Molecular analysis of reverse mutations from nonagouti (*a*) to black-and-tan (*a^t*) and white-bellied agouti (*A^w*) reveals alternative forms of agouti transcripts

Scott J. Bultman,^{1,2} Mitchell L. Klebig,¹ Edward J. Michaud,¹ Hope O. Sweet,³ Muriel T. Davisson,³ and Richard P. Woychik^{1,4}

¹Biology Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-8077 USA; ²The University of Tennessee, Oak Ridge Graduate School of Biomedical Sciences, Oak Ridge, Tennessee 37831-8077 USA; ³The Jackson Laboratory, Bar Harbor, Maine 04609 USA

The agouti gene regulates the differential production of eumelanin (black or brown) and phaeomelanin (yellow) pigment granules by melanocytes in the hair follicles of mice. The original nonagouti (*a*) allele, which confers a predominantly black coat color, has been shown to revert to two other more dominant agouti alleles, black-and-tan (*a^t*) and white-bellied agouti (*A^w*), with an exceptionally high frequency. The *a^t* and *A^w* alleles confer phenotypes in which the pigmentation is not uniformly distributed over the dorsal and ventral surfaces of the animal; in both cases the ventral surface of the animal is markedly lighter than the dorsal surface due to an increase in phaeomelanin production. To understand the unusually high reversion rate of *a* to *a^t* or *A^w*, and to decipher the molecular events associated with the different pigmentation patterns associated with these three agouti alleles, we have characterized *a*, *a^t* and *A^w* at the molecular level. Here, we report that insertions of 11, 6, and 0.6 kb are present at precisely the same position in the first intron of the agouti gene in *a*, *a^t*, and *A^w*, respectively. The *a* insertion consists of a 5.5-kb VL30 element that has incorporated 5.5 kb of additional sequence internally; this internal sequence is flanked by 526-bp direct repeats. The *a^t* allele contains only the VL30 element and a single, internal 526-bp repeat. The *A^w* allele has only a solo VL30 LTR. Based on the comparison of the structure of the *a^t* and *A^w* insertions, we propose that reverse mutations occur by excision of inserted sequences in *a* through homologous recombination, utilizing either the 526-bp direct repeats to generate *a^t* or the VL30 LTRs to generate *A^w*. Moreover, the analysis of these three alleles has allowed us to identify additional exons of the agouti gene that give rise to alternatively processed forms of agouti mRNA. We demonstrate that the distinct insertions in *a*, *a^t* and *A^w* cause pigmentation differences by selectively inactivating the expression of different forms of agouti transcripts.

[Key Words: agouti locus; *a*, *a^t*, *A^w*; reverse mutations; VL30; homologous recombination]

Received November 29, 1993; revised version accepted December 22, 1993.

The agouti locus in chromosome 2 of the mouse controls the differential production of eumelanin (black or brown, depending on alleles present at other loci) and phaeomelanin (yellow) pigment granules by the melanocytes within the hair follicle. Agouti pigmentation, which is the true wild-type coat color of mice, arises from the fact that the individual hair shafts contain a subapical yellow band on an otherwise black or brown background. At least 19 different alleles have been identified at the agouti locus that form a complex dominance hierarchy in which alleles associated with phaeomelanin production are generally dominant over alleles associated with eumelanin production (Silvers 1979; Green 1989). More-

over, a number of these alleles regulate the amount and distribution of phaeomelanin and eumelanin differently in different regions of the pelage, particularly across the dorsal and ventral surfaces (Silvers 1979; Green 1989).

The regional pigmentation differences observed with different agouti alleles are exemplified by the original nonagouti (*a*), black-and-tan (*a^t*), and white-bellied agouti (*A^w*) alleles. The *a/a* genotype gives rise to a predominantly black coat color, with relatively small amounts of yellow pigmentation restricted to the pinnae, mammae, and perineum. Nonagouti is recessive to the *A^w*, wild-type (*A*), and *a^t* alleles. The *a^t/a^t* genotype causes a unique pattern of pigmentation in which the dorsal surface is entirely black, the ventral surface and pinnae are completely yellow, and a sharp lateral line of demarcation separates the dorsal and ventral surfaces.

⁴Corresponding author.

Black-and-tan is dominant to *a* but recessive to *A^w*. It is dominant to *A* across the ventral surface of the animal, consistent with yellow being dominant to agouti, but is recessive to *A* across the dorsal surface of the animal, which is compatible with the observation that agouti is dominant to black. White-bellied agouti, which confers an agouti dorsal surface and a cream color over the ventral surface, is dominant to the *A*, *a^t*, and *a* alleles.

The *a*, *a^t*, and *A^w* alleles are also of interest because reverse mutations from *a* to *a^t* or *A^w* are more common than any other spontaneous agouti locus mutation (Dickie 1969). The incidence of 4.7×10^{-6} per generation (3.29 for *a^t* and 1.41 for *A^w*) represents one of the highest reverse mutation frequencies documented in the mouse (Schlager and Dickie 1967). This high reversion frequency implies that there may be some unusual structural features associated with the *a* mutation.

We and others previously cloned the agouti gene and determined that it contains four exons, the last three of which are coding (Bultman et al. 1992; Miller et al. 1993). The gene produces a 0.8-kb mRNA in *A/A* mice that encodes a novel 131-amino-acid protein that includes a putative signal peptide. The *A* protein has been hypothesized to be a secreted signaling molecule that functions as an antagonist to the α -melanocyte-stimulating hormone receptor on the surface of melanocytes (Bultman et al. 1992; Jackson 1993). Because the structure and expression of the agouti gene are altered by the most dominant agouti locus mutation (lethal yellow, *A^y*), mutations in the middle of the dominance hierarchy (*a* and *a^t*), and the most recessive member of the agouti allelic series (extreme nonagouti, *a^e*), we proposed that there is only one gene associated with agouti locus function (Bultman et al. 1992).

Here, we report that the *a*, *a^t*, and *A^w* alleles are associated with DNA insertions at precisely the same position in the first intron of the agouti gene. A portion of each inserted fragment corresponds to a member of the VL30 family of retrovirus-like transposable elements in the mouse. We describe how homologous recombination events within the integrated sequences of the *a* allele are associated with the reversions from *a* to *a^t* or *A^w*. Additionally, we have discovered that the agouti gene produces alternatively processed transcripts that are expressed in the *a^t* and *A^w* alleles but not in the *A* allele. We propose that the insertions in *a*, *a^t*, and *A^w* confer their unique coat-color phenotypes by selectively inactivating the expression of the different processed forms of agouti mRNA.

Results

Structural DNA rearrangements associated with the *a*, *a^t*, and *A^w* alleles

As part of our recent molecular characterization of the agouti gene, we identified DNA insertions within a 700-bp interval in the first intron both in *a* and in an *a^t* mutation called *a^{t-2Gso}* (formerly referred to as SB B + T; Bultman et al. 1992). This 700-bp region is defined by

EcoRI and *BglIII* restriction sites and is illustrated in Figure 1A. More recently, we determined that a new *A^w* mutation, called *A^{w-38f}*, also contains a DNA insertion in this same 700-bp interval (see below). To elucidate the molecular nature of the inserted sequences, overlapping DNA fragments spanning the insertions were cloned from the *a*, *a^{t-2Gso}*, and *A^{w-38f}* alleles. In all cases, the structure of the mutant fragments was compared with the corresponding cloned wild-type (*A*) region (Fig. 1). These analyses demonstrated that insertions of ~11, 6, and 0.6 kb are present in the first intron of the *a*, *a^{t-2Gso}*, and *A^{w-38f}* alleles, respectively (Fig. 1).

DNA sequence analysis of the cloned fragments revealed that the ends of the inserted DNA are identical in *a*, *a^{t-2Gso}*, and *A^{w-38f}* and correspond to the 5' and 3' regions of the long terminal repeat (LTR) of a retrovirus-like VL30 element (Figs. 1 and 2). Restriction mapping and DNA sequencing analyses of the inserted regions indicated that *a* and *a^t* each contain a complete 5.5-kb VL30 element, nearly identical to the element already reported (Adams et al. 1988; Hodgson et al. 1990) (Fig. 1). The VL30 sequences within both *a* and *a^{t-2Gso}* each contain an additional DNA segment that is unique to each allele (see below), whereas in *A^{w-38f}* only a single LTR is present (Fig. 1). The VL30 sequences in all three alleles are integrated in the opposite transcriptional orientation relative to agouti (Figs. 1 and 2).

For both the *a* and *a^{t-2Gso}* mutations, sequence analyses indicated that the additional segment of DNA within the VL30 element interrupts the VL30 sequence between nucleotides 2473 and 2474 of the published sequence (Hodgson et al. 1990; Fig. 1). The size of this additional sequence was determined to be 5.5 kb for the *a* allele and 526 bp for the *a^{t-2Gso}* allele. The complete nucleotide sequence of the 526-bp insert within the VL30 element of *a^{t-2Gso}* was determined and compared with the sequence of selected regions from the *a* allele. This analysis revealed that the extra 5.5-kb insert within the VL30 sequence in *a* contains direct terminal repeats of 526 bp in length, each of which corresponds to the single 526-bp insert in the *a^{t-2Gso}* allele (Fig. 1). Comparison of the sequence of this 526-bp direct repeat with GenBank indicated that it is 90% identical to a region between two T-cell receptor (TCR) γ variable region gene segments, V10.8A and V10.8B, which lie 3.5-kb apart in an opposite transcriptional orientation (Hayday et al. 1985; Garman et al. 1986) (Fig. 3). The functional significance of this homology is presently unclear because the 526-bp region does not correspond to a defined regulatory element or *cis*-acting sequence responsible for somatic recombination events that produce functional TCRs. Furthermore, hybridization of the 526-bp region to total mouse DNA revealed that it is repetitive in the mouse genome, suggesting that the 526-bp region is a member of an as yet uncharacterized family of DNA repetitive elements (data not shown). The nature of the other sequences within the inserted 5.5-kb of DNA in the *a* allele is currently under investigation.

Characterization of the sequence flanking the inserted VL30 sequences for the *a*, *a^{t-2Gso}*, and *A^{w-38f}* alleles

Figure 1. Identification of DNA insertions in the first intron of the agouti gene in the *a*, *a^t*, and *A^w* alleles. (A) Schematic representation of the genomic structure of the wild-type agouti gene with the 5' end enlarged immediately below. The four exons are depicted as solid boxes, and introns and flanking sequences are drawn as a solid, horizontal line. The insertion site in *a*, *a^t*, and *A^w* are indicated by a vertical arrow. The positions of two probes, 1.5 and 0.6, are indicated. (B) *Bam*HI, (R) *Eco*RI, (Bg) *Bgl*III. (B) Schematic representation of the 5' end of the agouti gene in the *a* and *a^{t-2Gso}* mutations. Insertions of 11 kb in *a* and 6 kb in *a^t* are represented by rectangles. The wide portion of the rectangles represents VL30 sequence with crosshatched regions corresponding to the LTRs. Numerals refer to nucleotides within VL30 in the 5' → 3' orientation and are based on the published sequence (Hodgson et al. 1990). The narrow portion of the rectangles represents additional sequence integrated between positions 2473 and 2474 of VL30. This additional sequence is ~5.5 kb in *a* and is bound by 526-bp direct repeats (shaded region with arrows above). All of this additional sequence is absent in the *a^t* insertion, except for one of the 526-bp direct repeats (indicated by shaded box). (C) Schematic representation of the 5' end of the agouti gene in the *a* and *A^{w-38f}* alleles. All of the 11-kb insertion in *a* is absent in the *A^w* allele, except for one VL30 LTR (indicated by the hatched box).

showed that the integrated DNA occurs at precisely the same position in each case (Fig. 2). Furthermore, a duplication of four nucleotides of host DNA at the integration site, and the fact that 2 bp are missing from the end of each LTR, suggests that the integration of VL30 into the agouti gene occurred through the same mechanism that is normally associated with the integration of retroviral sequences (Dhar et al. 1980; Shimotohno et al. 1980; Van Buren et al. 1980; Varmus 1982). This finding is in accord with VL30 integrations that have been characterized previously (Hodgson et al. 1983; Itin and Keshet 1983).

Molecular analysis of several new *a^t* and *A^w* reverse mutations

Based on the analysis of the *a^{t-2Gso}* and *A^{w-38f}* mutations,

we hypothesize that each arose by excision of inserted DNA from the *a* allele through homologous recombination. In the case of *a^{t-2Gso}*, we predict that homologous recombination between the 526-bp direct repeats occurred to remove most of the additional 5.5 kb of inserted sequence, and in *A^{w-38f}* we predict that homologous recombination between the VL30 LTRs removed most of the 11 kb of inserted sequence present in the *a* allele (Fig. 1). If homologous recombination is playing a role in the reversion from *a* to *a^t* or *A^w*, one would expect several independent *a^t* and *A^w* mutations that arise from *a* to exhibit the same structural alterations as those observed for *a^{t-2Gso}* and *A^{w-38f}*, respectively. To test this possibility, we analyzed DNA from 21 new mutant animals that arose in the C57BL/6J (*a/a*) colony at The Jackson Laboratory over the past 2 years. Eleven of these new

Figure 2. Nucleotide sequence of a region of DNA spanning the insertion site in *A*, *a*, *a^{t-2Gso}*, and *A^{w-38J}*. Wild-type agouti intron sequence is shown 5' → 3' in uppercase letters. The sequences at the ends of the insertions in *a*, *a^t*, and *A^w* are identical and correspond to the LTRs of a VL30 provirus, as shown below in lowercase letters. The shaded box highlights a 4-bp duplication of agouti sequence at the VL30 integration site in *a*, *a^t*, and *A^w*.

mutants exhibited a phenotype identical to *a^t*, and 10 others displayed a phenotype identical to *A^w* (Table 1). All of these mutations pass through the germ line and presumably represent spontaneous reverse mutations from *a* to *a^t* or *A^w*.

Figure 3. Nucleotide sequence comparison between the 526-bp direct terminal repeats in the *a* and *a^t* insertions (shaded region of the narrow rectangles in *a* and *a^t* in Fig. 1, top) and part of an intergenic regulatory region between two TCR γ variable gene segments (bottom). The sequence is presented in the same orientation as the arrows shown above the 526-bp repeats in Fig. 1B. The three capitalized nucleotides at the 5' end of the 526-bp repeat are present in *a^t* and in the repeat apposed with nucleotide 2474 of VL30 in *a* but are absent in the repeat apposed with nucleotide 2473 of VL30 in *a*.

Genomic DNA from each of the original mutant animals listed in Table 1 was analyzed utilizing probe 0.6, which lies just 3' to the insertion site in the *a*, *a^{t-2Gso}*, and *A^{w-38J}* mutations (see Fig. 1). This probe hybridizes to *Bgl*II fragments of 6, 7, and 2.6 kb in DNA from the *a*, *a^{t-2Gso}*, and *A^{w-38J}* mutations, respectively (Figs. 4 and 5). As expected, DNA from each of the 11 new *a^t* mutant animals contained a *Bgl*II fragment of 7 kb, which is identical in size to that of *a^{t-2Gso}* (Fig. 4). Similarly, DNA from each of the 10 new *A^w* mutant animals contained a *Bgl*II fragment of 2.6 kb, which is identical in size to that of *A^{w-38J}* (Fig. 5). The agouti alleles in the inbred strains 101 (*A^w/A^w*) and C3H (*A/A*), in contrast, are associated with a 2-kb *Bgl*II fragment, indicating the absence of any inserted sequence (Fig. 5). Further analysis of genomic DNA from each of the 21 new reversion mutants indicated that they also have the same size-altered *Bam*HI fragments as those observed for the *a^{t-2Gso}* and *A^{w-38J}* mutations when hybridized with a probe that lies just 5'

Table 1. New *A^w* and *a^t* mutations analyzed by Southern analysis

Remutation number	Strain of origin
<i>A^{w-34J}</i>	C57BL/6J
<i>A^{w-35J}</i>	B10.BL/SgSnJ
<i>A^{w-36J}</i>	C57BL/6J
<i>A^{w-37J}</i>	C57BL/6J
<i>A^{w-38J}</i>	C57BL/6J
<i>A^{w-39J}</i>	C57BL/6JIN-Sp
<i>A^{w-40J}</i>	C57BL/6J
<i>A^{w-41J}</i>	C57BL/KsJ- <i>dbm</i>
<i>A^{w-42J}</i>	C57BL/6J
<i>A^{w-43J}</i>	C57BL/6J
<i>A^{w-44J}</i>	C57BL/6J
<i>a^{t-46J}</i>	C57BL/6J
<i>a^{t-47J}</i>	C57BL/KsJ- <i>dbm</i>
<i>a^{t-48J}</i>	C57BL/6J
<i>a^{t-49J}</i>	C57BL/6J
<i>a^{t-50J}</i>	C57BL/6J
<i>a^{t-51J}</i>	C57BL/6J
<i>a^{t-52J}</i>	C57BL/6J
<i>a^{t-53J}</i>	C57BL/6J
<i>a^{t-54J}</i>	C57BL/6J
<i>a^{t-55J}</i>	C57BL/KsJ- <i>dbm</i>
<i>a^{t-56J}</i>	C57BL/6J × DBA/2J

Figure 4. Southern blot analysis of reverse mutations of *a* to *a^t*. Genomic DNA prepared from the *a^{t-2Gso}* stock (*a^t/a^t*), nine of the new *a^t* mutants summarized in Table 1 (*a^t/a*), and the C57BL/6J genetic background on which eight of these mutations occurred on (*a/a*) were digested with *Bgl*II, blotted, and hybridized with a ³²P-labeled fragment of DNA corresponding to probe 0.6 (see Fig. 1). Sizes of the DNA fragments detected by probe 0.6 are shown (right). The *a^{t-46f}* and *a^{t-47f}* mutants display the same 7-kb *Bgl*II fragment as the other *a^t* mutants (data not shown). The signal corresponding to the 7-kb *a^t* fragment is faint relative to the 6-kb *a* fragment in *a^{t-48f}* due to mosaicism.

to the DNA insertions (see probe 1.5 in Fig. 1; data not shown). Based on these data, we propose that all of the new reverse mutations contain the same DNA structural alterations that are present in the *a^{t-2Gso}* and *A^{w-38f}* mutations. Furthermore, because we analyzed each of the “original” mutants we were able to determine that the *a^t* reversions arose through intrachromosomal recombination. If the reversions had occurred through interchromosomal recombination, then we would have detected a size-altered restriction fragment associated with the reciprocal recombination product on the other chromosome, and we did not (data not shown). Because of the relatively large size of the DNA segment that is excised from *a* in *A^w* revertants, we have been unable to identify a diagnostic restriction enzyme necessary to perform a similar analysis for *A^w*.

Additional forms of agouti transcripts

As part of our recent molecular analysis of the lethal-yellow (*A^y*) mutation, we determined that *A^y* produces chimeric transcripts in which the first noncoding exon of agouti is replaced by the noncoding first exon of a closely linked gene called *Raly* (hnRNP associated with lethal yellow; Michaud et al. 1993). During the course of this analysis, a number of *A^y* cDNA clones were isolated that contained differentially spliced segments, 111 and 46 bp in length, integrated between the *Raly* and agouti exons (Michaud et al. 1993). These segments have been mapped to a region immediately 3' to a deletion breakpoint associated with the *A^y* mutation, which positions them ~110 kb upstream from the first coding exon of agouti (Michaud et al. 1994). We initially attributed these segments to cryptic splicing events associated with the *A^y* allele but could not rule out the possibility that the 111- and 46-bp segments actually represent additional 5' exons of the agouti gene.

Here, we demonstrate that the 111- and 46-bp segments are differentially incorporated into agouti-specific transcripts that are structurally different from the agouti

mRNA originally described (Bultman et al. 1992; Miller et al. 1993). The RNAs that incorporate the 111- and/or 46-bp segments, here referred to as form II transcripts, were detected with a reverse transcriptase–polymerase chain reaction (RT–PCR) assay using an oligonucleotide primer from the 111-bp region (primer 1) coupled with a primer from the third agouti exon (primer 3; Fig. 6A). These primers produced PCR products of 380 and 334 bp in size, depending on the presence or absence of the 46-bp exon in skin from the ventrum, but not the dorsum, of 4- and 5-day-old *A^w/a*, *a^t/a^t*, and *A^w/A^w* mice (Fig. 6B). Both of the form II transcripts are also present in *A^y/a* neonatal skin (and in adult tissues; see Michaud et al. 1993), which was included as a positive control. In contrast, neither of the form II transcripts could be detected as ethidium-stained RT–PCR fragments in the ventrum or the dorsum of age-matched neonatal *A/A* or *a/a* mice (Fig. 6B), which explains why these forms of agouti mRNA were not detected originally (Bultman et al. 1992).

Additionally, a similar RT–PCR assay was utilized to detect the presence of agouti transcripts that incorporate the 5' noncoding exon that we originally identified as the first agouti exon (Bultman et al. 1992). This form of agouti mRNA, here referred to as the form I transcript, was detected with a primer from the originally defined noncoding first exon (primer 2) coupled with same third exon primer as above (primer 3). With these primers, form I mRNA produced a single 315-bp PCR product of the expected size (Fig. 6A). Form I transcripts were detected in both the dorsum and the ventrum of 4- and 5-day-old *A/A*, *A^w/a*, and *A^w/A^w* mice, but not in the dorsum or the ventrum of age-matched *a/a* or *a^t/a^t* mice (Fig. 6C).

Overall, these RT–PCR analyses demonstrate that *A/A* mice produce form I transcripts in both their dorsal and ventral skin but do not produce form II transcripts. Conversely, *a^t/a^t* animals do not express form I transcripts and express form II transcripts but only in ventral skin and not in dorsal skin. Animals carrying the *A^w* allele (derived either from *A^{w-42f}* or 101) express both the form I and the ventral-specific form II transcripts, and *a/a* animals do not express detectable levels of either the form I or the form II transcripts (Table 2).

Discussion

DNA insertions in *a*, *a^{t-2Gso}*, and *A^{w-38f}*

As part of our recent molecular characterization of the agouti gene, we determined that the *a* and *a^{t-2Gso}* mutations each contain an extra fragment of DNA within an intervening sequence of the gene (Bultman et al. 1992). More recently, we determined that a new *A^w* mutation, called *A^{w-38f}*, also contains an extra fragment of DNA within this same region. Based on the analysis of genomic clones that span these inserted sequences, we now report that the insertions are 11, 6, and 0.6 kb in *a*, *a^{t-2Gso}*, and *A^{w-38f}*, respectively, and that they are all integrated at precisely the same position. The 11-kb inser-

Figure 5. Southern blot analysis of reverse mutations of *a* to *A^w*. Genomic DNA prepared from the 11 new *A^w* mutants described in Table 1 (*A^w/a*), the C57BL/6J genetic background on which 9 of these mutations occurred (*a/a*), and the inbred strains 101 (*A^w/A^w*) and C3H (*A/A*) were digested with *Bgl*III, blotted, and hybridized with a ³²P-labeled fragment of DNA corresponding to probe 0.6 (see Fig. 1). Sizes of the DNA fragments detected by probe 0.6 are shown (right). The signal corresponding to the 2.6-kb *A^w* fragment is faint relative to the 6-kb *a* fragment in *A^{w-35J}*, *A^{w-39J}*, and *A^{w-41J}* due to mosaicism. *A^{w-41J}*-1 (*A^w/a*) is a nonmosaic offspring of the original *A^{w-41J}* mutant that was included because the 2.6-kb *A^w* fragment is more apparent.

tion in *a* includes an entire 5.5-kb VL30 element plus an internal 5.5 kb of additional sequence flanked by direct terminal repeats of 526 bp. The 6-kb insertion in *a^{t-2Gso}* also includes an entire VL30 element but contains only one 526-bp direct repeat internally. The 0.6-kb insertion in *A^{w-38J}* corresponds to a single, intact VL30 LTR.

VL30 elements are retrovirus-like structures, ~5.5 kb in size, that are dispersed throughout rodent genomes at ~100–200 copies (Besmer et al. 1979; Keshet et al. 1980; Keshet and Itin 1982). Although VL30 elements have not been shown to encode any polypeptides and are replication-defective, RNA copies of their genome are capable

Figure 6. Identification of two forms of agouti transcripts by RT-PCR. (A) A schematic map of the agouti gene showing the approximate locations of exons (boxes labeled A, B, I, II, III, and IV) and oligonucleotide primers (arrows labeled 1, 2, and 3) used in the PCR analysis. The map is shown 5' → 3' (left–right), and the position of the chromosome 2 centromere is indicated (left); Michaud et al. 1994. The size of each exon or portion of exon amplified with the indicated primers is given in bp enclosed in parentheses (above). Each of the boxes, labeled A and B, may consist of one or more exons (see Discussion). The RT-PCR products detected are shown below the map. Primers 1 and 3 amplify 380- and 334-bp portions of form II transcripts that consist of the exons indicated, whereas primers 2 and 3 amplify a 315-bp segment of the form I transcript comprised of the indicated exons. (B, C) RT-PCR products amplified with primers 1 and 3 (B) or primers 2 and 3 (C) from total RNA treated with (+) or without (–) reverse transcriptase. All of the RNA samples were derived from the dorsal (D) or ventral (V) skin of 4- to 5-day-old neonatal mice of the following genotypes and strains: *A/A* (C3H/RI), *A^w/a* (*A^{w-42J}* stock), *a/a* (C57BL), *a^t/a^t*, *A^w/A* (101 × C3H), and *A^v/a*. Shown are photographs of RT-PCR products size-fractionated in agarose gels and stained with ethidium bromide. Numbers at left indicate the sizes in bp of the 1-kb ladder molecular weight marker (lane M); numbers at right indicate the sizes of the RT-PCR products. The 520-bp fragment was amplified using oligonucleotide primers specific for the β-actin mRNA and serves as a positive control. Although the 380-bp fragment of the form II transcript is not readily visible in the ventral skin of *A^w/A* (B), it and the other agouti-specific fragments hybridized to an agouti cDNA probe after the gel was Southern blotted to a nylon membrane (data not shown).

Table 2. Dorsal (D) and ventral (V) distribution of form I and II transcripts in neonatal skin from mice with different agouti genotypes

Genotype	Form I	Form II
<i>A/A</i>	D, V	—
<i>A^w/A</i>	D, V	V
<i>A^w/a</i>	D, V	V
<i>a^t/a^t</i>	—	V
<i>a/a</i>	—	—

of being copackaged by endogenous murine leukemia viruses. They can be retrotransposed when the pseudovirus infects new cells and the VL30 RNA undergoes reverse transcription and integrates into the host cell genome (Scolnick et al. 1979; Carter et al. 1986). Because retrovirus-like elements, particularly intracisternal A-particles, have been shown to affect the expression of numerous genes, even when integrated in the opposite transcriptional orientation (Kuff and Lueders 1988), we propose that the inserted VL30 elements in the *a* and *a^t* insertions are responsible for the alterations in agouti gene expression [see below].

Reverse mutations from *a* to *a^t* and *A^w* are mediated by homologous recombination

Comparison of the structure of the *a*, *a^t*, and *A^w* insertions suggests a model that explains the high rate of reverse mutations from *a* to *a^t* or *A^w*. Based on the finding that the VL30 element in *a^{t-2Gso}* contains a single copy of the 526-bp direct repeat associated with *a*, we propose that the reversion from *a* to *a^t* occurs by excision of 5 kb of DNA (the sequence integrated in the VL30) from the *a* allele, brought about by homologous recombination between the 526-bp direct terminal repeats. Likewise, because the insertion in *A^{w-38f}* is comprised of only a single VL30 LTR, we propose that the reversion from *a* to *A^w* occurs by excision of the entire insertion in *a* except for one VL30 LTR, brought about by homologous recombination between the VL30 LTRs. This model is supported by the fact that each of 11 new *a^t* mutations and 10 new *A^w* mutations appear to have the same structural rearrangements as those found in the *a^{t-2Gso}* or *A^{w-38f}* mutations. Additional support for a homologous recombination mechanism comes from the fact that the reversion of *a* to *a^t* and *A^w* occurs with a frequency of 4.7×10^{-6} , which is comparable to the reversion frequency of 3.3×10^{-6} for the original dilute (*d*) coat-color mutation (Schlager and Dickie 1967). The *d* mutation was shown to be caused by a retroviral insertion, and reversion to wild type occurs by homologous recombination between the LTRs (Copeland et al. 1983). Although we expect that this model will explain the genetics of *a*, *a^t*, and *A^w*, we cannot definitively rule out the possibility that other structural alterations might also be associated with these alleles.

Multiple forms of agouti mRNA

We demonstrated previously that the agouti gene produces a 0.8-kb mRNA in neonatal skin and that the abundance of this mRNA correlates well with the presence of pheomelanin (Bultman et al. 1992; Michaud et al. 1993). However, RT-PCR experiments presented here demonstrate clearly that instead of producing a single species of mRNA, the agouti gene produces a number of similarly sized transcripts that are different at their 5' ends. One class of transcript, referred to here as form I, consists of sequence derived from the first, second, third, and fourth agouti exons as described originally (Bultman et al. 1992; Miller et al. 1993). This form I transcript is expressed in skin from both the ventrum and the dorsum of 4- and 5-day-old *A/A*, *A^w/a*, and *A^w/A* mice but neither in the ventrum nor the dorsum of age-matched *a/a* or *a^t/a^t* mice (Table 2). Furthermore, expression of the form I transcript is regulated during early postnatal development such that it is produced while melanocytes are actively depositing pheomelanin into agouti hair follicles between days 3 and 7 (Galbraith 1964; Sakurai et al. 1975) but is absent both before and after this interval (data not shown). A second class of transcripts detected by RT-PCR, referred to here as form II, consists of differentially incorporated 111- and 46-bp segments spliced to the agouti coding exons. Both of these transcripts are expressed in skin from the ventrum, but not the dorsum, of 4- and 5-day-old *A^w/a*, *a^t/a^t*, and *A^w/A* mice. In contrast, form II transcripts are not detected in the ventrum or the dorsum of age-matched *A/A* or *a/a* mice (Table 2); this is the reason why these different forms of agouti mRNA were not identified originally in the analysis of the *A* allele, which was presumed to be wild type (Bultman et al. 1992). Furthermore, the form II transcripts are expressed throughout the hair-growth process in the ventrum of *A^w/A* mice (data not shown). Taken together, these findings provide compelling evidence that the form I transcript is associated with the production of the subapical band of pheomelanin in agouti hair throughout the pelage and that the form II transcripts are responsible for the predominating yellow pigmentation over the ventral surface in the *A^w* and *a^t* alleles.

The fact that *A/a^t* mice have a phenotype that is identical to that of *A^w* (Dunn 1928; Silvers 1979) is consistent with the expression data presented here. *A^w* produces both form I and form II transcripts, whereas *A* produces only the form I transcript and *a^t* produces only form II transcripts. *A/a^t* mice would be expected to produce both forms of transcripts, and as a result exhibit a phenotype identical to that of *A^w*.

From our analyses, form I and II agouti transcripts differ only in their 5' noncoding regions and apparently encode the same protein. Additional analysis of the 111-bp segment has revealed that it may be comprised of two separate exons (Michaud et al. 1994), neither of which contains a termination codon (Michaud et al. 1993). For this reason, we cannot exclude the possibility that the shorter form II mRNA that contains only the 111-bp segment actually has a 5' extension of the agouti open read-

ing frame. Also, because the 111- and 46-bp segments each contain a functional splice acceptor at its 5' junction (Michaud et al. 1993; unpubl.), we predict that there must be at least one more exon upstream of the 111-bp segment that is normally at the 5' end of form II transcripts. These matters are currently under investigation.

Insertions in a and a^t modulate the expression of form I and form II agouti transcripts

Comparison of the structure of the *a*, *a^t*, and *A^w* alleles suggests a model that explains how each insertion alters agouti gene expression. We propose that the VL30 element, along with the 5.5 kb of additional sequence, essentially shuts off agouti gene expression in the *a* mutation by negatively regulating the production of both form I and form II mRNA. This is presumably the case throughout the pelage, except for very restricted areas near the pinnae, mammae, and perineum, where phaeomelanin deposition still occurs. It is likely that any agouti transcripts produced from the mammae in the samples we analyzed (skin samples containing the pinnae and perineum were not used in our experiments) were present at a sufficiently low level as to be undetectable with our molecular techniques. Additionally, we propose that in the *a^t* mutation, the inserted segment of DNA (VL30 containing a single copy of the 526-bp direct repeat) down-regulates the production of form I transcripts but does not significantly affect the production of form II transcripts. Perhaps the inserted element in *a^t* is acting in a localized manner such that it does not influence the expression of the form II-specific exons, which are located ~100 kb upstream of the inserted VL30 sequences. The mechanism through which the inserted elements in *a* and *a^t* alter agouti expression is currently under investigation.

A^w is wild type

As demonstrated here, all of the *A^w* alleles that arise through reversion from *a* contain a solo VL30 LTR in the first intron of the agouti gene. It is possible that this solo LTR is redirecting agouti gene expression in a manner that is consistent with the production of form I and form II mRNAs in the *A^w* allele. Alternatively, based on the analysis of wild-type reversion alleles of the dilute mutation (Copeland et al. 1983), it is possible that the remaining LTR has no functional significance and does not interfere with the normal expression of the agouti gene. We favor the latter possibility because of our analysis of the *A^w* allele in the 101 inbred strain; this *A^w* allele does not contain a solo VL30 LTR in the same site as does *A^{w-38f}* or any of the other reversion mutants that we have analyzed thus far (S.J. Bultman, H.O. Sweet, M.T. Davisson, and R.P. Woychik, unpubl.). This implies that the original *a* mutation arose on a white-bellied agouti background. Furthermore, because form II transcripts are expressed in *A^w* but not in *A*, we propose that *A^w* is the true wild-type allele and that *A* actually represents a

loss-of-function mutation in which form II transcripts are not expressed. This is of particular interest because both *A* and *A^w* have traditionally been referred to as wild-type agouti alleles. The proposal that *A^w* is the true wild-type agouti allele is not only compatible with the molecular data presented here but is also consistent with reports that most populations of wild mice have a white-bellied agouti phenotype (Silvers 1979).

Materials and methods

Animals

All of the new *a^t* and *A^w* mutants reported here arose spontaneously in the production colony at The Jackson Laboratory (see Table 1), except for *a^{t-2Gso}* (formerly referred to as SB B+T), which arose spontaneously at the Oak Ridge National Laboratory in a cross between strains SEC/E (*a/a*, *b/b*, *c^{ch}/c^{ch}*) and C57BL/E (*a/a*). All animals were subsequently maintained at the Oak Ridge National Laboratory.

Cloning overlapping genomic DNA fragments that span the insertions in a, a^t, and A^w

A total of 250 µg of genomic liver DNA from wild-type *A/A* (C3H), *a/a* (C57BL/E), or *a^t/a^t* (*a^{t-2Gso}* stock) mice were digested with either *Bam*HI or *Bgl*II, and were size fractionated on a 10-40% sucrose gradient (Sambrook et al. 1989). Fractions containing appropriately sized fragments (see below) were ligated into the EMBL3 vector, packaged in vitro and screened using standard procedures (Ausubel et al. 1988; Sambrook et al. 1989). Utilizing a fragment that includes the first agouti exon plus some 5'-flanking sequence and part of the first intron (probe 1.5, Fig. 1), *Bam*HI fragments of 8, 16, and 11 kb were cloned from *A/A*, *a/a*, and *a^t/a^t* DNAs, respectively. Utilizing a 0.6-kb *Bam*HI-*Pst*I fragment in the first intron of agouti (probe 0.6, Fig. 1), *Bgl*II fragments of 2, 6, and 7 kb were cloned from *A/A*, *a/a*, and *a^t/a^t* DNAs, respectively. These fragments were subsequently subcloned into the pGEM (Promega) or pBluescript (Stratagene) vectors and purified as described (Ausubel et al. 1988; Sambrook et al. 1989). For *A^w*, a total of 250 µg of genomic liver DNA was prepared from the original *A^{w-38f}* mutant, partially digested with *Sau*3A, and size fractionated on a 10-40%-sucrose gradient (Sambrook et al. 1989). Fractions containing 15- to 20-kb fragments were ligated into the EMBL3 λ vector, packaged in vitro, and screened using standard procedures (Ausubel et al. 1988; Sambrook et al. 1989). Clones that hybridized to probes 1.5 and 0.6 (see Fig. 1) were purified, and subfragments were subsequently isolated and subcloned into pGem (Promega) or pBluescript (Stratagene) for further analysis.

DNA sequencing

Selected portions of genomic clones were sequenced by the dideoxynucleotide method (Sanger et al. 1977) using T7 DNA polymerase (U.S. Biochemical; Tabor and Richardson 1987). Analysis of the DNA sequence was conducted with the University of Wisconsin Genetics Computing Group sequence analysis programs (Devereux et al. 1984), and GenBank data base searches were performed using the BLAST algorithm (Altschul et al. 1990).

Southern blot analysis

Genomic adult liver or tail DNA (10 µg) was digested with restriction enzymes, electrophoresed through 0.8% agarose

gels, and blotted to GeneScreen (DuPont) with standard procedures [Ausubel et al. 1988; Sambrook et al. 1989]. Radiolabeled hybridization probes were prepared with the random hexamer labeling technique [Feinberg and Vogelstein 1984], and posthybridization filter washing was conducted under high stringency conditions (0.2× SSC, 0.1% SDS, at 68°C).

RT-PCR analysis

Total RNA (10 µg), prepared by the guanidine isothiocyanate procedure [Ausubel et al. 1988], was reverse transcribed [Kawasaki 1990], ethanol precipitated, and resuspended in 20 µl of H₂O. PCR was performed using 4 µl of the sample as described previously [Pieretti et al. 1991]. The sequences of the oligonucleotide primers in Fig. 6 are as follows: Primer 1, 5'-GAACTGGCATCAAAGTACCA; primer 2, 5'-CAATGCTCCTGCTCTGCCA; primer 3, 5'-TTCCGCTTCTCGGCTTCTTT. The β-actin primers were 5'-ATGGGTGAGAAGGACTCCTA (forward) and 5'-CAACATAGCACAGCTTCTCT (reverse). PCR products were electrophoresed through 3.0% agarose gels, transferred to GeneScreen Plus [DuPont], and hybridized to an agouti cDNA probe as described above.

Note added in proof

The sequence data described in this paper have been submitted to the GenBank data library under accession number L28700.

Acknowledgments

We are grateful to W.M. Generoso for providing the *a^{t-2Gso}* mutant stock, and W.M. Generoso, K.R. Johnson, J. Schimenti, E.M. Rinchik, and L.B. Russell for their constructive comments on this manuscript. We also thank many of our colleagues for helpful discussions and G. Miller for expert secretarial assistance. This research was jointly sponsored by the Office of Health and Environmental Research, U.S. Department of Energy (contract DE-AC05-84OR21400) with Martin Marietta Energy Systems, Inc. (R.P.W.), and by the National Institute of Environmental Health Sciences (IAG 222Y01-ES-10067) to R.P.W. as well as the National Science Foundation grant (BIR 8915728) to M.T.D. and H.O.S. This research was also supported in part by an Alexander Hollaender Distinguished Postdoctoral Fellowship to E.J.M., sponsored by the U.S. Department of Energy, Office of Health and Environmental Research, and administered by the Oak Ridge Institute for Science and Education.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

Adams, S.E., P.D. Rathjen, C.A. Stanway, S.M. Fulton, M.H. Malim, W. Wilson, J. Ogden, L. King, S.M. Kingman, and A.J. Kingsman. 1988. Complete nucleotide sequence of a mouse VL30 retro-element. *Mol. Cell. Biol.* 8: 2989–2998.

Altschul, S.F., W. Gish, W. Miller, E.W. Myers, and D.J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215: 403–410.

Ausubel, F.M., R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl. 1988. *Current protocols in molecular biology*. John Wiley and Sons, Greene, New York.

Besmer, P., U. Olshevsky, D. Baltimore, D. Dolberg, and H. Fan. 1979. Virus-like 30S RNA in mouse cells. *J. Virol.* 29: 1168–1176.

Bultman, S.J., E.J. Michaud, and R.P. Woychik. 1992. Molecular

characterization of the mouse agouti locus. *Cell* 71: 1195–1204.

Carter, A.T., J.D. Norton, Y. Gibson, and R.J. Avery. 1986. Expression and transmission of a rodent retrovirus-like VL30 gene family. *J. Mol. Biol.* 188: 105–108.

Copeland, N.G., K.W. Hutchison, and N.A. Jenkins. 1983. Excision of the DBA ecotropic provirus in dilute coat-color revertants of mice occurs by homologous recombination involving the viral LTRs. *Cell* 33: 379–387.

Devereux, J., P. Haeberli, and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* 12: 387–395.

Dhar, R., W.L. McClements, L.W. Enquist, and G.F. Van de Woude. 1980. Nucleotide sequences of integrated Moloney sarcoma provirus long terminal repeats and their host and viral junctions. *Proc. Natl. Acad. Sci.* 77: 3937–3941.

Dickie, M.M. 1969. Mutations at the agouti locus in the mouse. *J. Hered.* 60: 20–25.

Dunn, L.C. 1928. A fifth allelomorph in the agouti series of the house mouse. *Proc. Natl. Acad. Sci.* 14: 816–819.

Feinberg, A.P. and B. Vogelstein. 1984. Addendum: A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 137: 266–267.

Galbraith, D.B. 1964. The agouti pigment pattern of the mouse: A quantitative and experimental study. *J. Exp. Zool.* 155: 71–90.

Garman, R.D., P.J. Doherty, and D.H. Raulet. 1986. Diversity rearrangement, and expression of murine T cell gamma genes. *Cell* 45: 733–742.

Green, M.C. 1989. Catalogue of mutant genes and polymorphic loci. In *Genetic variants and strains of the laboratory mouse* (ed. M.F. Lyon and A.G. Searle), pp. 17–20. Oxford University Press, Oxford/New York.

Hayday, A.C., H. Saito, S. Gillies, D.M. Kranz, G. Tanigawa, H.N. Eisen, and S. Tonegawa. 1985. Structure, organization, and somatic rearrangement of T cell gamma genes. *Cell* 40: 259–269.

Hodgson, C.P., R.Z. Fisk, P. Arora, and M. Chotani. 1990. Nucleotide sequence of mouse virus-like (VL30) retrotransposon BVL-1. *Nucleic Acids Res.* 18: 673.

Hodgson, C.P., P.K. Elder, T. Ono, D.N. Foster, and M.J. Getz. 1983. Structure and expression of mouse VL30 genes. *Mol. Cell. Biol.* 3: 2221–2231.

Itin, A. and E. Keshet. 1983. Nucleotide sequence analysis of the long terminal repeat of murine virus-like DNA (VL30) and its adjacent sequences: Resemblance to retrovirus proviruses. *J. Virol.* 47: 656–659.

Jackson, I.J. 1993. More to colour than meets the eye. *Curr. Biol.* 3: 518–521.

Kawasaki, E.S. 1990. Amplification of RNA. In *PCR protocols: A guide to methods and applications* (ed. M.A. Innis, D.H. Gelfand, J.J. Sninsky, and T.J. White), pp. 21–27. Academic Press, San Diego, CA.

Keshet, E. and A. Itin. 1982. Patterns of genomic distribution and sequence heterogeneity of a murine "retrovirus-like" multigene family. *J. Virol.* 43: 50–58.

Keshet, E., Y. Shaul, J. Kaminchik, and H. Aviv. 1980. Heterogeneity of "virus-like" genes encoding retrovirus-associated 30S RNA and their organization with the mouse genome. *Cell* 20: 431–439.

Kuff, E.L. and K.K. Lueders. 1988. The intracisternal A-particle gene family: Structure and functional aspects. *Adv. Cancer Res.* 51: 183–276.

Michaud, E.J., S.J. Bultman, L.J. Stubbs, and R.P. Woychik. 1993. The embryonic lethality of homozygous lethal yellow mice (*A^y/A^y*) is associated with the disruption of a novel

- RNA-binding protein. *Genes & Dev.* **7**: 1203–1213.
- Michaud, E.J., S.J. Bultman, M.L. Klebig, M.J. van Vugt, L.J. Stubbs, L.B. Russell, and R.P. Woychik. 1994. A molecular model for the genetic and phenotypic characteristics of the mouse lethal yellow (A^y) mutation. *Proc. Natl. Acad. Sci.*
- Miller, M.M., D.M.J. Duhl, H. Vrieling, S.P. Cordes, M.M. Ollmann, B.M. Winkes, and G.S. Barsh. 1993. Cloning of the mouse agouti gene predicts a secreted protein ubiquitously expressed in mice carrying the lethal yellow mutation. *Genes & Dev.* **7**: 454–467.
- Pieretti, M., F. Zhang, Y.-H. Fu, S.T. Warren, B.A. Oostra, C.T. Caskey, and D.L. Nelson. 1991. Absence of expression of the FMR-1 gene in fragile X syndrome. *Cell* **66**: 817–822.
- Sakurai, T., H. Ochia, and T. Takeuchi. 1975. Ultrastructural change of melanosomes associated with agouti pattern formation in mouse hair. *Dev. Biol.* **47**: 466–471.
- Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: A laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Sanger, F., S. Nicklen, and A.R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci.* **74**: 5463–5467.
- Schlager, G. and M.M. Dickie. 1967. Spontaneous mutations and mutation rates in the house mouse. *Genetics* **57**: 319–330.
- Scolnick, E.M., W.C. Vass, R.S. Howk, and P.H. Duesberg. 1979. Defective retrovirus-like 30S RNA species of rat and mouse cells are infectious if packaged by type C helper virus. *J. Virol.* **29**: 964–972.
- Shimotohno, K., S. Mizutani, and H.M. Temin. 1980. Sequence of retrovirus provirus resembles that of bacterial transposable elements. *Nature* **285**: 550–554.
- Silvers, W.K. 1979. The agouti and extension series of alleles, umbrous, and sable. In *The coat colors of mice: A model for mammalian gene action and interaction*, pp. 6–44. Springer-Verlag, New York.
- Tabor, S. and C.C. Richardson. 1987. DNA sequence analysis with a modified form of T7 DNA polymerase. *Proc. Natl. Acad. Sci.* **84**: 4767–4771.
- Van Buren, C., J.G. Goddard, A. Berns, and I.M. Verma. 1980. Structure of Moloney murine leukemia viral DNA: Nucleotide sequence of the 5' long terminal repeats and adjacent cellular sequences. *Proc. Natl. Acad. Sci.* **77**: 3307–3311.
- Varmus, H.E. 1982. Form and function of retroviral proviruses. *Science* **216**: 812–820.

Molecular analysis of reverse mutations from nonagouti (a) to black-and-tan (a(t)) and white-bellied agouti (Aw) reveals alternative forms of agouti transcripts.

S J Bultman, M L Klebig, E J Michaud, et al.

Genes Dev. 1994, **8**:

Access the most recent version at doi:[10.1101/gad.8.4.481](https://doi.org/10.1101/gad.8.4.481)

References

This article cites 32 articles, 15 of which can be accessed free at:
<http://genesdev.cshlp.org/content/8/4/481.full.html#ref-list-1>

License

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).