DOI: 10.1534/genetics.107.073387

Rescue From Oculocutaneous Albinism Type 4 Using Medaka *slc45a2* cDNA Driven by Its Own Promoter

Shoji Fukamachi,*,†,[1] Masato Kinoshita,‡ Taro Tsujimura,* Atsuko Shimada,§ Shoji Oda,* Akihiro Shima,§ Axel Meyer,† Shoji Kawamura* and Hiroshi Mitani*

**Department of Integrated Biosciences, University of Tokyo, Kashiwa-no-ha, Kashiwa-shi, Chiba 277-8562, Japan, †Department of Biology, University of Konstanz, D-78457 Konstanz, Germany, ‡Division of Applied Biosciences, Graduate School of Agriculture, Kyoto University, Kyoto 606-8502, Japan and §Department of Biological Sciences, University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*

Manuscript received September 12, 2007
Accepted for publication November 30, 2007

ABSTRACT

Patients and vertebrate mutants with oculocutaneous albinism type 4 (OCA4) have mutations in the *solute carrier family 45 member 2* (*slc45a2*) gene. However, there is no empirical evidence for this gene–phenotype relationship. There is a unique OCA4 mutant in medaka (*b*) that exhibits albinism only in the skin, but the mechanism underlying this phenotype is also unknown. In this study, we rescued medaka OCA4 phenotypes, in both the eyes and the skin, by micro-injection of an *slc45a2*-containing genomic fragment or *slc45a2* cDNA driven by its own 0.9-kb promoter. We also identified a spontaneous nucleotide change of 339 bp in the promoter as the *b* mutation. There are multiple transcription start sites in medaka *slc45a2*, as in its human ortholog, and only the shortest and eye-specific mRNA is transcribed with the *b* mutation. Interestingly, we further revealed a conserved pyrimidine (Py)-rich sequence of ~10 bp in the promoter by medaka–pufferfish comparative genomics and verified that it plays an indispensable role for expression of *slc45a2* in the skin. Further studies of the 0.9-kb promoter identified in this study should provide insights into the *cis/trans*-regulatory mechanisms underlying the ocular and cutaneous expression of *slc45a2*.

THE colors and patterns on animal body surfaces are often important for visual communication in the wild and are determined primarily by pigment cells (chromatophores) in vertebrates. The chromatophores are distributed in the skin, and their types, sizes, densities, and physiological activities affect these colors and patterns. Although mouse mutants have contributed greatly to our knowledge of skin- and coat-color formation (see Coat Color Genes, http://www.espcr.org/micemut/), mammals possess only one type of chromatophore, the melanocyte. In fish, up to six chromatophore types (melano-, leuco-, erythro-, xantho-, irido-, and cyanophores) have been identified, and there are two distinctive model species to which molecular genetics can be feasibly applied, the zebrafish and the medaka. Chromatophore studies in these species have successfully provided novel clues to the development, regulation, and interaction of these chromatophores (*e.g.*, Parichy *et al.* 2000; Fukamachi *et al.* 2004a; Watanabe *et al.* 2006).

We have previously reported that the gene *slc45a2* of solute carrier family 45, member 2, which was formally called antigen isolated from immuno-selected melanoma 1 (aim-1) or membrane-associated transporter protein (matp), is mutated in *b*-locus mutants of medaka (Fukamachi *et al.* 2001). This gene is also mutated in human patients with oculocutaneous albinism type 4 (OCA4), *underwhite* mice, "cream" horses, and *silver* chickens (Newton *et al.* 2001; Mariat *et al.* 2003; Gunnarsson *et al.* 2007). The phenotypes in these different species are quite similar to one another in that melanin deposition is severely suppressed in the skin, whereas the eyes become slightly pigmented during maturation. Population genetic studies have indicated that polymorphisms of human *SLC45A2* are also related to population differences in human skin color (Nakayama *et al.* 2002; Soejima *et al.* 2006; Yuasa *et al.* 2006; Graf *et al.* 2007). *slc45a2* mRNA is expressed in melanocyte/melanophore precursors during embryonic development (Fukamachi *et al.* 2001; Baxter and Pavan 2002), and the activity of tyrosinase (melanin-synthesizing enzyme) is suppressed in *underwhite* melanocytes (Costin *et al.* 2003), but not in *b* melanophores (Shimada *et al.* 2002; Fukamachi *et al.* 2004b). Although these results strongly suggest that a deficiency in slc45a2 function causes the OCA4 phenotype (possibly, via the reduced activity of tyrosinase), evidence that directly supports this gene–phenotype correlation has not yet been obtained, such as knockout/knockdown of *slc45a2* or rescue from the mutant phenotype by the wild-type allele.

There is a unique OCA4 mutant in medaka; it has amelanotic skin, but its eyes are melanized. The fish is

[1]*Corresponding author:* Lehrstuhl für Zoologie und Evolutionsbiologie, Department of Biology, University of Konstanz, D-78457 Konstanz, Germany. E-mail: shoji.fukamachi@uni-konstanz.de

traditionally called an orange-red variant and has the "*b*" allele at the *b* locus (Aida 1921; Matsumoto and Hirose 1993; Shimada *et al.* 2002). Whereas other *b*-locus mutants with the typical OCA4 phenotype (b^{r1}, b^{d2}, b^{d4}, b^{d8}, b^{g8}, b^{g21}, and b^{p}) have mutations in the protein-coding region of *slc45a2*, a mutation has not been identified from the *b* allele. Considering that the *b* embryo does not transcribe *slc45a2* in the skin (Fukamachi *et al.* 2001), the *b* mutation probably occurs in a gene-regulating region that specifically controls the cutaneous expression of *slc45a2*.

In this study, we provide empirical *in vivo* evidence for the *slc45a2*–OCA4 causal relationship. Further, we identify (1) a promoter sequence that is sufficient for the oculocutaneous expression of *slc45a2*, (2) multiple transcription start sites that are used tissue specifically, and (3) the *b* mutation that causes the skin-specific albinism. We also report an intriguing instance in which comparative genomics successfully pinpoint a functional motif in the promoter.

MATERIALS AND METHODS

Construct preparation: *Cosmid (HC19):* We screened a bacterial artificial chromosome (BAC) library (Kondo *et al.* 2002) using the AlkPhos direct labeling and detection system (Amersham Pharmacia, Piscataway, NJ) and isolated a clone that contained *slc45a2*. The clone was subcloned into the SuperCos 1 cosmid vector (Stratagene, La Jolla, CA) after *Sau*3AI partial digestion using MaxPlax Lambda Packaging Extract (Epicenter Technologies, Madison, WI). We screened for it with colony PCR and isolated a cosmid that contained all seven exons of *slc45a2* (HC19; Figure 1A).

Promoter–cDNA fusion constructs (B and b constructs): We first amplified the *slc45a2* open reading frame (ORF) of the wild-type "*B*" allele (HNI inbred) by RT–PCR and cloned the products into the pCR4-TOPO vector (Invitrogen, San Diego). Then, the *Age*I–*Xba*I interval of the pEGFP-1 vector (Clontech, Palo Alto, CA), which contains the entire enchanced green fluorescent protein (EGFP) ORF, was substituted with the *Age*I–*Spe*I interval of the *slc45a2* ORF plasmid (the *Age*I site is within the first exon of *slc45a2* and the *Spe*I site is on the pCR4-TOPO vector). We then amplified the 5′ promoter region together with the first exon of the wild-type *B* allele (Sakura) or the mutant *b* allele (AA2 inbred) by genomic PCR. The products were inserted between *Age*I (in the multicloning site of the pEGFP-1 vector) and *Age*I (within the first exon of *slc45a2*) of the *slc45a2* plasmid (Figure 1B). The *dam*-methylase-sensitive *Xba*I site in the pEGFP-1 vector was demethylated in *Escherichia coli* strain ER2925 (New England Biolabs, Beverly, MA) for digestion.

For the ligation reactions, we resolved the digested inserts/vectors by electrophoresis on 1–2% agarose gels and recovered the fragments using the QIAquick gel extraction kit (QIAGEN, Valencia, CA). The vectors were dephosphorylated with shrimp alkaline phosphatase (United States Biochemical, Cleveland), when necessary. After further purification by phenol–chloroform extraction and isopropanol precipitation, the fragments were ligated using a DNA ligation kit Ver. 2 (Takara, Berkeley, CA).

Mutated promoter–cDNA fusion constructs (B-mut and B-del constructs): We PCR amplified two fragments from the wild-type "B construct" using four primers (one of them containing the mutation to be introduced; Figure 1C), which mutually overlapped by a 20-bp sequence, 5′-TTTCTTAAAAACACCCG CCC-3′, immediately upstream of the mutation. These fragments were joined by a second PCR using the 5′-most and 3′-most primers and were used to replace the *Eco*RI–*Eco*RI interval of the B construct.

Micro-injection: All the injection constructs were purified using the Plasmid Midi kit (QIAGEN), dissolved in double-distilled water to a final concentration of 10–20 ng/ml, and used for micro-injection. The injected embryos were incubated in ~0.001% methylene blue water at 25° until hatching.

Sequencing the 5′ upstream region of the medaka *slc45a2* locus: We isolated genomic DNA from the tail fins of adult fish. The primers for genomic PCRs were designed on the basis of the *amacr*–*slc45a2* intergenic sequence, which we had determined previously (Fukamachi *et al.* 2001), with melting points ~60°, calculated with the CPrimer software (http://iubio.bio.indiana.edu:7131/soft/molbio/Listings.html). All PCRs were performed using the following parameters: 94° for 1 min; 30 cycles of 98° for 20 sec, 60° for 1 min, 72° for 1–10 min; with a final extension at 72° for 10 min. We used a presequencing kit (United States Biochemical) for the preparation of the sequencing templates and sequenced the products using a BigDye terminator cycle sequencing kit (Applied Biosystems, Foster City, CA). Sequences were assembled and compared with SeqMan II software (DNAStar, Madison, WI).

RT–PCR: We used Isogen (Nippon Gene) to isolate the total RNA from embryos and the eyes of adult fish. First-strand cDNA was synthesized with ReverTra Ace reverse transcriptase (Toyobo) following the manufacturer's protocol. All PCRs were performed with the conditions described above.

5′ RNA-ligase-mediated RACE: We used the FirstChoice 5′ RNA-ligase-mediated RACE (RLM–RACE) kit (Ambion, Austin, TX). Total RNAs were extracted from the eyes and skin of wild-type (*color interfere*; Fukamachi *et al.* 2004a) and *b* (Hd-rR) fish using TRIzol reagent (Invitrogen). Nested PCR products were resolved by electrophoresis on a 1% agarose gel. Four major bands of two different sizes (see Figure 4B) were excised, and the DNAs were recovered as described above and cloned into the pCRII-TOPO vector (Invitrogen). A total of 61 positive clones (*i.e.*, 13–16 clones from each band) were PCR amplified and sequenced.

Electrophoretic mobility shift assay: The mouse B16 melanoma cell line and the mouse L929 fibroblast cell line were cultured at 37° with 10 mm HEPES-buffered L-15 medium containing 15% fetal calf serum. For the experimental use of logarithmically growing cells, 5×10^5 cells were inoculated into a 600-ml plastic flask and incubated for 2 days.

Cells (1×10^6) were harvested and washed with phosphate-buffered saline. The cell pellet was resuspended in 400 μl of buffer A [10 mm HEPES–KOH (pH 7.8), 1.5 mm $MgCl_2$, 10 mm KCl, 0.5 mm dithiothreitol (DTT), 0.5 mm phenylmethanylsulfonyl fluoride (PMSF), 0.1% NP-40], incubated on ice for 20 min, vortexed for 1 min, and then centrifuged. The pellet was resuspended with 100 μl of buffer C [20 mm HEPES–KOH (pH 7.8), 420 mm NaCl, 1.5 mm $MgCl_2$, 0.2 mm EDTA (pH 8.0), 25% glycerol, 0.5 mm DTT, 0.5 mm PMSF], gently mixed at 4° for 20 min, and centrifuged. The supernatant was used as the nuclear extract. The protein concentration was measured using the BCA method (Pierce, Rockford, IL).

Complementary oligonucleotide probes (see Figure 6A for the sequence of one strand) for electrophoretic mobility shift assay (EMSA) were annealed, their 5′ ends were labeled with [γ-^{32}P]ATP using T4 polynucleotide kinase (Toyobo), and they were purified using the QIAquick nucleotide removal kit (QIAGEN). For the binding reaction, 4 μg of nuclear extract were used. The binding reaction was carried out in 20 μl of a mixture containing 15 mm Tris-HCl (pH 7.4), 60 mm KCl,

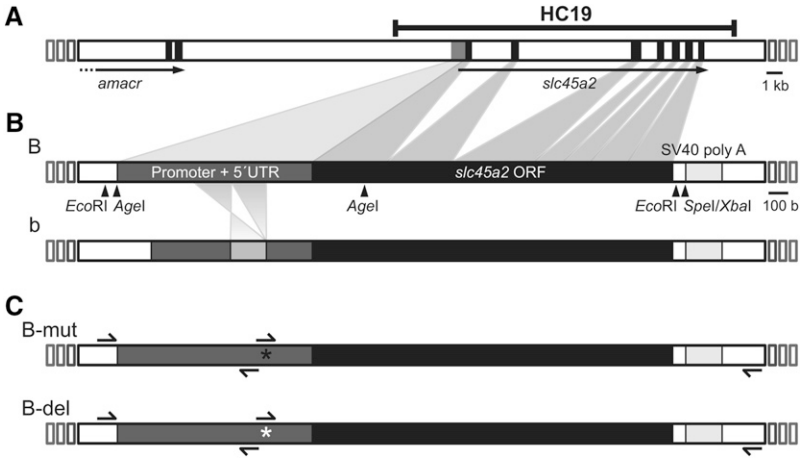

FIGURE 1.—Diagrams for injection constructs. (A) Genomic interval contained in HC19. The horizontal bar represents a part of the genomic sequence (linkage group 12), which contains the coding exons of the *slc45a2* and *amacr* genes (shown in solid boxes). (B) Structures of the promoter–cDNA fusion constructs: "B" with the wild-type promoter and "b" with the promoter of the *b* mutant containing the inv/ins/del sequence (indicated by a box with light shading). (C) Structures of the mutated promoter–cDNA fusion constructs: "B-mut" or "B-del" with a promoter in which the Py-rich motif (see Figure 5) is substituted with purines or deleted, respectively (indicated by asterisks). The positions of the restriction sites and primers used during the construct preparation (see MATERIALS AND METHODS) are indicated by arrowheads and arrows, respectively.

0.5 mM DTT, 7.5% glycerol, 0.25 μg/ml bovine serum albumin, 0.1 μg/ml poly(dI–dC), and 100 fmol of the ^{32}P-labeled probe. In a competition assay, 1 or 10 pmol (10-fold or 100-fold molar excess, respectively) of nonradiolabeled probe DNA was added to the standard mixture. After incubation at room temperature for 1 hr, the samples were loaded onto 0.5% polyacrylamide gels in 0.5× Tris–borate–EDTA buffer and resolved electrophoretically at 10 V/cm for 2 hr at room temperature. The gels were then dried, and images were obtained with a BAS-5000 imaging analyzer (Fuji Film).

RESULTS

An *slc45a2*-containing genomic fragment rescued melanin synthesis in the OCA4 medaka: We microinjected into medaka OCA4 mutant (b^{g8}) eggs a cosmid (HC19) containing all the exons/introns of medaka *slc45a2* and ~9 and 2 kb of the 5′- and 3′-flanking regions, respectively (Figure 1A). Among the *b*-locus mutants, melanin synthesis is suppressed most severely in the b^{g8} mutant (SHIMADA *et al.* 2002). We observed that 36 of 140 injected embryos exhibited chimeric deposition of melanin in the eye and/or skin (eye, 3; skin, 24; both, 9; Figure 2, A and B). Some fish retained the recovered skin melanophores until the adult stage, suggesting stable gene expression from HC19 (Figure 2, C and D). None of the mature fish successfully passed the transgene to their offspring, although one of the pairs occasionally (~10%) produced melanized but severely malformed and lethal embryos (data not shown). Therefore, HC19 must contain a gene and its regulatory regions that can rescue the medaka OCA4 phenotype.

***slc45a2* cDNA driven by its 0.9-kb promoter rescued melanin synthesis in the OCA4 medaka:** The rescuing gene is most likely *slc45a2*, but another gene might exist in HC19. A series of our published and preliminary experiments indicated that the transcription of *slc45a2* is not very strong (difficult to be detected) in medaka, *e.g.*, faint signals on whole-mount *in situ* hybridization (FUKAMACHI *et al.* 2001), no signal from embryos microinjected with an *slc45a2* promoter–EGFP plasmid (see below), no band on Northern hybridization, and no knockdown by morpholino injection (data not shown). The weak expression was similarly reported from mouse (BAXTER and PAVAN 2002), and it has also recently been reported that strong expression of *SLC45A2* may suppress melanin deposition in humans (GRAF *et al.* 2007). Therefore, we undertook to rescue the b^{g8} phenotype with *slc45a2* cDNA driven by its own promoter, rather than with a promoter for its ectopic expression.

We first sequenced the 5′-flanking region (~4.4 kb) of the medaka *slc45a2* locus and compared the sequences between two genetically divergent wild-type strains, HNI (northern inbred) and Sakura (southern noninbred). Besides many single-nucleotide polymorphisms (SNPs) and small insertions/deletions, we detected a large insertion of ~5 kb in the HNI allele 0.9 kb upstream from the translation initiation codon (Figure 3, A and B). Because the insertion of a fragment of several kilobases (*e.g.*, a transposon) into a promoter often abolishes the ordinary transcription of the gene (*e.g.*, IIDA *et al.* 2004), we expected that the region upstream from this insert may be dispensable for *slc45a2* expression.

Therefore, we constructed a plasmid in which 922 bp of the 5′-promoter region, which should also contain the 5′-untranslated region (UTR), were connected to the *slc45a2* ORF (1728 bp) and the 36 bp of the 3′-UTR (see Figure 1B). Micro-injection of the construct (B construct) successfully rescued the b^{g8} phenotype. From 127 injected eggs, nine embryos exhibited chimeric deposition of melanin in the eye and/or skin (eyes, two; skin, three; both, four; Figure 2, E and F). Therefore, the medaka OCA4 phenotype is caused by a defect in *slc45a2* and not in a neighboring gene in HC19. This result also shows that the 0.9-kb promoter can drive *slc45a2* expression at a level sufficient for melanin production in both the eyes and the skin.

The *b* mutation occurs in the promoter of *slc45a2*: Because the 0.9-kb promoter rescued both the ocular and the cutaneous phenotypes, we expected the *b* mutation (see Introduction) to occur in the promoter.

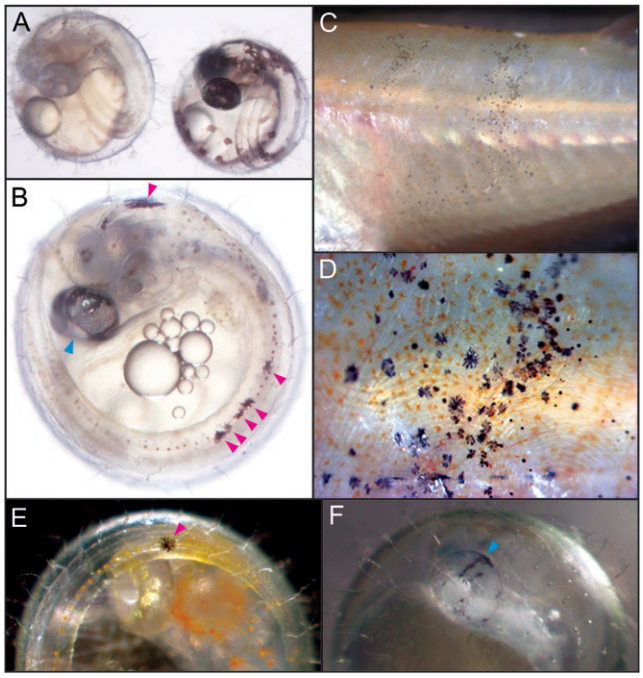

Figure 2.—Phenotypic rescue of medaka OCA4 mutants by *slc45a2* micro-injection. (A) Six-day embryos of wild-type (HNI, right) and b^{g8} (left) medaka. Melanin deposition in the eyeball, dorsal part of the head and trunk, and yolk sac is severely suppressed in b^{g8}. (B) A 6-day embryo of HC19-injected b^{g8}. Chimeric deposition of melanin is visible in the eye (light-blue arrowhead) and the dorsal part of the head and trunk (pink arrowheads). (C) Lateral view of b^{g8} adult fish injected with HC19. Rescued (melanized) melanophores are visible in clusters in the skin. (D) Higher magnification of C. Orange-colored cells are xanthophores. (E and F) b^{g8} (E) and SK^2 (F) embryos rescued by the B construct (see Figure 1B). Note that the pink- and silver-colored pigment cells (leucophores and iridophores, respectively) in b^{g8} are eliminated in SK^2 by the *leucophore-free* and *guanineless* genes (see text), which allows clearer detection of the rescued melanophores. The skin of adult SK^2 fish is also transparent, which allows the observation of the intact internal organs, as in the see-through strain (Wakamatsu *et al.* 2001).

Indeed, we found a unique nucleotide change in the promoter of the *b* mutant, an inversion of 167 bp, an insertion of 48 bp, and a deletion of 172 bp (Figure 3; see Figure 4C for details). The origin of the inserted 48 nucleotides is unknown, because the sequence cannot be aligned to neighboring sequences. This allele (inversion/insertion/deletion, inv/ins/del) was not found in at least 11 wild-type fish, which had been collected from various rivers and ponds in Japan and South Korea (Figure 3C; see Takehana *et al.* 2004 for references).

To assess the function of the *b* promoter, we substituted the 922-bp promoter of the B construct with the 795-bp promoter of the *b* mutant (b construct; Figure 1B) and micro-injected it into b^{g8} or SK^2 eggs (SK^2 is a triple mutant of the b^{g8}, *leucophore-free*, and *guanineless* loci recovered in this study; see Figure 2F). Fourteen of 284 injected embryos exhibited chimeric deposition of melanin, but only in the eyes [note that only 2/9 embryos rescued by the B construct had pigment only in the eye ($P < 0.001$, chi-square test)]. Therefore, we concluded that the *b* promoter is responsible for the mutant phenotype and that it is likely to be the inv/ins/del sequence.

Multiple transcription start sites in medaka *slc45a2*: Interestingly, we found that *slc45a2* mRNA in the *b* mutant lacks part of the 5′-UTR (Figure 4A). This indicates the existence of multiple transcription start sites (*i.e.*, the *b* mutation selectively abolishes the transcription of the full-length mRNA) or utilization of a weaker transcription start signal in the *b* mutant. Our results from RLM–RACE supported the former scenario. The wild-type medaka transcribe at least two major variants of *slc45a2* mRNA: the longer form in both the skin and the eyes and the shorter form only in the eyes (Figure 4B). The longer form was not detectable in the *b* mutant (even after 60 cycles of nested PCR), whereas the shorter form seemed to be expressed as usual. Thus, the *b* mutation specifically suppresses the transcription of the longer mRNA, which is necessary for melanization in the skin but not in the eyes, whereas the shorter mRNA is transcribed independently of the *b* mutation.

Medaka–pufferfish comparative genomics pinpoint a functional motif in the 0.9-kb promoter necessary for the cutaneous expression of *slc45a2*: The 339-bp region disrupted by the *b* mutation probably contains a functional motif, such as a transcription-factor-binding site (TFBS), that regulates the transcription of the longer form of *slc45a2* mRNA. However, we could not efficiently predict such motifs using online software, such as TFBIND (http://tfbind.ims.u-tokyo.ac.jp/) or TFSEARCH (http://molsun1.cbrc.aist.go.jp/research/db/TFSEARCH.html), because >100 TFBSs were identified within the 0.9-kb promoter. We also failed to make a prediction based on comparative genomics (phylogenetic footprinting), using human, mouse, rat, zebrafish, Takifugu, and Tetraodon genomes (Figure 5, A–C; other data not shown).

Intriguingly, however, a pairwise alignment of the medaka and pufferfish (Takifugu or Tetraodon) promoters using the mVISTA program (http://genome.lbl.gov/vista/index.shtml) revealed up to four conserved putative motifs in the 0.9-kb promoter (Figure 5C). This result allowed us to identify a pyrimidine (Py)-rich motif (motif D in Figure 5), which is located in very close proximity to the 5′-most transcription start site of *slc45a2*, at least in medaka (Figure 4A), zebrafish [expressed sequence tag (EST) CF266172 in GenBank], mouse (EST BB858311), and human (Graf *et al.* 2007). Although the sequences of the Py-rich motif are not strictly conserved among species, its functional importance in the skin is further suggested by two observations: (1) the Py-rich motif is interrupted by the 3′ edge of the *b* mutation in medaka (Figure 4A), and (2) cultured mouse B16 melanoma cells express nuclear proteins that specifically bind to the Py-rich motif (Figure 6).

To assess this hypothesis, we micro-injected promoter–cDNA constructs with a mutation exclusively in the

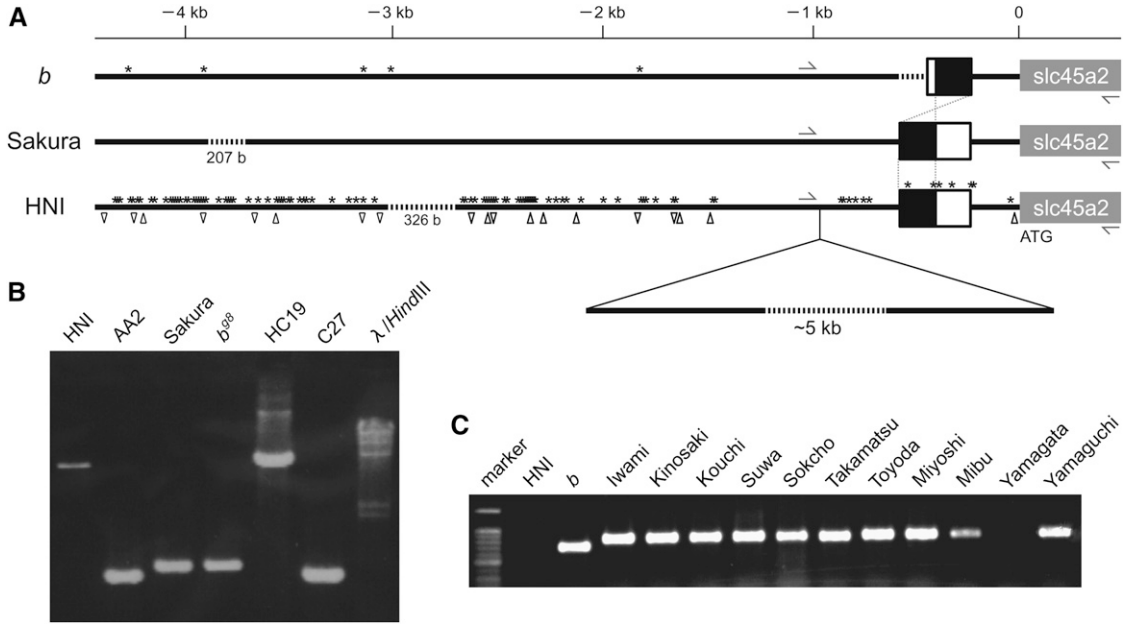

FIGURE 3.—Sequence comparison of the 5′-flanking genomic region of the medaka *slc45a2* locus. (A) The sequences of ~4.4 kb that correspond to the 5′ region contained in HC19 (see Figure 1A) of three medaka strains [two wild types (HNI and Sakura) and the *b* mutant] were compared. Shaded boxes on the right indicate the translated region of the first exon. Up to 108 SNPs and 96 allele-specific nucleotides (*i.e.*, insertions and deletions of 1–23 bp) were detected, which are indicated as asterisks, triangles, and inverted triangles. Additional deletions of 326 bp in HNI and 207 bp in Sakura are indicated by broken lines. The position of the large HNI-specific insert is indicated at the bottom. Solid and open boxes represent the inv/ins/del sequence found in the *b* allele (see text). Arrows show the approximate positions of the PCR primers used in B and C. (B) Electropherogram of genomic PCR products amplified by the primers indicated in A. Template DNAs are indicated at the top. Much longer products were amplified from HNI and the HNI-derived cosmid (HC19) and slightly shorter products were amplified from AA2 (*b*/*b* inbred) and the Hd-rR (*b*/*b* inbred)-derived cosmid (C27), relative to those amplified from Sakura and b^{g8} (b^{g8} was recovered from γ-irradiation mutagenesis of Sakura). (C) Electropherogram of genomic PCR products amplified by the same primers used in B. Template DNAs are indicated at the top. These fish are from the southern Japanese population, except for Yamagata and Sokcho, which belong to the northern Japanese and the eastern Korean populations, respectively. All the products were sequenced directly, but the inv/ins/del sequence occurred only in the *b* allele. Polymorphism information is summarized in Figure 4C.

Py-rich motif: the B-mut construct, in which the Py-rich motif (5′-CTTTCTCTCTTTCCTCTTTACT-3′) was substituted with purines (5′-GAAAGAGAGAAGGGAGAAA TGA-3′), and the B-del construct, in which the Py-rich motif was deleted (Figure 1C). The B-mut or the B-del construct rescued the OCA4 phenotype in 5 of 64 or 7 of 76 b^{g8} or SK^2 embryos, respectively, but only in the eyes [5/5 or 7/7 significantly ($P < 0.001$, chi-square test) differs from 2/9]. Thus, the Py-rich motif plays an indispensable role in melanin production in the skin, facilitating the transcription of the longer form of *slc45a2* mRNA in melanophores.

DISCUSSION

Vertebrate OCA4 mutants caused by mutations in *slc45a2*: Our results presented here provide robust *in vivo* evidence of an *slc45a2*–OCA4 causal relationship. We believe that the OCA4 phenotypes of other vertebrates are also caused by the loss of slc45a2 function and will be rescued by wild-type *slc45a2* nucleotides. However, this might not be the case in the *underwhite dominant brown* (uw^{Dbr}) mutant mouse, which has a missense point mutation in the coding region of *slc45a2* (NEWTON *et al.* 2001; DU and FISHER 2002). Dominant-negative slc45a2 protein has been proposed to explain the dominant inheritance of the uw^{Dbr} phenotype. However, the detailed function of slc45a2 protein in melanin synthesis is as yet not well understood (but see COSTIN *et al.* 2003).

There is another intriguing OCA4 mutant, *B*′ medaka, which exhibits "variegated" deposition of melanin in the adult skin (AIDA 1921). We found that the *B*′ mutant has no mutation in the coding region of *slc45a2*, but carries the inv/ins/del sequence (*i.e.*, the *b* mutation) in the promoter (our unpublished observation). The indistinguishable phenotypes of the *b* and *B*′ embryos can be explained by this finding. The mechanism for the later and incomplete enhancement of melanin deposition in the *B*′ skin remains unknown. We assume that other genomic regions in addition to the 0.9-kb promoter control the expression of medaka *slc45a2*. This is because we achieved a lower rescue frequency when using the B construct (9/127) than when using the cosmid HC19 (36/140; $P < 0.001$, chi-square test). Sequence comparisons of the entire *slc45a2* locus (*i.e.*, 5′/3′ genomic regions and introns contained in HC19; see Figure 1A) between the *b* and *B*′ alleles may reveal additional

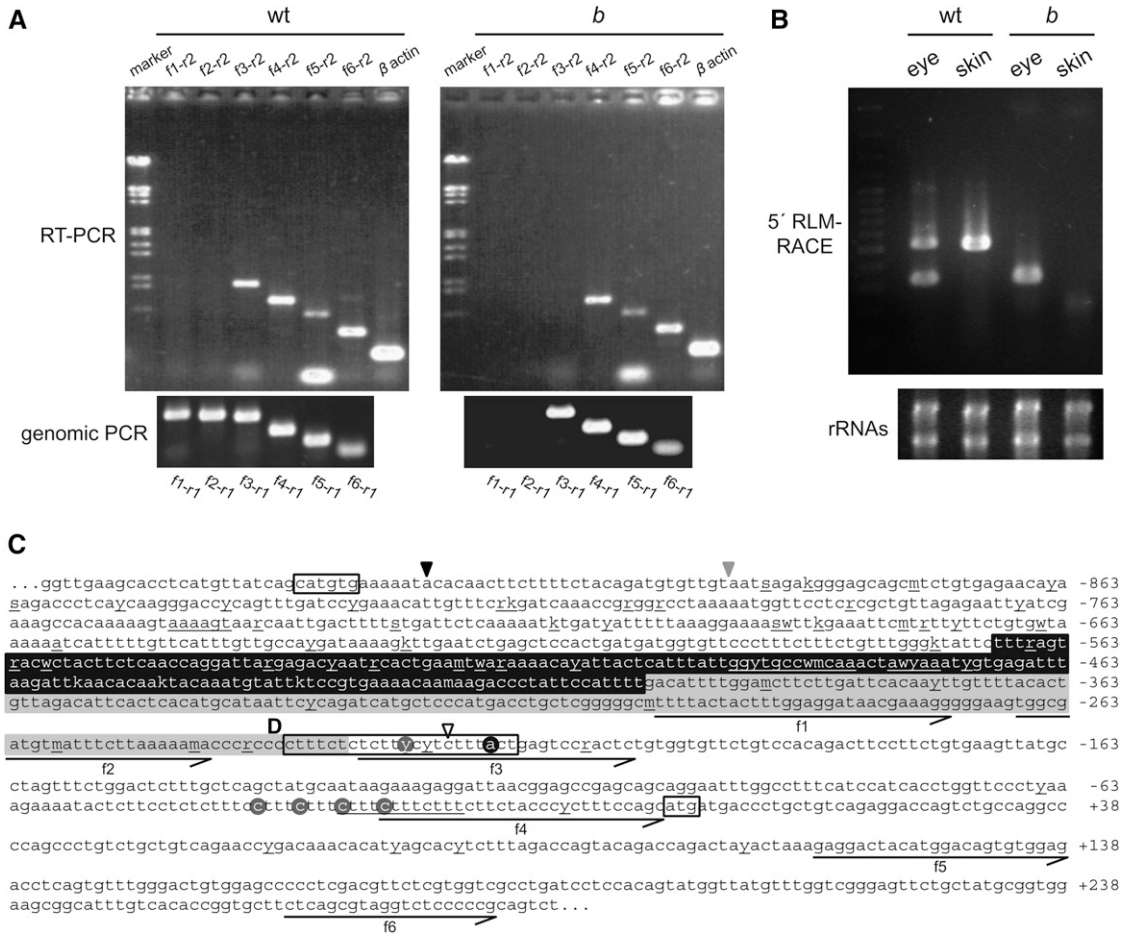

FIGURE 4.—The *b* mutation that suppresses the transcription of the longer mRNA variant of *slc45a2*. (A) Left: RT–PCR using various forward primers (f1–f6 in C) and a reverse primer (r2) designed to bind to the second exon (sequence not shown). Template cDNA was prepared from the eyes of wild-type adult medaka (HNI). Identical results were obtained using cDNAs from 1-, 2-, and 4-day HNI embryos (data not shown). The forward primers (f1–f6) and another reverse primer (r1) designed to bind to the first exon of *slc45a2* (sequence not shown) were used to perform the genomic PCR. Comparison of the RT–PCR and genomic PCR results indicates that *slc45a2* is transcribed from the region between f2 and f3. Right: the same analyses using cDNA and genomic DNA of the *b* mutant. Primers f1 and f2 did not amplify any fragment from either template, because their complementary sequences were destroyed by the *b* mutation (C). Primer f3 amplified products from the genomic DNA but not from the cDNA, indicating that *slc45a2* is transcribed from the region between f3 and f4 in the *b* mutant. Identical results were obtained using cDNAs from 1-, 2-, and 4-day *b* embryos (data not shown). (B) Electropherograms of 5′ RLM–RACE products. Two major bands of different sizes were obtained from the wild-type eyes. The longer mRNA was also transcribed in the skin, but the shorter mRNA was not. The *b* mutant lacked the longer product in both the eyes and the skin. (C) The *b* mutation. The 0.9-kb promoter sequence is shown. A solid or shaded arrowhead on top shows the 5′ end of the 0.9-kb promoter in the promoter–cDNA constructs (Figure 1) or the position of the large northern-population-specific insert (Figure 3), respectively. Residues highlighted by solid and shaded boxes are inverted and deleted sequences in the *b* allele, respectively. Underlined residues are polymorphic among the populations analyzed in Figure 3C. Residues highlighted by solid or shaded circles are transcription start sites in the skin or eyes, respectively, determined by 5′ RLM–RACE. The longer transcript is transcribed from different transcription start sites in the eyes and skin. The transcription start site of the shorter form seems not to be strictly controlled, possibly because of the repetitive sequence, $(CTTT)_n$. A transcription start site predicted by the Neural Network Promoter Prediction program (http://www.fruitfly.org/seq_tools/promoter.html) is also shown with an open arrowhead. Solid arrows indicate the positions of the PCR primers used in A. The closest E-box to the promoter (see DISCUSSION), the Py-rich motif D (see Figure 5), and the translation initiation codon are boxed.

mutations and provide new insight into the cutaneous expression of *slc45a2* in adult fish.

Medaka promoter for oculocutaneous expression of *slc45a2*: The 0.9-kb promoter isolated in this study can drive *slc45a2* transcription at a sufficient level for melanin production in both the eyes and the skin of medaka. This system might be useful in assessing the functional significance of the polymorphisms/mutations in *slc45a2 in vivo*; *i.e.*, if mammalian orthologs can similarly rescue the b^{g8} phenotype, we should be able to functionally assess the polymorphism/mutation of interest (see Introduction) by preparing constructs for micro-injection.

The promoter sequence should also be useful in investigating the molecular mechanisms that control *slc45a2* expression. Microphthalmia-associated transcription factor

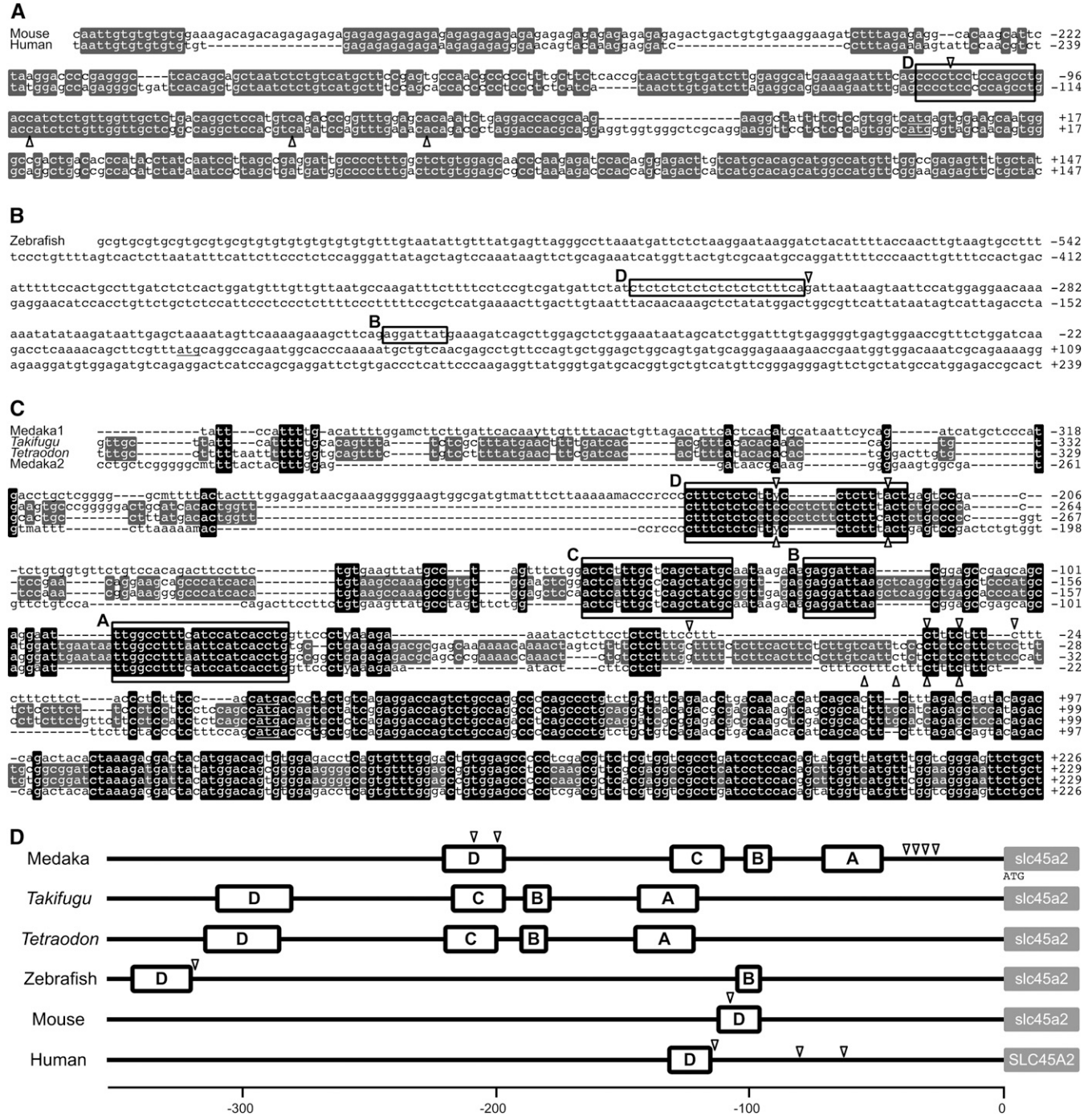

FIGURE 5.—Comparative genomic prediction of the functional motifs in the *slc45a2* promoter. (A) The promoters (~1 kb) and translated (~0.3 kb) regions of the mouse and human sequences were aligned with mVISTA (MAYOR *et al.* 2000). The translation start codon is underlined. Transcription start sites are indicated by arrowheads. Conserved nucleotides are indicated by shaded boxes. Note that many nucleotides are conserved not only in the translated region but also in the promoter region. Similar results were obtained from mouse–rat and Takifugu–Tetraodon comparisons (see C). The motif D, which was revealed by the medaka–pufferfish comparison (see C), is boxed. (B) The zebrafish sequence was only poorly aligned to any of the human, mouse, rat, Takifugu, Tetraodon, and medaka sequences (data not shown). Two of the four motifs predicted by the medaka–pufferfish comparison exist in zebrafish. (C) The medaka–pufferfish comparison. The mVISTA program aligned an identical medaka sequence differently to Takifugu and Tetraodon sequences, which are shown as Medaka1 and Medaka2, respectively. The upper regions (not shown here) rarely contain conserved nucleotides, even between the pufferfish genera. Nucleotides conserved only between the pufferfish genera are indicated by shaded boxes. Solid boxes show nucleotides conserved among the four sequences. Candidate motifs (where solid boxes of several nucleotides in length are found) are boxed (A–D). (D) Summary of the comparative genomics. The candidate motifs and translated regions are indicated with open and shaded boxes, respectively. Transcription start sites are shown with arrowheads.

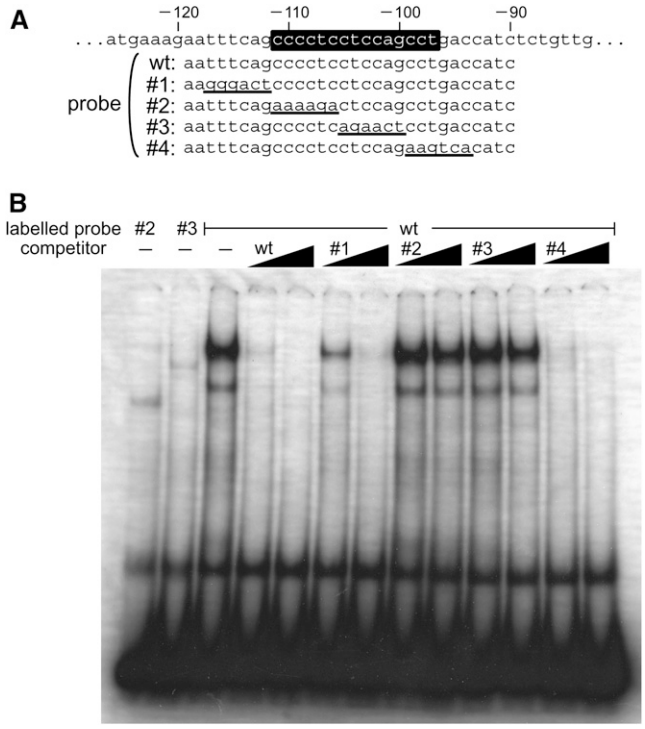

Figure 6.—Electrophoretic mobility shift assay (EMSA) using probes containing the Py-rich motif of mouse (motif D in Figure 5A) and nuclear extracts from cultured mouse melanoma cells. (A) Probe sequences. Part of the promoter sequence of mouse *slc45a2* is shown at the top. Residues are numbered with the translation start site as +1. A solid box indicates the Py-rich motif. Substituted residues in probes 1–4 are underlined. (B) EMSA results. Labeled probes and nonlabeled competitors are indicated at the top. The amount of competitor is 10× or 100× the labeled probe. Probes 1 and 4, but not 2 and 3, inhibited the binding of the labeled probe to the nuclear proteins.

(mitf) has been suggested to control *slc45a2* expression (Du and Fisher 2002). Indeed, mitf upregulates *slc45a2* transcription in both the mouse and medaka (Du and Fisher 2002; Béjar *et al.* 2003), and *mitf* mutant mice show reduced expression of *slc45a2* (Baxter and Pavan 2002). However, the 0.9-kb promoter, which does not possess the mitf-binding motif called E-box (CATGTG), rescued the OCA4 phenotype (see Figure 4C). Moreover, an immunoprecipitation assay has revealed that MITF does not bind to E-boxes 0.3–1.5 kb upstream from the translation start codon of human *SLC45A2* (Du and Fisher 2002). These observations suggest that mitf does not interact with the promoter of *slc45a2*. mitf upregulates *slc45a2* expression, probably by binding to other gene-regulating regions (as discussed above), or only indirectly by regulating the expression of other transcription factors that bind to the *slc45a2* promoter.

Medaka–pufferfish comparative genomics reveal an indispensable Py-rich motif: As an initial step in such a functional dissection of the 0.9-kb promoter, we used comparative genomics to predict TFBSs. Although phylogenetic footprinting is a good method with which to identify conserved and therefore functionally important sequences, the prediction of TFBSs in a promoter is often difficult. This is because (1) TFBSs are generally small (<10 bp); (2) sequences surrounding TFBSs are highly divergent, even in length, which disturbs the alignment of the TFBSs; (3) the TFBS itself is not strictly conserved among species; and (4) only limited numbers of genomic sequences are available, especially for vertebrates (see Boffelli *et al.* 2004; Wasserman and Sandelin 2004).

Interestingly, however, a comparison of the medaka and pufferfish genomes exceptionally predicted a functional sequence that plays an essential role, the Py-rich motif (Figure 5C). Although we do not believe that this medaka–pufferfish method will reveal a functional motif in all promoters, the recent publication of the medaka whole-genome sequence (Kasahara *et al.* 2007) should provide expanded opportunities to characterize promoter structures *in silico* before *in vitro*/*in vivo* experiments. However, the Py-rich motif should not be the only functional motif in the medaka *slc45a2* promoter; another motif that controls the shorter mRNA in the eye (Figure 4A) must also exist. To obtain a more complete understanding of the transcriptional regulation of *slc45a2*, further investigations will be necessary, including comparative genomic analyses of more species and *in vitro* screening for functional regions using deletion constructs.

We detected a nuclear protein that binds to the Py-rich motif in mouse melanoma cells (Figure 6). Identification of this protein (Stead and Mcdowall 2007) may provide new insight into *trans*-regulatory mechanisms for the cutaneous expression of *slc45a2*. However, we obtained an EMSA result identical to that shown in Figure 6 using mouse L929 fibroblast cells, which do not transcribe *slc45a2* or produce melanin (data not shown). Considering that the Py-rich motif is located very close to the transcription start site (Figure 5), it is possible that the protein that binds to the Py-rich motif is a general transcription factor that initiates transcription. To explain the tissue-specific transcription of *slc45a2*, other mechanisms must be considered, such as melanocyte-specific cofactors that activate the binding protein, another motif on the promoter to which a melanocyte-specific transcription factor binds, etc.

In summary, we rescued the medaka OCA4 phenotype with *slc45a2*, isolated the promoter sufficient for the oculocutaneous expression of *slc45a2*, revealed multiple mRNA variants that are tissue-specifically transcribed, identified the *b* mutation as the cause of the skin-specific albinism, and identified the Py-rich motif, which is essential for melanogenesis in the skin. Further studies of the 0.9-kb promoter sequence, including the Py-rich motif, will provide a more detailed understanding of the *cis*/*trans* regulatory mechanisms underlying the oculocutaneous expression of *slc45a2*.

The authors thank S. Asakawa and N. Shimizu of Keio University for the BAC library; M. Schartl of the University of Würzburg for the Hd-rR medaka; T. Suzuki of Nagoya University and S. Shibahara of Tohoku

University for advice on the EMSA; S. Takada and T. Yamashita of the University of Tokyo for fish care; and the Cell Resource Center for Biomedical Research, Institute of Development, Aging and Cancer, Tohoku University for providing the B16 and L929 cells. We also thank N. Kuzuno of the University of Tokyo for naming the SK^2 strain (Sukesuke). This work was supported by grants-in-aid for scientific research on the priority areas "Study of Medaka as a Model for Organization and Evolution of the Nuclear Genome" (no. 813) and "Comparative Genomics" (no. 015) from the Ministry of Education, Culture, Sports, Science and Technology of Japan. S.F. is supported by a research fellowship of the Japan Society for the Promotion of Science for Young Scientists (no. 17-10821) and by a long-term fellowship of the International Human Frontier Science Program Organization (no. 00059/2005-L).

LITERATURE CITED

Aida, T., 1921 On the inheritance of color in a fresh-water fish, *Aplocheilus latipes* Temmick and Schlegel, with special reference to sex-linked inheritance. Genetics **6:** 554–573.

Baxter, L. L., and W. J. Pavan, 2002 The oculocutaneous albinism type IV gene *Matp* is a new marker of pigment cell precursors during mouse embryonic development. Mech. Dev. **116:** 209–212.

Béjar, J., Y. Hong and M. Schartl, 2003 Mitf expression is sufficient to direct differentiation of medaka blastula derived stem cells to melanocytes. Development **130:** 6545–6553.

Boffelli, D., M. A. Nobrega and E. M. Rubin, 2004 Comparative genomics at the vertebrate extremes. Nat. Rev. Genet. **5:** 456–465.

Costin, G. E., J. C. Valencia, W. D. Vieira, M. L. Lamoreux and V. J. Hearing, 2003 Tyrosinase processing and intracellular trafficking is disrupted in mouse primary melanocytes carrying the *underwhite* (*uw*) mutation. A model for oculocutaneous albinism (OCA) type 4. J. Cell Sci. **116:** 3203–3212.

Du, J., and D. E. Fisher, 2002 Identification of Aim-1 as the underwhite mouse mutant and its transcriptional regulation by MITF. J. Biol. Chem. **277:** 402–406.

Fukamachi, S., A. Shimada and A. Shima, 2001 Mutations in the gene encoding B, a novel transporter protein, reduce melanin content in medaka. Nat. Genet. **28:** 381–385.

Fukamachi, S., M. Sugimoto, H. Mitani and A. Shima, 2004a Somatolactin selectively regulates proliferation and morphogenesis of neural-crest derived pigment cells in medaka. Proc. Natl. Acad. Sci. USA **101:** 10661–10666.

Fukamachi, S., S. Asakawa, Y. Wakamatsu, N. Shimizu, H. Mitani *et al.*, 2004b Conserved function of medaka pink-eyed dilution in melanin synthesis and its divergent transcriptional regulation in gonads among vertebrates. Genetics **168:** 1519–1527.

Graf, J., J. Voisey, I. Hughes and A. Van Daal, 2007 Promoter polymorphisms in the *MATP* (*SLC45A2*) gene are associated with normal human skin color variation. Hum. Mutat. **28:** 710–717.

Gunnarsson, U., A. R. Hellstrom, M. Tixier-Boichard, F. Minvielle, B. Bed'hom *et al.*, 2007 Mutations in *SLC45A2* cause plumage color variation in chicken and Japanese quail. Genetics. **175:** 867–877.

Iida, A., H. Inagaki, M. Suzuki, Y. Wakamatsu, H. Hori *et al.*, 2004 The tyrosinase gene of the *i(b)* albino mutant of the medaka fish carries a transposable element insertion in the promoter region. Pigment Cell Res. **17:** 158–164.

Kasahara, M., K. Naruse, S. Sasaki, Y. Nakatani, W. Qu *et al.*, 2007 The medaka draft genome and insights into vertebrate genome evolution. Nature. **447:** 714–719.

Kondo, M., A. Froschauer, A. Kitano, I. Nanda, U. Hornung *et al.*, 2002 Molecular cloning and characterization of *DMRT* genes from the medaka *Oryzias latipes* and the platyfish *Xiphophorus maculatus*. Gene **295:** 213–222.

Mariat, D., S. Taourit and G. Guerin, 2003 A mutation in the *MATP* gene causes the cream coat colour in the horse. Genet. Sel. Evol. **35:** 119–133.

Matsumoto, J., and E. Hirose, 1993 Deficiency of the gene *B* impairs differentiation of melanophores in the medaka fish, *Oryzias latipes*: fine structure studies. Pigment Cell Res. **6:** 45–51.

Mayor, C., M. Brudno, J. R. Schwartz, A. Poliakov, E. M. Rubin *et al.*, 2000 VISTA: visualizing global DNA sequence alignments of arbitrary length. Bioinformatics **16:** 1046–1047.

Nakayama, K., S. Fukamachi, H. Kimura, Y. Koda, A. Soemantri *et al.*, 2002 Distinctive distribution of AIM1 polymorphism among major human populations with different skin color. J. Hum. Genet. **47:** 92–94.

Newton, J. M., O. Cohen-Barak, N. Hagiwara, J. M. Gardner, M. T. Davisson *et al.*, 2001 Mutations in the human orthologue of the mouse *underwhite* gene (*uw*) underlie a new form of oculocutaneous albinism, OCA4. Am. J. Hum. Genet. **69:** 981–988.

Parichy, D. M., D. G. Ransom, B. Paw, L. I. Zon and S. L. Johnson, 2000 An orthologue of the *kit*-related gene *fms* is required for development of neural crest-derived xanthophores and a subpopulation of adult melanocytes in the zebrafish, *Danio rerio*. Development **127:** 3031–3044.

Shimada, A., S. Fukamachi, Y. Wakamatsu, K. Ozato and A. Shima, 2002 Induction and characterization of mutations at the *b* locus of the medaka, *Oryzias latipes*. Zool. Sci. **19:** 411–417.

Soejima, M., H. Tachida, T. Ishida, A. Sano and Y. Koda, 2006 Evidence for recent positive selection at the human *AIM1* locus in a European population. Mol. Biol. Evol. **23:** 179–188.

Stead, J. A., and K. J. Mcdowall, 2007 Two-dimensional gel electrophoresis for identifying proteins that bind DNA or RNA. Nat. Protoc. **2:** 1839–1848.

Takehana, Y., S. Uchiyama, M. Matsuda, S. R. Jeon and M. Sakaizumi, 2004 Geographic variation and diversity of the *cytochrome b* gene in wild populations of medaka (*Oryzias latipes*) from Korea and China. Zool. Sci. **21:** 483–491.

Wakamatsu, Y., S. Pristyazhnyuk, M. Kinoshita, M. Tanaka and K. Ozato, 2001 The see-through medaka: a fish model that is transparent throughout life. Proc. Natl. Acad. Sci. USA **98:** 10046–10050.

Wasserman, W. W., and A. Sandelin, 2004 Applied bioinformatics for the identification of regulatory elements. Nat. Rev. Genet. **5:** 276–287.

Watanabe, M., M. Iwashita, M. Ishii, Y. Kurachi, A. Kawakami *et al.*, 2006 Spot pattern of leopard *Danio* is caused by mutation in the zebrafish *connexin41.8* gene. EMBO Rep. **7:** 893–897.

Yuasa, I., K. Umetsu, S. Harihara, A. Kido, A. Miyoshi *et al.*, 2006 Distribution of the F374 allele of the *SLC45A2* (*MATP*) gene and founder-haplotype analysis. Ann. Hum. Genet. **70:** 802–811.

Communicating editor: D. J. Grunwald