ARTICLES

nature genetics

A natural allele of *Nxf1* suppresses retrovirus insertional mutations

Jennifer A Floyd[1], David A Gold[1], Dorothy Concepcion[2], Tiffany H Poon[2], Xiaobo Wang[3], Elizabeth Keithley[3], Dan Chen[4], Erica J Ward[2,10], Steven B Chinn[5,10], Rick A Friedman[5], Hon-Tsen Yu[6], Kazuo Moriwaki[7], Toshihiko Shiroishi[8] & Bruce A Hamilton[2,9]

Endogenous retroviruses have shaped the evolution of mammalian genomes. Host genes that control the effects of retrovirus insertions are therefore of great interest. The modifier-of-vibrator-1 locus (*Mvb1*) controls levels of correctly processed mRNA from genes mutated by endogenous retrovirus insertions into introns, including the *Pitpn*^{vb} tremor mutation and the *Eya1*^{BOR} model of human branchiootorenal syndrome. Positional complementation cloning identifies *Mvb1* as the nuclear export factor Nxf1, providing an unexpected link between the mRNA export receptor and pre-mRNA processing. Population structure of the suppressive allele in wild *Mus musculus castaneus* suggests selective advantage. A congenic *Mvb1*^{CAST} allele is a useful tool for modifying gene expression from existing mutations and could be used to manipulate engineered mutations containing retroviral elements.

Mus musculus is a complex species group with subpopulations that have diverged and hybridized since becoming commensal with humans some 10,000 years ago. Allopatric divergence and rehybridization of mouse lineages are thought to contribute to the diversity and activity of retroviral elements in the current mouse genomes[1,2]. Host genes that can influence the expression of newly introduced (or newly mobilized) viral elements might be under selection for variants that blunt these effects. Our results suggest that *Mvb1* is such a locus.

Mvb1 was originally identified as a strain-derived locus that modifies the neurological mutant vibrator[3]. The vibrator (*vb*) mutation is a hypomorphic allele of the gene *Pitpn*, encoding phosphatidylinositol transfer protein α (PITPα), caused by the insertion of an endogenous retrovirus (intracisternal A particle; IAP) into the fourth intron of the gene, resulting in a five- to ten-fold loss of PITPα expression. Homozygous *vb*/*vb* mice have severe action tremor (**Supplementary Video 1** online), progressive degeneration of interneurons in the brain stem and spinal cord, and uniform juvenile lethality. But vibrator mice carrying *Mvb1* alleles from the wild-derived CAST/Ei inbred strain have less severe tremors (**Supplementary Video 2** online) and survive to adulthood. In principle, this could be due to a change in physiological requirement for PITPα function or to a change in the steady-state expression level of PITPα derived from the mutated allele.

Modifier genes that act on retroviral insertions might be especially useful genetic tools, as such insertions comprise ~15% of spontaneous mutations in laboratory mice[2,4]. Here we examine the mechanism of *Mvb1*-mediated suppression by determining its effect on *vb* RNA expression, by examining its effects on other retrovirus-associated mutations (selected as examples of different retrovirus families and different classes of insertion sites, irrespective of family or class frequency), and by positionally identifying *Mvb1*. We show that suppressive *Mvb1* alleles elevate steady-state level of correctly processed *Pitpn* mRNA derived from *vb* alleles, creating an *in vivo* titration of this gene product. We show that *Mvb1* also modifies *Eya1*^{BOR}, a model of human branchiootorenal syndrome caused by insertion of an IAP element (the class most frequently associated with spontaneous mutations in mice) inserted into introns in the sense orientation[5]. We identify *Mvb1* by a positional complementation strategy as an allele of the gene *Nxf1*, encoding an mRNA nuclear export factor. Nxf1 protein is known to bind and mediate export of constitutive transport elements (CTEs) in the unspliced genomes of several retroviruses, including rodent IAPs[6], and other retroelements, including human LINEs[7]. Notably, we find a strong bias against CTE-containing insertions in introns and against sense orientation for elements that are found in introns in the public draft mouse genome. This suggests an orientation-sensitive selective pressure, which could also promote genetic variants that suppress the effects of such insertions, particularly in zones of hybridization among wild mouse populations. Consistent with this selection bias, we show that the IAP-suppressing CAST/Ei allele is the major allele in wild *M. m. castaneus* mice in

[1]Biomedical Sciences Graduate Program, [2]Department of Medicine, [3]Division of Otolaryngology and [4]Molecular Pathology Graduate Program, UCSD School of Medicine, 9500 Gilman Drive, La Jolla, California 92093-0644, USA. [5]House Ear Institute, Los Angeles, California 90057, USA. [6]Department of Zoology, National Taiwan University, Taipei, ROC 107, Taiwan. [7]RIKEN Bioresource Center, Kuoyadai 3-1-1, Tsukuba 305-0074, Japan. [8]Mammalian Genetics Laboratory, National Institute of Genetics, Mishima, Shizuoka-ken 411-8540, Japan. [9]Department of Cellular and Molecular Medicine and Rebecca and John Moores UCSD Cancer Center, UCSD School of Medicine, 9500 Gilman Drive, La Jolla, California 92093-0644, USA. [10]Present addresses: Program in Genetic Counseling, University of Michigan, Ann Arbor, Michigan, USA (E.J.W.); and Department of Human Genetics, Kresge Hearing Institute, University of Michigan, Ann Arbor, Michigan, USA (S.B.C.). Correspondence should be addressed to B.A.H. (bah@ucsd.edu).

Published online 28 September 2003; doi:10.1038/ng1247

Southeast Asia and a frequent minor allele in the related Japanese subspecies *Mus musculus molossinus*, thought to have arisen by hybridization between *M. m. castaneus* and *Mus musculus musculus* approximately 2,000–3,000 years ago[8].

RESULTS

Mvb1 is a dosage-sensitive modifier of *vb* RNA levels

Mvb1 modifies the severity of tremors in vibrator mice and the associated juvenile lethality[3]. To examine this effect in more detail, we monitored the longevity of *vb* mutant mice congenic on a C57BL/6J (B6) strain background with zero, one or two CAST/Ei alleles of *Mvb1* (**Fig. 1a**). Consistent with our behavioral observations, the longevity data indicate a semidominant mode of action and high penetrance.

To extend these observations to the cellular level, we examined the histology of vibrator mice homozygous with respect to either the B6 or CAST allele of *Mvb1* (**Fig. 1b**). Mutant mice on the B6 background had a high frequency of neuropathology in characteristic nuclei of the hindbrain and spinal cord. We observed vacuolated cells and cells with reduced staining of internal membrane structures in red nucleus, in deep cerebellar nuclei, in lateral vestibular nucleus, in the reticular formation of the pons and in the spinal cord. By contrast, mutants homozygous with respect to the CAST allele of *Mvb1* had essentially no vacuolated neurons and fewer (and less extreme) examples of neurons with other pathology.

To examine whether *Mvb1* acts by altering RNA levels from the mutated allele rather than by bypassing the requirement for this RNA, we examined RNA levels from *vb* mutant and wild-type littermates homozygous with respect to either B6 or CAST alleles (**Fig. 1c–e**). *Mvb1*$^{CAST/CAST}$ homozygotes accumulated approximately twice as much correctly processed *Pitpn* RNA from the *vb* allele as did their B6 littermates. In contrast, we observed no significant difference in RNA levels from wild-type alleles of *Pitpn* between mice of different *Mvb1* genotypes (data not shown). We obtained similar results using RNA from the liver, indicating that *Mvb1* activity is not restricted to the brain. In addition to increasing the level of normal *Pitpn* RNA derived from *vb* alleles, CAST alleles of *Mvb1* seem to decrease the level of a low-abundance, mutant-specific RNA detected by exons 5′ to the IAP insertion site (**Fig. 1e**). Western blots using antibodies specific for either PITPα or PITPβ showed that the changes in RNA levels correlate to changes in protein levels (**Fig. 1f**,**g**).

Mvb1 modifies *Eya1*^{BOR}

In principle, *Mvb1* alleles could act either as a locus-specific modifier of *Pitpn* (for example, in a feedback regulatory pathway) or as a more general modifier of a step in RNA biogenesis blocked by the IAP insertion. To test this idea, we crossed the suppressive (CAST) allele of *Mvb1* to several mutations that involve endogenous retrovirus insertions. The pattern of *Mvb1* interactions with these mutations should help to define the mechanism by which *Mvb1* acts because each mutation represents a distinct class of retroviral element, mechanism of interference with host gene function, or affected tissues (**Table 1**).

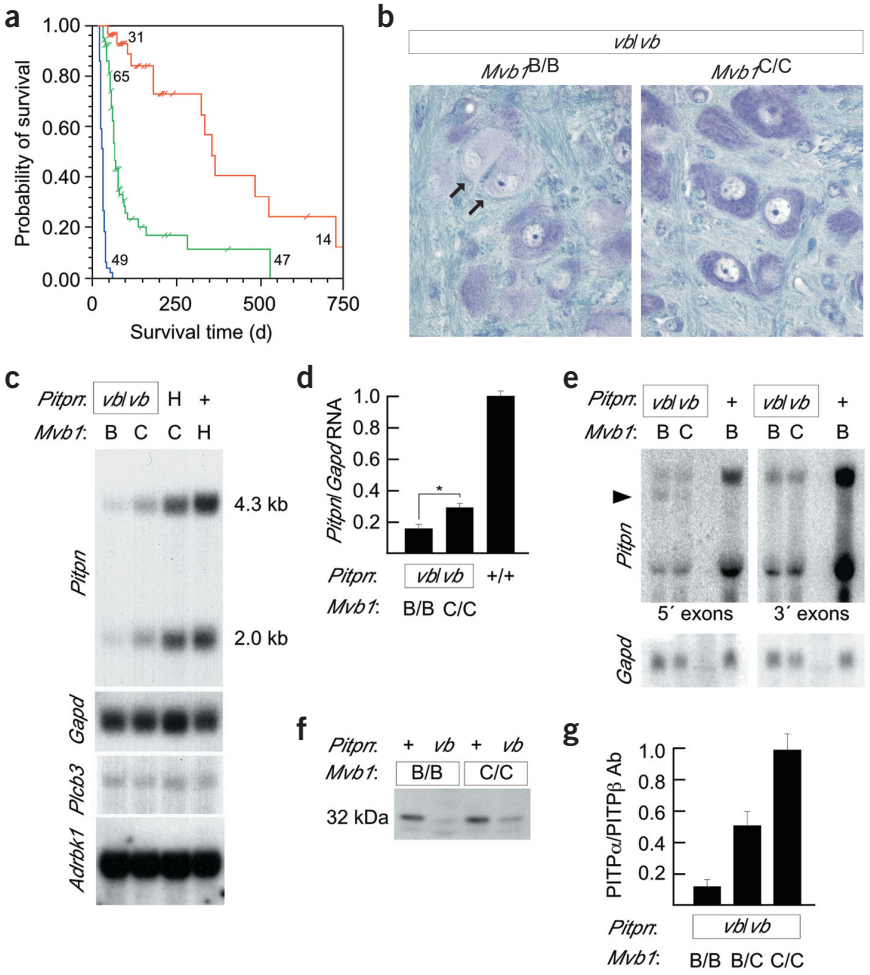

Figure 1 *Mvb1* is a semidominant modifier of vibrator mRNA level. (**a**) Dosage-sensitive impact of *Mvb1* alleles on lifespan of *vb* homozygotes. Product-limit analysis of censored survival data over two years is shown for each *Mvb1* genotype: B6/B6 in blue, B6/CAST in green and CAST/CAST in red. Initial and terminal cohort sizes are indicated; censored mice are indicated by a hash mark at the approximate time of removal. (**b**) Suppression of vibrator neuropathology. Sections from red nucleus show diminished staining of internal membrane structures and eccentric nuclei in a large fraction of cells from *vb*/*vb* *Mvb1*$^{B6/B6}$ mice (arrows). This pathology is notably absent in suppressed mice. (**c**) *Mvb1* alleles control level of RNA derived from *vb* alleles of *Pitpn*. *vb*/*vb* *Mvb1*$^{B6/B6}$ tissues express ~18% and *vb*/*vb* *Mvb1*$^{CAST/CAST}$ tissues express ~36% of normal *Pitpn* poly(A)$^+$ RNA levels after normalization to *Gapd* or to *Plcb3* or *Adrbk1*, two genes flanking the *Mvb1* interval on chromosome 19. H, heterozygote. (**d**) Aggregate data from several blots similar to that shown in **c** show quantitatively consistent effects of *vb* and *Mvb1* genotypes. For comparisons between blots, all bands were normalized to the average of nonmutant control samples within a blot. (**e**) An aberrant RNA containing the 5′ exons of *Pitpn* is present in total RNA extracts from *vb* mutants but not littermate controls. (**f**) Chemiluminescent detection illustrates *Mvb1* control of PITPα protein level in *vb* mutant tissue. (**g**) Ratio of PITP proteins from triplicate samples on a single blot detected by ^{125}I-labeled protein A. B, B6 allele; C, CAST allele.

Table 1 Mutations tested for suppression by *Mvb1*^{CAST}

Gene	Allele	Insertion element	Site	Orientation	Phenotypes	Mechanism	Suppression by *Mvb1*	Source	Ref.
Pitpn	*vb*	IAP	Intron	Sense	Tremor, lethality	Decreased expression	Yes	Lab colony	3
Eya1	*BOR*	IAP	Intron	Sense	Circling, deafness	Decreased expression	Yes	K. Johnson	5
Agouti	*iy*	IAP	Intron	Antisense	Coat color, obesity	Ectopic expression	No	Jackson Lab	15
Axin	*Fu*	IAP	Intron	Antisense	Tail kink	Dominant truncation	No	Jackson Lab	16
Dab1	*scm*	(IAP)	Intron	Antisense	Ataxia	Splicing into parental IAP	No	Jackson Lab	20, 21
Hairless	*hr*	MuLV	Intron	Sense	Complete hair loss	Decreased expression	No	Jackson Lab	11
Myo5a	*d*	MuLV	Intron	Sense	Coat color	Abnormal RNA expression	No	DBA/2J	10, 33
Agouti	*a*	VL30	Intron	Antisense	Coat color	Decreased expression (isoform-specific)	No	C57BL/6J	14

Eya1^{BOR} is a recessive model of human branchiootorenal syndrome. An IAP insertion into an intron reduces the level of host gene expression by ~50% (ref. 5). Homozygous mice (*Eya1*$^{BOR/BOR}$) have a pronounced cochlear malformation and can be recognized behaviorally by characteristic head bobbing, circling and failure to startle in response to loud noise. In some genetic backgrounds, frequent kidney agenesis and dysgenesis are associated with perinatal and juvenile death.

Mvb1^{CAST} suppresses several *Eya1*^{BOR} phenotypes. We evaluated phenotypes that depend on development of the inner ear in mutant mice from a digenic cross (F_2 from C3H–*Eya1*^{BOR} × B6.CAST–*Mvb1*^{CAST}; **Fig. 2**). Behaviorally, *Eya1*^{BOR} *Mvb1*$^{CAST/CAST}$ mice had noticeably less head bobbing and circling than *Eya1*$^{BOR/BOR}$ mice with B6 or C3H alleles at *Mvb1*. More quantitatively, we assessed auditory brainstem responses to a calibrated sound played in one ear of mutant mice

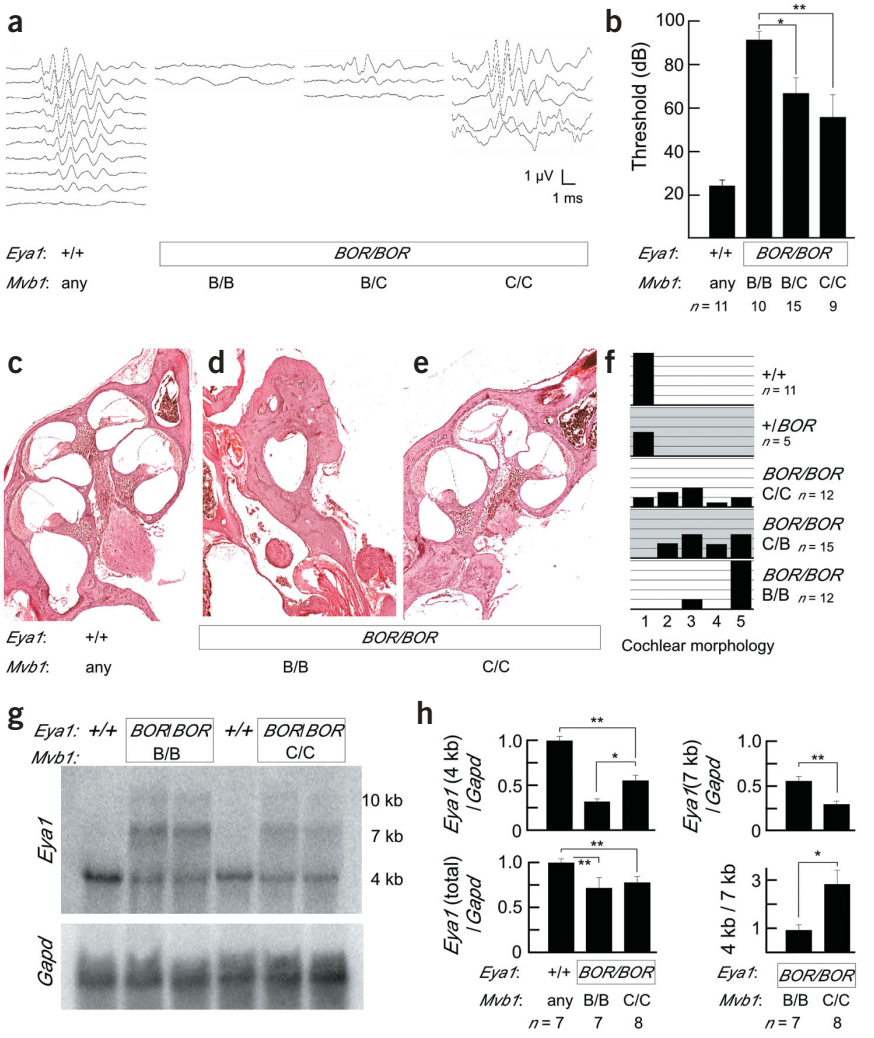

Figure 2 *Mvb1* modifies *Eya1*^{BOR} phenotypes and RNA level. (**a**) Typical auditory brainstem responses to clicks presented at progressively decreasing amplitudes. For each mouse, top line represents response to 100 dB, followed by 85, 70, 60, 55, 50, 45, 40, 35 and 30 dB or until wave form of the response is lost. (**b**) Distribution of ABR data for each genotype is plotted as decibel level at which response loses waveform. Group differences between *Mvb1*$^{B6/B6}$ mutants and the other two genotypes were significant at $P < 0.05$ (*) or 0.01 (**) by Student's *t*-test. (**c–e**) Cochlear malformations in *Eya1*^{BOR} homozygotes are less severe in the presence of *Mvb1*^{CAST} alleles. Typical histology at the level of the eighth cranial nerve is shown for (**c**) nonmutant, (**d**) *Eya1*$^{BOR/BOR}$ *Mvb1*$^{B6/B6}$ and (**e**) *Eya1*$^{BOR/BOR}$ *Mvb1*$^{CAST/CAST}$ mice. (**f**) Histogram of cochlear morphology scores. 1 = normal cochlea; 3 = nearly normal cochlear turns but abnormal number of ganglion cells; 5 = cochlea with only the basal turn. (**g**) Representative northern blot shows increased level of correctly processed 4-kb message and reduced level of the mutant-specific 7-kb and 10-kb messages in *Eya1*$^{BOR/BOR}$ *Mvb1*$^{CAST/CAST}$ mice compared with *Eya1*$^{BOR/BOR}$ *Mvb1*$^{B6/B6}$ mice. (**h**) Normalized values in replicate experiments showed significant RNA differences among indicated genotypes using two-sided *t*-tests. The level of the wild-type 4-kb RNA was lower in mutants than littermate controls ($^{**}P \leq 0.001$) and lower in mutants with the B6 allele than those with the CAST allele at *Mvb1* ($^{*}P \leq 0.015$). The level of mutant-specific 7-kb RNA was lower in mutants carrying the CAST allele of *Mvb1*. The total amount of all *Eya1* messages was smaller in mutants than controls but not significant between *Mvb1* genotypes among mutant mice. Differences in the ratio of correct 4-kb product to aberrant 7-kb product were significant at $P < 0.011$. B, B6 allele; C, CAST allele.

(**Fig. 2a**). The difference in average threshold response between *Eya1*$^{BOR/BOR}$ *Mvb1*$^{B6/B6}$ and *Eya1*$^{BOR/BOR}$ *Mvb1*$^{CAST/CAST}$ mutant mice was highly significant (ANOVA, $P < 0.001$), and the data were consistent with a semidominant effect (**Fig. 2b**). Histology of the inner ear was then analyzed on a scale of 1–5 for quality of the cochlear structure by an investigator blind to genotype. Nonmutant and heterozygous mice had normal cochleas irrespective of *Mvb1* genotype (**Fig. 2c**). In contrast, *Eya1*^{BOR} mutants had different cochlear phenotypes depending on their *Mvb1* genotypes, and the data suggest a semidominant modifying effect. *Eya1*^{BOR} mice without the *Mvb1*^{CAST} allele (**Fig. 2d**) had only the basal turn and fewer spiral ganglion cells (10 of 12 mice). In contrast, *Eya1*^{BOR} mice homozygous with respect to *Mvb1*^{CAST} varied: some had normal cochlear structure (3 of 12 mice); some had normal structure with a variable loss of spiral ganglion cells in the apical turn (6 of 12 mice; **Fig. 2e**); and others had more severe phenotypes (3 of 12 mice). A histogram of scores by genotype is shown in **Figure 2f**.

Mvb1 also modifies *Eya1*^{BOR} mortality caused by kidney dysgenesis or agenesis on the C3H genetic background[5,9]. Among 93 F_2 *Eya1*$^{BOR/BOR}$ mice from a second cross (C3H–*Eya1*^{BOR} × CAST/Ei), we observed significantly fewer *Mvb1*$^{C3H/C3H}$ homozygotes (14 of 93; $P > 0.01$, χ^2 test) and significantly more *Mvb1*$^{CAST/CAST}$ homozygotes (35 of 93) relative to *Mvb1* heterozygotes (*Mvb1*$^{C3H/CAST}$; 44 of 93). By contrast, we observed no significant mortality or kidney agenesis in the cross (B6C3–*Eya1*$^{BOR/+}$ × B6.CAST–*Mvb1*) that we used to test suppression of inner ear phenotypes. Taken together, these observations indicate that *Mvb1* acts on a second IAP-associated mutation, that this activity is present in the kidney and the developing inner ear and that additional *Eya1* modifiers exist in B6 and C3H.

To test whether *Mvb1*^{CAST} also suppresses *Eya1*^{BOR} by elevating RNA levels, we examined RNA isoforms by northern blotting (**Fig. 2g**,**h**). We used adult skeletal muscle as an accessible source of *Eya1* RNA from a tissue that remains grossly normal in the mutant. We observed correctly spliced product and two mutant-specific bands of higher molecular weight in samples from mutant mice. Probes corresponding to exons 5′ or 3′ to the insertion site both detected these aberrant bands, suggesting interstitial retention of some inserted sequences. Sequencing of RT–PCR products across exons flanking the insertion site confirmed that IAP-specific sequences were retained. Notably, quantification of bands by phosphorimage analysis showed that suppressive alleles of *Mvb1* effect a quantitative shift in RNA isoforms from high-molecular-weight mutant-specific bands to the correctly processed wild-type product. Together with our results on the *Pitpn*^{vb} mutations, this indicates that *Mvb1* acts on mutations in distinct pathways in a broad range of tissues and suggests that the mechanism of suppression involves at least one step in RNA processing.

Mvb1 is selective for retroviral and mutational class

To further define the mechanism of *Mvb1*, we tested its effect on other retrovirus-associated mutations, chosen to probe different mechanisms of mutation, irrespective of their frequencies among spontaneous mouse mutations. This group includes mutations with similar structures to *vb* and *Eya1*^{BOR} but caused by different classes of insertion element and IAP insertions that act through other mechanisms (**Table 1**).

Myo5a^{d} (ref. 10) and hairless (*hr*; refs. 11–13) are murine leukemia virus (MLV) insertions into introns that inhibit expression of the inserted host gene. Both insertions are in the same transcriptional orientation as the host gene (sense orientation). *Mvb1* alleles did not affect the coat color of *Myo5a*^{d} mutants or the timing or completeness of hair loss in *hr* mutants (data not shown). Similarly, the nonagouti (*a*) allele of Agouti is an inactivating VL30 retrotransposon insertion present in the B6 strain[14]. *Mvb1* alleles did not affect follicle and shaft coloration of hairs taken from either dorsal or ventral locations (data not shown).

The intermediate yellow allele of Agouti (*A*^{iy}) is an antisense-orientation IAP insertion 5′ to the first coding exon. Ectopic expression of *A*^{iy} from the LTR promoter results in hypopigmentation and obesity[15]. *Mvb1* did not affect the extent or timing of pigmentation or weight gain in *A*^{iy} mutant mice (data not shown). Agouti phenotypes are sensitive to expression level and pattern, and so the absence of phenotypic difference here strongly suggests that *Mvb1* does not alter transcription from this LTR promoter.

Mvb1 shows orientation selectivity for IAP elements. The Fused mutation (*Axin*^{Fu}) is an IAP-element insertion in the antisense orientation into the eighth intron of *Axin*[16], a gene required for axial patterning[17]. A proportion of *Axin*^{Fu} transcripts retain a part of the IAP element, introducing a stop codon. This fusion RNA encodes the RGS signaling domain of axin but lacks the DIX regulatory domain[18,19]. *Axin*^{Fu} produces dominant tail kinks with somewhat variable penetrance on several genetic backgrounds we examined. We observed no systematic shift in severity or frequency of tail kinks between *Mvb1*$^{B6/B6}$ and *Mvb1*$^{CAST/CAST}$ homozygotes carrying one copy of *Axin*^{Fu} (data not shown).

The *Dab1*^{scm} mutation causes inappropriate retention of a nonmutagenic, antisense-oriented IAP element in an intron, which is spliced out from the premutation normal allele of the parental strain[20,21]. We observed no difference in the severity of gait ataxia or frequency of falls between mutant mice homozygous for either allele of *Mvb1* (data not shown).

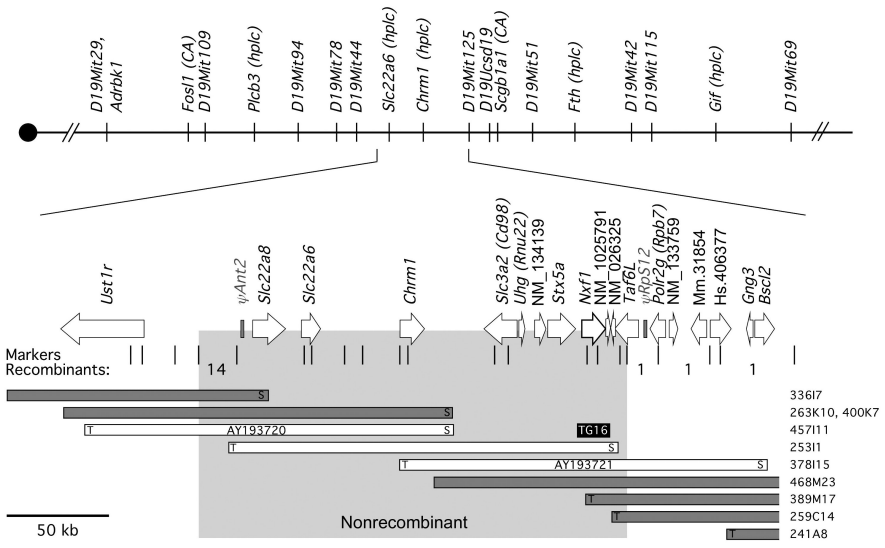

Figure 3 Recombination and physical map of the *Mvb1* interval. Recombination markers used to position the *Mvb1* interval in the cross *vb*/+ *Mvb1*$^{B6/CAST}$ × *vb*/+ *Mvb1*$^{B6/CAST}$ are indicated. Exclusion mapping resulted in a 210-kb nonrecombinant interval. BACs spanning the interval were sequenced to identify positional candidate genes, drawn approximately to scale. Polymorphic microsatellites derived from the genome sequence and SNPs derived from comparative cDNA resequencing used in fine mapping are indicated below the genes, as well as the number of recombinants identified in each interval.

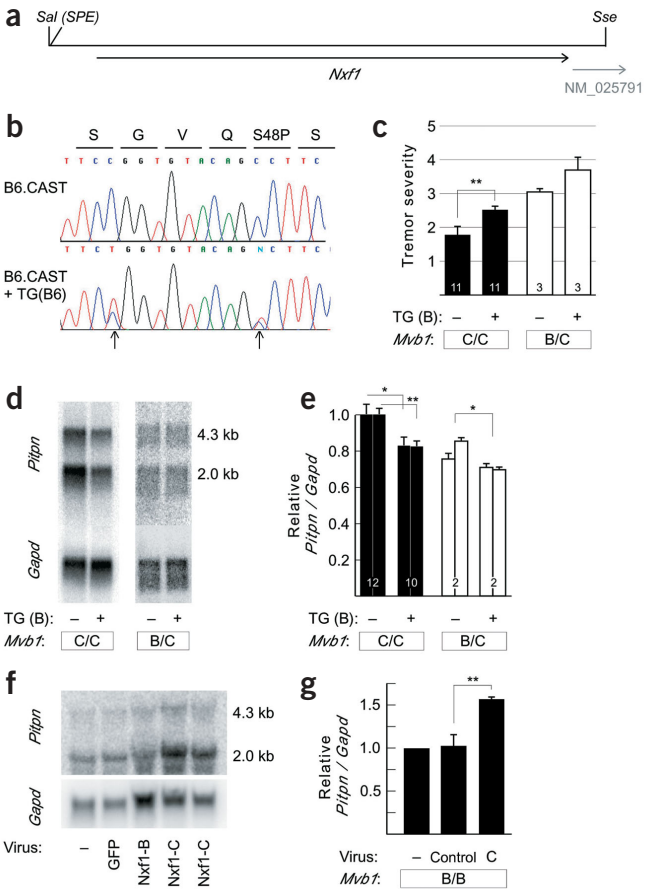

Figure 4 Transgenic complementation of *Mvb1*. (**a**) The genomic transgene contains the full-length *Nxf1* cDNA and flanking sequences as well as the first exon of a predicted gene with EST support. (**b**) RT–PCR sequencing at variant sites shows nearly equal expression from transgenic and endogenous alleles. (**c**) More severe tremors in mice carrying the permissive allele transgene in comparison to littermate controls. Observers were blind to genotype. $n = 11$ for both groups. The difference is significant at $P < 0.01$ by one-tailed *t*-test. A similar sample from *Mvb1* heterozygous mice is not independently significant due to limited sample size; combined *P* value for both groups is 0.0018. (**d**) Reduced *Pitpn* RNA expression in mice carrying the permissive allele transgene in comparison to littermate controls. (**e**) Quantification of bands indicates that the effect of the transgene is quantitatively similar to heterozygosity at the endogenous locus. For each genotype, the two bars indicate separate measures of the two normal mRNA bands. The number of independent samples is indicated. *t*-test *P* values are significant for effect of the transgene on each band in the $Mvb1^{CAST/CAST}$ background (0.02 and 0.0026) and for both bands combined in the $Mvb1^{B6/CAST}$ background (0.026), for a combined *P* value of 1.4×10^{-6}. (**f**) Northern blot from primary cultures of *vb/vb* $Mvb1^{B6/B6}$ cells from a single mouse untreated (–) or infected with pLenti/V5 carrying EGFP, *Nxf1* B6 allele or *Nxf1* CAST allele, hybridized to full-length *Pitpn* or *Gapd*. (**g**) Normalized *Pitpn*/*Gapd* ratios indicate specific elevation of correctly processed *Pitpn* mRNA in *Nxf1* CAST virus–infected cultures compared to either control infection or uninfected cells. B, B6 allele; C, CAST allele.

Taken together, these results suggest that *Mvb1* acts on sense-oriented insertions including at least some class D elements, such as IAPs, but not class C insertions, such as MLVs, and that the activity might require presence of the retrovirus in the nascent RNA. Notably, class C and class D viruses differ in how they balance production of spliced mRNA with export of unspliced genome RNA, a feature relevant to the RNA-level suppression of both $Pitpn^{vb}$ and $Eya1^{BOR}$.

Mapping of *Mvb1*

To identify the gene responsible for the modifying activity of *Mvb1*, we mapped the region by recombination in ~7,790 informative meioses from an F_1 intercross (**Fig. 3**). We bred mice heterozygous with respect to both *vb* and *Mvb1* alleles and genotyped the resulting offspring with simple sequence-length polymorphism (SSLP) or single-nucleotide polymorphism (SNP) markers flanking the locus. Mice carrying *Mvb1*-recombinant chromosomes were either scored for *vb* tremor severity or bred an additional generation to produce homozygous *vb/vb* mice for scoring. Additional SSLP and SNP markers generated in the course of sequencing the interval allowed us to refine the interval to 210 kb. Despite the large number of meioses, we were unable to narrow the region any further by recombination owing to a hot spot at the proximal end that accounted for most of the recombination events observed in an interval of ~500 kb around *Mvb1*.

To identify candidate genes in this interval, we sequenced two BACs that span the interval and identified genes using BLAST homology searches and conservation with the draft human genome; these predictions are now replicated in the public mouse genome sequence[4]. To identify polymorphic candidate genes, we compared brain RNA expression levels and cDNA sequences from B6 and B6.CAST congenic mice. None of the genes had marked differences in expression or changes in size of RNA bands on northern blots, but the neutral amino-acid transporter gene *Slc3a2* (also called *Cd98*, *Mdu1* or *4F2*), the mRNA export factor *Nxf1* (also called *Tap*) and the chromatin-remodeling complex component *Taf6l* (also called *Paf65a*) had nonsynonymous substitutions in coding sequences or codon deletions (*Taf6l*) that implicate them as positional candidate genes. Notably, Nxf1 protein is known to mediate nuclear export of unspliced, CTE-containing RNAs[22]. In principle, this could explain both class-specific and orientation-specific effects of *Mvb1*, making *Nxf1* a strong biological as well as positional candidate gene.

Nxf1 transgenes complement *Mvb1*

To test whether *Nxf1* could complement *Mvb1 in vivo*, we injected a 16.4-kb genomic fragment containing *Nxf1* (B6 allele) into wild-type mouse pronuclei and bred the resulting transgenic mice to our *vb* $Mvb1^{CAST}$ congenic line (**Fig. 4a**). We assayed transgene expression and complementation in *vb/vb* $Mvb1^{CAST/CAST}$ TG/+ mice and nontransgenic controls. Transgenic mice from line 1 expressed the transgene at levels comparable to those transcribed from the endogenous locus, as shown by RT–PCR sequence assays of polymorphic sites (**Fig. 4b**).

We assayed tremor severity and *Pitpn* RNA expression level in matched pairs of transgenic and nontransgenic mice. Because we placed the transgene carrying the permissive B6 allele into the background of the suppressive CAST allele, we expected transgenic mice to have more severe tremor and less *Pitpn* RNA (despite potential overexpression of *Nxf1*). Tremor and movement phenotype, as assessed by three independent investigators blind to genotype using a numerical scale of severity based on the distinctive behaviors of *vb* mutants with different endogenous *Mvb1* alleles (**Fig. 4c**), showed that *Nxf1* modifies vibrator. Tremor was scored several times during the life of the mouse. Transgenic mice were consistently scored as more severely affected than controls by each investigator. The *Nxf1* transgene also affected *Pitpn* RNA levels from *vb* alleles (**Fig. 4d,e**). We measured *Pitpn* expression from several matched pairs of transgenic and control mice by northern-blot analysis of brain RNA. *Pitpn* RNA level averaged 15% less in transgenic mice than in nontransgenic mice after normalization to *Gapd* and was

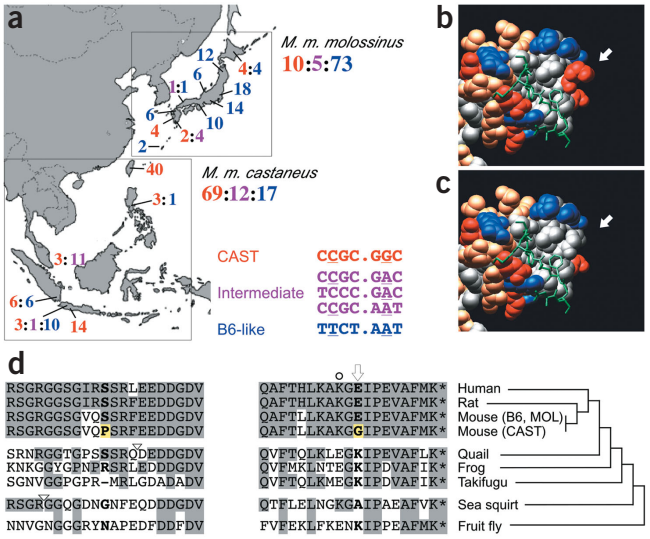

Figure 5 *Nxf1* alleles in inbred strains and wild mice. (**a**) Geographic distribution of *Nxf1* haplotypes by site of collection. Alleles are indicated for four SNPs around exon 2, including the S48P position (underlined) and three SNPs around exon 21, including E610G (underlined). Not shown are 14 additional polymorphic sites in between that increase the diversity of Glu610 haplotypes but not the diversity Gly610 haplotypes (see **Table 2**). (**b**) Model of the Nxf1 UBA domain from Grant *et al.*[24,25], PDB accession 1aoi. Acidic residues are in red, basic in blue and polar in copper. The structure is oriented to place the FG-binding face facing out of the page. Bound FXFG peptide is shown in green. Arrow indicates the position of mouse Glu610 site. (**c**) The E610G alteration changes a surface charge adjacent to a reported nuclear pore interaction surface. (**d**) Alignment of Nxf1 amino acid sequences around the two variant sites in three mammals, three non-mammal vertebrates, a urochordate and an insect. Tree structure indicates only topology among but not distance between species. Low-complexity sequence around S48P (bold type; position relative to the mouse sequence) shows several nonconservative replacements and insertions (triangles, inserted residues not shown). Sequence around E610G shows fewer and more conservative changes. Note E610K in non-mammal vertebrates and fruit flies is accompanied by K608E (circle), conserving an adjacent charge pair in the structure.

lower in transgenic mice in all nine paired comparisons ($P = 0.002$). This is roughly comparable to the *Pitpn* level expressed by *Mvb1* heterozygotes (**Fig. 4e**).

To determine whether expression of the CAST allele of *Nxf1* could also increase the level of *vb*-derived *Pitpn* RNA in a B6 background, we infected primary skin fibroblast cultures from a single B6-*vb*/*vb* mouse with recombinant lentiviral vectors expressing EGFP or $Nxf1^{B6}$ (control infections) or $Nxf1^{CAST}$ and compared them with an uninfected control (**Fig. 4f,g**). Cells infected with virus expressed ~60% more *Pitpn* than did the three controls, whereas cells infected with control viruses expressed *Pitpn* RNA at levels comparable to those expressed in uninfected cultures.

$Nxf1^{CAST}$ is a natural allele and polymorphic in wild mice

The C57BL6/J and CAST/Ei alleles of *Nxf1* differ by two amino acid polymorphisms, S48P and E610G. Genetic variations that distinguish inbred laboratory strains may reflect either natural polymorphisms present in the wild population or new polymorphisms that have been fixed in a lineage after cultivation and under altered selective pressures. To test whether $Nxf1^{CAST}$ represents a natural variation, we resequenced both of these polymorphisms and selected other sites from 49 *M. m. castaneus* and 45 *M. m. molossinus* mice trapped across southeast Asia and Japan as well as several inbred strains (**Fig. 5a** and **Table 2**). The two amino acid polymorphisms are in strong linkage disequilibrium with each other and with noncoding polymorphisms in this population. Notably, the suppressive CAST allele (Pro48, Gly610) defines the major haplotype in *M. m. castaneus* but is a minor allele in *M. m. molossinus* mice. Among inbred strains, the suppressive haplotype is present only in *M. m. castaneus* derivatives.

The *M. m. castaneus* allele of *Nxf1* may have arisen under selective pressure related to retroviral insertions. Although one cannot actually prove selection from a retrospective analysis of a single region, several features of the data suggest this. Ser48 sits in a poorly conserved region of Nxf1 that is not essential to Nxf1 biochemical function[23]. In contrast, Glu610 is highly conserved among vertebrates, sits in a UBA domain thought to mediate nuclear export through physical interaction with the nuclear pore and has a significant chemical shift after binding FG peptide[24,25] (**Fig. 5b–d**). The two amino acid polymorphisms (and their haplotypes) are not in Hardy-Weinberg equilibrium across any nonfixed population we examined, consistent with ongoing natural selection, extensive selectively neutral population admixture or both. More persuasively, Gly610 is present on only one

Table 2 *Nxf1* haplotypes in wild mice

Nxf1 haplotype (S48P, E610G underlined)	Block structure	Residues 48, 610	*M. m. molossinus*	*M. m. castaneus*	Geographic distribution	Inbred strains
CCGC.ACGTT40ATG3G.GGC	1.1.1	Pro, Gly	6	69	Indonesia, Japan, Malaysia, Philippines, Taiwan	CAST, CASA
TCCC.ACGTT40ATG3G.GAC	3.1.2	Pro, Glu		1	Philippines	
TCCC.ATGTT40ATG3G.GAC	3.2.2	Pro, Glu	5	6	Indonesia (Bandar Lanpung), Japan	
CCGC.ATGTT44ATG3A.GAC	1.3.2	Pro, Glu		11	Malaysia	
CCGC.ACGTT40ATG3G.GAC	1.1.2	Pro, Glu	4		Japan	
CCGC.ACGTT40ATG3G.AAT	1.1.3	Pro, Glu		1	Indonesia (Bandung)	
TTCC.ACACC34G0G2A.AAT	2.4.3	Ser, Glu		10	Indonesia (Bandung)	129/SvImJ, A/J, AKR/J, BALB/cJ, 3HeB/FeJ, C57BL/6J, DBA/2J, FVB/NJ
TTCT.GCACC34GTA2A.AAT	4.5.3	Ser, Glu	66		Japan	MOLC, MOLD, MOLE, MOLF, MOLG
TTCT.GCACC34GTG2A.AAT	4.6.3	Ser, Glu	5		Japan	
Total haplotypes examined:			86	98		15

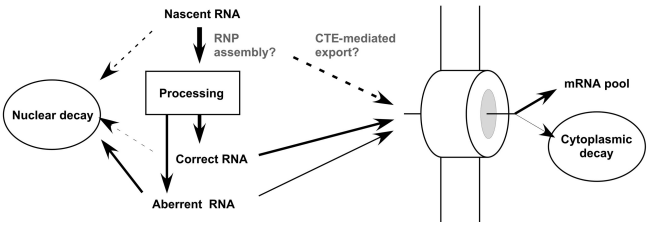

Figure 6 Model for *Nxf1* modifier activity on retrovirus insertional mutations. Flow diagram shows possible fates for newly synthesized pre-mRNA. Variant alleles of *Nxf1* could influence the steady-state ratio of aberrant to correctly processed mRNA by altering a kinetic balance between RNP assembly and CTE-mediated export or by influencing the sites or composition of RNP assembly.

19-marker haplotype across *Nxf1* and is the only allelic variant in haplotype to show this pattern. This indicates that Gly610 is encoded by a 'young' allele that has become the major allele only comparatively recently, a hallmark of positive selection[26].

We further reasoned that if *Nxf1* alleles arose under selective pressure due to retroviral insertions, then retained provirus sites in existing genomes should have a pattern consistent with selection against sense-strand insertions into introns. We identified 179 sites containing strong matches to a well characterized IAP-derived CTE[6] in the draft C57BL/6 genome[4] and looked for evidence of nearby genes. Eighteen of the CTE sites occurred in introns of known or predicted genes with experimental evidence (spliced ESTs in at least one organism or conservation of multiple exons with the Takifugu genome). The orientation of these sites with respect to host genes was highly biased, with 15 of 18 arranged in the antisense orientation ($P = 0.005$, two-tailed *t*-test). This provides further evidence that sense-oriented retroviral insertions exert a selective pressure, to which the *M. m. castaneus* allele of *Nxf1* may be a response.

DISCUSSION

Here we described the genetic mechanism, properties and identity of a novel suppressor of retrovirus insertional mutations. Retrovirus insertions can alter host gene function by altering transcriptional regulation, interrupting coding sequence or reducing the processing efficiency of nascent RNA. Suppressive alleles of *Mvb1* derived from CAST/Ei seem to act specifically on hypomorphic mutations caused by insertions into introns that result in reduced expression of a normal mRNA. *Mvb1* does not act on IAP insertions known to affect transcriptional initiation. In both *vb*/*vb* and *Eya1*$^{BOR/BOR}$ mice, we identified aberrant RNAs in mutant tissues that are reduced in favor of normally processed mRNAs by the presence of *Mvb1*^{CAST} alleles, strongly implying that *Mvb1* has a role in mRNA processing. This activity is broadly expressed: we observed effects in brain, liver, kidney, muscle and otic development.

We found that this altered mRNA processing phenotype results from an allelic alteration at *Nxf1*, encoding an mRNA export factor, which is homologous to yeast *Mex67*. Recent work has identified a few key protein factors that link sequential steps in mRNA biogenesis, such as Aly/REF1. Aly forms part of the TREX complex, which links transcription to export in yeast and mammalian cells[27]. Aly acts as a transcriptional cofactor[28], is independently recruited to nascent RNAs along with splicing factors[29] and acts in export, in part through interactions with Nxf1 (ref. 30). Although Nxf1 has been shown to interact with other factors in mRNA biogenesis, including Aly, these interactions have been previously interpreted as recruitment of the export machinery to processed mRNA. Our results are the first to show an influence of the canonical export receptor on preceding steps in pre-mRNA processing. This could occur through kinetic competition among alternate pathways for nascent RNA or through altered RNP configuration on nascent RNAs to which Nxf1 is bound (**Fig. 6**). The location and nature of the E610G polymorphism in the UBA domain suggest a possible alteration in the interaction of Nxf1 with the nuclear pore, which could affect RNP assembly and processing by altering the latency of nascent or partially processed RNA for transport or otherwise moving the transcript away from processing compartments in the nucleus.

The distribution of *Nxf1* alleles in wild mouse populations indicates that this allele has undergone substantial recent fixation among *M. m. castaneus* mice in southeast Asia, a hallmark of recent selection. Although this could also result from hitchhiking or neutral drift, the fact that this site alters a conserved charge on a functional domain of the protein, yet is the only site for which the major allele occurs on a single extended haplotype, argues that this may be a site undergoing adaptive selection. In addition, we show that the class of event on which the CAST *Nxf1* allele acts is significantly underrepresented in the C57BL/6 genome (which lacks this compensating mechanism), providing another level of evidence to suggest at least a contextual selective advantage in populations where new retroviral insertion arise with some frequency.

The properties of *Mvb1* suggest that it could be used to engineer a binary system for titrating gene expression *in vivo*. This could be particularly useful for creating mouse models of human disorders in which loss of gene expression is a target of therapeutic efforts and dose-responsiveness of functional recovery is unknown (such as Fragile X syndrome and certain cancers) or for titrating the effects of mutations created by insertional elements reported to use an NXF1-binding CTE, such as human LINEs[31]. Titrating expression from the endogenous promoter by a processing step should allow level-of-expression manipulation independent of timing or cell-type distribution of expression.

METHODS

Mice and genetic typing. We obtained mutant mice from the Jackson Laboratory and maintained them locally as stocks. We made the vibrator mutation and the *Mvb1* interval from strain CAST/Ei simultaneously congenic on C57BL/6J (B6). We derived B6.CAST–*Mvb1* mice, bred to other mutants, from our colony of B6–*Pitpn*^{vb} *Mvb1*^{CAST} mice as incipient congenics at N6 or later. B6.CAST–*Mvb1* mice bred to transgenics were N15 or later. We determined genotypes by PCR assays specific for each mutation or for unique microsatellites flanking the mutant locus.

Molecular biology. We isolated RNA using Trizol (Invitrogen) according to the manufacturer's protocol. We prepared northern blots by separating RNA treated with formaldehyde or glyoxal by electrophoresis and capillary transfer to nylon membranes. We verified cloned probe fragments by DNA sequencing, isolated them by PCR and labeled them using random primer DNA synthesis in the presence of ^{32}P-dCTP. We quantified specific binding after hybridization and washing by phosphorimage analysis on a Molecular Dynamics Storm imager.

We prepared total protein for western blots in RIPA buffer, quantified it by Bradford assay and subjected it to SDS–PAGE. We detected PITPα and PITPβ with polyclonal antibodies directed against unique C-terminal peptides[3] followed either by secondary antibody coupled to horseradish peroxidase and ECL detection (**Fig. 1e**) or by ^{125}I-protein A and phosphorimage analysis (**Fig. 1f**).

We constructed recombinant lentiviral genomes from PCR-amplified fragments cloned into pLenti6/V5-TOPO (Invitrogen). Virions were packaged with VSV-G in the UCSD Gene Therapy Program Vector Development Lab.

We infected primary cells from *vb* mutant mice with viral supernatants and selected for blasticidin resistance before collection.

Electrophysiology. Hearing threshold was measured using auditory brainstem response (ABR) to click stimuli for the right ear by an investigator blind to *Mvb1* genotype. We anesthetized mice (with 50 mg ketamine hydrochloride per kg body weight, 5 mg xylazine hydrochloride per kg body weight and 1 mg acepromazine maleate per kg body weight intraperitoneally) and placed them in a single-walled acoustic booth (Industrial Acoustics) on a heating pad. We inserted subdermal electrodes (Astro-Med, Grass Instrument Division) at the vertex (active electrode), the mastoid (reference) and the hind leg (ground). Click stimuli (0.1 ms, 10 per s) were delivered to a Beyer DT 48 200-ohm speaker fitted with an ear speculum for placement in the external auditory meatus. The recorded ABR was amplified and digitized by a battery-operated preamplifier and input to an ABR recording system that provides computer control of the stimulus, recording and averaging functions (Tucker Davis Technology). We presented successively decreasing amplitude stimuli in 5-dB steps to the mouse and averaged and displayed the recorded stimulus-locked activity (n = 512). We defined threshold as the stimulus level between the record with no visibly detectable response and the record with a clearly identifiable response.

Histology. Vibrator mice were killed and perfused with saline and then Bouin's fixative. We removed brains, left them in fixative for several days, dehydrated them and embedded them in paraffin. We cut sagittal sections (10 μm) and stained them with luxol fast blue and cresyl violet. *Eya1* mice were killed after we measured ABR by cardiac perfusion with warm saline and phosphate-buffered, 4% paraformaldehyde. We dissected the cochleas, left in fixative overnight (at 4 ºC, decalcified them in 10% EDTA for 1 week, dehydrated them and embedded them in paraffin. We cut sections (7 μm) parallel to the modiolus and stained them with hematoxylin and eosin. Histology scores from 1 to 5 were assessed by a trained investigator blind to genotypes.

Transgenic mouse production. We subcloned an 18-kb *Spe*I fragment from B6-derived BAC clone 253I1 into λBlueStar (Novagen). We gel-isolated a 16.4-kb fragment containing 16,406 bp from the *Spe*I site adjacent to the *Stx5a* polyadenylation signal through *Nxf1* to a *Sse*8387I site in the first intron of the non-polymorphic gene encoding cDNA clone NM_1025791 and approximately 10 bp of pBlueStar cloning site sequence and injected this fragment into pronuclei of B6D2 hybrid mice. We screened pups for presence of the transgene by a competitive three-primer PCR assay using a forward primer corresponding to the cloning site of the vector at the 5′ end of the gene competing with a second forward primer outside the transgenic segment in the endogenous locus and a reverse primer in common to both transgenic and endogenous copies. Neither parental strain nor several derived hybrids had any observed *vb*-modifying effects in prior reports[3,32] or in non-transgenic littermates tested here.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We thank A. Ryan for advice and assistance with electrophysiology; X-D. Fu and C. J. Wills for discussions; M. Rosenfeld, R. Kolodner and A. Wynshaw-Boris for comments on draft manuscripts; A. Miyanohara for assistance with viral packaging; the UCSD Cancer Center Transgenic Mouse Facility for transgenic mouse production; and I. Kalcheva for assistance with BAC sequencing. This work was supported by grants from the US National Institutes of Health (B.A.H. and E.K.) and the Medical Research Service of the US Department of Veterans Affairs (E.K.). B.A.H. is a Pew Scholar in the Biomedical Sciences.

COMPETING INTERESTS STATEMENT

The authors declare competing financial interests (see the *Nature Genetics* website for details).

Received 20 May; accepted 10 September 2003
Published online at http://www.nature.com/naturegenetics/

1. Boeke, J.D. & Stoye, J.P. Retrotransposons, endogenous retroviruses, and the evolution of retroelements. in *Retroviruses* (eds. Coffin, J.M., Hughes, S.H. & Varmus, H.E.) 343–436 (Cold Spring Harbor Laboratory Press, Plainview, New York, 1997).
2. Hamilton, B.A. & Frankel, W.N. Of mice and genome sequence. *Cell* **107**, 13–16 (2001).
3. Hamilton, B.A. *et al.* The vibrator mutation causes neurodegeneration via reduced expression of PITPα: positional complementation cloning and extragenic suppression. *Neuron* **18**, 711–722 (1997).
4. Waterston, R.H. *et al.* Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**, 520–562 (2002).
5. Johnson, K.R. *et al.* Inner ear and kidney anomalies caused by IAP insertion in an intron of the *Eya1* gene in a mouse model of BOR syndrome. *Hum. Mol. Genet.* **8**, 645–653 (1999).
6. Tabernero, C. *et al.* Identification of an RNA sequence within an intracisternal-A particle element able to replace Rev-mediated posttranscriptional regulation of human immunodeficiency virus type 1. *J. Virol.* **71**, 95–101 (1997).
7. Lindtner, S., Felber, B.K. & Kjems, J. An element in the 3′ untranslated region of human LINE-1 retrotransposon mRNA binds NXF1(TAP) and can function as a nuclear export element. *RNA* **8**, 345–356 (2002).
8. Yonekawa, H. *et al.* Hybrid origin of Japanese mice *Mus musculus molossinus*: evidence from restriction analysis of mitochondrial DNA. *Mol. Biol. Evol.* **5**, 63–78 (1988).
9. Xu, P.X. *et al.* Eya1-deficient mice lack ears and kidneys and show abnormal apoptosis of organ primordia. *Nat. Genet.* **23**, 113–117 (1999).
10. Mercer, J.A., Seperack, P.K., Strobel, M.C., Copeland, N.G. & Jenkins, N.A. Novel myosin heavy chain encoded by murine dilute coat colour locus. *Nature* **349**, 709–713 (1991).
11. Cachon-Gonzalez, M.B. *et al.* Structure and expression of the hairless gene of mice. *Proc. Natl. Acad. Sci. USA* **91**, 7717–7721 (1994).
12. Potter, G.B. *et al.* The hairless gene mutated in congenital hair loss disorders encodes a novel nuclear receptor corepressor. *Genes Dev.* **15**, 2687–2701 (2001).
13. Stoye, J.P., Fenner, S., Greenoak, G.E., Moran, C. & Coffin, J.M. Role of endogenous retroviruses as mutagens: the hairless mutation of mice. *Cell* **54**, 383–391 (1988).
14. Bultman, S.J. *et al.* 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. *Genes Dev.* **8**, 481–490 (1994).
15. Duhl, D.M., Vrieling, H., Miller, K.A., Wolff, G.L. & Barsh, G.S. Neomorphic agouti mutations in obese yellow mice. *Nat. Genet.* **8**, 59–65 (1994).
16. Vasicek, T.J. *et al.* Two dominant mutations in the mouse fused gene are the result of transposon insertions. *Genetics* **147**, 777–786 (1997).
17. Zeng, L. *et al.* The mouse Fused locus encodes Axin, an inhibitor of the Wnt signaling pathway that regulates embryonic axis formation. *Cell* **90**, 181–192 (1997).
18. Fagotto, F. *et al.* Domains of axin involved in protein-protein interactions, Wnt pathway inhibition, and intracellular localization. *J. Cell Biol.* **145**, 741–756 (1999).
19. Hsu, W., Zeng, L. & Costantini, F. Identification of a domain of Axin that binds to the serine/threonine protein phosphatase 2A and a self-binding domain. *J. Biol. Chem.* **274**, 3439–3445 (1999).
20. Sheldon, M. *et al.* Scrambler and yotari disrupt the disabled gene and produce a reeler-like phenotype in mice. *Nature* **389**, 730–733 (1997).
21. Ware, M.L. *et al.* Aberrant splicing of a mouse disabled homolog, *mdab1*, in the scrambler mouse. *Neuron* **19**, 239–249 (1997).
22. Gruter, P. *et al.* TAP, the human homolog of Mex67p, mediates CTE-dependent RNA export from the nucleus. *Mol. Cell* **1**, 649–659 (1998).
23. Bachi, A. *et al.* The C-terminal domain of TAP interacts with the nuclear pore complex and promotes export of specific CTE-bearing RNA substrates. *RNA* **6**, 136–158 (2000).
24. Grant, R.P., Hurt, E., Neuhaus, D. & Stewart, M. Structure of the C-terminal FG-nucleoporin binding domain of Tap/NXF1. *Nat. Struct. Biol.* **9**, 247–251 (2002).
25. Grant, R.P., Neuhaus, D. & Stewart, M. Structural basis for the interaction between the Tap/NXF1 UBA domain and FG nucleoporins at 1Å resolution. *J. Mol. Biol.* **326**, 849–858 (2003).
26. Sabeti, P.C. *et al.* Detecting recent positive selection in the human genome from haplotype structure. *Nature* **419**, 832–837 (2002).
27. Strässer, K. *et al.* TREX is a conserved complex coupling transcription with messenger RNA export. *Nature* **417**, 304–308 (2002).
28. Bruhn, L., Munnerlyn, A. & Grosschedl, R. ALY, a context-dependent coactivator of LEF-1 and AML-1, is required for TCRα enhancer function. *Genes Dev.* **11**, 640–653 (1997).
29. Zhou, Z. *et al.* The protein Aly links pre-messenger-RNA splicing to nuclear export in metazoans. *Nature* **407**, 401–405 (2000).
30. Gatfield, D. & Izaurralde, E. REF1/Aly and the additional exon junction complex proteins are dispensable for nuclear mRNA export. *J. Cell Biol.* **159**, 579–588 (2002).
31. Ostertag, E.M. *et al.* A mouse model of human L1 retrotransposition. *Nat. Genet.* **32**, 655–660 (2002).
32. Weimar, W.R., Lane, P.W. & Sidman, R.L. *Vibrator* (*vb*): a spinocerebellar system degeneration with autosomal recessive inheritance in mice. *Brain Res.* **251**, 357–364 (1982).
33. Seperack, P.K., Mercer, J.A., Strobel, M.C., Copeland, N.G. & Jenkins, N.A. Retroviral sequences located within an intron of the dilute gene alter dilute expression in a tissue-specific manner. *EMBO J.* **14**, 2326–2332 (1995).