In vivo *Atoh1* targetome reveals how a proneural transcription factor regulates cerebellar development

Tiemo J. Klisch^{a,b,c,1}, Yuanxin Xi^{d,1}, Adriano Flora^{a,c}, Ligu Wang^d, Wei Li^{d,2}, and Huda Y. Zoghbi^{a,b,c,e,2}

^aDepartment of Molecular and Human Genetics, ^bThe Howard Hughes Medical Institute, ^dDivision of Biostatistics, The Dan L. Duncan Cancer Center, Department of Molecular and Cellular Biology, and ^eDepartments of Neuroscience and Pediatrics, and Program in Developmental Biology, Baylor College of Medicine, Houston, TX 77030; and ^cThe Jan and Dan Duncan Neurological Research Institute at Texas Children's Hospital, Houston, TX 77030

Contributed by Huda Y. Zoghbi, January 5, 2011 (sent for review November 16, 2010)

The proneural, basic helix–loop–helix transcription factor *Atoh1* governs the development of numerous key neuronal subtypes, such as cerebellar granule and brainstem neurons, inner ear hair cells, and several neurons of the proprioceptive system, as well as diverse nonneuronal cell types, such as Merkel cells and intestinal secretory lineages. However, the mere handful of targets that have been identified barely begin to account for *Atoh1*'s astonishing range of functions, which also encompasses seemingly paradoxical activities, such as promoting cell proliferation and medulloblastoma formation in the cerebellum and inducing cell cycle exit and suppressing tumorigenesis in the intestine. We used a multipronged approach to create a comprehensive, unbiased list of over 600 direct *Atoh1* target genes in the postnatal cerebellum. We found that *Atoh1* binds to a 10 nucleotide motif (AtEAM) to directly regulate genes involved in migration, cell adhesion, metabolism, and other previously unsuspected functions. This study expands current thinking about the transcriptional activities driving neuronal differentiation and provides a framework for further neurodevelopmental studies.

transcriptional regulation | chromatin immunoprecipitation | deep sequencing | E-box motif

Proneural transcription factors play a key role in the specification of a multitude of neural progenitors in the developing brain. Several studies have suggested that the dozen or so proneural genes, which encode a family of closely related basic helix–loop–helix (bHLH) transcription factors (1), are able to be responsible for such protean developmental programs by activating “secondary” transcription factors, which over time induce effector genes (2, 3). Unfortunately, it has been difficult to identify the direct target genes of these proneural factors, which can exert an impressive range of effects.

For example, *Atoh1*, the mouse homolog of *Drosophila atonal* (4, 5), governs the differentiation of key populations of the nervous system (cerebellar granule and brainstem neurons, inner ear hair cells, and numerous components of the proprioceptive and interoceptive systems) (6–10), as well as diverse nonneuronal cell types (e.g., Merkel cells and intestinal secretory lineages) (11–14). Although studies based on candidate gene approaches have revealed a few putative targets of *Atoh1* (10, 15–18), this handful of targets barely begins to account for this astonishing range of functions, which encompasses seemingly paradoxical activities, such as promoting cell proliferation (and medulloblastoma) in the cerebellum and promoting cell cycle exit and suppressing tumorigenesis in the intestine (17, 19, 20).

To better understand *Atoh1*'s role in neurogenesis, we set out to identify its molecular functions by defining its direct target genes in vivo in the cerebellar granule neuron precursors (CGP), a key population in the developing brain. The CGPs are specified during early embryonic development [embryonic day (E) 11–14] and, after a clonal expansion during the first postnatal week in mouse, differentiate to give rise to the granule cells of the cerebellum (21, 22). Interestingly, postnatal day 0 (P0) CGPs are not completely committed to a neuronal fate, although the majority of

later CGPs (around P5) are already committed (23). *Atoh1* is expressed in the proliferating CGPs at both times, providing an ideal opportunity to study its various functions.

We took a three-pronged approach to define *Atoh1* targets. First, we determined an *Atoh1* DNA-binding signature by identifying genomic *Atoh1* binding sites with in vivo ChIP using the *Atoh1* FLAG-tagged knock-in mouse model (17). Second, we generated a CGP chromatin signature by identifying the global histone H3 lysine 4 methylation status (Histone-seq). Third, we established an *Atoh1* expression signature in CGPs by mapping gene expression differences between WT and *Atoh1*-null cerebellar tissue (RNA-seq). Analysis of these results identified hundreds of *Atoh1* targets. By analyzing this *Atoh1* “targetome,” we found that *Atoh1*'s function in CGP development is much broader than anticipated.

Results

We took advantage of the *Atoh1*-floxed conditional knockout mouse model (24), as well as a hormonally inducible CrePR allele knocked into the *Atoh1* locus (*Atoh1*^{CrePR}) that has a restricted and low recombination activity in CGPs (25). This mouse model enables us to delete *Atoh1* specifically in the cycling progenitors of the outer external granule layer (EGL). We crossed *Atoh1*^{CrePR/flox} mice with mice carrying the Cre-inducible YFP gene in the Rosa locus (R26YFP) (26). After activation of Cre by RU486, specific cells in *Rosa*^{YFP/YFP}; *Atoh1*^{CrePR/flox} animals (designated here as *Atoh1*^Δ) lose *Atoh1* expression and turn on the reporter gene. We deleted *Atoh1* either at P0 or P5 and used the inward migration of differentiating CGPs as a read-out for differentiation (27). Twenty-four hours after Cre recombination, we collected the tissue and counted YFP⁺ cells with respect to their position to the cerebellar surface. YFP⁺ cells located in the EGL were classified as undifferentiated, and YFP⁺ cells in the internal granule layer (IGL) were classified as differentiated (Fig. 1). After *Atoh1* deletion at P0, a large number of cells remained undifferentiated (Fig. 1A); *Atoh1* deletion at P5 had less effect on differentiation, as most *Atoh1*-null cells migrated into the IGL (Fig. 1B).

To quantify the data, we generated a differentiation index by calculating the ratio of cells located in the IGL relative to the EGL. The differentiation index at P0 was much lower than at P5, suggesting that at an early time point many cells require *Atoh1* to

Author contributions: T.J.K., W.L., and H.Y.Z. designed research; T.J.K. and A.F. performed research; Y.X., L.W., and W.L. contributed new reagents/analytic tools; T.J.K., Y.X., A.F., L.W., W.L., and H.Y.Z. analyzed data; and T.J.K., A.F., W.L., and H.Y.Z. wrote the paper.

The authors declare no conflict of interest.

Data deposition: All primary sequence data of RNA-seq, Histone-seq, and *Atoh1* ChIP-seq have been deposited in the National Center for Biotechnology Information Gene Expression Omnibus repository, www.ncbi.nlm.nih.gov/geo (accession no. GSE22111).

¹T.J.K. and Y.X. contributed equally to this work.

²To whom correspondence may be addressed. E-mail: WL1@bcm.edu or hzoghbi@bcm.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1100230108/-DCSupplemental.

Fig. 1. Loss of *Atoh1* leads to an early differentiation halt. (A and B) Using *Atoh1*^{CrePR1/flox}; *Rosa*^{YFP/YFP} mice, we depleted *Atoh1* by administering RU486 to the cerebellum at P0 (A) or P5 (B) and collected the cerebellum 24 h later. In the 20- μ m sagittal sections, *Atoh1*-depleted cells were visualized by YFP expression (green channel). Nuclei are stained with DAPI (blue channel). (Scale bars, 50 μ m.) (C) Quantification of differentiating cells measured by their location. Shown is the ratio of cells located in the IGL in respect to the EGL.

differentiate (Fig. 1C). To confirm that these results are caused by *Atoh1* deletion from dividing CGPs situated on the outer EGL, we used an organotypic culture system in which we applied GFP-containing or Cre-IRES-GFP adenoviruses on the surface of cultured *Atoh1*^{flox/flox} P0 or P5 cerebella to infect only the outermost cell layers of the EGL (Fig. S1A). CGPs in which *Atoh1* was deleted at P0 still resided on the outer surface (Fig. S1A). CGPs in which *Atoh1* was deleted at P5, however, migrated inwards.

These data suggest that although *Atoh1* is highly expressed in the EGL at both P0 and P5, perhaps some of its key targets that are robustly expressed at P5 are able to drive EGL differentiation. We thus set out to identify the direct *Atoh1* target genes in vivo.

Generation of Three Global Cerebellar Signatures. We defined a gene as a direct target when (i) its locus is bound by *Atoh1*, (ii) it is transcriptionally active in the developing cerebellum, and (iii) its expression is misregulated after *Atoh1* deletion. First, we identified the in vivo DNA-binding sites of *Atoh1* using ChIP followed by deep sequencing (ChIP-seq) at P5, using a FLAG-tagged *Atoh1* knock-in mouse model to increase the efficiency and specificity of the *Atoh1* ChIP-seq. (17). These ChIP-seq experiments defined 19,227 putative *Atoh1* binding sites in the genome (Dataset S1 and Dataset 2). The high-quality of the data were confirmed by the good reproducibility of the two samples (Fig. S1B), high sequencing saturation (Fig. S2A), and high validation rate: 27 of 30 randomly picked positive regions and 0 of 24 randomly picked negative regions were verified by ChIP-qPCR (Fig. S2B). Moreover, we independently identified and validated four of the five described *Atoh1* target regions [*Atoh1* (10), *Barhl1* and *Barhl2* (15), and *Gli2* (17)] and found additional *Atoh1* binding sites in *Atoh1*, *Barhl1*, and *Gli2* (Fig. S2 C and D and Dataset S2).

To identify *Atoh1* loci, which are associated with active transcription, we investigated the histone H3 lysine 4 monomethylation (H3K4me1) and trimethylation (H3K4me3) state at P5. H3K4me1 marks active distal regulatory regions, and H3K4me3 is present at active promoter regions (28). Our Histone-seq data were robust, as evidenced by good reproducibility (Fig. S1C)

and sequence saturation (Fig. S2 E and F and Dataset S1). Interestingly, *Atoh1* binding peaks were highly correlated with H3K4me1 but not H3K4me3 (Fig. 2 A and B, Fig. S1 B and C, and Dataset S1), indicating that most *Atoh1* binding regions were located in enhancer regions.

As a third criterion, we wanted to analyze the gene-expression changes associated with the loss of *Atoh1*. Because *Atoh1*-null mice die at birth (9), we could not use P5 cerebella but compared the transcriptional profiles of E18.5 WT and *Atoh1*-null cerebellar tissue instead. Having obtained high-quality data [as judged by its high reproducibility (Fig. S1D), equal transcript coverage, and high sequence saturation (Fig. S3 A–C and Dataset S1)], we created a differentially expressed transcript list and identified 4,064 differentially expressed transcripts with an adjusted *P* value less than 0.01, of which most were down-regulated in the *Atoh1*-null tissue (Fig. S3D and Dataset S3). Importantly, the overall histone methylation status of the gene bodies at P5 was highly enriched in the RNA transcripts identified at E18.5 compared with transcripts that are not expressed, suggesting that these transcripts are still actively transcribed (Fig. 2 C and D).

***Atoh1* E-Box Associated Motif Characterizes the *Atoh1* Genomic Signature.** Because proneural factors bind to a loose E-Box consensus sequence, CANNTG (29), we searched for this motif in our target regions. We could identify at least one E-Box in 91% (17,568) of *Atoh1*-bound regions, consistent with the idea that these could be true *Atoh1*-binding sites. Little is known

Fig. 2. H3K4 ChIP-seq analyses. (A and B) Spearman rank correlation of the histone methylation levels (normalized peak intensities at the *Atoh1* binding sites) with *Atoh1* DNA-binding levels. Although the *Atoh1* DNA-binding regions were less correlated to H3K4me3 (B), they possessed H3K4me1 marks (A). (C and D) Average ChIP-seq profiles for H3K4me1 (C) and H3K4me3 (D) in relation to the E18.5 transcript signature. The normalized read numbers per 10 million reads are shown in a 5-kb window, in which each detected transcript is normalized to 3 kb in length, with 1 kb upstream of transcriptional start site (TSS) and 1 kb downstream of transcriptional end site (TES). H3K4me1 marks are enriched over the entire region (C); H3K4me3 show two strong peaks, roughly one nucleosome up- and downstream of the TSS (D).

about sequences flanking the E-Box that dictate the specificity of proneural factors binding to DNA in vivo. Exploiting the ChIP-seq information, we identified a unique 10-nucleotide palindrome (Fig. 3A) with an E-Box at its core, which we named the Atoh1 E-Box Associated Motif, or AtEAM. Most Atoh1-bound regions possessed at least one AtEAM (Fig. 3A), over half of which (53%) were conserved in mammals. Electrophoretic mobility-shift assays demonstrated the specificity of the motif, which bound Atoh1 much more strongly than Ascl1 (Fig. 3B, lanes 2 and 6), a related proneural factor that binds generic E-Boxes but is not expressed in the cerebellum (30). Atoh1 and Ascl1 were similarly expressed in the cellular extracts used for the EMSA based on Western blot analysis.

Luciferase reporter assays using a single AtEAM motif in front of a minimal promoter (which more closely mirrors the endogenous binding characteristics) demonstrated that Atoh1 activated the WT reporter over twofold (Fig. 3C). Introducing single-point mutations for individual nucleotides leaving the E-Box core intact, Atoh1 still strongly activated the reporter (Fig. 3C). Two mismatches, however, whether consecutive or interspersed, abolished Atoh1 binding (Fig. 3C). Only Atoh1, and not Ascl1 or

the proneural factor Ngn1, activated a luciferase reporter construct bearing a single AtEAM (Fig. 3D), confirming Atoh1's selectivity. It thus appears that the AtEAM is a key determinant of Atoh1 binding to its target sequences.

Identifying Atoh1 Targets. We combined the Atoh1-binding (ChIP-seq), histone methylation (Histone-seq), and expression signatures (RNA-seq) to generate a meaningful Atoh1 targetome, which is unique in being a comprehensive direct target gene list for a proneural factor to be created under physiological conditions (Dataset S4) (see *SI Materials and Methods* for methodology). From the list of targets, we selected 633 transcripts corresponding to 601 genes with a *P* value of less than 0.01 (Dataset S5). Among these were the previously reported Atoh1 targets *Atoh1* (10), *Barhl1* (15), and *Gli2* (17), but also crucial neuronal differentiation factors that have not previously been shown to be direct targets of Atoh1, such as *NeuroD1*, *NeuroD2*, and *NeuroD6*, as well as *Nhlh1* and *Nhlh2*.

To validate our targetome, we first focused on genes that have been implicated in cerebellar development by consulting the mouse-mutant database (31). Of 601 Atoh1 targets, 22 genes were associated either with "abnormal cerebellar granule layer" (15 genes, *P* value = 2.61E-17, 11-fold enrichment), "abnormal cerebellar development" (16, genes, *P* value = 1.06E-16, 15-fold enrichment), or both (9 genes). All 22 genes were validated by ChIP-PCR for Atoh1 binding in P5 CGPs (Fig. 4A). Moreover, the expression of *Barhl1* and five previously unrecorded Atoh1 targets (*Ccnd2*, *Pax6*, *Plxn2b*, *Cxcr4*, *Sema7a*) was lost in *Atoh1*-null E18.5

Fig. 3. AtEAM discovery and characterization. (A) Distribution analysis of the AtEAM shown as pie charts, including the absolute numbers for each category and percentile. The motif is shown on top. The plus sign indicates that the category includes the shown number of motifs or more. The majority of Atoh1-bound regions possess at least one AtEAM. (B) EMSA validating the Atoh1 binding efficiency to AtEAM; Neuro2a cells and labeled oligo (lane 1); transfected with Atoh1-FLAG (lanes 2–5) or Ascl1-FLAG expression plasmids (lane 6). Although there is a weak shift in the Ascl1-FLAG transfected sample (lane 6), Atoh1-FLAG transfection results in a strong shift (lane 2) compared with the untransfected control (lane 1). Competitors are in 100× molar excess. (C) Dual luciferase reporter assay using a firefly reporter construct harboring one AtEAM in front of a minimal promoter either WT or mutated in designated positions. Data are presented as mean $\pm$ SEM after normalization to *Renilla* control. Cotransfection of Atoh1-FLAG (250 ng) activates the WT reporter [0 mismatch mutations (MM), green bar] over twofold. One mismatch mutation (orange) had little effect on the induction capacity of Atoh1; 2 MM (red) abolished the activation. Mutation positions are indicated on the x axis. ANOVA with Tukey's post hoc *t* test revealed a statistical significance of $*P < 0.01$. (D) Dual luciferase reporter assay using WT luciferase construct in DAOY cells, cotransfected with either Atoh1-FLAG (green bars), Ascl1-FLAG (red bars), or Ngn1 (blue bars). Data are presented as mean $\pm$ SEM. Triangle indicative for increasing amounts of transfected DNA (0, 50, 100, 250 ng). Only Atoh1 is able to activate the reporter ($*P < 0.01$).

Fig. 4. The Atoh1 targetome is enriched for genes relevant to Atoh1 null phenotypes. (A) Validation of the 22 genes identified by the mouse mutant analysis using Atoh1 ChIP-PCR. All regions are detected in the input (I) and elution (E) fractions of the cerebellum (Cb) ChIP, but not in the elution of the forebrain (Fb) ChIP, where Atoh1 is not expressed. The negative region corresponds to the Atoh1 TSS region not bound by Atoh1 in ChIP-seq. (B) In situ hybridization of Atoh1 target genes in E18.5, 20- μ m sagittal sections of Atoh1 wildtype (*Atoh1*^{+/+}) and Atoh1-null (*Atoh1*^{-/-}) tissue. Like Atoh1, all targets are expressed in the EGL. (Scale bar, 200 μ m.)

cerebellar anlage tissue (Fig. 4B), strongly supporting the notion that our *Atoh1* targetome reflects true in vivo *Atoh1* targets.

***Atoh1* Regulates Proliferation and Directly Influences CGP Differentiation.** To better understand the function of *Atoh1* through the analysis of its target genes, we exploited the gene ontology database DAVID (32). We found that *Atoh1* regulates not only genes associated with transcription (105 genes, P value = $7.35\text{E-}10$) but also genes involved in cell cycle (33 genes, P value = $5.85\text{E-}04$) and chromosomal organization (33 genes, P value = $4.36\text{E-}07$). *Atoh1* also directly activates genes involved in primary metabolic processes and their regulation (132 genes, P value = $1.03\text{E-}03$ and 110 genes, P value = $5.35\text{E-}05$, respectively). In addition, we found an enrichment in ribonucleoprotein complex genes (30 genes, P value = $2.13\text{E-}04$) and genes associated with the mitochondrion (64, P value = $6.95\text{E-}04$). All P values were false-discovery rate-adjusted (32). We validated a subset of these genes by ChIP-PCR (Fig. S4A). These data indicate that *Atoh1* helps regulate proliferation and maintenance of CGPs, pointing to a broad role for *Atoh1* in progenitor biology.

Although we expected *Atoh1* to regulate some genes involved in signal transduction cascades, as CGPs are exposed to numerous external signals (33), a literature-based pathway analysis discovered at least 100 *Atoh1* targets associated with signaling pathways. Besides the already known pathways involved in cerebellar development (33), such as Sonic Hedgehog (15 genes), Notch (15 genes), TGF- β (14 genes), retinoic acid (6 genes), and Wnt signaling (15 genes), we identified many components of MAP kinase pathways. Among our targets, 22 genes were associated with ERK signaling, 15 with NF- κ B signaling, and 11 with JNK signaling (Dataset S5). Using ChIP-PCR, we tested and validated a subset of these genes (Fig. S4B). Our data thus point to an unexpectedly prominent role for *Atoh1* in allowing CGPs to respond to external cues.

During postnatal development, CGPs undergo a rapid clonal expansion, start their differentiation program, and migrate inwards (21). The *Atoh1* target list reflected these key features, as it was enriched for the terms proliferation (36 genes; P value = $9.11\text{E-}07$), differentiation (73 genes; P value = $2.83\text{E-}04$), cell adhesion (37 genes; P value = $4.32\text{E-}04$), and migration (23 genes; P value = $1.37\text{E-}06$) (Dataset S5). We validated a subset of these by ChIP-PCR (Fig. S4C). To further validate our results, we analyzed the colocalization of six differentiation proteins (Pax6, NeuroD6, NeuroD1, Nhlh1/2, and Sema7a) (red channel in Fig. 5A) in P0 and P5 cerebellar tissue with *Atoh1* using the GFP-tagged *Atoh1* knock-in mouse (green channel in Fig. 5A) (25). All six proteins were colocalized with *Atoh1* at either P0 or P5 (white in Fig. 5A). Intriguingly, the transcription factors NeuroD1 and Nhlh1/2 were expressed in the EGL at high levels only at P5, suggesting that they accumulate over time in *Atoh1*-expressing CGPs or are activated by *Atoh1* at this later time point (Fig. 5A).

***Atoh1* Primes CGP Differentiation by Early Gene Activation.** The fact that EGL cells, which normally express *Atoh1*, can differentiate even in the absence of *Atoh1* suggests that the level of some *Atoh1* targets might increase over time to reach a threshold that is sufficient to sustain the differentiation process. If this assumption is correct, loss of *Atoh1* at a very early stage of cerebellar development should halt the CGP differentiation process, whereas deletion at a later stage would have less effect because the downstream factors have already accumulated. Note that this hypothesis is in agreement with the results shown in Fig. 1, where conditional deletion of *Atoh1* at P0 had a more pronounced effect on cell migration than deletion at P5. To explore whether the migratory defect was a reflection of a differentiation deficit, we performed lineage-tracing experiments using *Atoh1*^{CrePR/flox}; *Rosa*^{YFP/YFP} animals. We focused on P5 cerebellum, where we

Fig. 5. *Atoh1* activates and colocalizes with CGP differentiation markers. (A) Colocalization of Pax6, NeuroD6, NeuroD1, Nhlh1/2, and Sema7a (red channel) with *Atoh1* (green channel) in 20- μ m sagittal sections of P0 and P5 *Atoh1*^{GFP/GFP} cerebella. (B) Using *Atoh1*^{CrePR/flox}; *Rosa*^{YFP/YFP} mice, we depleted *Atoh1* at P5 and collected the cerebellum 24 h later. In the 20- μ m sagittal sections, *Atoh1* depleted cells were visualized by YFP expression (green channel). The differentiation marker Tuj1 (red channel) colocalized with *Atoh1*-independent cells located in the IGL, but it was absent from *Atoh1*-dependent cells located in the EGL. Although the expression of Pax6 and NeuroD6 is *Atoh1*-independent, the expression of NeuroD1, Nhlh1/2, and Sema7a is *Atoh1*-dependent while the CGPs still reside in the EGL. Colocalization is shown in white. (Scale bars, 20 μ m.)

could identify a minority of cells whose migration was still dependent on *Atoh1* (i.e., did not migrate out of the EGL when *Atoh1* was deleted) (Fig. 1B) and a consistent number of cells whose migration was *Atoh1*-independent (i.e., migrated into the IGL in spite of *Atoh1* deletion) (Fig. 1B).

If *Atoh1* were purely a proliferation factor, as suggested by its exclusive expression in the EGL, its deletion should cause premature differentiation. Instead, we find that the EGL-residing cells that deleted *Atoh1* were negative for the differentiation marker Tuj1, suggesting that they require *Atoh1* to initiate the neuronal differentiation program (Fig. 5B). The lack of differentiation is also reflected in the loss of target-gene expression of additional differentiation markers, such as NeuroD1, Nhlh1/2, and Sema7a (Fig. 5B). The IGL-residing, *Atoh1*-deleted cells were positive for these markers, however, indicating that these cells did not require *Atoh1* for differentiation (Fig. 5B).

Discussion

We generated an extensive list of *Atoh1* primary targets in the postnatal cerebellum by combining *Atoh1* in vivo binding data, histone methylation signatures, and expression data to uncover the roles of *Atoh1* in cerebellar development. The list of genes that we generated (Dataset S5) included the majority of the

previously identified targets (10, 15, 17) and was highly enriched in genes reported to have loss-of-function phenotypes similar to *Atoh1*-null mice, validating our approach. This targetome provides insight into the role of *Atoh1* in cerebellar granule genesis.

AtEAM Is a Unique, *Atoh1*-Specific DNA Binding Sequence. Proneural bHLH transcription factors recognize a generic canonical 6-mer E-Box motif (CANNTG) with further specification imparted by the amino acid sequence of the bHLH domains of the factors themselves (30, 34). Recent genome-wide studies have shown that the binding motifs identified using *in vitro* methods are sometimes remarkably different from the actual *in vivo* motifs (35).

Using *in vivo* *Atoh1* binding regions, we identified a unique 10-mer *Atoh1* binding motif, AtEAM, present in the known genomic loci regulated by *Atoh1* that have been identified by a candidate gene approach (10, 15–17). The AtEAM is highly enriched in the *Atoh1* targetome. Despite the importance of proneural transcription factors in mammalian development, their *in vivo* DNA binding motifs have yet to be mapped. In *Drosophila* it has been shown that several proneural transcription factors have different binding affinities, attributable to subtle preferences in nucleotides flanking the E-Box core (36), but distinct sequence motifs have been elusive. Our identification of this highly specific target sequence explains, at least in part, the differences in the function of proneural factors and highlights the utility of global *in vivo* ChIP-seq analyses in identifying DNA-binding motifs.

Diversity of *Atoh1* Target Genes. Two studies reported that *Atoh1* modulates Sonic Hedgehog (Shh) signaling to regulate CGP proliferation (17, 34); our data confirm this function as the *Gli2* locus has 18 additional *Atoh1* binding regions (Dataset S2). Moreover, *Atoh1* controls 57 genes associated with cell cycle regulation and proliferation (Dataset S5), including 15 additional *Atoh1* targets involved in Shh signaling, such as *Cend2*, *Ptchd2*, and *Mycn* (Dataset S5). We not only identified *Mycn* as a direct target of *Atoh1* but also *Mxd3*, a gene needed for CGP proliferation (37) and *Mxd4*. We propose that *Atoh1* might also influence the Myc-Max-Mad network, which plays a role in CGP proliferation (38).

Atoh1's function is not restricted to the regulation of genes directly involved in cell cycle, but extends to the regulation of genes involved in cell metabolism and ribosome biogenesis (Dataset S5). These genes have been shown to be necessary to ensure the high energy and biosynthesis demand of cycling precursors is met (39, 40). Our data thus suggest a major role for *Atoh1* in regulating the biological processes supporting proliferation in the postnatal developing cerebellum and point to *Atoh1* as a key contributor to tumors derived from CGPs (17, 19).

Not only does *Atoh1* regulate the proliferative state of the CGP, but it also plays a key role regulating genes critical for granule cell migration (Fig. S4 and Dataset S5). We show that several migratory genes are expressed in the EGL and lost in the *Atoh1*-null (Fig. 4B) and identified a number of genes involved in semaphorin signaling, such as *Plxn2b* and *Sema6a*. Although granule migration is halted in the *Sema6a* knockout, the *Plxn2b* knockout displays a migratory defect of cycling CGPs (41, 42). By directly regulating both of these factors, *Atoh1* might be one factor controlling the transition between CGP proliferation and migration. It will be interesting to explore the crosstalk of semaphorins during migration and the interaction of these factors with other migratory pathways regulated by *Atoh1*, such as *Cxcr4*. *Atoh1* might tie the migratory machinery to underlying structural rearrangements of the cytoskeleton, as *Atoh1* directly regulates, among other factors implicated in migration, *Itgb1*, *Actb*, and *Myh9* (43–45).

***Atoh1* and House-Keeping Genes.** It is somewhat surprising that *Atoh1* regulates a number of “house-keeping” genes. Recent

data have shown that such “ubiquitously” expressed genes can be differentially regulated (46, 47). Broadly expressed genes could be regulated by at least two mechanisms, one that induces a ubiquitous and basal expression, and another more firmly regulated mechanism designed to meet the demands of particular cell types. This idea is supported by the observation that a lot of ribosomal targets, such as *Rpl35a*, *Mrpl11*, *Rps19*, *Rpl41*, and *Mrpl18*, follow the expression profile of *Atoh1*, which gradually decreases from E18 to P21 as shown in the Cerebellar Development Transcriptome Database (48).

***Atoh1* Primes the Differentiation Process.** The observation that deletion of *Atoh1* can have different effects on proliferating EGL cells indicates that *Atoh1*-expressing cells in the developing cerebellum exist in at least two different states: one that still requires *Atoh1* to differentiate, and another whose differentiation has advanced to become *Atoh1*-independent. Despite the fact that both populations are present at an early (P0) and a later (P5) stage of development, the *Atoh1*-dependent population is much more abundant at P0. The transition between an *Atoh1*-dependent to an *Atoh1*-independent state over time suggests that *Atoh1* might be important to initiate a differentiation program—a function that we refer to here as “priming”—that later becomes self-supporting and no longer requires *Atoh1*. A mechanistic explanation of *Atoh1*'s priming function is suggested by its targetome, which reveals numerous genes involved in multiple steps of the differentiation program. We propose that, at an early stage, *Atoh1* is the driving force of gene activation of both downstream differentiation transcription factors, such as *NeuroD1*, *Nhlh1*, and *Nhlh2*, and their effector genes, such as genes coding for cell-adhesion molecules, cytoskeleton rearrangement factors, and proteins involved in MAPK pathways. Later in development, the downstream transcription factors take over the cell's differentiation program, thereby allowing development to proceed even though *Atoh1* is down-regulated. Interestingly, the *Drosophila* homolog *atonal* does not seem to regulate genes directly linked to differentiation (49).

Consistent with these two roles of *Atoh1* in the developing cerebellum, we find that *Atoh1* deletion at an early stage (P0) strongly inhibits differentiation (Fig. 1A). At a later phase of cerebellar development, however, the main function of *Atoh1* seems to be in regulating cell proliferation, and its deletion results in exit from the cell cycle and differentiation. This model raises the question as to how the cell strikes the balance between proliferative and differentiative *Atoh1* activities during CGP development. An attractive explanation is that a subset of transcription factors regulated by *Atoh1* might accumulate over time and eventually activate more differentiation genes, thereby attenuating *Atoh1*'s function in proliferation (50), as has been proposed for the bHLH factor *Xngnr1* in *Xenopus* (51).

In summary, by focusing on the developing cerebellum we have found that *Atoh1* occupies a central position in the regulation of diverse steps of CGP development, from cell proliferation to differentiation and migration. Further insight into the dynamics of *Atoh1* binding throughout CGP development and identification of target genes of other transcription factors will uncover the similarities and differences in their genetic programs during development of one organ or even across organs, and allow the generation of an exhaustive molecular map of mammalian embryogenesis.

Materials and Methods

Mouse Strains, Timed Mating, and Genotyping. Animal housing, husbandry, and euthanasia were conducted under the guidelines of the Center for Comparative Medicine, Baylor College of Medicine. For mouse strains refer to *SI Materials and Methods*.

Next Generation Sequencing. *Atoh1* ChIP was performed as described (17), pooling four P5 cerebella. Histone-seq [using H3K4me1 (ab8895; Abcam) and

H3K4me3 (ab8580; Abcam)] was performed as described (28), using four P5 cerebella. RNA-seq was performed using poly-A enriched RNA. Solexa sequencing was performed at the Center for Cancer Epigenetics Solexa Sequencing Core at M D Anderson (Houston, TX), where the sequencing libraries were prepared. The details of preparation and the bioinformatics analyses are presented in *SI Materials and Methods*.

Immunofluorescence and Cell Quantification. Immunofluorescence was performed as previously described (17) using the following antibodies: rabbit anti-GFP (1:250; GTX26556; Gene Tex Inc.); rabbit anti-Pax6 (1:250; PRB-278P; Covance); rabbit anti-NeuroD1 (1:250; ab16508; Abcam); rabbit anti-Math2/NeuroD6 (1:250; ab85824; Abcam); rabbit anti-Nhlh1/2 (1:250; ab72601; Abcam); and rabbit anti-Sema7a (1:250; ab23578; Abcam). To generate the differentiation index, we quantified YFP⁺ cells from three different animals at each age (P0 and P5). For each animal we analyzed at least 15 sections at

a distance of 50 μ m each to span the entire length of the cerebellum from the rostral to the caudal extremities. We scored a total of 3,047 cells for P0 and 3,269 cells for P5 animals. The ratio between IGL (differentiated) and EGL (nondifferentiated) cells was calculated for each section and used to generate an average differentiation index.

ACKNOWLEDGMENTS. We thank Louis S. Ramagli from the MD Anderson Cancer Center for sequencing, Christina Thaller from the Gene Expression and Microscopy Cores of the Baylor Intellectual and Developmental Disabilities Research Center (HD024064) for in situ hybridization support, and the H.Y.Z. laboratory members and Vicky Brandt for critical input on the manuscript. This work was supported by The Howard Hughes Medical Institute (H.Y.Z.) and in part by the Duncan Scholar Award (to W.L.). H.Y.Z. is an investigator and T.J.K. a research associate of The Howard Hughes Medical Institute.

- Powell LM, Jarman AP (2008) Context dependence of proneural bHLH proteins. *Curr Opin Genet Dev* 18:411–417.
- Anderson DJ (1999) Lineages and transcription factors in the specification of vertebrate primary sensory neurons. *Curr Opin Neurobiol* 9:517–524.
- Bally-Cuif L, Hammerschmidt M (2003) Induction and patterning of neuronal development, and its connection to cell cycle control. *Curr Opin Neurobiol* 13:16–25.
- Jarman AP, Grau Y, Jan LY, Jan YN (1993) Atonal is a proneural gene that directs chordotonal organ formation in the *Drosophila* peripheral nervous system. *Cell* 73:1307–1321.
- Akazawa C, Ishibashi M, Shimizu C, Nakanishi S, Kageyama R (1995) A mammalian helix-loop-helix factor structurally related to the product of *Drosophila* proneural gene atonal is a positive transcriptional regulator expressed in the developing nervous system. *J Biol Chem* 270:8730–8738.
- Helms AW, Johnson JE (1998) Progenitors of dorsal commissural interneurons are defined by MATH1 expression. *Development* 125:919–928.
- Birmingham NA, et al. (2001) Proprioceptor pathway development is dependent on Math1. *Neuron* 30:411–422.
- Machold R, Fishell G (2005) Math1 is expressed in temporally discrete pools of cerebellar rhombic-lip neural progenitors. *Neuron* 48:17–24.
- Ben-Arie N, et al. (1997) Math1 is essential for genesis of cerebellar granule neurons. *Nature* 390:169–172.
- Helms AW, Abney AL, Ben-Arie N, Zoghbi HY, Johnson JE (2000) Autoregulation and multiple enhancers control Math1 expression in the developing nervous system. *Development* 127:1185–1196.
- Yang Q, Birmingham NA, Finegold MJ, Zoghbi HY (2001) Requirement of Math1 for secretory cell lineage commitment in the mouse intestine. *Science* 294:2155–2158.
- Ben-Arie N, et al. (2000) Functional conservation of atonal and Math1 in the CNS and PNS. *Development* 127:1039–1048.
- Marich SM, et al. (2009) Merkel cells are essential for light-touch responses. *Science* 324:1580–1582.
- Birmingham NA, et al. (1999) Math1: An essential gene for the generation of inner ear hair cells. *Science* 284:1837–1841.
- Kawauchi D, Saito T (2008) Transcriptional cascade from Math1 to Mbf1 and Mbf2 is required for cerebellar granule cell differentiation. *Dev Biol* 322:345–354.
- Gazit R, Krizhanovsky V, Ben-Arie N (2004) Math1 controls cerebellar granule cell differentiation by regulating multiple components of the Notch signaling pathway. *Development* 131:903–913.
- Flora A, Klisch TJ, Schuster G, Zoghbi HY (2009) Deletion of Atoh1 disrupts Sonic Hedgehog signaling in the developing cerebellum and prevents medulloblastoma. *Science* 326:1424–1427.
- Krizhanovsky V, Soreq L, Kliminski V, Ben-Arie N (2006) Math1 target genes are enriched with evolutionarily conserved clustered E-box binding sites. *J Mol Neurosci* 28:211–229.
- Ayrault O, et al. (2010) Atoh1 inhibits neuronal differentiation and collaborates with Gli1 to generate medulloblastoma-initiating cells. *Cancer Res* 70:5618–5627.
- Bossuyt W, et al. (2009) Atonal homolog 1 is a tumor suppressor gene. *PLoS Biol* 7:e39.
- Hatten ME, Alder J, Zimmerman K, Heintz N (1997) Genes involved in cerebellar cell specification and differentiation. *Curr Opin Neurobiol* 7:40–47.
- Vaillant C, Monard D (2009) SHH pathway and cerebellar development. *Cerebellum* 8:291–301.
- Pogoriler J, Millen K, Utset M, Du W (2006) Loss of cyclin D1 impairs cerebellar development and suppresses medulloblastoma formation. *Development* 133:3929–3937.
- Shroyer NF, et al. (2007) Intestine-specific ablation of mouse atonal homolog 1 (Math1) reveals a role in cellular homeostasis. *Gastroenterology* 132:2478–2488.
- Rose MF, Ahmad KA, Thaller C, Zoghbi HY (2009) Excitatory neurons of the proprioceptive, interoceptive, and arousal hindbrain networks share a developmental requirement for Math1. *Proc Natl Acad Sci USA* 106:22462–22467.
- Badea TC, Wang Y, Nathans J (2003) A noninvasive genetic/pharmacologic strategy for visualizing cell morphology and clonal relationships in the mouse. *J Neurosci* 23:2314–2322.
- Chédotal A (2010) Should I stay or should I go? Becoming a granule cell. *Trends Neurosci* 33:163–172.
- Robertson AG, et al. (2008) Genome-wide relationship between histone H3 lysine 4 mono- and tri-methylation and transcription factor binding. *Genome Res* 18:1906–1917.
- Massari ME, Murre C (2000) Helix-loop-helix proteins: Regulators of transcription in eucaryotic organisms. *Mol Cell Biol* 20:429–440.
- Nakada Y, Hunsaker TL, Henke RM, Johnson JE (2004) Distinct domains within Mash1 and Math1 are required for function in neuronal differentiation versus neuronal cell-type specification. *Development* 131:1319–1330.
- Blake JA, Bult CJ, Eppig JT, Kadin JA, Richardson JE; Mouse Genome Database Group (2009) The Mouse Genome Database genotypes:phenotypes. *Nucleic Acids Res* 37(Database issue):D712–D719.
- Huang W, et al. (2007) The DAVID Gene Functional Classification Tool: A novel biological module-centric algorithm to functionally analyze large gene lists. *Genome Biol* 8:R183.
- Behesti H, Marino S (2009) Cerebellar granule cells: Insights into proliferation, differentiation, and role in medulloblastoma pathogenesis. *Int J Biochem Cell Biol* 41:435–445.
- Bertrand N, Castro DS, Guillemot F (2002) Proneural genes and the specification of neural cell types. *Nat Rev Neurosci* 3:517–530.
- Massie CE, Mills IG (2008) ChIPping away at gene regulation. *EMBO Rep* 9:337–343.
- Powell LM, Deaton AM, Wear MA, Jarman AP (2008) Specificity of Atonal and Scute bHLH factors: Analysis of cognate E box binding sites and the influence of Senseless. *Genes Cells* 13:915–929.
- Barisone GA, Yun JS, Diaz E (2008) From cerebellar proliferation to tumorigenesis: New insights into the role of Mad3. *Cell Cycle* 7:423–427.
- Wey A, Cerdeno VM, Pleasure D, Knoepfler PS (2010) c- and N-myc regulate neural precursor cell fate, cell cycle, and metabolism to direct cerebellar development. *Cerebellum* 9:537–547.
- Dang CV, Le A, Gao P (2009) MYC-induced cancer cell energy metabolism and therapeutic opportunities. *Clin Cancer Res* 15:6479–6483.
- Jones RG, Thompson CB (2009) Tumor suppressors and cell metabolism: A recipe for cancer growth. *Genes Dev* 23:537–548.
- Renaud J, et al. (2008) Plexin-A2 and its ligand, Sema6A, control nucleus-centrosome coupling in migrating granule cells. *Nat Neurosci* 11:440–449.
- Friedel RH, et al. (2007) Plexin-B2 controls the development of cerebellar granule cells. *J Neurosci* 27:3921–3932.
- Vicente-Manzanares M, Choi CK, Horwitz AR (2009) Integrins in cell migration—The actin connection. *J Cell Sci* 122:199–206.
- Farwell AP, Dubord-Tomasetti SA, Pietrzykowski AZ, Stachek SJ, Leonard JL (2005) Regulation of cerebellar neuronal migration and neurite outgrowth by thyroxine and 3,3',5'-triiodothyronine. *Brain Res Dev Brain Res* 154:121–135.
- Ma X, Kawamoto S, Hara Y, Adelstein RS (2004) A point mutation in the motor domain of nonmuscle myosin II-B impairs migration of distinct groups of neurons. *Mol Biol Cell* 15:2568–2579.
- Zainuddin A, Chua KH, Abdul Rahim N, Makpol S (2010) Effect of experimental treatment on GAPDH mRNA expression as a housekeeping gene in human diploid fibroblasts. *BMC Mol Biol* 11:59.
- Røge R, et al. (2007) Commonly used reference genes are actively regulated in in vitro stimulated lymphocytes. *Scand J Immunol* 65:202–209.
- Sato A, et al. (2008) Cerebellar development transcriptome database (CDT-DB): Profiling of spatio-temporal gene expression during the postnatal development of mouse cerebellum. *Neural Netw* 21:1056–1069.
- Aerts S, et al. (2010) Robust target gene discovery through transcriptome perturbations and genome-wide enhancer predictions in *Drosophila* uncovers a regulatory basis for sensory specification. *PLoS Biol* 8:e1000435.
- Rowan S, et al. (2010) Precise temporal control of the eye regulatory gene Pax6 via enhancer-binding site affinity. *Genes Dev* 24:980–985.
- Seo S, Lim J-W, Yellajoshyl D, Chang L-W, Kroll KL (2007) Neurogenin and NeuroD direct transcriptional targets and their regulatory enhancers. *EMBO J* 26:5093–5108.