The regulation of gene expression in hair cells

Allen F. Ryan^{a,b,*}, Ryoukichi Ikeda^a, and Masatsugu Masuda^a

^aDepartments of Surgery/Otolaryngology, University of California, San Diego – School of Medicine, 9500 Gilman Drive, La Jolla, CA 92093, USA

^bDepartments of Neurosciences, University of California, San Diego – School of Medicine, 9500 Gilman Drive, La Jolla, CA 92093, USA

Abstract

No genes have been discovered for which expression is limited only to inner ear hair cells. This is hardly surprising, since the number of mammalian genes is estimated to be 20–25,000, and each gene typically performs many tasks in various locations. Many genes are expressed in inner ear sensory cells and not in other cells of the labyrinth. However, these genes are also expressed in other locations, often in other sensory or neuronal cell types. How gene transcription is directed specifically to hair cells is unclear.

Key transcription factors that act during development can specify cell phenotypes, and the hair cell is no exception. The transcription factor ATOH1 is well known for its ability to transform nonsensory cells of the developing inner ear into hair cells. And yet, ATOH1 also specifies different sensory cells at other locations, neuronal phenotypes in the brain, and epithelial cells in the gut. How it specifies hair cells in the inner ear, but alternate cell types in other locations, is not known.

Studies of regulatory DNA and transcription factors are revealing mechanisms that direct gene expression to hair cells, and that determine the hair cell identity. The purpose of this review is to summarize what is known about such gene regulation in this key auditory and vestibular cell type.

1. Genes and cell fate regulation

It is the differential expression of genes that give rise to the unique characteristics of each cell type in the body and determine their behavior. Because almost all cells share the same genes, differences in regulation determine the patterns of genes that give rise to different cellular types and control their biochemical and physiological characteristics. It is therefore not surprising that gene expression is highly regulated by multiple processes.

The core of a gene is its expressed sequence, transcribed into mRNA, edited to generate one or more splice variants, and translated into protein. However, genes typically span a much greater sequence range than their expressed domain. Within the total area of the gene are the regulatory elements that determine in what cells, at what times, and in what amounts a gene will be expressed (Ben-Tabou de-Leon and Davidson, 2007). These regulatory elements can

*Corresponding author. ENT-0666, UCSD School of Medicine, 9500 Gilman Drive, La Jolla, CA 92093, USA. Tel.: +1 858 534 4659; fax: +1 858 534 5319. afryan@ucsd.edu (A.F. Ryan)..

be very close to the expressed sequence, embedded within it, or located many thousands of bases distant. They include the gene promoter, located upstream from and adjacent to the expressed sequence, which contains the binding sites for the transcriptional complex that produces RNA from the DNA template. They also include DNA regions known as enhancers, activation of which may be required for gene expression to occur, or may increase expression. In the obverse, they include repressors that can block or reduce gene expression.

Promoters, enhancers and repressors are activated by DNA binding proteins that recognize specific motifs in the DNA sequence (Ben-Tabou de Leon and Davidson, 2007). These include RNA polymerase II (Pol-II) that forms the core of the transcription initiation complex, and transcription factors (TFs) that interact with Pol-II either directly or via intermediary proteins including transcription-associated proteins and the mediator complex (Riethoven, 2010; Poss et al., 2013). It is the combination of TFs present in a cell and binding to regulatory elements within a gene that helps to determine the genes expressed within that cell (Ma, 2006). This combinatorial TF regulation has been called a “second genetic code” (Hallikas et al., 2006).

In order to activate gene transcription, Pol-II and TFs must first access the regulatory DNA. Control of this access is another layer of gene regulation that helps to define cell phenotypes and behavior (Reik, 2007). Nuclear DNA is coiled around histones, and acetylation of these protein complexes enhances the uncoiling of DNA, providing access for transcriptional and regulatory factors. Methylation of histones can have more complex effects, either enhancing or suppressing DNA coiling depending upon methylation degree and the histone targets. Methylation of DNA in the region of the promoter can also block access for Pol-II and other key factors, rendering a gene silent. These epigenetic changes enable cells to alter the basic genetic code in response to intrinsic or environmental conditions (Jaenisch and Bird, 2003).

There are also many post-transcriptional mechanisms of gene regulation that occur after the generation of mRNA transcripts (Day and Tuite, 1998). These include mRNA splicing, changes in mRNA stability, and short regulatory RNAs that can inactivate mRNA, primarily by degradation, thereby blocking translation. In addition, many proteins are regulated post-translationally by enzymatic cleavage and/or chemical modifications such as phosphorylation.

Obviously, these proximate gene regulatory processes can be activated by many extrinsic and intrinsic factors. Thus, to gene regulation is a very broad phenomenon. The expression of many inner ear genes are known to be influenced by extra- and intra-cellular receptors, cellular activation or cell damage, to name a few potential factors. Such influences also participate in gene regulation. However, for the purposes of this review we concentrate upon proximate gene regulation, with an emphasis on TFs and their interactions with regulatory sequences within genomic DNA. Thus, we discuss genes for which analysis of gene regulatory regions important for HC gene expression have been identified and analyzed.

2. Regulation of genes in hair cells

No genes have been discovered with expression that is limited entirely to inner ear hair cells (HCs). Since the number of mammalian genes is limited to 20–25,000 (Pennisi, 2012), and since each gene typically performs multiple tasks in a variety of locations, this is perhaps not surprising. Genes expressed in the sensory cells, but not in other cells of the inner ear, are also expressed in other cell types at other locations, sometimes in other sensory or neuronal cells but also in completely different classes of cells. How the expression of such genes is directed specifically to HCs is not well defined. However, as noted above, TFs contribute significantly to the direction and activation of gene transcription (Weingarten-Gabbay and Segal, 2014), and this can be presumed to be true for HCs.

3. Regulation of *Atoh1*

Key TFs that act during development can play a particularly important role in specifying cell phenotypes. Many such TFs belong to the basic helix-loop-helix (bHLH) family of TFs, and in particular the Class II subfamily (Jones, 2004). Thus, the MYOD family is associated with the muscle cell phenotype (Olson et al., 1991), while NEUROD1 can specify a neural cell phenotype. The bHLH TFs bind to DNA motifs that contain an E-box motif (CANNTG), with some factors requiring quite specific bases within and around the E-box (e.g. Klisch et al., 2011). All bHLH factors dimerize before binding, either as homodimers or in combination with another bHLH factor. However, Class II bHLH factors cannot bind to an E-box as homodimers. They require heterodimerization with specific members of the Class I family of bHLH factors to bind to their recognition sites in a gene regulatory element (Jones, 2004).

The Class II bHLH TF ATOH1 (MATH1) was initially identified as essential for HC generation by Bermingham et al. (1999), since mice lacking the *atoh1* gene fail to develop cells with any of the physical or molecular characteristics of HCs. Soon thereafter, Zheng and Gao (2000) reported that overexpression of the *atoh1* gene induces the formation of extra HCs in the neonatal organ of Corti *in vitro*. *In vivo* overexpression was also found to induce ectopic HCs in the cochlea (e.g. Kawamoto et al., 2003). An independent mechanism that can specify HC identity prior to fate determination was recently identified by Ahmed et al. (2012), who found that co-expression of the TF SIX1 and the transcriptional coactivator EYA1 was sufficient to induce a HC phenotype in nonsensory cochlear cells. While SIX1/EYA1 induced ATOH1 expression in some of these cells, others did not express ATOH1, suggesting an alternative induction pathway.

While ATOH1 is well known for its ability to transform non-sensory cells of the developing inner ear into HCs, it also specifies sensory cells at other locations including the eye and olfactory system, neuronal phenotypes in brain, and epithelial cells in the gut. How it specifies HCs in the inner ear, but alternate cell types at other locations, is not known. However, the participation of additional TFs as partners with ATOH1 is one possibility. Differential combinatorial coding is a common mechanism by which TFs can exert differential effects (Weingarten-Gabbay and Segal, 2014).

In the inner ear, expression of ATOH1 is limited primarily to embryonic precursors of both supporting cells and HCs, with later limitation to developing HCs (Driver et al., 2013). However, expression does not continue into adulthood (Lumpkin et al., 2003). The regulatory regions of the *atoh1* gene that control this expression have been explored by Helms and colleagues. In an initial study (Helms and Johnson, 1998), they constructed a transgenic mouse using 15 kb 5' to the coding sequence of the *atoh1* gene plus 6 kb 3' to the coding sequence to drive beta galactosidase expression. Expression closely, but not entirely, reproduced the native expression of ATOH1, including expression in developing HCs. In contrast, a construct containing 2.5 kb of 5' DNA and 1.5 kb of 3' DNA failed to produce any reporter expression in mice.

In a second study, Helms, et al. (2000) used deletion analysis to narrow down the region of expression regulation. They found that while 5' DNA was insufficient to drive reporter gene expression, 6 kb of 3' DNA again substantially reproduced the normal pattern of ATOH1. Additional deletions revealed that this regulatory activity was contained within a 1.4 kb fragment located 3.4 kb 3' to the *atoh1* coding sequence. Homology analysis with human genomic DNA identified two regions of high homology, 561 bp (termed enhancer A) and 404 bp (termed enhancer B) which were separated and flanked by regions of low homology (Fig. 1A). There was no sequence similarity between these two enhancers.

Transgenic mice generated only with multiple copies of enhancer A exhibited expression in developing HCs. Mice with multiple copies of enhancer B also directed expression to HCs, although ectopic expression at other locations was noted, suggesting that repressor activity within enhancer A is important for limiting expression of the native gene. This analysis demonstrates that the two enhancers display overlapping, redundant function as far as HC expression is concerned, but may also cooperate in regulating *atoh1*.

Interestingly, Helms et al. found that when the transgenic line including 6 kb 3' to the *atoh1* coding sequence, containing both enhancer A and enhancer B, was crossed onto an *atoh1*-null background, no reporter gene expression was noted. This suggests that these enhancers may depend upon ATOH1 itself for expression.

Additional studies on the enhancers identified by Helms et al. (2000) were performed by Ahmed et al. (2012). These investigators found that co-transfection of the greater epithelial ridge (GER) with expression constructs encoding the homeodomain transcription factor SIX1 and its transcriptional co-activator EYA1 was sufficient to induce nonsensory cells to adopt a HC phenotype. This combination also induced expression of the *atoh1* gene, and this induction was enhanced by the SOX family TF SOX2. To determine whether this activation of *atoh1* was mediated by the enhancers identified by Helms et al. (2000), they performed co-transfection studies using reporters in which different *atoh1* enhancer constructs drove expression of green fluorescent protein (GFP) via a minimal promoter.

When a construct containing both enhancer A and B was used (Fig. 1B), SIX1/EYA1 transfection robustly activated the GFP reporter, and this effect was increased by about 10% when SOX2 was included. When enhancer A constructs were co-transfected with EYA1 and SIX1, there was minimal induction of GFP expression. However, SOX2 alone induced

moderate GFP and this was enhanced by the addition of SIX1 plus EYA1 (Fig. 1C). In contrast, while enhancer B constructs responded robustly to SIX1/EYA1 (Fig. 1D), this effect was unaffected by the addition of SOX2. Consistent with these results, enhancer A was found to contain a binding site for SOX2, and enhancer B a binding site for SIX1, that were highly conserved in the three mammalian and the chick *atoh1* genes. Electrophoretic mobility shift and chromatin immunoprecipitation (ChIP) assays confirmed binding of the respective TFs to these sites. Moreover, their mutation abolished the effects of the matching TF on reporter constructs.

These results indicate that while each enhancer can induce the expression of the *atoh1* gene, the two enhancers can interact to regulate expression differentially in response to distinct TF combinations via direct binding to the regulatory DNA.

4. Regulation of *pou4f3*

Regulation of *atoh1* is important for the induction of HCs and for expression during development. The *pou4f3* gene encodes a POU-domain TF required for the late differentiation of HCs (Erkman et al., 1996; Xiang et al., 1997), and is expressed in auditory and vestibular HCs just after the appearance of ATOH1 (Erkman et al., 1996; Pan et al., 2012), suggesting that it may be a direct target for ATOH1 regulation. Pan et al. (2012) constructed a transgenic mouse in which *atoh1* expression terminates soon after its initial onset. This led to dramatic decreases in *pou4f3* expression, consistent with regulation by ATOH1. Hu et al. (2010) co-transfected cochlear neural progenitor cells with an ATOH1 expression plasmid and a reporter construct consisting of 1.24 kb of 5' murine *pou4f3* genomic DNA driving green fluorescent protein (GFP). They found that ATOH1 transfection enhanced GFP expression from the reporter, further supporting regulation. Unlike the case for *atoh1*, *pou4f3* gene expression by HCs continues throughout life. The regulation of this gene therefore has implications for adult HC gene expression.

Regulation of the *pou4f3* gene in HCs was further evaluated by Masuda et al. (2011) using a transgenic mouse model. They found that 8.5 kb of genomic DNA, 5' to the initial codon of the *pou4f3* coding sequence, was sufficient to drive expression of either beta galactosidase or GFP in HCs of the embryonic and neonatal inner ear (Fig. 2). Transgene expression was highly similar to that of the native gene (Erkman et al., 1996). HC expression of GFP was initiated at embryonic day 12.5 (e12.5) in vestibular HCs and e14.5 in cochlear HCs. Vestibular and inner HC expression continued into adulthood. However, expression did not completely recapitulate the normal pattern of *pou4f3* inner ear expression. While inner ear expression of the *pou4f3* gene is limited to HCs, the 8.5 kb *pou4f3*/GFP transgene was also expressed in embryonic inner ear ganglion neurons. Moreover, expression in outer HCs faded in the weeks after birth, so that unlike the *pou4f3* gene (Erkman et al., 1996) it was not observed in adult outer HCs.

The 8.5 kb genomic fragment contained three smaller regions that were highly conserved across divergent mammalian species (mouse, human, cow and dog) (Fig. 3). These included a proximal conserved region ~400 bp immediately 5' to the *pou4f3* ATG, a first distal conserved region of ~60 bp starting 1.3 kb 5' to the ATG, and a second distal conserved

region of ~260 bp starting ~8.2 kb 5' to the ATG in the mouse (and at 4.7, 5.4 and 7.1 kb 5' in the human, dog and cow genes). Based on their conservation, these regions seemed likely to contain regulatory sequences that control expression of the *pou4f3* gene.

Interestingly, the 8.5 kb region did not contain any conserved binding sites for POU4F3 itself. This was unexpected, since TFs like POU4F3 that are expressed throughout life often bind to their own promoters to maintain expression by autoregulation (Deneris and Hobert, 2014). To determine whether POU4F3 was required to maintain expression, the GFP transgenic was bred onto a *pou4f3* null mouse background. This resulted in rapid loss of GFP expression in HCs after birth. To confirm autoregulation, the transgene construct was co-transfected with an expression plasmid encoding human POU4F3 (hPOU4F3) into HEK293 cells. hPOU4F3 transfection enhanced GFP expression. Taken together, these data indicate that the *pou4f3* gene does maintain expression via autoregulation. However, this regulation could be indirect, mediated by other TFs that do bind to the 8.5 kb regulatory fragment. Of course, there is also the possibility that POU4F3 binds to an unrecognized sequence in the transgene. However, given the extensive characterization of POU class TF sequence binding characteristics (e.g. Xiang et al., 1995), this seems unlikely.

Because the *pou4f3* gene is expressed immediately after ATOH1 in HCs and given the data of Hu et al. (2010) confirming upstream regulation, Masuda, et al. (2011) evaluated the conserved regions of the transgene for ATOH1 preferred binding sites as defined by Klisch et al. (2011). Three such sites were identified, two in the second distal conserved region, and one in the proximal conserved region (Fig. 3). It should be noted that ATOH1 could not bind to these sites alone. It would need to dimerize with another class of bHLH TF to achieve binding. To determine whether ATOH1 can in fact bind to these sites, ChIP was performed. HEK293 cells were co-transfected with the *pou4f3*-GFP transgene and a human FLAG-tagged ATOH1 expression plasmid (hATOH1-FLG). This resulted in GFP expression substantially greater than that seen with the transgene alone, indicating that ATOH1 had activated the transgene. An antibody against FLAG was then used to precipitate ATOH1 from the cells, followed by PCR using primers specific for the proximal, first distal or second distal conserved regions. The results were compared to DNA that had not been selected by immunoprecipitation. Only DNA from the second distal conserved region was enriched by ChIP, confirming binding to that regulatory region. These results establish that ATOH1 regulates the *pou4f3* gene by direct interaction with an enhancer located ~8.2 kb 5' to the murine start codon. The identity of its Class I bHLH binding partner was not determined.

TFs typically do not function in isolation, but rather in combinations that can include many factors on a given gene and regulatory area. The specific combination of TFs present on gene regulatory areas often influences the site and amount of gene transcription. This combinatorial effect helps to explain why a relatively small number of TFs (~2600 in the mammal; Babu et al., 2004) can control the expression of so many genes in so many cell types. To explore the possibility that other TFs might co-operate with ATOH1 on the second distal conserved region to control HC *pou4f3* expression, Masuda, et al. (2011) also evaluated the four mammalian sequences to identify other highly-conserved TF binding sites. Sites for GATA, the Class I bHLH TCF3 (also known as TFE2, E2a or E12/E47) and

SP1 TFs were present in all four species, in the same order, near or in the ATOH1 sites (Fig. 3). This suggests that ATOH1 may partner with one or more additional factors to influence HC expression of the *pou4f3* gene.

In a second study, Masuda, et al. (2012) assessed this hypothesis, as well as the possibility that TFs with binding sites near ATOH1 might cooperate more generally in regulating the HC phenotype. Explants of postnatal day 1.5 (p1.5) organ of Corti were harvested from *pou4f3*/GFP transgenic mice. Explants were transfected by electroporation with ATOH1, which preferentially entered cells in the GER. Prior studies have shown that this procedure results in the induction of HCs in cochlear explants (e.g. Jones et al., 2006). In the days following transfection, GFP-positive GER cells emerged, indicating that ATOH1 had activated the transgene. Moreover, most of these cells were also positive for myosin VIIA (MYO7A), indicating at least partial induction of the HC phenotype.

Masuda et al. (2012) then co-transfected p1.5 organ of Corti explants with ATOH1 plus GATA3, SP1 or TCF3 to determine whether any of these TFs would cooperatively enhance the activity of ATOH1. Both TCF3 and GATA3 significantly increased the number of GFP-positive GER cells, while SP1 co-transfection had no effect. In addition, TCF3 and GATA 3 enhanced the expression of MYO7A in most of the GFP-positive cells. These results indicate that ATOH1 combines with other TFs to regulate not only the *pou4f3* gene, but also the HC phenotype. They also suggest that TCF3 may heterodimerize with ATOH1 on its E-box site, although other E-boxes that can bind TCF3 as a homodimer were also conserved in the 5' DNA.

Co-operation of TFs is not limited to those that bind in close proximity. TFs bound to widely separated regulatory elements can interact to jointly control gene expression. This is thought to be accomplished in part by the action of the mediator complex, proteins that link TFs bound to near and distant regulatory elements to the transcription initiating machinery at the transcription start site or promoter (Poss et al., 2013). To determine whether TFs bound to more distant sites might interact with ATOH1 at the *pou4f3* promoter, Ikeda, et al. (2014) performed a more detailed analysis of TF binding site conservation in the DNA 5' to the *pou4f3* coding sequence. They identified binding sites for 22 TFs that were highly conserved within the proximal conserved region, the first distal conserved region and the second distal conserved region across the four mammalian species, and that are expressed in the embryonic sensory epithelium (Fig. 3).

They then performed co-transfection of each of the identified TFs plus ATOH1 into the p1.5 GER from *pou4f3*/GFP mice, and quantified cells expressing GFP and MYO7A. This revealed additional TFs that were able to influence the ability of ATOH1 to induce *pou4f3* gene expression and induce a HC phenotype: ETV4, NMYC (Fig. 4), and ETS2 enhanced the generation of GFP- and MYO7A-positive GER cells by ATOH1; HES1, HES5 and NEUROD1 reduced the effects of ATOH1. The authors also used CHIP to determine whether the ATOH1-enhancing TFs identified in their study or by Masuda et al. (2012) bound to the 8.5 kb 5' *pou4f3* DNA (Fig. 5). HEK 293 cells were co-transfected with a FLAG-tagged TF expression vector and the 8.5 kb *pou4f3*/GFP reporter plasmid. DNA from these cultures was unselected (input control) or immunoprecipiated with antibody for that

TF, and PCR performed for each conserved region. ChIP enrichment for the proximal conserved region, immediately upstream from the *pou4f3* AT, was observed for ETS2, NMYC and ETV4. Like ATOH1, TCF3, GATA3 and NMYC bound to the second distal conserved region located 8.2 kb 5' to the *pou4f3* ATG. These data implicate the proximal and second distal conserved regions as containing enhancers that influence the expression of the *pou4f3* gene in HCs. They also indicate that combinatorial TF coding plays a role in the regulation of the *pou4f3* gene, and more generally in determination of the HC phenotype, by direct interaction on regulatory DNA.

5. Regulation of myosin VIIA

MYO7A is expressed in HCs beginning at approximately the same time as the *pou4f3* gene (e.g., Xiang, et al., 1997), also suggesting the possibility of direct regulation by ATOH1. Also like POU4F3, MYO7A is expressed throughout the life of the HC. It is a protein critical to the function of HCs, playing a major role in the transduction complex at the tips of the stereocilia. As with other HC proteins, it is expressed at other sites, including the photoreceptors of the eye, and mutations in the *myo7a* gene cause Usher syndrome. Boëda et al. (2001) analyzed regulation of *myo7a* by sequential deletion of 5' genomic DNA in transgenic mice. They started with a yeast artificial chromosome (YAC) construct in which GFP expression was controlled by ~200 kb of human *myo7a* DNA 5' to the transcription start site in exon 1 and the majority of intron 1. Transgenic mice generated using this construct expressed GFP in vestibular HCs, inner HCs, and apical turn outer HCs, and ectopic expression in a few spiral ligament fibrocytes. Expression in other organs in which the *myo7a* gene is expressed was not observed.

They then generated a series of mice with decreasing amounts of 5' DNA. A construct containing 2100 bp 5' to the transcription initiation site as well as all of intron 1 showed expression virtually identical to that observed in the YAC transgenic. Deletion of a few hundred bases from the 3' end of intron 1 produced similar expression, except for reduced expression in HCs of the striolar region of the maculae and no expression in outer HCs. Shorter constructs resulted in very reduced HC expression. However, when an insulator sequence was added to the 5' end of a construct containing only 118 bp 5' to the transcription initiation site plus most of intron 1, expression in extrastriolar vestibular HCs and inner HCs was observed, indicating that inhibitory sequences upstream of the 118 bp sequence were likely responsible for reduced expression in the intermediate 5' constructs. Thus, the most proximal 118 bp of 5' DNA, which was the most highly conserved *myo7a* sequence between mouse and human, was sufficient to drive expression in the majority of inner ear HCs. However, when the first intron was entirely deleted, HC expression was completely abolished. Their study is consistent with the presence of a HC enhancer in the first intron of the mouse *myo7A* gene, which directs expression through the proximal promoter immediately 5' to the transcription start site.

Of note, this is the only gene regulatory sequence yet identified that directs expression only to HCs and not to other organs that express the *myo7a* gene. The enhancers that direct expression to these other sites are presumably present at other locations in or near (in genomic terms) the gene.

6. Comparison of HC gene enhancers

It is possible that the regulation of different genes with inner ear expression limited to HCs occurs via common mechanisms, based either on similar regulatory sequences or upon the binding of common TFs. To assess the relationship between enhancers from different genes that direct expression to HCs, we compared the sequences of HC enhancers from the *pou4f3*, *myo7A* and *atoh1* genes in the mouse. Comparisons of the various identified enhancers were performed using MacVector (MacVector, Inc., Cary, NC, USA) and rVISTA (<http://rvista.decode.org>) software to detect homologous sequence or orders of TF binding sites. No sequence or TF binding site order was detected between the various enhancers (data not shown). This suggests that there is no unique sequence signature or TF binding order that identifies HC-targeting regulatory DNA.

We then evaluated the *myo7A* and *atoh1* murine HC enhancers for TF-binding motifs, irrespective of order. This was performed using TRANSFAC Match 1.0 software (Qaigen, Valencia, CA USA) as well as MacVector search for binding sites not present in the database. Interestingly, we found multiple, closely-spaced ATOH1 binding motifs (Klisch et al., 2011) in the *myo7A* enhancer. These would also require a co-TF like E2F1 for high-affinity binding. Sites for GATA3 (Chen et al., 2012) and for NMYC (Murphy et al., 2009) were also present. Similarly, in the 1.4 kb, *atoh1* regulatory region, we found multiple ATOH1, NMYC and GATA3 binding sequences clustered in enhancer A. As noted above, clustered binding sites for these same TFs were found within the second distal conserved region of the *pou4f3* gene. Interestingly, none of these sites were present in *atoh1* enhancer B. Since this enhancer was also able to regulate ATOH1 expression, this may occur independent of or downstream from direct DNA binding by these TFs. The data suggest that, rather than a unique sequence signature, a combination of specific TF-binding motifs may be common to some enhancers that direct gene expression to HCs. They may also play a role in determining the HC phenotype via regulation of and with ATOH1.

Ahmed et al. (2012) speculated that SIX1 and EYA1 might regulate the *pou4f3* gene. We therefore assessed the conserved regions of the *pou4f3* gene for the SIX1 and SOX2 binding sites noted by them in the *atoh1* enhancers A and B. These were not detected. Although an alternative SOX2 binding site had been identified in distal conserved region 2 of the *pou4f3* gene (Ikeda et al., 2014, see Fig. 3), no alternative SIX1 binding sites (Liu et al., 2012) were found to be conserved. Thus, SIX1 and EYA1 do not appear to directly regulate the *pou4f3* gene within the 8.5 kb 5' region.

7. Conclusions

A small number of regulatory sequences that direct gene expression to HCs have been identified. Analysis of the sequence of these regulatory elements reveals no obvious homology. However, there is similarity in the TFs that can bind to some of these enhancers. This is similar to the situation in muscle, where muscle gene enhancers share little or no sequence homology and yet bind similar sets of TFs (Ordahl, 1992).

None of the enhancers identified to date produce HC expression that is completely faithful to that of the native gene. Thus, one or more HC type may not express under the control of the enhancer, temporal expression is not exactly reproduced, or ectopic expression is observed. This likely reflects the presence of additional enhancers at other locations in the genes.

Enhancers identified for the *atoh1* and *pou4f3* genes were not found to direct gene expression exclusively to HCs, since each also directs expression to other cell types. This is consistent with the observation, mentioned above, that no genes are expressed exclusively in this cell type. The unique effects of these genes in the HC may be explained by factors other than enhancer sequence. This could include unique TF combinations on the enhancers (Ma, 2006), differential epigenetic modifications of genes (Jaenisch and Bird, 2003), post-transcriptional regulatory processes including splicing and microRNAs (Lewis and Steel, 2010; Friedman et al., 2009), or post-translational processes such as differential cell signaling (Lanford et al., 1999). Alternatively, silencing sequences (Imagawa, 1996) at other locations in the gene may limit expression in other organs. In contrast, the enhancer in intron 1 of the *myo7a* gene does direct expression exclusively to HCs. Thus, there are also unique regulatory sequences for this cell type.

The set of genes expressed preferentially in HCs for which promoter and enhancer analysis has been performed is small. Evaluation of other such genes may identify additional common TF combinations that regulate genes in this cell type. Continued analysis of the three genes described here and of additional genes could also yield regulatory sequences and TFs important for gene expression in different HC types. This information would have not only scientific value, but could also be useful for designing cell-specific gene therapies or derivation of HC subtypes from stem cells.

Acknowledgments

Supported by grants from the Veterans Administration Research Service (BX001205), the NIH/NIDCD (DC000139) and the National Organization for Hearing Research.

References

- Ahmed M, Wong EY, Sun J, Xu J, Wang F, Xu PX. Eya1-Six1 interaction is sufficient to induce hair cell fate in the cochlea by activating Atoh1 expression in cooperation with Sox2. *Dev. Cell.* 2012; 22:377–390. [PubMed: 22340499]
- Babu MM, Luscombe NM, Aravind L, Gerstein M, Teichmann SA. Structure and evolution of transcriptional regulatory networks. *Curr. Opin. Struct. Biol.* 2004; 14:283–291. [PubMed: 15193307]
- Ben-Tabou de-Leon S, Davidson EH. Gene regulation: gene control network in development. *Annu Rev. Biophys. Biomol. Struct.* 2007; 36:191. [PubMed: 17291181]
- Bermingham NA, Hassan BA, Price SD, Vollrath MA, Ben-Arie N, Eatock RA, Bellen HJ, Lysakowski A, Zoghbi HY. Math1: an essential gene for the generation of inner ear hair cells. *Science.* 1999; 284:1837–1841. [PubMed: 10364557]
- Boëda B, Weil D, Petit C. A specific promoter of the sensory cells of the inner ear defined by transgenesis. *Hum. Mol. Genet.* 2001; 10:1581–1589. [PubMed: 11468276]
- Chen Y, Bates DL, Dey R, Chen PH, Machado AC, Laird-Offringa IA, Rohs R, Chen L. DNA binding by GATA transcription factor suggests mechanisms of DNA looping and long-range gene regulation. *Cell. Rep.* 2012; 29:1197–1206. [PubMed: 23142663]

- Day DA, Tuite MF. Post-transcriptional gene regulatory mechanisms in eukaryotes: an overview. *J. Endocrinol.* 1998; 157:361–371. [PubMed: 9691970]
- Deneris ES, Hobert O. Maintenance of postmitotic neuronal cell identity. *Nat. Neurosci.* 2014; 17:899–907. [PubMed: 24929660]
- Driver EC, Sillers L, Coate TM, Rose MF, Kelley MW. The Atoh1-lineage gives rise to hair cells and supporting cells within the mammalian cochlea. *Dev. Biol.* 2013; 376:86–98. [PubMed: 23318633]
- Erkman L, McEvilly RJ, Luo L, Ryan AK, Hooshmand F, O'Connell SM, Keithley EM, Rapaport DH, Ryan AF, Rosenfeld MG. Role of transcription factors Brn-3.1 and Brn-3.2 in auditory and visual system development. *Nature.* 1996; 381:603–606. [PubMed: 8637595]
- Friedman LM, Dror AA, Mor E, Tenne T, Toren G, Satoh T, Biesemeier DJ, Shomron N, Fekete DM, Hornstein E, Avraham KB. MicroRNAs are essential for development and function of inner ear hair cells in vertebrates. *Proc. Natl. Acad. Sci. U. S. A.* 2009; 106:7915–7920. [PubMed: 19416898]
- Hallikas O, Palin K, Sinjushina N, Rautiainen R, Partanen J, Ukkonen E, Taipale J. Genome-wide prediction of mammalian enhancers based on analysis of transcription-factor binding affinity. *Cell.* 2006; 124:47–59. [PubMed: 16413481]
- Helms AW, Abney AL, Ben-Arie N, Zoghbi HY, Johnson JE. Autoregulation and multiple enhancers control Math1 expression in the developing nervous system. *Development.* 2000; 127:1185–1196. [PubMed: 10683172]
- Helms AW, Johnson JE. Progenitors of dorsal commissural interneurons are defined by MATH1 expression. *Development.* 1998; 125:919–928. [PubMed: 9449674]
- Hu X, Huang J, Feng L, Fukudome S, Hamajima Y, Lin J. Sonic hedgehog (SHH) promotes the differentiation of mouse cochlear neural progenitors via the Math1-Brn3.1 signaling pathway in vitro. *J. Neurosci. Res.* 2010; 88:927–935. [PubMed: 19908278]
- Ikeda R, Pak K, Ryan AF. Transcription factors with conserved binding sites near ATOH1 on the *pou4f3* gene enhance the induction of cochlear hair cells. *Molec Neurobiol.* Jul 12.2014 2014. [Epub ahead of print] PMID: 25015561.
- Imagawa M. Negative regulation of gene expression in eukaryotes. *Neurochem. Int.* 1996; 29:565–572. [PubMed: 9113124]
- Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat. Genet.* 2003; 33(Suppl.):245–254. [PubMed: 12610534]
- Jones JM, Montcouquiol M, Dabdoub A, Woods C, Kelley MW. Inhibitors of differentiation and DNA binding (Ids) regulate Math1 and hair cell formation during the development of the organ of Corti. *J. Neurosci.* 2006; 26:550–558. [PubMed: 16407553]
- Jones S. An overview of the basic helix-loop-helix proteins. *Genome Biol.* 2004; 5:226. [PubMed: 15186484]
- Kawamoto K, Ishimoto S, Minoda R, Brough DE, Raphael Y. Math1 gene transfer generates new cochlear hair cells in mature guinea pigs in vivo. *J. Neurosci.* 2003; 23:4395–4400. [PubMed: 12805278]
- Klisch TJ, Xi Y, Flora A, Wang L, Li W, Zoghbi HY. In vivo Atoh1 targetome reveals how a proneural transcription factor regulates cerebellar development. *Proc. Natl. Acad. Sci. U. S. A.* 2011; 108:3288–3293. [PubMed: 21300888]
- Lanford PJ, Lan Y, Jiang R, Lindsell C, Weinmaster G, Gridley T, Kelley MW. Notch signalling pathway mediates hair cell development in mammalian cochlea. *Nat. Genet.* 1999; 21:289–292. [PubMed: 10080181]
- Lewis MA, Steel KP. MicroRNAs in mouse development and disease. *Semin. Cell. Dev. Biol.* 2010; 21:774–780. [PubMed: 20152923]
- Liu Y, Nandi S, Martel A, Antoun A, Ioshikhes I, Blais A. Discovery, optimization and validation of an optimal DNA-binding sequence for the Six1 homeodomain transcription factor. *Nucleic Acids Res.* 2012; 40:8227–8239. [PubMed: 22730291]
- Lumpkin EA, Collisson T, Parab P, Omer-Abdalla A, Haeberle H, Chen P, Doetzlhofer A, White P, Groves A, Segil N, Johnson JE. Math1-driven GFP expression in the developing nervous system of transgenic mice. *Gene Expr. Patterns.* 2003; 3:389–395. [PubMed: 12915300]

- Ma Q. Transcriptional regulation of neuronal phenotype in mammals. *J. Physiol.* 2006; 575:379–387. [PubMed: 16825304]
- Masuda M, Dulon D, Pak K, Mullen LM, Li Y, Erkman L, Ryan AF. Regulation of *POU4F3* gene expression in hair cells by 5' DNA in mice. *Neuroscience.* 2011; 197:48–64. [PubMed: 21958861]
- Masuda M, Pak K, Chavez E, Ryan AF. TFE2 and GATA3 enhance induction of *POU4F3* and myosin VIIa positive cells in nonsensory cochlear epithelium by ATOH1. *Dev. Biol.* 2012; 372:68–80. [PubMed: 22985730]
- Murphy DM, Buckley PG, Bryan K, Das S, Alcock L, Foley NH, Prenter S, Bray I, Watters KM, Higgins D, Stallings RL. Global MYCN transcription factor binding analysis in neuroblastoma reveals association with distinct E-box motifs and regions of DNA hypermethylation. *PLoS One.* 2009; 4:e8154. [PubMed: 19997598]
- Olson EN, Brennan TJ, Chakraborty T, Cheng TC, Cserjesi P, Edmondson D, James G, Li L. Molecular control of myogenesis: antagonism between growth and differentiation. *Mol. Cell. Biochem.* 1991; 104:7–13. [PubMed: 1922004]
- Ordahl CP. Developmental regulation of sarcomeric gene expression. *Curr. Top. Dev. Biol.* 1992; 26:145–168. [PubMed: 1563276]
- Pan N, Jahan I, Kersigo J, Duncan JS, Kopecky B, Fritsch B. A novel Atoh1 “self-terminating” mouse model reveals the necessity of proper Atoh1 level and duration for hair cell differentiation and viability. *PLoS One.* 2012; 7:e30358. [PubMed: 22279587]
- Pennisi E. ENCODE project writes eulogy for junk DNA. *Science.* 2012; 337:1159–1160. [PubMed: 22955811]
- Poss ZC, Ebmeier CC, Taatjes DJ. The mediator complex and transcription regulation. *Crit. Rev. Biochem. Mol. Biol.* 2013; 48:575–608. [PubMed: 24088064]
- Reik W. Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature.* 2007; 447(7143):425–432. [PubMed: 17522676]
- Riethoven JJ. Regulatory regions in DNA: promoters, enhancers, silencers, and insulators. *Methods Mol. Biol.* 2010; 674:33–42. [PubMed: 20827584]
- Weingarten-Gabbay S, Segal E. The grammar of transcriptional regulation. *Hum. Genet.* 2014; 133:701–711. [PubMed: 24390306]
- Xiang M, Gan L, Li D, Chen ZY, Zhou L, O'Malley BW Jr, Klein W, Nathans J. Essential role of POU-domain factor Brn-3c in auditory and vestibular hair cell development. *Proc. Natl. Acad. Sci. U. S. A.* 1997; 94(17):9445–9450. [PubMed: 9256502]
- Xiang M, Zhou L, Macke JP, Yoshioka T, Hendry SH, Eddy RL, Shows TB, Nathans J. The Brn-3 family of POU-domain factors: primary structure, binding specificity, and expression in subsets of retinal ganglion cells and somatosensory neurons. *J. Neurosci.* 1995; 15:4762–4785. [PubMed: 7623109]
- Zheng JL, Gao WQ. Overexpression of Math1 induces robust production of extra hair cells in postnatal rat inner ears. *Nat. Neurosci.* 2000; 3:580–586. [PubMed: 10816314]

Fig. 1. Combinatorial regulation of GFP expression in the GER mediated by *atoh1* enhancers (A), after co-transfection of SIX1, EYA1 and/or SOX2. Each embryonic GER was co-transfected with two reporter constructs; one in which a constitutive CMV promoter drove expression of red fluorescent protein (RFP) to identify all transfected cells, and another in which an *atoh1* enhancer region was upstream from GFP. TF expression plasmids were also co-transfected at the same time. The proportion of yellow cells in the GFP/RFP images (given in %) represents the degree to which TF co-transfection induced GFP expression for each reporter. The 1.4 kb enhancer (B) was highly responsive to SIX1/EYA1, and this effect was enhanced by SOX2. Enhancer A (C) was moderately responsive to SOX2, and this effect was modestly enhanced by EYA1. Enhancer B was highly responsive to SIX1/EYA1 (D), but this effect was not influenced by SOX2 (not shown). From Ahmed et al. (2012), with permission.

Fig. 2. Postnatal expression of GFP in the inner ear of a transgenic mouse in which 8.5 kb of DNA 5' to the start codon of the *pou4f3* gene was used to drive expression of the reporter. **a.** All inner ear sensory epithelia express GFP. Within the cochlea (**b–d**) and vestibular labyrinth (**e & f**), only HCs express GFP. **g.** Phalloiden labeling (red) of GFP-positive HCs. Adapted from Masuda et al., 2012, with permission.

Fig. 3.

Bioinformatic analysis of the 5' region of the *pou4f3* gene identified three homology regions (red) that are highly-conserved across four widely-separated mammalian species. The mouse sequence (8.5 kb) is illustrated. Highly conserved are a proximal region (Prox) immediately 5' to the coding sequence and two distal regions (Dist I, Dist II) located up to 8.4 kb 5'. Within these regions, binding sites for 22 TFs that are expressed in the sensory epithelium of the embryonic inner ear at around the time of ATOH1 expression were also conserved between mouse, human, cow and dog. Question marks indicate short, highly-conserved sequences that do not correspond to known TF binding sites. Adapted from Ikeda et al. (2014), with permission.

Fig. 4.

GFP and MYO7A expression in the GER induced by hATOH1 electroporation is enhanced by co-transfection with hNMYC. Transfection with hATOH1 alone induced ectopic GFP expression primarily in the GER. The array of native HCs is identified by arrows, and sequential imaging was used to identify any native HCs that might have migrated out of their normal positions during culture. Co-transfection of hATOH1 plus hNMYC enhanced induction of GFP in the GER by 68%. Moreover, the majority of GFP-positive cells were also immunopositive for MYO7A (arrows indicate a few GFP-positive cells that were negative for MYO7A), suggesting a broader adoption of a HC phenotype. Enhancement by NMYC is illustrated, but similar increases were also observed for GATA3, TCF3, ETC4 and ETS2. Adapted from Ikeda et al. (2014), with permission.

Fig. 5.

a. Chromatin immunoprecipitation (ChIP) of TF DNA binding to conserved *pou4f3* regions by TFs that enhance the ability of ATOH1 to induce GFP and MYO7A expression in cells of the early postnatal GER. TFs bound to the proximal (Prox) and/or second distal (Dist II) conserved regions of the 8.5 kb *pou4f3* 5' sequence. No TFs bound to the first distal conserved region (Dist I), providing an internal PCR control. **b.** Semi-quantification of the PCR suggests especially strong relative binding of TFs to the second distal conserved region. Adapted from Ikeda et al. (2014), with permission.