**Identification of MITF regulated transcript isoforms of ubiquitously expressed genes**

Stephen M Ostrowski1, Eleanor Klibaner-Schiff1, Erin Hamlin1Judith R Boozer1, Kiran Ebrahimi, Konrad Kaminiow1, and David E Fisher1\*

1 Cutaneous Biology Research Center, Department of Dermatology, Massachusetts General Hospital, Harvard Medical School, Charlestown, MA 02129, USA; Department of Dermatology, Harvard Medical School, Massachusetts General Hospital, Boston, MA 02114, USA

**\*** Correspondence: [sostrowski@mgh.harvard.edu](mailto:sostrowski@mgh.harvard.edu), dfisher3@mgh.harvard.edu

Microphthalmia Transcription Factor (MITF) is a key transcriptional regulator in the melanocytic lineage. MITF belongs to the MiT subfamily of transcription factors (TFEB, TFE3, and TFEC) that bind consensus E-Box sequences (CA[C/T]GTG) to initiate transcription1. MITF regulates the expression of genes involved in melanocyte survival, melanin synthesis and melanosomal transport1. Additionally, MITF plays critical roles in melanoma survival via regulation of anti-apoptotic, cell cycle, oxidative stress, and cellular metabolism genes, and has been implicated in melanoma therapy response1.

Multi-omic approaches have been employed to characterize MITF-dependent transcriptional regulation. ChIP-seq analysis has identified sites of MITF occupancy on gene regulatory regions in melanoma cells and primary human melanocytes2. However, the number of genes bound by transcription factor far exceeds the number of possible direct target genes, making it challenging to distinguish functional from non-functional binding. Correlating transcription factor expression with potential target gene expression across large microarray or RNA-sequencing datasets has proven effective in identifying transcription factor regulated genes3. Similarly, RNA-sequencing after MITF knockdown or MITF overexpression identifies genes that are regulated (directly or indirectly) by MITF4. Overall, integrating chromatin occupancy with gene expression enhances the identification of functional transcriptional targets.

Although RNA-sequencing inherently quantifies transcript-level expression, the analyses described above were performed at the gene level. This is largely because gene level analyses are both more statistically robust and experimentally tractable, as knockdown of single isoforms is difficult. However, understanding transcript level dynamics remains important as transcript isoforms from the same gene can have diverse and non-overlapping functions. For example, MITF itself has several transcript isoforms that vary from widely expressed to tissue specific; the M-isoform of MITF (MITF-M) is expressed almost exclusively in melanocytes and melanoma1. Importantly, even transcripts that account for a relatively low percentage of total gene expression can have important cellular functions. For example, most cases of Timothy Syndrome result from a mutation affecting a minority isoform (~20% of total gene expression) of the CACNA1C gene5.

We wished to test whether MITF gene level analysis might obscure evidence of MITF regulation of transcript isoforms due to dilution of isoform specific associations. We first computed Spearman and Pearson correlation coefficients between expression of MITF-M (ENST00000394351) and each protein coding transcript and each gene in melanoma samples from the Cancer Cell Line Encyclopedia (CCLE) and the Tsoi et al 6,7. We retained transcripts with strong correlation with MITF in both datasets (Spearman or Pearson coefficient ≥ 0.5 in both CCLE and Tsoi), yielding 817 “MITF correlated” transcripts from 398 genes (Supplemental Table 1). Gene ontology (GO) and pathway analysis revealed associations with pigmentation-related genes and genes known to be regulated by MITF-M, validating that correlation with MITF enriches for MITF target genes (Supplemental Figure 3).

Finally, we filtered this list to include only transcripts where transcript-MITF correlation showed a stastically significant stronger correlation than the gene-MITF correlation was using CoCor(Figure 1A, Supplemental Figure 1)8. This final list of “discordantly correlated transcripts” comprised 86 transcripts from 80 genes (Figure 1A, Supplemental Table 1). Discordantly correlated transcripts showed enrichment in melanoma in both CCLE and TCGA melanoma data sets compared to expression of their parent genes (Supplemental Figure 2). However, GO and pathway analysis did not reveal any association with pigmentation or MITF, suggesting that these are novel associations with MITF (Supplemental Figure 3).

To determine whether discordantly correlated transcripts exhibit evidence of direct regulation by MITF, we compiled features of MITF regulation for all transcripts by processing data from published studies. Specifically, we quantified the number of MITF ChIP peaks within 1kb of the transcription start site (TSS), number of E-box sites within 1kb of the transcription start site, and effect of MITF knockdown on transcript expression (Supplemental File 1). We then compared these features across three transcript categories: 1) all transcripts, 2) transcripts showing MITF correlation only (817 transcripts), and 3) discordantly correlated transcripts (86 transcripts). Compared to all transcripts, discordantly correlated transcripts exhibited significant enrichment of number of E-box sites, number of MITF ChIP peaks, and reduction of transcript expression after MITF knockdown (Figure 2A). Compared to MITF correlated transcripts, discordantly correlated transcripts demonstrated enrichment in the number of MITF ChIP peaks and in reduced expression after MITF knockdown (Figure 2A). These findings support the idea that discordantly correlated transcripts are under direct MITF regulation.

Finally, we sought to explore how MITF may regulate specific transcripts without affecting total expression of the parent genes. Transcription can initiate at multiple TSSs within a gene. While some TSSs are separated by only a few nucleotides and likely share regulatory elements, others are more distantly spaced and regulated by alternative promoters9. Using a recently published annotation of human promoters, we found that while 33% of all human protein coding transcripts and 35% MITF correlated transcripts are associated with a unique promoter, 51% of discordantly correlated transcripts are associated with a unique promoter 9 (Figure 2B). Overall, our findings suggest that discordantly correlated transcripts are preferentially regulated by MITF through localized cis-regulatory elements at unique promoters, while other transcripts from the same genes are not regulated by MITF, as their expression is initiated from alternative regulatory sequences.

In summary, we identified a list of transcripts whose expression is strongly correlated with MITF and for which features of direct MITF regulation are enriched. Because these regulatory relationships are not apparent at the gene level, they likely represent novel MITF targets that have gone unrecognized in prior analyses. Functional characterization of these transcripts the melanocyte and melanoma lineages can be examined using CRISPRi strategies for isoform-specific perturbation10. We also provide a fully annotated dataset of MITF-regulated transcript features to support future studies on MITF-dependent transcript level regulation (Supplemental File 1).

**Data Availability Statement**

CCLE gene and transcript expression (CCLE 2024 Q4) were downloaded from <https://depmap.org/portal/>; data from melanoma samples were subsetted for melanoma specific analysis. For Tsoi cell lines, FASTQ files were downloaded, and processed with Kallisto to generate t

Genomic coordinates (Hg38) of Ebox sequences (CACATG/CATGTG and CACGTG) were compiled using Biostrings and BSgenome libraries in R. This was then cross-referenced with genomic coordinates of the transcription start sites transcripts, to quantify the number of EBox sites within 1KB of transcription start site for each transcript.

For ChIP-Seq (GSE77437, GSE61967) and Cut&Run (GSE153020) analysis, FastQ files from for MITF and IgG or input control samples were downloaded analysis were downloaded, aligned to Hg38 with Bowtie, followed by generation of indexed BAM files with Samtools. After peak calling with Macs2, DiffBind package was used to assess differential binding. Peaks significant enriched in MITF samples were extracted and saved in BED format. ChIPseeker was used to annotate MITF-bound regions and associate them to regions proximal to transcription start sites.

For transcript level analysis of RNA sequencing FASTQ files were downloaded from with MITF knockdown (GSE283655, GSE163646, GSE115845, PRJNA704810, PRJEB30337) or MITF overexpression, (PRJNA704810, GSE163646). PCA analysis was performed and outlier samples were removed, these outliers correlated with lack of expected effect on MITF (knockdown or overexpression). FASTQ files were processed with Kallisto to generate transcript level count tables. For each data set, a value of siMITF expression/siCON expression was calculated for each transcript (excluding values where expression was zero in the siCON condition). A mean value siMITF/siCON across all four data sets was calculated for each transcript.

Data and R scripts used to generate figures available on Github; Supplemental\_Table 1 and Supplemental File\_1 are available on Github (<https://github.com/SteveOMGH/Discordant_transcript_project>)

**ORCIDs**  
Stephen Ostrowski: <https://orcid.org/0000-0002-0669-7502>

Eleanor Klibaner-Schiff: https://orcid.org/0000-0003-1473-5123

Konrad Kaminiów: https://orcid.org/0000-0002-7481-2621

David Fisher: https://orcid.org/0000-0002-5506-0575

**Conflict of Interest Statement**

DEF has a financial interest in Soltego, a company developing salt-inducible kinase inhibitors for topical skin-darkening treatments that might be used for a broad set of human applications. The interests of DEF were reviewed and are managed by Massachusetts General Hospital and Partners HealthCare in accordance with their conflict-of-interest policies. DEF also discloses consulting relationships with Pierre Fabre, Coherent Medicines, Biocoz, Swiss Rockets, and Tasca.

**Acknowledgements**

We thank Hank Adelmann and Jun Song for critical reading of the manuscript

**Author Contributions Statement**

Conceptualization: DEF, SMO; Data Curation: KK, JB; Investigation: SMO, EK-S, EH, DEF; Writing - Original Draft Preparation: SMO. Writing – review & editing: SMO, DEF

**References**

1 Kawakami, A. & Fisher, D. E. The master role of microphthalmia-associated transcription factor in melanocyte and melanoma biology. *Lab Invest* **97**, 649-656 (2017). <https://doi.org:10.1038/labinvest.2017.9>

2 Laurette, P. *et al.* Transcription factor MITF and remodeller BRG1 define chromatin organisation at regulatory elements in melanoma cells. *Elife* **4** (2015). <https://doi.org:10.7554/eLife.06857>

3 Hoek, K. S. *et al.* Novel MITF targets identified using a two-step DNA microarray strategy. *Pigment Cell Melanoma Res* **21**, 665-676 (2008). <https://doi.org:10.1111/j.1755-148X.2008.00505.x>

4 McGill, G. G. *et al.* Bcl2 regulation by the melanocyte master regulator Mitf modulates lineage survival and melanoma cell viability. *Cell* **109**, 707-718 (2002). <https://doi.org:10.1016/s0092-8674(02)00762-6>

5 Dick, I. E., Joshi-Mukherjee, R., Yang, W. & Yue, D. T. Arrhythmogenesis in Timothy Syndrome is associated with defects in Ca(2+)-dependent inactivation. *Nat Commun* **7**, 10370 (2016). <https://doi.org:10.1038/ncomms10370>

6 Tsoi, J. *et al.* Multi-stage Differentiation Defines Melanoma Subtypes with Differential Vulnerability to Drug-Induced Iron-Dependent Oxidative Stress. *Cancer Cell* **33**, 890-904 e895 (2018). <https://doi.org:10.1016/j.ccell.2018.03.017>

7 Ghandi, M. *et al.* Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature* **569**, 503-508 (2019). <https://doi.org:10.1038/s41586-019-1186-3>

8 Diedenhofen, B. & Musch, J. cocor: a comprehensive solution for the statistical comparison of correlations. *PLoS One* **10**, e0121945 (2015). <https://doi.org:10.1371/journal.pone.0121945>

9 Demircioglu, D. *et al.* A Pan-cancer Transcriptome Analysis Reveals Pervasive Regulation through Alternative Promoters. *Cell* **178**, 1465-1477 e1417 (2019). <https://doi.org:10.1016/j.cell.2019.08.018>

10 Davies, R. *et al.* CRISPRi enables isoform-specific loss-of-function screens and identification of gastric cancer-specific isoform dependencies. *Genome Biol* **22**, 47 (2021). <https://doi.org:10.1186/s13059-021-02266-6>

**Figure 1 –** **A subset of transcripts in melanoma have high correlation with MITF, while the parent genes of the same transcripts show lower correlation with MITF.** A) Average of Pearson and Spearman correlation between each transcript (Y-axis) and its parent gene (X-axis) and MITF. Highlighted in pink are transcripts whose expression is correlated with MITF (Pearson correlation or Spearman correlation> 0.5) highlighted in red are transcripts with discordant MITF correlation (Pearson or Spearman correlation with MITF >0.5 and transcript is at least 20% more correlated with MITF than its parent gene). B) Transcript map of protein-coding *ABR* transcripts expressed in CCLE melanoma, with location of EBOX and MITF ChIP peaks in relation to TSS. C) Scatter plot of M-MITF expression against *ABR* transcript ENST00000544853 expression (red) or summed expression of all other ABR transcripts (blue).

**Figure 2 –** **Discordantly correlated transcripts are enriched for features associated with MITF regulation and are more likely to be associated with a unique promoter**. A) Number of EBox within 1kB of TSS, number of MITF ChIP peak within 1kB of TSS and effect of MITF knockdown on transcript expression in each transcript group (all transcripts, MITF correlated transcripts, discordantly correlated transcripts B) Proportion of transcripts in each group that are associated with a unique promoter. For EBox, ChIP peak, binned expression and promoter analysis, Fisher's exact test used for pairwise comparisons. Wilcoxon signed-rank test used for transcript expression analysis (n.s.: Not significant \*: p < 0.05, \*\*: p < 0.01 ,\*\*\*: p < 0.001, \*\*\*\*:p<.0001)

**Supplemental Figure 1. Outline of transcript filtering strategy.**

**Supplemental Figure 2 – Expression of discordantly correlated genes is enriched in melanoma** A) Skin cutaneous melanoma (SKCM) ranked expression of discordantly correlated transcripts (red) and their parent genes (blue) in CCLE and TCGA. B) Expression of *ABR* total gene expressioninCCLE tumor types (expression in melanoma is highlighted in blue) and ABR transcript ENST00000544853 expression inCCLE tumor types (expression in melanoma highlighted in red.

**Supplemental Figure 3 – Gene Ontology and Pathway analysis of MITF correlated transcripts and discordantly correlated transcripts**

**Supplemental Figure 4 – Discordantly correlated transcript group has lower average expression in CCLE melanoma as compared to MITF correlated transcripts**. Expression values in CCLE melanoma for each transcript group.