268-275, doi:10.1016/j.tcb.2009.03.007 (2009).

6 Hayward, N. K. *et al.* Whole-genome landscapes of major melanoma subtypes. *Nature* **545**, 175-180, doi:10.1038/nature22071 (2017).

7 Petrelli, F. *et al.* Prognostic Survival Associated With Left-Sided vs Right-Sided Colon Cancer: A Systematic Review and Meta-analysis. *JAMA Oncol* **3**, 211-219, doi:10.1001/jamaoncol.2016.4227 (2017).

8 Rabbie, R., Ferguson, P., Molina-Aguilar, C., Adams, D. J. & Robles-Espinoza, C. D. Melanoma subtypes: genomic profiles, prognostic molecular markers and therapeutic possibilities. *J Pathol* **247**, 539-551, doi:10.1002/path.5213 (2019).

9 Roychoudhuri, R., Putcha, V. & Moller, H. Cancer and laterality: a study of the five major paired organs (UK). *Cancer Causes Control* **17**, 655-662, doi:10.1007/s10552-005-0615-9 (2006).

10 Bello, D. M. *et al.* Prognosis of Acral Melanoma: A Series of 281 Patients. *Annals of Surgical Oncology* **20**, 3618-3625, doi:10.1245/s10434-013-3089-0 (2013).

11 Klemen, N. D. *et al.* Survival after checkpoint inhibitors for metastatic acral, mucosal and uveal melanoma. *J Immunother Cancer* **8**, doi:10.1136/jitc-2019-000341 (2020).

12 Shoushtari, A. N. *et al.* The efficacy of anti-PD-1 agents in acral and mucosal melanoma. *Cancer* **122**, 3354-3362, doi:10.1002/cncr.30259 (2016).

13 Cerami, E. *et al.* The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* **2**, 401-404, doi:10.1158/2159-8290.CD-12-0095 (2012).

14 Cheng, D. T. *et al.* Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets (MSK-IMPACT): A Hybridization Capture-Based Next-Generation Sequencing Clinical Assay for Solid Tumor Molecular Oncology. *J Mol Diagn* **17**, 251-264, doi:10.1016/j.jmoldx.2014.12.006 (2015).

15 Gao, J. *et al.* Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* **6**, pl1, doi:10.1126/scisignal.2004088 (2013).

16 Zehir, A. *et al.* Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nat Med* **23**, 703-713, doi:10.1038/nm.4333 (2017).

17 Liang, W. S. *et al.* Integrated genomic analyses reveal frequent TERT aberrations in acral melanoma. *Genome Res* **27**, 524-532, doi:10.1101/gr.213348.116 (2017).

18 Yeh, I. *et al.* Targeted Genomic Profiling of Acral Melanoma. *J Natl Cancer Inst* **111**, 1068-1077, doi:10.1093/jnci/djz005 (2019).

19 Luo, L. Y. & Hahn, W. C. Oncogenic Signaling Adaptor Proteins. *J Genet Genomics* **42**, 521-529, doi:10.1016/j.jgg.2015.09.001 (2015).

20 Bentires-Alj, M. *et al.* A role for the scaffolding adapter GAB2 in breast cancer. *Nat Med* **12**, 114-121, doi:10.1038/nm1341 (2006).

21 Cheung, H. W. *et al.* Amplification of CRKL induces transformation and epidermal growth factor receptor inhibitor resistance in human non-small cell lung cancers. *Cancer Discov* **1**, 608-625, doi:10.1158/2159-8290.CD-11-0046 (2011).

22 Hemmeryckx, B. *et al.* Crkl enhances leukemogenesis in BCR/ABL P190 transgenic mice. *Cancer Research* **61**, 1398-1405 (2001).

23 Kostrzewska-Poczekaj, M. *et al.* Copy number gains of the putative CRKL oncogene in laryngeal squamous cell carcinoma result in strong nuclear expression of the protein and influence cell proliferation and migration. *Sci Rep* **10**, 24, doi:10.1038/s41598-019-56870-5 (2020).

24 Chernoff, K. A. *et al.* GAB2 amplifications refine molecular classification of melanoma. *Clin Cancer Res* **15**, 4288-4291, doi:10.1158/1078-0432.CCR-09-0280 (2009).

25 Horst, B. *et al.* Gab2-mediated signaling promotes melanoma metastasis. *Am J Pathol* **174**, 1524-1533, doi:10.2353/ajpath.2009.080543 (2009).

26 Cloutier, R. *et al.* Elpistostege and the origin of the vertebrate hand. *Nature* **579**, 549-554, doi:10.1038/s41586-020-2100-8 (2020).

27 Nakamura, T., Gehrke, A. R., Lemberg, J., Szymaszek, J. & Shubin, N. H. Digits and fin rays share common developmental histories. *Nature* **537**, 225-228, doi:10.1038/nature19322 (2016).

28 Shubin, N. H., Daeschler, E. B. & Jenkins, F. A., Jr. The pectoral fin of Tiktaalik roseae and the origin of the tetrapod limb. *Nature* **440**, 764-771, doi:10.1038/nature04637 (2006).

29 Tanaka, M. Fins into limbs: Autopod acquisition and anterior elements reduction by modifying gene networks involving 5'Hox, Gli3, and Shh. *Dev Biol* **413**, 1-7, doi:10.1016/j.ydbio.2016.03.007 (2016).

30 Xu, B., Feng, X. & Burdine, R. D. Categorical data analysis in experimental biology. *Dev Biol* **348**, 3-11, doi:10.1016/j.ydbio.2010.08.018 (2010).

31 Philippidou, P. & Dasen, J. S. Hox genes: choreographers in neural development, architects of circuit organization. *Neuron* **80**, 12-34, doi:10.1016/j.neuron.2013.09.020 (2013).

32 Panganiban, G. & Rubenstein2, J. L. R. Developmental functions of the Distal-less: Dlx homeobox genes. *Development* **129**, 4371-4386 (2002).

33 Petit, F., Sears, K. E. & Ahituv, N. Limb development: a paradigm of gene regulation. *Nat Rev Genet* **18**, 245-258, doi:10.1038/nrg.2016.167 (2017).

34 Takeuchi, J. K. *et al.* Tbx5 and Tbx4 trigger limb initiation through activation of the Wnt/Fgf signaling cascade. *Development* **130**, 2729-2739, doi:10.1242/dev.00474 (2003).

35 Sheth, R. *et al.* Distal Limb Patterning Requires Modulation of cis-Regulatory Activities by HOX13. *Cell Rep* **17**, 2913-2926, doi:10.1016/j.celrep.2016.11.039 (2016).

36 Li, S. *et al.* Cistrome-GO: a web server for functional enrichment analysis of transcription factor ChIP-seq peaks. *Nucleic Acids Res* **47**, W206-W211, doi:10.1093/nar/gkz332 (2019).

37 Chablais, F. & Jazwinska, A. IGF signaling between blastema and wound epidermis is required for fin regeneration. *Development* **137**, 871-879, doi:10.1242/dev.043885 (2010).

38 Dealy, C. N. & Kosher, R. A. IGF-I and Insulin in the Acquisition of Limb- Forming Ability by the Embryonic Lateral Plate. *Developmental Biology* **177**, 291-299 (1996).

39 Dealy, C. N. & Kosher, R. A. IGF-I, insulin and FGFs induce outgrowth of the limb buds of amelic mutant chick embryos. *Development* **122**, 1323-1330 (1996).

40 Ye, F. *et al.* Overexpression of insulin-like growth factor-1 attenuates skeletal muscle damage and accelerates muscle regeneration and functional recovery after disuse. *Exp Physiol* **98**, 1038-1052, doi:10.1113/expphysiol.2012.070722 (2013).

41 Ye, L., Robertson, M. A., Mastracci, T. L. & Anderson, R. M. An insulin signaling feedback loop regulates pancreas progenitor cell differentiation during islet development and regeneration. *Dev Biol* **409**, 354-369, doi:10.1016/j.ydbio.2015.12.003 (2016).

42 Zhang, Y. M. *et al.* Distant Insulin Signaling Regulates Vertebrate Pigmentation through the Sheddase Bace2. *Dev Cell* **45**, 580-594 e587, doi:10.1016/j.devcel.2018.04.025 (2018).

43 Carboni, J. M. *et al.* BMS-754807, a small molecule inhibitor of insulin-like growth factor-1R/IR. *Mol Cancer Ther* **8**, 3341-3349, doi:10.1158/1535-7163.MCT-09-0499 (2009).

44 Garcia-Echeverria, C. *et al.* In vivo antitumor activity of NVP-AEW541—A novel, potent, and selective inhibitor of the IGF-IR kinase. *Cancer Cell* **5**, 231-239 (2004).

45 Vlahos, C. J., Matter, W. F., Hui, K. Y. & Brown, R. F. A Specific Inhibitor of Phosphatidylinositol 3=Kinase, 2-(4-Morpholinyl)-8-phenyl-4H-l-benzopyran-4-one (LY294002). *The Journal of Biological Chemistry* **289**, 5241-5248 (1994).

46 Wada, M., Horinaka, M., Yamazaki, T., Katoh, N. & Sakai, T. The dual RAF/MEK inhibitor CH5126766/RO5126766 may be a potential therapy for RAS-mutated tumor cells. *PLoS One* **9**, e113217, doi:10.1371/journal.pone.0113217 (2014).

47 Kim, K. *et al.* Clinicopathologic characteristics of early gastric cancer according to specific intragastric location. *BMC Gastroenterol* **19**, 24, doi:10.1186/s12876-019-0949-5 (2019).

48 Razumilava, N. & Gores, G. J. Cholangiocarcinoma. *The Lancet* **383**, 2168-2179, doi:10.1016/s0140-6736(13)61903-0 (2014).

49 Tang, Q. *et al.* Anatomic mapping of molecular subtypes in diffuse glioma. *BMC Neurol* **17**, 183, doi:10.1186/s12883-017-0961-8 (2017).

50 White, R. M. *et al.* Transparent adult zebrafish as a tool for in vivo transplantation analysis. *Cell Stem Cell* **2**, 183-189, doi:10.1016/j.stem.2007.11.002 (2008).

51 Kaufman, C. *et al.* A zebrafish melanoma model reveals emergence of neural crest identity during melanoma initiation. **351**, aad2197, doi:10.1126/ (2016).

52 White, R. M. *et al.* DHODH modulates transcriptional elongation in the neural crest and melanoma. *Nature* **471**, 518-522, doi:10.1038/nature09882 (2011).

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

54 Dankort, D. *et al.* Braf(V600E) cooperates with Pten loss to induce metastatic melanoma. *Nat Genet* **41**, 544-552, doi:10.1038/ng.356 (2009).

55 Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760, doi:10.1093/bioinformatics/btp324 (2009).

56 Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589-595, doi:10.1093/bioinformatics/btp698 (2010).

57 Cibulskis, K. *et al.* Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol* **31**, 213-219, doi:10.1038/nbt.2514 (2013).

58 Shen, R. & Seshan, V. E. FACETS: allele-specific copy number and clonal heterogeneity analysis tool for high-throughput DNA sequencing. *Nucleic Acids Res* **44**, e131, doi:10.1093/nar/gkw520 (2016).

59 Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21, doi:10.1093/bioinformatics/bts635 (2013).

60 Li, B. & Dewey, C. N. RSEM- accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 1471-2105 (2011).

61 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).

62 Korotkevich, G., Sukhov, V. & Sergushichev, A. Fast gene set enrichment analysis. *bioRxiv*, doi:10.1101/060012 (2019).

63 Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* **38**, 576-589, doi:10.1016/j.molcel.2010.05.004 (2010).

64 Khan, A. *et al.* JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. *Nucleic Acids Res* **46**, D260-D266, doi:10.1093/nar/gkx1126 (2018).

65 Grossman, R. L. *et al.* Toward a Shared Vision for Cancer Genomic Data. *N Engl J Med* **375**, 1109-1112 (2016).

66 Hoadley, K. A. *et al.* Cell-of-Origin Patterns Dominate the Molecular Classification of 10,000 Tumors from 33 Types of Cancer. *Cell* **173**, 291-304 e296, doi:10.1016/j.cell.2018.03.022 (2018).

67 Cancer Genome Atlas Research Network. Electronic address, e. d. s. c. & Cancer Genome Atlas Research, N. Comprehensive and Integrated Genomic Characterization of Adult Soft Tissue Sarcomas. *Cell* **171**, 950-965 e928, doi:10.1016/j.cell.2017.10.014 (2017).

68 Cancer Genome Atlas Research Network. Electronic address, w. b. e. & Cancer Genome Atlas Research, N. Comprehensive and Integrative Genomic Characterization of Hepatocellular Carcinoma. *Cell* **169**, 1327-1341 e1323, doi:10.1016/j.cell.2017.05.046 (2017).

69 Robertson, A. G. *et al.* Comprehensive Molecular Characterization of Muscle-Invasive Bladder Cancer. *Cell* **171**, 540-556 e525, doi:10.1016/j.cell.2017.09.007 (2017).

70 Fishbein, L. *et al.* Comprehensive Molecular Characterization of Pheochromocytoma and Paraganglioma. *Cancer Cell* **31**, 181-193, doi:10.1016/j.ccell.2017.01.001 (2017).

71 Cancer Genome Atlas Research, N. Comprehensive molecular profiling of lung adenocarcinoma. *Nature* **511**, 543-550, doi:10.1038/nature13385 (2014).

72 Cancer Genome Atlas Research, N. Comprehensive molecular characterization of gastric adenocarcinoma. *Nature* **513**, 202-209, doi:10.1038/nature13480 (2014).

73 Cancer Genome Atlas Research, N. Comprehensive molecular characterization of urothelial bladder carcinoma. *Nature* **507**, 315-322, doi:10.1038/nature12965 (2014).

74 Cancer Genome Atlas, N. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**, 330-337, doi:10.1038/nature11252 (2012).

75 Cancer Genome Atlas, N. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61-70, doi:10.1038/nature11412 (2012).
76 Cancer Genome Atlas Research, N. Comprehensive genomic characterization of squamous cell lung cancers. *Nature* **489**, 519-525, doi:10.1038/nature11404 (2012).
77 Ciriello, G. *et al.* Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer. *Cell* **163**, 506-519, doi:10.1016/j.cell.2015.09.033 (2015).
78 Cancer Genome Atlas Research, N. The Molecular Taxonomy of Primary Prostate Cancer. *Cell* **163**, 1011-1025, doi:10.1016/j.cell.2015.10.025 (2015).
79 Cancer Genome Atlas Research, N. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* **455**, 1061-1068, doi:10.1038/nature07385 (2008).
80 Cancer Genome Atlas, N. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature* **517**, 576-582, doi:10.1038/nature14129 (2015).
81 Cancer Genome Atlas Research, N. *et al.* Integrated genomic characterization of endometrial carcinoma. *Nature* **497**, 67-73, doi:10.1038/nature12113 (2013).
82 Cancer Genome Atlas Research, N. Comprehensive molecular characterization of clear cell renal cell carcinoma. *Nature* **499**, 43-49, doi:10.1038/nature12222 (2013).
83 Davis, C. F. *et al.* The somatic genomic landscape of chromophobe renal cell carcinoma. *Cancer Cell* **26**, 319-330, doi:10.1016/j.ccr.2014.07.014 (2014).
84 Cancer Genome Atlas Research, N. *et al.* Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med* **368**, 2059-2074, doi:10.1056/NEJMoa1301689 (2013).
85 Cancer Genome Atlas Research, N. Integrated genomic analyses of ovarian carcinoma. *Nature* **474**, 609-615, doi:10.1038/nature10166 (2011).
86 Brennan, C. W. *et al.* The somatic genomic landscape of glioblastoma. *Cell* **155**, 462-477, doi:10.1016/j.cell.2013.09.034 (2013).
87 Cancer Genome Atlas Research, N. Integrated genomic characterization of papillary thyroid carcinoma. *Cell* **159**, 676-690, doi:10.1016/j.cell.2014.09.050 (2014).
88 Tothova, Z. *et al.* Multiplex CRISPR/Cas9-Based Genome Editing in Human Hematopoietic Stem Cells Models Clonal Hematopoiesis and Myeloid Neoplasia. *Cell Stem Cell* **21**, 547-555 e548, doi:10.1016/j.stem.2017.07.015 (2017).
89 Lindsay, H. *et al.* CrispRVariants charts the mutation spectrum of genome engineering experiments. *Nat Biotechnol* **34**, 701-702, doi:10.1038/nbt.3628 (2016).
90 Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120, doi:10.1093/bioinformatics/btu170 (2014).
91 DeLuca, D. S. *et al.* RNA-SeQC: RNA-seq metrics for quality control and process optimization. *Bioinformatics* **28**, 1530-1532, doi:10.1093/bioinformatics/bts196 (2012).
92 Hu, Y. *et al.* An integrative approach to ortholog prediction for disease-focused and other functional studies. *BMC Bioinformatics* **12**, 1471-2105 (2011).
93 Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-359, doi:10.1038/nmeth.1923 (2012).
94 Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079, doi:10.1093/bioinformatics/btp352 (2009).
95 Ramirez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* **44**, W160-165, doi:10.1093/nar/gkw257 (2016).

96 Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137, doi:10.1186/gb-2008-9-9-r137 (2008).

97 Freese, N. H., Norris, D. C. & Loraine, A. E. Integrated genome browser: visual analytics platform for genomics. *Bioinformatics* **32**, 2089-2095, doi:10.1093/bioinformatics/btw069 (2016).