Discovery of transcriptional regulators and signaling pathways in the developing pituitary gland by bioinformatic and genomic approaches

Michelle L. Brinkmeier, Shannon W. Davis, Piero Carninci, James W. MacDonald, Jun Kawai, Debashis Ghosh, Yoshihide Hayashizaki, Robert H. Lyons, Sally A. Camper

A R T I C L E   I N F O

Article history:
Received 13 December 2007
Accepted 24 November 2008
Available online 11 February 2009

Keywords:
Cap trapper
Homeobox gene
Prop1
Gene expression
Ames dwarf

A B S T R A C T

We report a catalog of the mouse embryonic pituitary gland transcriptome consisting of five cDNA libraries including wild type tissue from E12.5 and E14.5, Prop1<sup>df/df</sup> mutant at E14.5, and two cDNA subtractions: E14.5 WT–E14.5 Prop1<sup>df/df</sup> and E14.5 WT–E12.5 WT. DNA sequence information is assembled into a searchable database with gene ontology terms representing 12,009 expressed genes. We validated coverage of the libraries by detecting most known homeobox gene transcription factor cDNAs. A total of 45 homeobox genes were detected as part of the pituitary transcriptome, representing most expected ones, which validated library coverage, and many novel ones, underscoring the utility of this resource as a discovery tool. We took a similar approach for signaling-pathway members with novel pituitary expression and found 157 genes related to the BMP, FGF, WNT, SHH and NOTCH pathways. These genes are exciting candidates for regulators of pituitary development and function.

Introduction

The pituitary gland is a pivotal component of the endocrine system due to its role in controlling a wide range of fundamental bodily activities including growth, puberty development, thyroid gland function, and the capacity to cope with stress. These critical functions are regulated by six different hormones secreted by five specialized cell types within the anterior pituitary gland, in response to local stimulatory and inhibitory peptides released by the hypothalamus.

Pituitary development in the mouse begins at embryonic day (E) 9.5 and continues after birth. All of the hormone-secreting cells in the anterior lobe derive from a single primordial tissue, Rathke’s pouch, which forms from the oral ectoderm adjacent to the base of the developing brain. Formation of Rathke’s pouch occurs in contact with a specialized region of the neuroectoderm, the infundibulum, which will later form the posterior lobe of the pituitary gland.

The molecular mechanisms underlying pituitary organ growth and cell differentiation are complex. Signaling molecules including FGF, BMP, NOTCH, WNT and SHH family members have roles in establishing regional identity within Rathke’s pouch and contribute to its growth [1–5]. The overall effect of these signaling molecules is to induce the regional expression of numerous critical transcription factors, which control the differentiation of the hormone secreting cells between E13.5 and birth.

Failed or incomplete differentiation of anterior pituitary cells during embryonic development can result in the genetically heterogeneous disorder, Multiple Pituitary Hormone Deficiency (MPHD). Molecular genetic analysis of spontaneous and genetically engineered mouse mutants with pituitary defects has revealed a functional role for many homeobox transcription factors in the etiology of MPHD in mouse and man. These include Prop1, Pou1f1, Hex1, Lhx3, Lhx4, Otx2, Pitx1, and Pitx2 [reviewed by [6–8]]. Lesions in these genes only account for approximately half of the known genetic MPHD, suggesting that many of the key factors involved in the process of pituitary cell proliferation and differentiation remain unknown.

Mutations in <i>PROP1</i> are the most common known genetic cause of MPHD in humans [9]. Patients with <i>PROP1</i> mutations commonly exhibit deficiencies in Growth Hormone (GH), Prolactin (PRL), Thyroid Stimulating Hormone (TSH) and gonadotropins, with evolving adrenocorticotropic deficiency (ACTH). Similarly, the spontaneous Ames dwarf mouse model (<i>Prop1</i><sup>df/df</sup>), that harbors a homeodomain mutation in Prop1 displays a phenotype that includes hypoplasia of the anterior pituitary accompanied by nearly complete loss of the hormone-secreting cell types that produce GH (somatotropes), PRL (lactotropes), and TSH (thyrotropes) [10–13].
Interestingly, genetic background has a more significant impact on the viability of Prop1 deficient mice than any functional difference between the Prop1<sup>fl/fl</sup> and Prop1<sup>null</sup> alleles, suggesting that genetic modifiers may influence the clinical features of some MPHDI patients [13]. Prop1 expression is first detectable in Rathke’s pouch at E10–10.5. It reaches a peak at E12.5 after which time expression declines rapidly [14]. A time of intense cell proliferation occurs in the developing pituitary from E12.5–E14.5, which precedes the appearance of the differentiated pituitary cell types. The effects of mutations in Prop1 and other critical transcription factors including Pitx2, Lhx3 and Hesx1 are evident at this time [15–17]. This suggests that potential Prop1 targets and other important as yet unknown transcription factors may be expressed around E12.5 and/or E14.5.

At E12.5 the anterior pituitary is composed primarily of proliferating, undifferentiated cells. There is a very small group of non-proliferating cells that have begun to differentiate and express Foxa2 and the alpha subunit of the gonadotropins and thyrotropin, Cga [18]. At this developmental time Prop1<sup>fl/fl</sup> and WT pituitaries are indistinguishable in size and morphology [13,19]. By E14.5 there are more cells in the anterior lobe of normal mice, and they appear loosely packed and glandular. In contrast, the pituitary glands of Prop1<sup>fl/fl</sup> mutants are highly dysmorphic at E14.5 because cells apparently fail to migrate away from the dorsal aspect of the gland and colonize the anterior lobe. They exit the cell cycle but fail to differentiate [13,19,20].

We exploited the Prop1<sup>fl/fl</sup> model to gain insights into the embryonic pituitary transcriptome. We used cDNA subtractive hybridization between Prop1<sup>fl/fl</sup> and wild type pituitary gland primordia at E14.5 to identify differentially expressed genes in mutant and normal embryonic pituitary [21]. We took a differential display approach to define an expression profile for the developing anterior pituitary at E12.5 and E14.5 by comparing Rathke’s pouch tissue from wild type samples taken at these times [22]. Collectively, these studies represented the first expression profiles of the developing pituitary gland and also paved the way for demonstrating the role of the Wnt signaling pathway in regulating pituitary growth. More specifically, these studies identified Wnt5a, Aes and Tcf7l2 in the developing pituitary. We have now demonstrated functional roles for each of these genes in pituitary development through the analysis of corresponding mutant mouse models [23–25].

Numerous gene expression profiling studies have been carried out using adult pituitary tissue and cell lines, but only a few studies have addressed developmental changes in pituitary gene expression aside using adult pituitary tissue and cell lines, but only a few studies have addressed developmental changes in pituitary gene expression [19,23]. The same sequences were also subjected to BLAST analysis against the mouse-specific Reference Sequence (RefSeq) (Release #20, www.ncbi.nlm.nih.gov/RefSeq/). Clones were identified based on the first BLAST identity recorded with an expect value of 1x10e-5 or lower. This analysis identified 51,352 clones present in the encyclopedia. 5364 clones were not identified by BLAST analysis. While some of these clones have poor sequence quality through the polyA stretch at the 3’ end that interferes with BLAST analysis, we anticipate that some of these clones will be identified when compared to more recent UniGene and Refseq builds. Further analysis to determine the number of unique genes represented revealed 12,009 different cDNA clones distributed between the five libraries.

Table 1. Number and distribution of cDNA clones in the embryonic pituitary encyclopedia

<table>
<thead>
<tr>
<th></th>
<th>E12.5 WT</th>
<th>E14.5 WT</th>
<th>E14.5 Prop&lt;sup&gt;fl/fl&lt;/sup&gt;</th>
<th>Sub1&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Sub2&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total UniGene entries/library</td>
<td>4192</td>
<td>4339</td>
<td>4268</td>
<td>4710</td>
<td>5191</td>
</tr>
<tr>
<td>E12.5 WT&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1246&lt;sup&gt;d&lt;/sup&gt;</td>
<td>165</td>
<td>285</td>
<td>182</td>
<td>146</td>
</tr>
<tr>
<td>E14.5 WT</td>
<td>165</td>
<td>642</td>
<td>150</td>
<td>294</td>
<td>354</td>
</tr>
<tr>
<td>E14.5 Prop&lt;sup&gt;fl/fl&lt;/sup&gt;</td>
<td>285</td>
<td>150</td>
<td>1256</td>
<td>147</td>
<td>185</td>
</tr>
<tr>
<td>Sub1</td>
<td>182</td>
<td>294</td>
<td>147</td>
<td>629</td>
<td>326</td>
</tr>
<tr>
<td>Sub2</td>
<td>146</td>
<td>354</td>
<td>185</td>
<td>623</td>
<td>1284</td>
</tr>
</tbody>
</table>

<sup>a</sup> E14.5 WT-E14.5 Prop<sup>fl/fl</sup>  <sup>b</sup> E14.5 WT-E12.5 WT.  <sup>c</sup> Example interpretation: clones sequenced from the E12.5 WT library matched 4192 different UniGene entries, 1246 were unique to this library. 3029 were found in all libraries.  
<sup>d</sup> Library-specific transcripts are on the diagonal, shaded in gray.
To gain greater insight into the unique cDNA distribution between the three non-subtracted libraries, we compared the relative distribution of 12,894 cDNA clones between the E12.5 WT, E14.5 WT and E14.5 Prop1\(^{df/df}\) libraries (Fig. 2a). Overall, there is little overlap between transcripts detected in these three libraries, where any two of the three libraries considered shared at least 11–12% clone representation. Each of the three libraries contains a significant proportion of apparently library-specific genes (14%), suggesting that each library contributes an important collection of unique genes to the encyclopedia that may be dependent on developmental stage and genotype.

The two subtracted libraries each add a dimension of genetic novelty to the encyclopedia

We compared the relative distribution of 13,035 cDNA clones between the E14.5 WT, E14.5 Prop1\(^{df/df}\) and Sub1 libraries (Fig. 2b). The Sub1 library (E14.5 WT minus E14.5 Prop1\(^{df/df}\)) should theoretically represent those genes expressed at E14.5 in normal pituitaries but not in Prop1\(^{df/df}\) mutants. This trend is observed, suggesting that the subtraction was effective. Gene representation in the E14.5 WT and Sub1 libraries overlaps by 16% compared to a 13% overlap between E14.5 Prop1\(^{df/df}\) and Sub1. The percentage of genes expressed in both E14.5 WT and E14.5 Prop1\(^{df/df}\) is also 10%. This implies that the Sub1 and E14.5 Prop1\(^{df/df}\) libraries are as different as the E14.5 WT and E14.5 Prop1\(^{df/df}\) libraries in terms of the expressed genes being detected or sampled. In addition to those genes that were detected in both E14.5 WT and Sub1, there are also genes detected in only one library. Uniquely represented genes constitute 10–14% of the total expressed genes, which equates to 1612 genes in the case of the Sub1 library, demonstrating that it adds a dimension of genetic novelty to the encyclopedia.

We evaluated the E12.5 WT, E14.5 WT and Sub2 libraries, comparing 14,509 clones (Fig. 2c). The Sub2 and E14.5 WT libraries should share expression of a larger subset of genes compared to E14.5 WT and E12.5 WT, and this trend was observed. E14.5 WT and Sub2 common transcripts overlap by 17%, compared to an overlap of 10% between E14.5 WT and E12.5 WT. This overlap is equivalent to that observed between Sub2 and E12.5 WT (13%) suggesting that this subtraction was also effective. The Sub2 library adds new information to the encyclopedia as 2283 (16%) unique cDNA clones present in the Sub2 library were not sampled in the E14.5 WT or E12.5 WT libraries.

**Gene Ontology (GO) terminology reveals novel library-specific biological processes**

We wanted to determine the unique biological processes represented in each library as a means of identifying the genes that contribute to the morphological and functional differences in pituitary development [13,19]. Using Gene Ontology (GO) terms we probed the five sets of genes uniquely sampled from each library for enriched processes and also analyzed the set of genes that were detected universally. Twelve GO terms were significantly enriched (unadjusted \(p\)-value < 0.01) in the universally expressed transcript set (Table 2). Enriched terms include ‘translation’, ‘rRNA splicing’, ‘protein folding’ and ‘protein transport’, fundamental biological processes that are active in all tissues, irrespective of developmental stage or genotype.

The E14.5 WT library contains three significantly over-represented GO terms; ‘spermidine metabolic process’, ‘hydrogen peroxide catabolic process’, and ‘defense response to bacterium’ (Table 2). While these were unexpected, it is not difficult to envision how genes in these categories could be engaged in other known processes that occur predominantly at this time. For example, spermidine markedly increases DNA binding activity of certain transcription factors in the pituitary gland [44]. It may, therefore, play an important role in the changing transcription characteristic of the developing pituitary at
this time. The ‘defense response to bacterium’ GO term is also significantly enriched in the Prop1df/df library (Table 2) and is represented by more genes, compared to the E14.5 library (data not shown).

The E14.5 Prop1df/df library sampling was enriched for six GO terms (Table 2), five of which are not over-represented in any of the other library samples, providing clues regarding the types of processes occurring in the Prop1df/df pituitary that could differ from the normal state. For example, ‘Response to oxygen radical’ was enriched, a process that is involved in apoptosis, which is elevated in Prop1df/df [19,45]. Two significantly over-represented GO terms emerged from the E12.5 WT library sample (Table 2). ‘Positive regulation of calcium signaling’ precedes the catabolism of spermidine in the pituitary, which was represented at E14.5, and the role for calcium signaling is well documented [46,47]. The Sub1 and Sub2 library samples collectively contain four over-represented GO terms, which do not appear in any of the other samples (Table 2). ‘Exocytosis’ and ‘response to pH’ may provide a platform on which to launch future functional studies for some of the genes represented.

Validation of the embryonic pituitary encyclopedia

To validate the five libraries that comprise the embryonic pituitary encyclopedia we queried the database for the presence of several ‘knowns’, consisting of some homeobox genes with well-characterized roles in the developing pituitary gland and related gene family members. Searches for gene name, UniGene ID, and DNA or peptide sequences identified nine homeobox genes in the database (Table 3). Pou1f1, a target of Prop1 [14] was detected in the E14.5 WT library, and not in the E14.5 Prop1df/df library, as expected. Lhx3, Pitx1 and Pitx2, and two members of the Six gene family, were identified in several libraries [6,48–51]. Other expected genes detected included Otx2, Prop1, and Pax6 [6,7,52]. Three of the expected homeobox genes, Lhx4, Hesx1, and Pou3f4, were not detected in the database following in silico searching. Absence from the database can be an artifact of the sampling strategy for sequencing. For example, Hesx1 was detected by PCR at the expected developmental stages, suggesting that deeper sequencing would have revealed the presence of these genes (data not shown).

To confirm the identity and integrity of the clones identified computationally from end sequence, at least one clone representing each homeobox gene identified in Table 3 was sequenced from both ends and examined to determine whether the libraries were enriched for full-length clones. In most cases the clones were full-length compared to the NCBI Mus musculus UniGene database (Build #159) validating the Cap-trapper technique for producing full-length clones. This validation strategy revealed that the high throughput first pass

![Fig. 2. Venn diagrams illustrate the relationships between the five libraries in the encyclopedia. (A) Comparison of a total of 12,894 genes from the E12.5 WT, E14.5 WT and E14.5 Prop1df/df. (B) Comparison of a total of 13,035 genes from the E14.5 WT, E14.5 Prop1df/df and Sub1 libraries. (C) Comparison of a total of 14,509 genes from the E12.5 WT, E14.5 WT and Sub2 libraries.](image-url)
sequencing correctly identifies greater than 95% of the clones accurately. The misidentified clones tended to be ones with poor expect values from the initial BLAST analysis.

Identification of additional homeobox genes in the embryonic pituitary encyclopedia

Given the critical roles of numerous homeobox genes in the developing pituitary gland, we used the database as a tool for discovering additional homeobox genes expressed in the pituitary gland. The database is searchable using a combination of text searching with the terms ‘homeo’, ‘homebox’ and ‘homeodomain’ or querying the database using the 60 amino acid homeodomain consensus sequence [53]. The most comprehensive method was to search using gene names that are included in a list of homeodomain containing proteins at Prosite (http://ca.expasy.org/prosite/). We recovered numerous additional homeobox genes expressed in the encyclopedia (Table 4).

Differential expression analysis confirms transcription of novel homeobox genes

We analyzed 23 of these genes with respect to expression in the developing pituitary gland at E12.5, E14.5, and E18.5 using a combination of in situ hybridization analysis, RT-PCR, and quantitative PCR. Fig. 3 shows a subset of the in situ hybridization results. Dlx3 and Rax expression was confirmed at E14.5 (Fig. 3a and c) and Dlx1 and Zfh1 at E14.5 (Fig. 3b and d). For PCR analyses cDNA was generated from E12.5, E14.5, and E18.5 WT, and E14.5 Prop1<sup>df/df</sup> pituitary primordia. The quality of these cDNAs was confirmed using a series of primers specific for Prop1, Pou1f1, Hex1, and Hprt1 (Fig. 4, data not shown). Using standard RT-PCR expression of 12 genes was confirmed during pituitary gland development. Meis1, Adnp, Prrx2, Lhx2, Zfhx3, Pkmox2, Pbx2, and Meis2 were expressed at similar levels at all time points analyzed in wild type and in Prop1<sup>df/df</sup> mutant Rathke's pouch cDNA. Mecoc2, Em2oX, Otx2, and Zeb2 appeared to exhibit developmentally regulated expression, with the peak expression of Mecoc2 at E12.5, Em2oX at E14.5, and Zeb2 at E18.5 (Fig. 4a). In addition, expression of Mecoc2 and Otx2 appeared elevated in Prop1<sup>df/df</sup> mutant Rathke's pouch compared to wild type controls. Quantitative Real Time PCR was used to confirm the results suggested by standard RT-PCR (Fig. 4b), cDNA was generated from additional E12.5, E14.5, E18.5 WT, and E14.5 Prop1<sup>df/df</sup> pituitary primordia and the quality confirmed using primer sets for Prop1 and Pou1f1. All samples were done in triplicate and normalized to Hprt1. The fold activation was determined by comparison to E14.5 WT, which was set to 1. Both Mecoc2 and Otx2 peak in expression at E12.5 and are elevated 7.9 and 6 fold, respectively, relative to E14.5 WT. The peak expression of Em2oX is between E12.5 and E14.5, while Zeb2, Adnp, and Meis1 are decreased in the Prop1<sup>df/df</sup> mutant Rathke's pouch compared to wild type controls. Taken together these results suggest that Mecoc2, Otx2, Zeb2, Adnp, and Meis1 could merit further evaluation as downstream targets of Prop1.

The identification of individual homeobox genes in the specific libraries does not necessarily correlate with the actual expression

<p>| Table 2 |</p>
<table>
<thead>
<tr>
<th>Enriched gene ontology (GO) terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>GO term&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Universal expression set</td>
</tr>
<tr>
<td>GO:0006457</td>
</tr>
<tr>
<td>GO:0001567</td>
</tr>
<tr>
<td>GO:0008330</td>
</tr>
<tr>
<td>GO:0006397</td>
</tr>
<tr>
<td>GO:0009408</td>
</tr>
<tr>
<td>GO:0009586</td>
</tr>
<tr>
<td>GO:0015001</td>
</tr>
<tr>
<td>GO:0049007</td>
</tr>
<tr>
<td>GO:0066334</td>
</tr>
<tr>
<td>GO:0003039</td>
</tr>
<tr>
<td>GO:0030198</td>
</tr>
</tbody>
</table>

<sup>a</sup> www.geneontology.org.

<sup>b</sup> GO terms with an unadjusted p-value < 0.01 are listed.

<p>| Table 3 |</p>
<table>
<thead>
<tr>
<th>Libraries contain homeobox genes implicated in pituitary development</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene symbol</td>
</tr>
<tr>
<td>Lhx3</td>
</tr>
<tr>
<td>Otx2</td>
</tr>
<tr>
<td>Pox6</td>
</tr>
<tr>
<td>Ptx1</td>
</tr>
<tr>
<td>Ptx2</td>
</tr>
<tr>
<td>Pou1f1</td>
</tr>
<tr>
<td>Prop1</td>
</tr>
<tr>
<td>SIX1</td>
</tr>
</tbody>
</table>

<sup>a</sup> NCBI UniGene accession number.

<sup>b</sup> Number of cDNA clones identified in each library corresponding to indicated gene. Blank indicates none.

<sup>c</sup> E14.5 WT–E14.5 Prop1<sup>df/df</sup>.

<sup>d</sup> E14.5 WT–E12.5 WT.
profile of that gene. For instance, Meis1 was only identified in the E12.5 library; however, our RT-PCR data show that it is expressed at E12.5, E14.5, and E14.5 Prop1<sup>+/−</sup> embryos. Therefore, it is misleading to assume that the identification of a gene in a specific library or even a subtracted library provides accurate differential expression information.

**Database analysis implicates additional signaling molecules in pituitary development**

Fundamental to the current understanding of the genetic mechanisms driving pituitary development has been the identification of numerous signaling pathways that exert their effects on the pituitary primordia. Previous studies have demonstrated the importance of individual members of the BMP, FGF, NOTCH, SHH and WNT signaling pathways in directing normal pituitary development [1–3,5,23–25,54,55]. A continuing challenge in the field is to identify additional members of these pathways and to determine how these various signaling pathways interact with each other during pituitary organogenesis to promote development of the organ.

We used Genomatix software ([www.genomatix.de/](http://www.genomatix.de/)) to build functional genetic networks for signaling pathways known to affect pituitary development and queried the database for the expression of these genes. For example, in the case of the BMP pathway, a group of 78 Bmp-related genes emerged, and we identified 20 of them in the encyclopedia. This analysis was conducted for additional pathways, revealing a total of 61 genes expressed in the developing pituitary gland related to the BMP (N = 20), FGF (N = 13), NOTCH (N = 2), SHH (N = 1) or WNT (N = 25) signaling pathways. A subset corresponding to the FGF pathway is shown in Table 5, and the complete list is available in Supplementary Table 1. 17 additional genes were identified that are involved in functional crosstalk between two or more signaling pathways (Table 5 and Supplementary Table 1).

As an alternative strategy we searched the library for GO terms related to signaling and for specific gene names for genes known to be involved in the BMP, FGF, NOTCH, SHH, and WNT pathways. This search identified an additional 72 genes (Supplement Table 1). This list has not been verified by end sequencing yet.

**Spatial expression of three Bmp-related genes in the developing pituitary**

We searched the set of 61 signaling genes described above for those with previously demonstrated roles in the developing pituitary. Twelve “known” genes are in the set including Aes and Tle4 [23], Bmpr1a and Bmpr1b [56], Chrd [55], Fgf10 [1] Nkx2-1 [4], Hey1 and Notch2 [3], Wnt5a [2,25,55], Axin2 [2,57] and Fzd5 [21] (Supplementary Table 1). Previously we presented the expression pattern for the BMP signaling related genes Nkx1 and Fstl1, found in our developmental library, at E12.5 and E14.5, and for Id3 at E10.5 [54]. We expanded the analysis of Id3 and examined the expression of Id2 and Tgf<sub>b</sub> (Fig. 5). Id2 expression was detected in the dorsal aspect of Rathke's pouch at E12.5 (Fig. 5a). Expression persists at E14.5 in the
expanding anterior pituitary (Fig. 5b) but appears restricted to those cells closest to the lumen of the pouch. \(ld2\) does not appear to be expressed in the proliferating cells of the expanding Rathke’s pouch. By E16.5 \(ld2\) expression is barely detectable in the pituitary (data not shown). \(ld3\) expression was detected in Rathke’s pouch at E12.5, enriched in the dorsal part of the pouch and in the intermediate lobe (Fig. 5c). Similar to \(ld2\), \(ld3\) expression does not appear in the proliferating cells of the expanding Rathke’s pouch. Expression of \(ld3\) continues at E14.5 (Fig. 5d) in a similar pattern, restricted to the dorsal pouch and intermediate lobe and is barely detectable by E16.5 (data not shown). \(Tgfbi\) showed a striking expression pattern. While there are no transcripts in Rathke’s pouch itself, the gene is expressed in the mesenchymal tissue immediately surrounding the pouch at E12.5 (Fig. 5e). At E14.5, this expression continues more strongly and is also detectable surrounding the posterior pituitary (Fig. 5f). Like \(ld2\) and \(ld3\), \(Tgfbi\) expression wanes by E16.5. These studies confirm expression of these genes in the pituitary and support the encyclopedia as reflection of pituitary gene expression.

**Interplay of signaling pathways in the developing pituitary**

We surveyed the signaling genes represented in the encyclopedia using Genomatix software to identify those genes that displayed a functional overlap between two or more of the five pathways defined in Supplementary Table 1. This analysis revealed the presence of an additional 16 genes in our encyclopedia that are involved in functional cross talk during organ development, including pituitary (Supplementary Table 1). A subset of these (4/16) is shown in Table 5.

4/16 genes of the pathway interaction genes identified belong to the helix–loop–helix transcription factor family (Table 5). These include \(Ascl1\) and \(Id1\). The zebrafish \(Ascl1\) orthologue (\(Ascl1a\)) is a critical gene for pituitary cell differentiation and cell survival [58]. Zebrafish \(Fgf3\) mutants fail to express \(Ascl1a\), suggesting that \(Ascl1a\) acts downstream of Fgf. Zebrafish \(Fgf3\) is orthologous to mammalian \(Fgf10\), known for its function in mammalian pituitary development, implying a conserved mechanism for FGF signaling and \(Ascl1\) expression [1]. Additionally, \(Id1\) mediated expression of \(Bmp2\) leads to reduced \(Ascl1\) expression [59]. \(Ascl1\), \(Id1\), \(Id2\), \(Id3\), and \(ld4\) are all implicated as differentially expressed in pituitary tumors [28,60–62]. Together these studies suggest a functional link between FGF and BMP signaling and identify the \(ld\) gene family and \(Ascl1\) as interesting candidates in mediating cell differentiation in the developing pituitary. This example demonstrates the utility of the embryonic pituitary encyclopedia in providing a foundation for furthering our understanding of signaling in the developing pituitary.

**Discussion**

In previous studies we described a partial expression profile of the developing pituitary gland and paved the way for understanding the role of WNT signaling in pituitary development using a combination of differential display PCR [22] and subtractive hybridization [21]. The datasets generated in these studies were fairly small (~400 genes in each), but the gene discovery was significant, providing proof of principle that generation of a more comprehensive pituitary transcriptome profile would likely enable significant gene discovery and provide insight into the transcription factors and signaling molecules expressed in the developing pituitary.

In this study we generated a full-length cDNA encyclopedia of the developing pituitary at wild type E12.5 and E14.5 and in the \(Prop1\) mutant at E14.5 using an approach designed to enrich for full-length clones, maximize gene discovery, and permit high throughput gene recovery [38,40–43]. Using sequence data from these libraries we built a comprehensive database through which to access transcript information. It is searchable by Gene Ontology (GO) terminology; direct sequence query, including nucleic and amino acid sequence; gene name or text-string options; motifs for transcription factor binding; and UniGene ID and database-specific clone number. This encyclopedia contains 12,009 unique transcripts, represented by 56,716 cDNA clones, and accounts for over half of the known genes in the mouse genome (22,723) [http://www.ensembl.org/Mus_musculus/index.html]. Thus, the developing pituitary transcriptome has depth and complexity, and our study provides clones ideal for use in functional studies such as cell culture transfection and the generation of transgenic mice.

Each of the libraries contained within the encyclopedia includes a set of library-specific cDNA clones representing between 629 (Sub1) and 1264 (Sub2) unique transcripts. Exclusivity in any one library is not a true readout of the pituitary transcriptome, but rather, a reflection of the sequencing approach taken. Novelty and enrichment for rare transcripts were favored in place of the complete sequencing of any one library. Despite this limitation, the GO terminology analysis suggests that certain biological processes are significantly enriched in a library-specific manner. The genes represented by these processes could provide the basis for future novel mechanistic studies of embryonic pituitary development.

We identified a total of 45 homeobox genes that are expressed in the developing pituitary gland, expanding significantly on the subset of homeobox genes with well-characterized roles. Transcripts for seven of these newer genes have been detected in pituitary cell lines (\(Lhx2\), \(Meis2\), \(Pknax1\), \(Pou4f1\)), pituitary adenomas (\(Pou4f1\), \(Zeb2\)), or other...
developing vertebrates (Rax) [37,60,63–71]. These genes correspond to over fifteen different homeodomain classes including the LIM, Paired (Prd), Paired-like (Prd-like) and POU homeobox families [reviewed by [72]]. Given the importance of known genes from these gene families in pituitary organogenesis such as Prop1, Pou1f1, Pitx1, and Pitx2, the identification of Adnp (POU-like), Prrx2 (Prd), and Prrx1 (Prd), suggests that these genes may also play important roles in driving cell proliferation and/or differentiation during the formation of the pituitary gland. Mutations in some of these genes may underlie some cases of MPHD, where the current genetic cause is unknown.

Interestingly, Meox2, Otx2, Emx2, and Zeb2 are differentially expressed in the developing Rathke’s pouch between E12.5 and E18.5. No obvious differences in pituitary cell specification are noted in Emx2 mutant mice [87]. Meox2 and Otx2 are elevated in the Prop1[+/-] E14.5 pituitary relative to E14.5 wild type, and Prop1[+/-] newborns have ectopic, elevated expression of Otx2 compared to wild type littermates (Mortensen, unpublished observation). Meis1 and Adnp are reduced in the Prop1[+/-] E14.5 pituitary relative to E14.5 WT. The dynamic expression patterns of these homeobox genes in the developing pituitary raise the question of whether they are regulated by Prop1 and whether they have functional roles in pituitary development.

Some of the remaining homeobox genes identified represent novel classes of homeobox genes, not previously found in the developing pituitary. For example, three zinc finger (ZF) homeobox genes were identified in the encyclopedia (Table 4). Zeb2 expression was detected at low levels at E12.5, expression levels increased at both E14.5 and E18.5 in normal mice and expression was reduced in the Prop1[+/-] E14.5 pituitary (Fig. 4), and Zfx1 has been detected at e12.5 by in situ hybridization (data not shown). Mutations in one of these genes, ZEB2 or ZFHX1B, cause Mowat Wilson Syndrome, a human disorder characterized by a distinctive facial phenotype in association with mental retardation, microcephaly, and short stature [73]. Because this gene is expressed in human pituitary, it is intriguing to posit that affected individuals with this syndrome have a pituitary defect [60].

Several genes belonging to the TALE family of homeobox genes were identified in the embryonic pituitary (Pbx2, Meox2, Meis1, Meis2, Pknox2, Tgif1, and Tgif2). Genes from this family are required to specify cell fate in numerous organs including the spleen, and they may play a similar role in the developing pituitary [74]. The TALE gene Pbx1 regulates expression of the Eshb gene in pituitary gonadotrope-like cells and Gnrh1 in the hypothalamus [64,75]. MEIS2 interacts with LHX2 to regulate Cga expression in pituitary gonadotrope-like cells [66]. Our expression analysis of Meox2 suggests that Prop1 may repress its expression during pituitary development.

One of the current challenges in the field is to determine how the various signaling pathways interact with each other during pituitary organogenesis. BMP and FGF signals influence anterior pituitary cell differentiation [155], and SHH signaling may participate in boundary formation with FGF and BMP signaling in the ventral diencephalon [54,55]. Cross talk between the WNT, BMP, FGF and SHH pathways is evident in Tgf12, Wnt5a, and noggin mutants, which exhibit pituitary developmental abnormalities [2,24,54]. Using the embryonic pituitary encyclopedia we identified 79 additional BMP, FGF, NOTCH, SHH or WNT pathway members with novel expression in the developing pituitary. One such example is Fgf13, which is critical for neural
Members of both the Sprouty-related, EVH1 domain containing 2
Tbx3 T-box 3 transcription factor 3
Spry4 Jun

The expression patterns of these were studied by in situ hybridization.

**Library generation**

Pituitary tissue for library generation was prepared as described [38]. Briefly, starting from 50 μg of total RNA, first strand cDNA was prepared with an anchored oligo-dT primer adapter [80] and the cDNA was cap-selected [81] and subsequently cloned in a Lambda FLC vector [41]. Libraries were excised into plasmids and sequenced. E14.5 WT was given the library ID K7, E14.5 Prop1df/+ was K8, and E12.5 WT was K9 [38]. Additionally, subtracted libraries were prepared from E14.5 WT minus E14.5 Prop1df/+ and E14.5 WT minus E12.5 WT, using established procedures [82].

**Database generation**

Mouse UniGene Build #159 (Mm.seq and Mm.seq.uniq) and Mouse RefSeq Release 20 were obtained from NCBI FTP servers, along with the ‘blastall’ program used to perform the BLAST identifications. Each cDNA sequence was first BLASTed against the Mm.seq.uniq dataset, and then against the full mouse UniGene set (Mm.seq). All cDNA sequences were also BLASTed against RefSeq to provide an alternative identification.

A database was constructed to contain the sequences themselves, along with Riken identification and a ‘library’ identification, indicative of the tissue source along with the UniGene and RefSeq identifications, where obtainable. Web-based search functions were developed to allow exploration of the contents of the various libraries by sequence, gene name, etc. From the GO Consortium (www.geneontology.org) we downloaded a database of GO identifiers in hierarchical format (10/31/06 release). This data, along with files from the NCBI ‘Gene’ database (‘gene2go’, ‘gene2unigene’) allowed us to query our UniGene identifications using GO terminology.

**Gene Ontology (GO) terminology analysis**

Over-represented Gene Ontology (GO) terms in the set of unique sequences from each library as well as the intersecting sequences for all libraries were detected using a modification of the standard Fisher’s exact test [83]. Briefly, we selected those sequences that were unique to each library based on the UniGene ID and mapped these IDs to Entrez Gene IDs using the mouseLMMappings package of Bioconductor (http://www.bioconductor.org). We mapped the union of all sequences in each library to Entrez Gene to use as the ‘universe’ from which the unique sequences were selected. In other words, the usual heuristic explanation of a Fisher’s exact test is the idea of selecting colored balls from an urn. In our case, the unique sequences from each library are the balls that were selected from the urn, and the union of all sequences in each library is the ‘universe’ of balls in the urn. We computed over-representation of GO terms using the topGO package of Bioconductor, which uses the directed acyclic graph (DAG) structure of the GO ontologies to decorrelate significant subordinate terms from consideration when estimating the significance of higher-order terms.

**In situ hybridization**

RNA in situ hybridization was performed as described previously [84]. Dlx1, Dlx3, Rax, Zfhx1, Id2, Id3 and Tgfbi cDNAs were isolated from the embryonic pituitary encyclopedia [38]. Id2, Id3 and Tgfbi were linearized with NotI, and Dlx1, Dlx3, Rax and Zfhx1 were linearized

### Table 5

Identification of signaling-related genes in the embryonic pituitary encyclopedia

<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>Gene name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fgf10</td>
<td>Fibroblast growth factor 10</td>
</tr>
<tr>
<td>Fgf13</td>
<td>Fibroblast growth factor 13</td>
</tr>
<tr>
<td>Fgf14</td>
<td>Fibroblast growth factor 14</td>
</tr>
<tr>
<td>Fgf17</td>
<td>Fibroblast growth factor 17</td>
</tr>
<tr>
<td>Fgf1</td>
<td>Fibroblast growth factor receptor 1</td>
</tr>
<tr>
<td>Fgf2</td>
<td>Fibroblast growth factor receptor 2</td>
</tr>
<tr>
<td>Fibp</td>
<td>Fibroblast growth factor intracellular binding protein</td>
</tr>
<tr>
<td>Frz3</td>
<td>Fibroblast growth factor receptor substrate 3</td>
</tr>
<tr>
<td>Spry1</td>
<td>Sprouty homolog 1 (Drosophila)</td>
</tr>
<tr>
<td>Spry2</td>
<td>Sprouty homolog 2 (Drosophila)</td>
</tr>
<tr>
<td>Spry4</td>
<td>Sprouty homolog 2 (Spry2)</td>
</tr>
<tr>
<td>Spred2</td>
<td>Sprouty-related, EVH1 domain containing 2</td>
</tr>
<tr>
<td>Sp3</td>
<td>Trans-acting transcription factor 3</td>
</tr>
<tr>
<td>Tbx3</td>
<td>T-box 3</td>
</tr>
<tr>
<td>Id2</td>
<td>Inhibitor of DNA binding 2</td>
</tr>
<tr>
<td>Id3</td>
<td>Inhibitor of DNA binding 3</td>
</tr>
<tr>
<td>Tgfβ4</td>
<td>Transforming growth factor, beta induced</td>
</tr>
<tr>
<td>Ascl1*</td>
<td>Achaete-scute complex homolog-like 1 (Drosophila)</td>
</tr>
<tr>
<td>Id1*</td>
<td>Inhibitor of DNA binding 1</td>
</tr>
<tr>
<td>Jun*</td>
<td>Jun oncogene</td>
</tr>
<tr>
<td>Neurod1*</td>
<td>Neurogenic differentiation 1</td>
</tr>
</tbody>
</table>

* A subset of the 63 genes identified is shown. The complete list is available in Supplementary Figure 1.

† Of the 63 genes detected, only Fgf10 was previously known to be expressed in the pituitary gland.

* The expression patterns of these were studied by in situ hybridization.

† The asterisk indicates genes in the basic helix-loop-helix family.

differentiation in Xenopus [76]. Additionally, three Sprouty genes were identified in the encyclopedia (Spry1, Spry2 and Spry4). Spry1, Spry2, and Spry4 expression was previously reported in E14.5 Rathke’s pouch [77]. Sprouty genes have highly conserved roles as antagonists of FGF signaling, which can affect differentiation [78,79]. Thus, several members of the Sprouty family may regulate FGF signaling during differentiation in Xenopus [76]. Additionally, three Sprouty genes were identified in the encyclopedia (Spry1, Spry2 and Spry4). Spry1, Spry2, and Spry4 expression was previously reported in E14.5 Rathke’s pouch [77]. Sprouty genes have highly conserved roles as antagonists of FGF signaling, which can affect differentiation [78,79]. Thus, several members of the Sprouty family may regulate FGF signaling during pituitary organogenesis.

An important application of our studies lies in the ability to utilize the full-length cDNA clones. Some examples are the full-length cDNA clones for the Fstl1, Id3, and Six6 genes that were used for expression analysis by in situ hybridization in pituitary development in demonstrating the important role of noggin [54]. Full-length cDNA clones for the forkhead transcription factor Fox2 and numerous members of both the Wnt and Frizzled gene families, (Fzd6, Axin2 and Dvl2, see Supplementary Table 1) have been invaluable [2,18].

In conclusion, the embryonic pituitary encyclopedia represents a valuable gene discovery tool that can be used to identify novel genes and pathways in the developing wild type and mutant pituitary (Prop1df/df) and to initiate functional studies. We have just scratched the surface and anticipate that this full-length cDNA encyclopedia will act as an ongoing resource for studying pituitary development and disease.

### Materials and methods

**Mice**

Prop1df/+ mouse stocks were obtained from Dr. A. Bartke (Southern Illinois University, Carbondale, IL) in 1988 and maintained at the University of Michigan according to NIH guidelines. Heterozygote carriers were intercrossed to generate timed pregnancies and the morning after mating was designated as E0.5. Developing pituitary glands were dissected from E14.5 embryos with the aid of a dissecting microscope and were individually stored in RNA Later (Ambion, Austin, TX) at −20 °C. Genomic DNA was extracted from the yolk sac, genotyping was performed as described [21], and pituitaries were pooled for RNA extraction and cDNA generation following genotyping. CD1 mice were purchased from Charles River Laboratories (Wilmington, MA) as timed pregnancies and pituitary glands were dissected from E12.5, E14.5, and E18.5 embryos. Pituitaries were pooled for RNA extraction and cDNA generation. The University of Michigan Committee on Use and Care of Animals approved all experiments.
with Sacll. T3 polymerase was used in all cases to generate antisense probes.

RT-PCR and PCR

RNA was extracted from pituitary tissue dissected from CD1 mouse embryos at E12.5, E14.5, and E18.5 and stored in RNAlater (Ambion, Foster City, CA) following the recommended commercial protocol. RNA was further purified via RNAqueous 4PCR (Ambion, Foster City, CA). cDNA was prepared from total RNA using the Superscript First Strand Synthesis System for RT-PCR (Invitrogen, Carlsbad, CA). Briefly, 5 μg total RNA was incubated at 42 °C for 50 min in a 10 μl reaction containing 2 μl cDNA buffer, 4 μl MgCl₂, 2 μl 0.1 μl DTT, and 1 μl of Superscript II RT enzyme. Standard PCR reaction components at final concentrations were 1 × PCR Reaction buffer (Roche, Indianapolis, IN), 1.5 mM MgCl₂ (Perkin Elmer, Waltham, MA), 200 mM dNTP mix (Roche), and 0.4 U GoTaq polymerase (Perkin Elmer) per 25 μl reaction. Standard reactions contained 25–50 ng template cDNA.

All PCR reactions were performed using the following conditions: 94 °C; 4 min, followed by 30 cycles of 94 °C; 30 s, PCR-specific annealing temperature; 30 s and 72 °C; 30 s, followed by 1 cycle of 72 °C; 10 min. PCR-specific annealing temperatures and primers are listed in Supplementary Table 2.

Quantitative Real Time PCR was performed using TaqMan Gene Expression Assays On Demand (Applied Biosystems, Foster City, CA). Briefly 25–50 ng of template cDNA was added to 10 μl Taqman Universal PCR Master Mix (Applied Biosystems), 8 μl double distilled H₂O, and 1 μl TaqMan gene specific primer set. All samples were done

![Fig. 5. Developmental expression of 3 novel Bmp-related genes in the embryonic pituitary gland. (A) Expression of Id2 is predominantly in the intermediate lobe (IL) and around the lumen of the pituitary and is excluded from the forming anterior lobe (delineated by the black dashed line). Expression is also detected in the ventral diencephalon (VD). This expression pattern persists at E14.5 (B). (C) At E12.5 Id3 is expressed in the IL and dorsal part of the cells surrounding the pituitary lumen (indicated by the horizontal black line) and this expression pattern continues at E14.5 (D). (E) At E12.5 Tgfb1 was detected in the mesenchymal tissue immediately surrounding the developing pituitary (shown by the black arrows). By E14.5 (F), Tgfb1 expression becomes stronger in the mesenchyme and is also detected surrounding the posterior pituitary (PP) (shown by the black arrows).](image-url)
In triplicate and standardized to Hprt1, Real Time PCR was performed using the Prism 7000 Sequence Detection System (Applied Biosystems). Fold activation and standard deviations were determined as previously described [85].

Acknowledgments

This work was supported by the following grants: National institutes of Health R37HD30428 and HD34283 (SAC), NIH Grant R01GM72007 (JM, DG), University of Michigan Center for Computational Medicine and Biology (SAC, RHL, DG), University of Michigan DNA Sequencing Core Facility (RHL) and Research Grant for the RIKEN Genome Exploration Research Project from the Ministry of Education, Culture, Sports, Science and Technology, Japan (YH). We thank A.H. Mortensen, M.A. Potok, and N.M. Solomon for their contributions to this work.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ygeno.2008.11.010.

References


