Identification of Gene Variants Associated with Melanocyte Stem Cell Differentiation in Mice Predisposed for Hair Graying

Allison C. Fialkowski,*,1 Denise J. Levy,† Dawn E. Watkins-Chow,‡ Joseph W. Palmer,§ Roshan Darji,†,2 Hemant K. Tiwari,*, William J. Pavan,† and Melissa L. Harris‡,3

*Department of Biostatistics, †Department of Biology University of Alabama at Birmingham, Birmingham, AL, and ‡Genetic Disease Research Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD

ABSTRACT Age-related hair graying is caused by malfunction in the regenerative potential of the adult pigmentation system. The retention of hair color over the life of an organism is dependent on the ability of the melanocyte stem cells and their progeny to produce pigment each time a new hair grows. Age-related hair graying is variable in association with genetic background suggesting that quantitative trait loci influencing this trait can be identified. Identification of these quantitative trait loci may lead to the discovery of novel and interesting genes involved in stem cell biology and/or melanogenesis. With this in mind we developed previously a sensitized, mouse modifier screen and discovered that the DBA/1J background is particularly resistant to melanocyte stem cell differentiation in comparison to the C57BL/6J background. Melanocyte stem cell differentiation generally precedes hair graying and is observed in melanocyte stem cells with age. Using quantitative trait loci analysis, we have now identified three quantitative trait loci on mouse chromosomes 7, 13, and X that are associated with DBA/1J-mediated variability in melanocyte stem cell differentiation. Taking advantage of publicly-available mouse sequence and variant data, in silico protein prediction programs, and whole genome gene expression results we describe a short list of potential candidate genes that we anticipate to be involved in melanocyte stem cell biology in mice.

KEYWORDS hair graying melanocyte stem cell QTL modifier

In humans, the visible phenotype of graying scalp hair is associated primarily with aging and varies among people of different ethnic or ancestral geographic origin. In people 45-65 years of age, the intensity (percent per individual) and frequency (percent of individuals) of gray hair ranges from 15 to 42% and 43–93%, respectively (Panhard et al. 2012). The fact that these differences are associated with ancestry, or genetic background, suggest that hair graying is, in part, a genetic trait. The quantitative nature of hair graying further suggests that this trait is influenced by the combination of effects at multiple genetic loci. Gray hair is easy to dismiss as a purely cosmetic phenotype, however, research into the molecular etiology of this phenotype has contributed to our basic understanding of stem cell biology, tissue regeneration and homeostasis, melanocyte-related diseases like vitiligo and melanoma, and the role of stem cells in aging phenotypes. Accordingly, we are interested in identifying the genetic modifications that contribute to hair graying variation in the hopes of discovering novel genes that participate in these processes.

Unfortunately, genetic studies of gray hair in humans can be difficult to perform due to the confounding variable associated with hair dyeing. Accurate phenotyping can only be performed in individuals and regions of the scalp (near the roots) that maintain natural hair color. Not surprisingly, to date there is only one reported genome-wide association...
study that has successfully identified a genetic locus, IRF4, that is involved in age-related hair graying in humans (Adhikari et al. 2016). As an alternate to studying pigmentation in humans, the mouse model system has long been used to investigate genetic and molecular mechanisms related to melanocyte biology (Jackson 1994, 1997; Steingrímsson et al. 2006). Thus, we sought to assess whether mice, as an alternate to humans, could help in the identification of genetic variants that contribute the phenotypic diversity of hair graying. Previously we reported the development of a sensitized screen to evaluate genetically diverse inbred mouse strains for their ability to influence hair graying (Harris et al. 2015). In this screen we employed the Tg(DctSox10) transgenic mouse line to predispose mice to hair graying. The cellular mechanism responsible for hair graying in these mice is melanocyte stem cell (McSC) differentiation. This phenomenon precedes hair graying and is positively associated with hair graying severity (Harris et al. 2013). Differentiated McSCs have also been observed in human hairs and their number increases with age (Nishimura et al. 2005), which makes McSC differentiation a relevant cellular phenotype to evaluate for genetic loci that may modify the extent of age-related hair graying in both mouse and humans. Mechanistically we predict that modifier genes that effect McSC differentiation may directly regulate the process of melanogenesis, but could also be involved in initial McSC establishment, proliferation or migration of McSC progenitors.

McSCs that are undergoing differentiation produce visible ectopic pigmentation when viewed by light microscopy and the number of hairs that contain these ectopically pigmented McSCs (EPMs) varies in animals of different genetic backgrounds (Harris et al. 2015). Mice that are hemizygous for a conditional, Sox10-expressing transgene, designated C57BL/6J-Tg(DctSox10)/0, are extremely susceptible to McSC differentiation. In contrast, progeny derived from mating C5BL/6J-Tg(DctSox10)/0 mice to other inbred genetic backgrounds (C3H/HeJ, 129SvEvTac, FVB/NTac, DBA/1J, BALB/CJ) produced F1 hybrids that all exhibit reduced numbers of hairs with EPMs in response to the transgene (Harris et al. 2015). Tg(DctSox10)/0 hybrid animals produced by mating C57BL/6J-Tg(DctSox10)/0 mice to DBA/1J mice exhibit a particularly low level of transgene-mediated EPMs. Reduction of EPMs in these F1 hybrids suggests a dominant mode of inheritance for EPM resistance and we sought to identify these DBA/1J-associated resistance loci using QTL linkage mapping.

**METHODS**

**Mice**

C57BL/6J and DBA/1J mice were obtained from the Jackson Laboratory. The Tg(Dct-Sox10) transgenic line Tg(Dct-Sox10)CF1-10Pav (Hakami et al. 2006)) was generated previously, established on the FVB/N background and maintained through a combination of backcrossing to C57BL/6J and by intercross. All (DBA/1J x C57BL/6J-Tg(DctSox10)/0) F1-Tg(DctSox10)/0 animals (abbreviated as D1B6F1-Tg(DctSox10)/0) used in this study were generated by mating one C57BL/6J-Tg(DctSox10)/0 male to several DBA/1J females. D1B6F1-Tg(DctSox10)/0 females were then mated to individual C57BL/6J males to generate N2 backcross progeny.

**Phenotype Analysis**

Assessment of hairs with EPMs was performed as described previously (Harris et al. 2015). Briefly, between 9-11 weeks of age, hair along a 2x2 cm region of the lower back was plucked by hand to synchronize and initiate hair regrowth. Hairs in the plucked region were allowed to regrow for seven days (equivalent to hair cycle stage anagen III/IV). Skin from these animals was dissected and processed for cryosectioning. Using light microscopy, approximately four 10 μm sections were analyzed in total skipping at least three sections between those analyzed to prevent counting the same hair twice. EPMs occur in the hair bulge at the insertion point of the arrector pili muscle, thus only sectioned hair follicles that spanned the entire region from the sebaceous gland and past the junction of the dermal/subcutis were counted. Between 50-100 hairs were examined for the presence or absence of EPMs. The percentage of hairs with EPMs was calculated by dividing the number of hairs with EPMs by the total number of hairs analyzed. 122 N2 mice were phenotyped. For QTL analysis, selective genotyping was performed on 79 of the N2 mice exhibiting the highest (> 50%, n = 39) and lowest (<20%, n = 40) percentage of hairs exhibiting EPMs. The EPM phenotype was converted to a binary trait with the high EPM phenotype scored as 1 and referred to as ‘affected’, while the low EPM phenotype was scored as 0 and referred to as ‘unaffected’. Graphing was performed using Graphpad Prism (Graphpad Software). Brightfield microscopy was performed on a Zeiss Observer.D1 compound microscope. Images were obtained with an Axioscam Hrc camera (Zeiss) using the ZEN software (Zeiss) and processed with Adobe Photoshop (Adobe).

**Genotyping**

Presence of the Tg(Dct-Sox10) transgene was determined by PCR using primers that generate an amplicon spanning the Dct promoter and the Sox10 cDNA: 5’-AGCAGTATGGCTTGAGCCT-3’; 5’-TCCAGTCGTAGCCCGTGAC-3’. PCR cycling was performed as published previously (Harris et al. 2013). SNP genotyping was performed on a custom panel of 1449 SNPs (equivalent to the Mouse Medium Density Linkage Panel, Illumina) using the GoldenGate Genotyping Universal-32 Assay Kit with UDG (Illumina). Complete SNP genotyping data are available in Supplemental File 1 (sheet name-Original Sample Genotypes). In preparation for QTL analysis using R/qtl, 559 non-informative SNPs and SNPs with a high number of no call (NC) values or were omitted. Sample genotypes were also recoded such that homozygote genotypes matching the parental C57BL/6J genotype were designated AA, and heterozygote genotypes were designated AB. SNPs with unknown chromosomal coordinates in the original genotyping data (listed as chr 0) were identified in the genome. The chromosomal coordinates of each SNP were then converted to Sex-Averaged cM-G2F1 centimorgan positions using Mouse Map Converter (http://cgd.jax.org/mousemapconverter/). The final data matrix used for QTL analysis is included in Supplemental File 1 (sheet name- Converted Sample Genotypes).

**Statistical analysis**

A total of 79 mice were initially evaluated, and 3 removed for low-quality genotyping results. QTL linkage analysis of 76 mice (36 with the high and 40 with the low EPM phenotype, 37 males and 39 females) and 890 markers using EPM as a binary trait was performed using the R/qtl software (Rv.3.4.4, qtl v.1.42-8). One-dimensional scans were conducted without and with sex as an additive covariate using logistic regression with the EM algorithm (Dempster et al. 1976; Lander and Botstein 1989). Separate LOD significance thresholds were obtained from 1,000 permutations for autosomal SNPs and 18,850 permutations for X chromosome SNPs (see R/QTL documentation for explanation of the permutation counts applied within the scanone function). Two-dimensional scans were conducted with sex as an additive covariate using logistic regression with the EM algorithm. Separate LOD significance thresholds were obtained from
1,000 permutations for autosomal SNP pairs, 355,289 permutations for X chromosome SNP pairs, and 9,425 permutations for autosomal:X chromosome SNP pairs (see R/QTL documentation for explanation of the permutation counts applied within the scantwo function). For the QTL found significant at the 0.05 level in the two-QTL analyses, multiple-QTL analyses were performed with sex as an additive covariate using logistic regression with multiple imputation (Sen and Churchill 2001). The locations of the QTL were updated based on maximum likelihood (Zeng et al. 1999) using the refineqtl function within R/QTL.

Identification of candidate genes
Whole skin RNASeq data comparing C57BL/6J to DBA/1J was retrieved at NCBI GEO using the accession # GSE86315. This dataset included read counts previously generated by aligning RNASeq reads to the mouse genome (GRCm38/mm10) using TopHat2, assessment of mapping quality using RSeQC and RNA-SeqQC, and read counting using HTSeq (Swindell et al. 2017). HTSeq reads from the C57BL/6J and DBA/1J control skin samples (2 males and 2 females per strain all treated with a non-toxic lanolin-derived occlusion cream) were used to generate normalized read counts (median ratio method) and compared to obtain differential expression values using DESeq2. The complete differential expression data including base mean (mean of the normalized counts), log2 fold change, and adjusted p values is provided in Supplemental File 2. Wildtype C57BL/6J McSC RNASeq data were retrieved at NCBI GEO using the accession # GSE102271. RNASeq reads were aligned to the Ensembl GRCm38.p5 primary DNA assembly using STAR (v2.5.2b) and normalized read counts (median ratio method) determined using DESeq2. Normalized read counts with a value of 0 were omitted from further analysis. The complete McSC expression data are provided in Supplemental File 3. Variants between C57BL/6J and DBA/1J were obtained from the Mouse Genome Project (REL-1505; ftp://ftp-mouse.sanger.ac.uk/). The ‘genome variants’ function of PROVEAN (v1.1.3 and GRCm38 Ensembl 74; http://provean.jcvi.org) was used to score variants based on predicted protein function. PROVEAN summary and detailed results are provided in Supplemental File 4.

Reagent and Data Availability
All data associated with this manuscript is available within the manuscript or as supplemental files. Supplemental Files 1-4 are available via the GSA Figshare portal. Supplemental File 1 provides the SNP genotyping data of the N2 animals. Supplemental File 2 provides the differential mRNA expression data comparing DBA/1J and C57BL/6J skin from the reanalysis of GSE86315. Supplemental File 3 provides the mRNA expression data of wildtype C57BL/6J McSCs from the reanalysis of GSE102271. Supplemental File 4 provides the summary and detailed output from PROVEAN. Supplemental material available at Figshare: https://doi.org/10.25387/g3.7489625.

RESULTS
Distribution of EPM susceptibility in progeny derived from backcrossing D1B6F1-Tg(DctSox10)/0 to C57BL/6J
To identify the genetic determinants from DBA/1J that promote EPM resistance we set out to map loci that modify the production of EPMs in progeny produced from backcrossing D1B6F1-Tg(DctSox10)/0 females to C57BL/6J males (these progeny are hereafter referred to as N2 mice). At approximately eight weeks of age, the hair was plucked along the lower back of N2 mice to induce and synchronize hair growth. One week later, skin from these mice was obtained from the plucked region and assessed for EPMs using histological methods (Figure 1a). 122 N2 mice that carry the Tg(DctSox10) transgene were evaluated phenotypically, and exhibited a range of EPM measurements extending between the C57BL/6J-Tg(DctSox10)/0 and D1B6F1-Tg(DctSox10)/0 parental phenotypes (Figure 1b). A statistically significant gender effect (t-test, p-value = 0.009) is also observed in N2 mice with the phenotypic mean of the female N2 animals skewed toward more resistant to EPMs suggesting the need for including sex as covariate during QTL mapping.

QTL analysis provides evidence for three QTL loci associated EPM variability
To map genetic modifiers that affect resistance to McSC differentiation, we used a selective genotyping approach where only the animals with the most extreme phenotypes are genotyped for linkage analysis (based on (Lander and Botstein 1989)). A total of 79 N2 mice were genotyped and represent those animals with the highest (>50%, n = 39) and lowest (<20%, n = 40) percentage of hairs exhibiting EPMs (Figure 1b). These
| Chr | Pos | LOD | Pval |
|-----|-----|-----|------|
| rs6312657 | 1 | 30.99 | 0.60 |
| CEL-2_168586738 | 2 | 79.43 | 1.42 |
| rs13477487 | 3 | 68.08 | 1.63 |
| rs3659791 | 4 | 38.57 | 1.37 |
| rs3664617 | 5 | 0.16 | 0.45 |
| rs6389420 | 6 | 57.63 | 1.46 |
| rs6160140 | 7 | 28.39 | 2.60 |
| rs3656875 | 8 | 29.32 | 2.01 |
| rs13480345 | 9 | 42.94 | 1.70 |
| rs3699409 | 10 | 3.37 | 1.29 |
| rs11109548 | 11 | 18.85 | 2.40 |
| rs3706330 | 12 | 6.60 | 1.12 |
| c13.loc20 | 13 | 21.79 | 3.56 |
| rs6156908 | 14 | 31.24 | 0.98 |
| rs711814 | 15 | 3.58 | 0.46 |
| c16.loc38 | 16 | 39.39 | 2.36 |
| c17.loc3 | 17 | 3.00 | 0.28 |
| rs13483244 | 18 | 6.53 | 0.30 |
| rs6236348 | 19 | 0.68 | 0.40 |

**Table 1: Autosomal single QTL analysis for QTL loci linked to EPM variability**

1. Results from the single-QTL analysis. Cells shaded in gray highlight loci described in the Results. chr, chromosome; pos, centimorgan position; LOD, LOD value; pval, p-value.
mice were genotyped using a panel of 1449, evenly-dispersed, mouse-specific SNP loci assays (Illumina, Supplemental File 1). Among the 1449 SNPs evaluated, 890 were found to be reliable and informative between C57BL/6J and DBA/1J. 3 of the original 79 mice had low-quality genotyping scores and were removed prior to QTL analysis. Identification of individual QTL was first performed using a single-QTL genome scan approach (R/qtl; Broman et al. 2003). Only the genotyped animals mentioned above were included in these analyses and thus the high and low EPM percentage were treated as a binary trait. Results from the single-QTL analysis, without sex as a covariate and a 5% significance threshold, indicates the presence of one QTL on chr 7 when sex is included as an additive covariate (lod.add- 2.91, p-value- 0.09; Table 1). Using a 10% significance threshold, there is also weak support for an additional QTL on chr 7 at sex = 1.25. Using a 10% significance threshold, there is also weak support for an additional QTL on chr 7 when sex is included as an additive covariate (lod.add- 2.91, p-value- 0.09; Table 1). To look for additional QTL that may contribute to DBA/1J-mediated EPM resistance, a two-QTL genome scan approach was performed. Results from the two-QTL analyses focusing on pairs of autosomal QTL (with or without sex as an additive covariate) does not identify any pairs of that significantly improve the two-QTL model (lod.full or lod.add) above that of a single QTL (lod.fv1 or lod.av1) when using a 5% significance threshold (Table 3). However, when assessing for pairs of QTL between the autosomes and the X chromosome (with sex as an additive covariate), there is significant evidence for a pair on chr 7 and X chromosome (Table 3). The full model (lod.full) containing QTL at chr 7 and X chromosome provides a better fit to the data than both the best single-QTL model (lod.fv1) and the additive model (lod.add). Interaction between these QTL is also significant (lod.int) suggesting that these QTL are epistatic. When the chr 7 and X chromosome genotypes are considered together (Figure 2c), the effect of the chr 7 QTL in females is the same; a low proportion of individuals exhibit the affected, high EPM phenotype when heterozygous for the C57BL/6J allele. No statistically significant interaction of these two QTL help explain the sex difference observed in the N2 distribution (Figure 1a). The chr 7 QTL in females is the same; a low proportion of individuals exhibit the affected, high EPM phenotype when heterozygous for the C57BL/6J allele. No statistically significant interaction of these two QTL help explain the sex difference observed in the N2 distribution (Figure 1a). The chr

---

Table 2 Sex-stratified single-QTL analysis for QTL loci linked to EPM variability

| position | chr | pos | lod | pval | position | chr | pos | lod | pval |
|----------|-----|-----|-----|------|----------|-----|-----|-----|------|
| rs6312657 | 1 | 30.99 | 1.13 | 0.99 | rs13475706 | 1 | 0.48 | 0.59 | 1.00 |
| CEL-2_135876979 | 2 | 60.99 | 0.94 | 1.00 | rs3726974 | 2 | 80.78 | 1.40 | 0.94 |
| rs13477487 | 3 | 68.08 | 1.51 | 0.86 | c3.loc44 | 3 | 44.22 | 1.20 | 0.98 |
| rs3659791 | 4 | 38.57 | 1.56 | 0.78 | rs13477643 | 4 | 14.92 | 0.42 | 1.00 |
| mCV23386455 | 5 | 27.75 | 0.33 | 1.00 | rs3660964 | 5 | 23.12 | 0.27 | 1.00 |
| rs6389420 | 6 | 57.63 | 1.15 | 0.98 | rs6238771 | 6 | 19.90 | 1.12 | 0.99 |
| rs3700608 | 7 | 0.26 | 0.39 | 1.00 | c7.loc29 | 7 | 29.26 | 3.54 | 0.03 |
| CEL-8_33812776 | 8 | 16.34 | 1.31 | 0.95 | c8.loc44 | 8 | 47.68 | 2.55 | 0.20 |
| rs13480271 | 9 | 35.75 | 1.04 | 1.00 | c9.loc57 | 9 | 57.03 | 2.23 | 0.40 |
| CEL-10_58149652 | 10 | 23.51 | 0.46 | 1.00 | c10.loc45 | 10 | 23.51 | 2.69 | 0.15 |
| rs6197743 | 11 | 38.94 | 1.98 | 0.44 | c11.loc15 | 11 | 18.55 | 1.77 | 0.68 |
| rs3706330 | 12 | 6.60 | 1.81 | 0.62 | rs3023711 | 12 | 59.74 | 0.28 | 1.00 |
| rs6283319 | 13 | 40.58 | 2.29 | 0.27 | c13.loc19 | 13 | 20.79 | 1.98 | 0.55 |
| gnf14.055.608 | 14 | 23.52 | 0.39 | 1.00 | c14.loc33 | 14 | 31.87 | 0.58 | 1.00 |
| rs3715857 | 15 | 4.87 | 0.95 | 1.00 | rs6188239 | 15 | 10.26 | 0.33 | 1.00 |
| rs4201998 | 16 | 35.46 | 1.52 | 0.80 | c16.loc36 | 16 | 37.39 | 1.09 | 0.99 |
| rs6312657 | 1 | 30.99 | 1.13 | 0.99 | rs13483525 | 19 | 3.11 | 0.19 | 1.00 |

---

*Results from the single-QTL analysis stratified by sex. Cells shaded in gray highlight loci described in the Results. chr, chromosome; pos, centimorgan position; lod, LOD value; pval, p-value.*
13 QTL, on the other hand, has an effect opposite to that of the QTL at chr 7 and is associated with EPM susceptibility.

Identification of candidate genes that influence the EPM phenotype

As the first step to prioritizing candidate genes for the QTL identified above, we used 1.5-LOD support intervals (the chromosomal region where the LOD score is within 1.5 of its maximum) to determine the bounds of the QTL linkage region (Figure 3, Table 4). In order to consider the chr 7 and chr X QTL simultaneously we performed multiple-QTL analysis to fit these QTL into one model, refine the QTL locations, and use these LOD estimates to define the 1.5-LOD support intervals (Figure 3a). The chr 13 intervals were derived from the single-QTL analysis (Figure 3b). For these QTL, the 1.5-LOD support intervals were relatively large and encompassed a significant number of genes; the chr 7 QTL interval covered 22 Mbp with 444 genes, the chr 13 QTL covered 58 Mbp with 603 genes, and the X chromosome QTL interval covered 44 Mbps with 377 genes.

As an approach to stratify the genes represented in these intervals we characterized them using publicly-available gene expression data, mouse resequencing data, an online program for predicting the effects of coding variants on protein function, and a curated list of genes known to be involved in pigmentation. First, any genes that are differentially expressed in skin from DBA/1J and C57BL/6J animals highlights cell-specific or systemic changes that may impact McSC function. For instance, genes that are upregulated in DBA/1J animals may be dominant drivers of EPM resistance that promote McSC stemness over differentiation, while those that are downregulated in DBA/1J may highlight proteins responsible for heightened melanogenesis in C57BL/6J animals. Using NCBI Gene Expression Omnibus (www.ncbi.nlm.nih.gov/geo/) we identified a study that compared the transcriptomes of skin derived from 11-week-old DBA/1J and C57BL/6J animals using RNAseq. This study evaluated strain-specific effects of Imiquimod on skin, and included expression data from control subjects (two males and two females for each strain) that were treated only with a non-toxic lanolin-derived occlusion cream (Swindell et al. 2017). Using DESeq2 to compare these DBA/1J and C57BL/6J control subjects we find a number of genes that exhibit a significant twofold difference in expression (p-adjusted < 0.05; Supplemental File 2).

Second, we were also interested to capture genes that are expressed at detectable levels in McSCs that are not up or down-regulated in DBA/1J and C57BL/6J skins. We anticipate these genes represent candidates that may alter intrinsic McSC dynamics at the protein level, either through differential transcript expression or as the consequence of genetic variation in their coding sequence. Using
a previously published RNAseq dataset (Harris et al. 2018), we generated a list of expression values from C57BL/6J wildtype McSCs isolated from dormant (telogen-stage) hairs. These gene expression values represent the read counts per gene after normalization using the median ratio method (DESeq2). Genes that were not expressed (value = 0) were excluded and the remaining genes ranked by percent (Supplemental File 3).

Third, we evaluated the genes within the linkage intervals for those that exhibit genetic variation between C57BL/6J and DBA/1J as well as those that are known to participate in pigmentation. Mouse resequencing and associated variant call data were used to identify genes that contain coding variants between C57BL/6J and DBA/1J (Mouse Genomes Project, Sanger; GRCm38)(Yalcin et al. 2011; Keane et al. 2011). These include missense, insertion and deletion mutations that are predicted to have deleterious protein consequences by PROVEAN (Choi and Chan 2015) as well as nonsense and frameshift mutations (Supplemental File 4). Genes known to be involved in pigmentation were derived from a comprehensive and curated gene list (Baxter et al. 2018).

Using the four criteria described above, we developed a short list of potential candidate genes. This short list includes any genes that exhibit a statistically-significant, twofold difference in gene expression between DBA/1J and C57BL/6J skins, any genes with a percent rank of expression greater than 50% in C57BL/6J McSCs that also contain a deleterious coding mutation, and any known pigmentation genes. While this approach may selectively filter out some candidate genes that are expressed nonautonomous to the McSC and those genes that are under the influence of differential transcript regulation or post-translational modifications, this abridged list provides a reasonable starting point for future follow-up (Table 5).

**DISCUSSION**

Modifier genes are recognized for their contribution to phenotypic variation observed in disease and it’s been suggested that their identification may help predict “sensitive pathways and nodes for therapeutic intervention” (Hamilton and Yu 2012). With this in mind, we were interested in identifying naturally-occurring genetic variants in mouse that could contribute to variability in the phenotype of hair graying. Hair graying is the consequence of disrupting the regenerative activity of McSCs within the hair follicle or the function of their progeny. We developed a sensitized screen using mouse inbred lines as an unbiased approach to identify novel genes that participate in these mechanisms. Using the Tg(DctSox10) transgene as a condition to predispose mice to hair graying, we generated a list of expression values from C57BL/6J wildtype McSCs isolated from dormant (telogen-stage) hairs. These gene expression values represent the read counts per gene after normalization using the median ratio method (DESeq2). Genes that were not expressed (value = 0) were excluded and the remaining genes ranked by percent (Supplemental File 3).

Using the four criteria described above, we developed a short list of potential candidate genes. This short list includes any genes that exhibit a statistically-significant, twofold difference in gene expression between DBA/1J and C57BL/6J skins, any genes with a percent rank of expression greater than 50% in C57BL/6J McSCs that also contain a deleterious coding mutation, and any known pigmentation genes. While this approach may selectively filter out some candidate genes that are expressed nonautonomous to the McSC and those genes that are under the influence of differential transcript regulation or post-translational modifications, this abridged list provides a reasonable starting point for future follow-up (Table 5).

**DISCUSSION**

Modifier genes are recognized for their contribution to phenotypic variation observed in disease and it’s been suggested that their identification may help predict “sensitive pathways and nodes for therapeutic intervention” (Hamilton and Yu 2012). With this in mind, we were interested in identifying naturally-occurring genetic variants in mouse that could contribute to variability in the phenotype of hair graying. Hair graying is the consequence of disrupting the regenerative activity of McSCs within the hair follicle or the function of their progeny. We developed a sensitized screen using mouse inbred lines as an unbiased approach to identify novel genes that participate in these mechanisms. Using the Tg(DctSox10) transgene as a condition to predispose mice to hair graying, we generated a list of expression values from C57BL/6J wildtype McSCs isolated from dormant (telogen-stage) hairs. These gene expression values represent the read counts per gene after normalization using the median ratio method (DESeq2). Genes that were not expressed (value = 0) were excluded and the remaining genes ranked by percent (Supplemental File 3).

Using the four criteria described above, we developed a short list of potential candidate genes. This short list includes any genes that exhibit a statistically-significant, twofold difference in gene expression between DBA/1J and C57BL/6J skins, any genes with a percent rank of expression greater than 50% in C57BL/6J McSCs that also contain a deleterious coding mutation, and any known pigmentation genes. While this approach may selectively filter out some candidate genes that are expressed nonautonomous to the McSC and those genes that are under the influence of differential transcript regulation or post-translational modifications, this abridged list provides a reasonable starting point for future follow-up (Table 5).

**DISCUSSION**

Modifier genes are recognized for their contribution to phenotypic variation observed in disease and it’s been suggested that their identification may help predict “sensitive pathways and nodes for therapeutic intervention” (Hamilton and Yu 2012). With this in mind, we were interested in identifying naturally-occurring genetic variants in mouse that could contribute to variability in the phenotype of hair graying. Hair graying is the consequence of disrupting the regenerative activity of McSCs within the hair follicle or the function of their progeny. We developed a sensitized screen using mouse inbred lines as an unbiased approach to identify novel genes that participate in these mechanisms. Using the Tg(DctSox10) transgene as a condition to predispose mice to hair graying, we generated a list of expression values from C57BL/6J wildtype McSCs isolated from dormant (telogen-stage) hairs. These gene expression values represent the read counts per gene after normalization using the median ratio method (DESeq2). Genes that were not expressed (value = 0) were excluded and the remaining genes ranked by percent (Supplemental File 3).

Using the four criteria described above, we developed a short list of potential candidate genes. This short list includes any genes that exhibit a statistically-significant, twofold difference in gene expression between DBA/1J and C57BL/6J skins, any genes with a percent rank of expression greater than 50% in C57BL/6J McSCs that also contain a deleterious coding mutation, and any known pigmentation genes. While this approach may selectively filter out some candidate genes that are expressed nonautonomous to the McSC and those genes that are under the influence of differential transcript regulation or post-translational modifications, this abridged list provides a reasonable starting point for future follow-up (Table 5).
While we were particularly interested in identifying QTL that could explain reduced McSC differentiation, we found that the DBA/1J genetic background provides resistance to this cellular phenotype (Harris et al. 2015). In search of modifier genes that could mitigate McSC differentiation we performed QTL analysis on progeny derived from backcrossing (C57BL/6J x DBA/1J)F1-Tg(DctSox10)/0 animals to C57BL/6J. In brief, we identified three linkage regions across three chromosomes: chr 7, 13 and the X chromosome.

Two-QTL analysis also revealed a novel, epistatic interaction between the QTL on chr 7 and the X chromosome. The effect of the X chromosome QTL is only observed in males and functions to toggle the effects of the chr 7 QTL from one of EPM resistance to EPM susceptibility when the allele at the X chromosome QTL is C57BL/6J-derived (Figure 2c). This unique interaction helps to explain the reduced effects of the DBA/1J allele at the chr 7 QTL in males (Figure 2a), as well as sex-specific skewing of the EPM phenotype in the N2 population as a whole (Figure 1).

Likely due to the relatively small number of N2 progeny evaluated for linkage, the 1.5-LOD intervals for these QTL were large and encompassed a number of genes. These were selectively filtered to include only those with differential gene expression in whole skin between C57BL/6J and DBA/1J, those with evidence to support expression within McSCs but with altered protein function, and genes associated with pigmentation phenotypes (Table 5). In considering the function of these candidate genes, we anticipated that all three of the QTL could influence McSC behavior at any number of timepoints, and include both autonomous and non-autonomous mechanisms. In general, McSCs reside in a specialized niche within the hair follicle and are activated in coordination with hair follicle stem cells. At each new hair cycle McSCs proliferate and produce progeny that will colonize the hair bulb. These progeny cells will differentiate into melanocytes and initiate the process of melanogenesis, which includes the synthesis of melanin within melanosomes and the trafficking of these melanosomes for deposition into keratinocytes of the growing hair shaft (Osawa 2009). EPMs, like those in Tg(DctSox10) mice, are generated when McSCs within the stem cell niche do not self-renew properly and instead differentiate prematurely (Nishimura et al. 2005; Harris et al. 2013). Thus, resistance or susceptibility to EPMs could be the result of mitigating or exacerbating this process as early as when McSCs are deciding their fate or as late as the final stages of pigment production. In addition, because the initial EPM phenotype is driven by the Tg(DctSox10) transgene as a consequence of Sox10 overexpression, these QTL could also act as regulators of SOX10.

As one example, there are a number of genes within the linkage intervals for chr 7, 13 and the X chromosome that have known roles in cellular mechanisms associated with the function of the melanosome organelle. These include Oca2, Trpm1, Bloc1s5, Dtnbp1, Klf13a, and Apit/α. Since our phenotypic assessment of EPMs is dependent on the production of visual pigmentation, variability in the EPM phenotype may reflect changes in melanosome biogenesis or maturation. In particular, Trpm1 encodes a protein called transient receptor potential cation channel, subfamily M, member 1 (also known as melastatin) and is highly expressed in C57BL/6J McSCs (Table 5). TRPM1 localizes to non-melanosomal vesicles and its activity is associated with increased intracellular melanin content (Oancea et al. 2009). Analysis of mouse resequencing data demonstrates that the Trpm1 gene contains DBA/1J-related missense and frameshift coding mutations that

---

**Table 4 1.5-LOD support intervals for the QTL identified by single-QTL and two-QTL analysis**

| chr | Marker (nearest SNP) | pos (cM) | pos (bp) | lod |
|-----|---------------------|----------|----------|-----|
| 7   | c7.loc22 (CEL-7_36725559) | 22.255   | 49625007  | 6.59|
| 7   | c7.loc29 (rs6160140) | 28.255   | 66060155  | 8.57|
| 7   | c7.loc33 (rs3676254) | 33.255   | 71221998  | 6.98|
| 13  | rs3707097 | 10.799   | 34176778  | 4.05|
| 13  | c13.loc20 (rs6411274) | 21.790   | 48544161  | 3.56|
| 13  | rs6316213 | 41.589   | 92107548  | 1.90|
| X   | cX.loc35 (gnfX.076.619) | 39.169   | 96113472  | 4.28|
| X   | cX.loc43 (rs13483929) | 46.169   | 106916191 | 5.78|
| X   | cX.loc51 (rs3697198) | 56.169   | 139862370 | 4.17|

*Coordinates of 1.5 LOD intervals by chromosome (chr) provided in centimorgan position (pos (cM)) and GRCm38 chromosomal coordinate (pos (bp)) with their associated LOD score (lod).*
| CHR | ENS GENE ID | TX START | GENE SYMBOL | DBA/B6 BASE MEAN | DBA/B6 LOG2FC | DBA/B6 PADJ | B6 McSC BASE MEAN | B6 McSC %RANK | DELETERIOUS MUTATION (PROVEAN) | PIGMENT GENE |
|-----|-------------|----------|-------------|-----------------|---------------|-------------|-----------------|--------------|--------------------------------|-------------|
| 7   | ENSMUSG00000030500 | 51622005 | Slc17a6 | 0.41 | 0.85 | NA | NA | NA | yes | |
| 7   | ENSMUSG00000030450 | 56239759 | Oca2 | 0.76 | -2.02 | NA | 86.28 | 57.1 | yes | |
| 7   | ENSMUSG00000025324 | 58568245 | Atp10a | 355.93 | -0.34 | 0.3831 | 921.66 | 74.9 | missense | |
| 7   | ENSMUSG00000070526 | 59307923 | Snhg14 | 48.79 | 1.20 | 0.0485 | 201.43 | 52.4 | |
| 7   | ENSMUSG00000033510 | 62461870 | Otc7a | 31.31 | 1.97 | 0.0004 | 14.26 | 45 | |
| 7   | ENSMUSG00000030523 | 64153834 | Trpm1 | 31.66 | -0.79 | 0.2737 | 9110.73 | 97.5 | missense, frameshift | |
| 7   | ENSMUSG00000030519 | 64501705 | Apba2 | 74.38 | 1.48 | 0.0118 | 478.52 | 68.5 | |
| 7   | ENSMUSG00000030516 | 65296164 | Tjp1 | 3137.77 | 0.32 | 0.2594 | 8348.55 | 97.2 | |
| 7   | ENSMUSG00000030554 | 67730159 | Synn | 487.29 | -0.14 | 0.6855 | 4258.52 | 92.6 | |
| 7   | ENSMUSG0000005333 | 69528268 | Igf1r | 1615.04 | 0.06 | 0.8773 | 7788.35 | 96.9 | |
| 13  | ENSMUSG00000038982 | 38602708 | Boc1s5 | 498.19 | 0.08 | 0.8225 | 552.71 | 69.7 | missense | |
| 13  | ENSMUSG00000047094 | 40001881 | Ofcc1 | 20.68 | -1.40 | 0.0118 | 478.52 | 68.5 | missense | |
| 13  | ENSMUSG00000021359 | 40715674 | Tfap2a | 2043.05 | 0.18 | 0.6246 | 2590.07 | 87.4 | |
| 13  | ENSMUSG00000021367 | 42301269 | Edn1 | 315.69 | 0.39 | 0.4820 | 331.90 | 65.6 | yes | |
| 13  | ENSMUSG00000021375 | 44920786 | Dtnbp1 | 1142.13 | 0.04 | 0.9163 | 1226.71 | 78.2 | |
| 13  | ENSMUSG00000021385 | 46749086 | Kif13a | 2042.58 | 0.23 | 0.4269 | 6899.94 | 96.3 | |
| 13  | ENSMUSG00000021392 | 49421310 | Ippk | 553.23 | 0.40 | 0.2038 | 695.89 | 71.8 | yes | |
| 13  | ENSMUSG00000021403 | 54949410 | Unc5a | 70.35 | -2.55 | 0.0000 | 76.41 | 56.3 | |
| 13  | ENSMUSG00000021410 | 56438354 | Slc25a48 | 221.79 | 1.23 | 0.0054 | 92.77 | 57.6 | |
| 13  | ENSMUSG00000021461 | 63304708 | Fgfr4 | 231.28 | -1.08 | 0.0000 | 5089.12 | 94.2 | |
| 13  | ENSMUSG00000021466 | 63508327 | Fancc | 616.94 | 1.14 | 0.0003 | 581.19 | 70.2 | missense, frameshift | |
| 13  | ENSMUSG00000021462 | 65278813 | Zfp369 | 123.77 | 1.06 | 0.0003 | 911.45 | 74.7 | |
| 13  | ENSMUSG00000021594 | 67389308 | Zfp429 | 128.05 | 1.06 | 0.0003 | 581.19 | 70.2 | missense, frameshift | |
| 13  | ENSMUSG00000021514 | 67617000 | Zfp729a | 616.94 | -1.10 | 0.0000 | 1547.58 | 81 | yes | |
| 13  | ENSMUSG00000021589 | 67896000 | Srd5a1 | 1121.55 | 1.06 | 0.0003 | 911.45 | 74.7 | |
| 13  | ENSMUSG00000021626 | 71957920 | Irx1 | 882.46 | -0.46 | 0.3316 | 120.96 | 59.2 | deletion | |
| 13  | ENSMUSG00000021647 | 72628819 | Irx2 | 818.44 | -0.57 | 0.0920 | 144.65 | 60.2 | missense | |
| 13  | ENSMUSG00000021611 | 73627000 | Tert | 315.69 | 0.32 | 0.2594 | 8348.55 | 97.2 | |
| 13  | ENSMUSG00000021594 | 75869336 | Rhobtb3 | 697.60 | -0.17 | 0.5492 | 1774.22 | 82.7 | |
| 13  | ENSMUSG00000021594 | 83504033 | Mef2c | 3516.98 | -0.20 | 0.6329 | 3271.05 | 89.9 | yes | |

(continued)
| Chr | ENS GENE ID | SYMBOL | LOG2FC | PADJ  | BASE MEAN | MUTATION (PROVEAN) | TYPE |
|-----|-------------|--------|--------|-------|-----------|-------------------|------|
| 7   | ENSMUSG0000000031214 | Ophn1 | 583.13 | 0.10  | 0.8114    | 1988.49 | frameshift | yes  |
|     | ENSMUSG00000079487 | Med12  | 1154.22 | 0.12  | 0.7437    | 10169.50 | frameshift | yes  |
|     | ENSMUSG00000051159 | Cited1 | 7.43   | 2.07  | 0.0982    | 544.11  | frameshift | yes  |
| 13  | ENSMUSG00000033792 | Atp7a  | 1791.99 | 0.61  | 0.0044    | 2306.47 | frameshift | yes  |
|     | ENSMUSG00000033777 | Tlr13  | 265.61  | -1.04 | 0.0201    | 8.77    | frameshift | yes  |
|     | ENSMUSG00000079428 | Tceal7 | 172.12  | -1.71 | 0.0000    | NA      | frameshift | yes  |
|     | ENSMUSG00000057439 | Kir3dl2 | 29.28  | 0.35  | 6.6       | NA      | frameshift | yes  |
|     | ENSMUSG00000044550 | Tceal3  | 45.61   | 0.04  | 0.9784    | 168.31  | frameshift | yes  |
|     | ENSMUSG00000087368 | BC065397 | 23.56  | 0.60  | 0.4724    | 419.10  | frameshift | yes  |
|     | ENSMUSG00000042515 | Mum1l1 | 32.08   | 0.04  | 0.9784    | 168.31  | frameshift | yes  |

**Candidate gene list for the QTL on chr 7, 13, and the X chromosome.** The DBA/B6 columns refer to differential expression from RNAseq data comparing C57BL/6J and DBA/1J whole skin. Those values in bold represent genes that have both an adjusted p-value of <0.05 and a twofold change in expression in either direction. The B6 McSC columns refer to the ranked RNAseq data from C57BL/6J McSCs. Deleterious mutations were determined using PROVEAN, and the type of mutation is provided. Known pigment genes are marked with a ‘yes’. ENS gene ID, Ensembl gene ID; TX START, bp position of the transcription start site (GRCm38); BASE MEAN, mean of the normalized RNAseq counts; LOG2FC, log2 fold change; PADJ, adjusted p-value; %RANK, rank of expression for the indicated gene as a percentage of the entire genes within the dataset.

**LITERATURE CITED**

Adhikari, K., T. Fontanil, S. Cal, J. Mendoza-Revilla, M. Fuentes-Guajardo et al., 2016 A genome-wide association scan in admixed Latin Americans identifies loci influencing facial and scalp hair features. Nat. Commun. 7: 10815. https://doi.org/10.1038/ncomms10815

**ACKNOWLEDGMENTS**

We thank Arturo Incao for indispensable help with mouse colony maintenance and the remaining members of the Pavan Lab for their general support. In addition, we recognize Marypat Jones and Ursula Harper in the NHGRI Genomics Core for assistance with SNP genotyping.
Agarwal, P., M. P. Verzi, T. Nguyen, J. Hu, M. L. Ehlers et al., 2011 The MADS box transcription factor MEF2C regulates melanocyte development and is a direct transcriptional target and partner of SOX10. Development 138: 2555–2565. https://doi.org/10.1242/dev.056804
Baxter, L. L., D. E. Watkins-Chow, W. J. Pavan, and S. K. Loftus, 2018 A curated gene list for expanding the horizons of pigmentation biology. Pigment Cell Melanoma Res. (epub ahead of print) https://doi.org/10.1111/pcmr.12743
Broman, K. W., H. Wu, S. Sen, and G. A. Churchill, 2003 R/qtl: QTL mapping in experimental crosses. Bioinformatics 19: 889–890. https://doi.org/10.1093/bioinformatics/btg112
Choi, Y., and A. P. Chan, 2015 PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. Bioinformatics 31: 2745–2747. https://doi.org/10.1093/bioinformatics/btv195
Dempster, A. P., N. M. Laird, and D. B. Rubin, 1976 Maximum Likelihood from Incomplete Data via the EM Algorithm. Ann. Stat. 3(1): 1–22. https://doi.org/10.1214/aos/1176343750
Hakami, R. M., L. Hou, L. L. Baxter, S. K. Loftus, E. M. Southard-Smith et al., 2006 Genetic evidence does not support direct regulation of EDNRB by SOX10 in migratory neural crest and the melanocyte lineage. Mech. Dev. 123: 124–134. https://doi.org/10.1016/j.mod.2005.11.004
Hamilton, B. A., and B. D. Yu, 2012 Modifier genes and the plasticity of genetic networks in mice. PLoS Genet. 8: e1002644. https://doi.org/10.1371/journal.pgen.1002644
Harris, M. L., K. Buac, O. Shakhova, R. Hakami, M. Wegner et al., 2013 A dual role for SOX10 in the maintenance of the postnatal melanocyte lineage and the differentiation of melanocyte stem cell progenitors. PLoS Genet. 9: e1003644. https://doi.org/10.1371/journal.pgen.1003644
Harris, M. L., T. D. Fufa, J. W. Palmer, S. S. Joshi, D. M. Larson et al., 2018 A direct link between MITF, innate immunity, and hair graying. PLoS Biol. 16: e2003648. https://doi.org/10.1371/journal.pbio.2003648
Harris, M. L., D. J. Levy, D. E. Watkins-Chow, and W. J. Pavan, 2015 Ectopic differentiation of melanocyte stem cells is influenced by genetic background. Pigment Cell Melanoma Res. 28: 223–228. https://doi.org/10.1111/pcmr.12344
Jackson, I. J., 1997 Homologous Pigmentation Mutations in Human, Mouse and Other Model Organisms. Hum. Mol. Genet. 6: 1613–1624. https://doi.org/10.1093/hmg/6.10.1613
Jackson, I. J., 1994 Molecular and Developmental Genetics of Mouse Coat Color. Annu. Rev. Genet. 28: 189–217. https://doi.org/10.1146/annurev.ge.28.120194.001201
Keane, T. M., L. Goodstadt, P. Danecek, M. A. White, K. Wong et al., 2011 Mouse genomic variation and its effect on phenotypes and gene regulation. Nature 477: 289–294. https://doi.org/10.1038/nature10413
Koike, C., T. Obara, Y. Uriu, T. Numata, R. Sanuki et al., 2010 TRPM1 is a component of the retinal ON bipolar cell transduction channel in the mGluR6 cascade. Proc. Natl. Acad. Sci. USA 107: 332–337. https://doi.org/10.1073/pnas.0912730107
Lander, E. S., and D. Botstein, 1989 Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121: 185–199.
Nishimura, E. K., S. R. Grant, and D. E. Fisher, 2005 Mechanisms of Hair Graying: Incomplete Melanocyte Stem Cell Maintenance in the Niche. Science 307: 720–724. https://doi.org/10.1126/science.1099593
Oancea, E., J. Vriens, S. Brauchi, J. Jun, I. Splawski et al., 2009 TRPM1 forms ion channels associated with melanin content in melanocytes. Sci. Signal. 2: ra21. https://doi.org/10.1126/scisignal.2000146
Osawa, 2009 Melanocyte stem cells. StemBook.
Panhard, S., I. Lozano, and G. Loussouarn, 2012 Graying of the human hair: a worldwide survey, revisiting the ‘50’ rule of thumb. Br. J. Dermatol. 167: 865–873. https://doi.org/10.1111/j.1365-2133.2012.11095.x
Seberg, H. E., E. V. Otterloo, S. K. Loftus, H. Liu, G. Bonde et al., 2017 TFAP2 paralogs regulate melanocyte differentiation in parallel with MITF. PLoS Genet. 13: e1006636. https://doi.org/10.1371/journal.pgen.1006636
Sen, S., and G. A. Churchill, 2001 A statistical framework for quantitative trait mapping. Genetics 159: 371–387.
Steinigrsson, E., N. G. Copeland, and N. A. Jenkins, 2006 Mouse coat color mutations: from fancy mice to functional genomics. Dev. Dyn. Off. Publ. Am. Assoc. Anat. 235: 2401–2411.
Swindell, W. R., K. A. Michaels, A. J. Sutter, D. Diaconu, Y. Fritz et al., 2017 Imiquimod has strain-dependent effects in mice and does not uniquely model human psoriasis. Genome Med. 9: 24. https://doi.org/10.1016/j.mod.2005.11.004
Yalcin, B., K. Wong, A. Agam, M. Goodson, T. M. Keane et al., 2011 Sequence-based characterization of structural variation in the mouse genome. Nature 477: 326–329. https://doi.org/10.1038/nature10432
Zeng, Z. B., C. H. Kao, and C. J. Basten, 2001 Estimating the genetic architecture of quantitative traits. Genet. Res. 74: 279–289. https://doi.org/10.1017/S0016672399004255

Communicating editor: F. Pardo-Manuel de Villena