Characterization of hormone-producing cell types in the teleost pituitary gland using single-cell RNA-seq

Khadeeja Siddique, Eirill Ager-Wick, Romain Fontaine, Finn-Arne Weltzien ✉, & Christiaan V. Henkel ✉

The pituitary is the vertebrate endocrine gland responsible for the production and secretion of several essential peptide hormones. These, in turn, control many aspects of an animal’s physiology and development, including growth, reproduction, homeostasis, metabolism, and stress responses. In teleost fish, each hormone is presumably produced by a specific cell type. However, key details on the regulation of, and communication between these cell types remain to be resolved. We have therefore used single-cell sequencing to generate gene expression profiles for 2592 and 3804 individual cells from the pituitaries of female and male adult medaka (Oryzias latipes), respectively. Based on expression profile clustering, we define 15 and 16 distinct cell types in the female and male pituitary, respectively, of which ten are involved in the production of a single peptide hormone. Collectively, our data provide a high-quality reference for studies on pituitary biology and the regulation of hormone production, both in fish and in vertebrates in general.

Background & Summary

The pituitary is a master endocrine gland in vertebrates, which is involved in the control of a variety of essential physiological functions including growth, metabolism, reproduction, and response to stress. These functions are modulated by the secretion of several peptide hormones, produced by different endocrine cell types of the adenohypophysis. While in mammals the endocrine cells are distributed throughout the adenohypophysis, the pituitary of teleost fish is highly compartmentalized, with specialized hormone-producing cell types located in specific regions. As the teleost pituitary produces at least eight peptide hormones in distinct cell types, the physiology of the pituitary gland is relatively complex. The rostral pars distalis contains lactotropes and corticotropes, which produce prolactin (Prl) and adrenocorticotropic hormone (Acth), respectively. Somatotropes (producing growth hormone, Gh), gonadotropes (luteinizing hormone, Lh and follicle-stimulating hormone, Fsh) and thyrotropes (thyroid-stimulating hormone, Tsh) are located in the proximal pars distalis, and somatolactotropes (somatolactotropin, Sl) and melanotropes (α-melanocyte stimulating hormone, α-Msh) in the pars intermedia. In addition, in tetrapods, the gonadotropins Lh and Fsh can be produced in the same cell, while fish gonadotropes are generally thought to secrete either Fsh or Lh, but not both.

In recent years, transcriptome sequencing (RNA-seq) has emerged as a powerful technology to study the expression and regulation of pituitary genes. For example, RNA-seq has been used to study gene expression in the zebrafish pituitary during maturation, in prepubertal female silver European eel pituitary glands, and in FACS-selected Lh cells from the Japanese medaka pituitary. Although these studies provide a valuable perspective on the regulation of hormone production, RNA-seq is a bulk technology, which averages gene expression over the entire tissue sample. As a result, it does not provide information on which hormones are produced in which cells. Knowledge on the physiology and development of the individual pituitary cell types, as well as on the regulatory mechanisms involved in hormone production, therefore remains limited.
In this study, we employed single-cell RNA-seq (scRNA-seq) technology, which allows the study of biological questions in which cell-specific changes in the transcriptome are important, e.g. cell type identification, variability in gene expression, and heterogeneity within populations of genetically identical cells. scRNA-seq has recently been used to describe the heterogeneity within the pituitary cell populations in mammals and zebrafish. Unlike bulk RNA-seq, scRNA-seq promises to disentangle the processes at work in pituitary hormone production.

We used medaka (Oryzias latipes) to construct a transcriptomic snapshot of the cell types in the adult pituitary gland. Medaka is an emerging model species, which has a small genome (800 Mbp) and is a representative of the largest radiation within the teleosts (euteleosts, Cohort Euteleostomorpha). As such, it is a complementary model to zebrafish, which belongs to another major teleost clade (Cohort Ototomorpha).

We used the 10x Genomics scRNA-seq platform to analyze female and male pituitaries separately and obtained 2592 and 3804 high-quality cellular transcriptome profiles, respectively. This dataset provides a comprehensive transcriptomic reference on the teleost pituitary, as well as on its constituent cell types. After bioinformatics analysis, we found these profiles to belong to 15 and 16 distinct cell populations in the female and male pituitary, respectively, of which ten are involved in the production of a single peptide hormone. This suggests the teleost pituitary is subject to a higher degree of division of labour than its mammalian counterpart, in which many cells express multiple hormone-encoding genes.

**Methods**

**Fish strain and animal husbandry.** For this study, we used a transgenic line of Japanese medaka derived from the d-R genetic background, in which the Lhβ gene (lhβ) promoter drives the expression of the green fluorescent protein (hrGFP) gene. Fish were kept in a re-circulating system (up to 10–12 fish per 3-liter tank) at 28 °C on a 14/10 h light/darkness cycle and were fed three times a day with a mix of dry food and *Artemia*. The animal experiments performed in this study were approved by the Norwegian University of Life Sciences, following guidelines for the care and welfare of research animals.

**Pituitary sampling and cell dissociation.** Pituitaries were collected from 24 female and 23 male adult medaka of approximately 8 months old (eggs fertilized 20 October 2017, sampling 6 June 2018). At this age, medaka are fully mature and sexually dimorphic. Fish were sacrificed between 07.30–09.00 h in the morning (around the time of spawning induced by the start of the artificial light period) by hypothermic shock (immersion in ice water) to minimize distress. The pituitary was dissected immediately after severing the spinal cord. In order to reduce the influence of sampling time on the results, fish were processed in alternating batches of ten individuals of each sex at a time.

After sampling, pituitaries were pooled per sex and were dissociated as described. Briefly, fresh pituitaries were put in an Eppendorf tube with modified PBS (phosphate buffered saline, pH 7.75, 290 mM sodium, 0.05% bovine serum albumin) and kept on ice until processing. Pituitaries were treated with 0.1% w/v trypsin type II S for 30 minutes at 26 °C, then with 0.1% w/v trypsin inhibitor type I S supplemented with 2 µg/mL DNase I for 20 minutes at 26 °C, and subsequently gently mechanically dissociated by repeated aspiration with a strained glass pipette. The cell solution was filtered through a 35-µm cell strainer (BD Pharningen) to remove potentially remaining clumps of cells. Dissociated pituitary cells were resuspended in 90 µL modified PBS, counted and visually inspected for good cell dissociation and quality. Samples were kept on ice for approximately 30 minutes until scRNA-seq library preparation, which commenced at 12.00 h on the same day.

**10x Genomics library preparation.** We prepared scRNA-seq libraries on the 10x Genomics Chromium Controller at the Genomics Core Facility of the Radium Hospital (part of Oslo University Hospital). For the initial step of the library preparation (GEM generation and barcoding) 35 µL of the cell suspension prepared in the previous step was used, containing approximately 10,000 (female) and 11,000 (male) cells. The 10x Genomics Single Cell 3′ Reagents Kits v2 were used according to the manufacturer’s guidelines to prepare cDNA libraries for a target of 4,000 cells.

**Sequencing.** Per sample, four redundant sequencing libraries (without additional sequencing controls) were analyzed on an Illumina NextSeq 500 at the Genomics Core Facility. Read 1 was sequenced for 28 cycles (covering the 26 nt cellular barcode and UMI), read 2 for 96 cycles (covering the cDNA insert). For downstream analyses, the resulting FASTQ files were collated per sample.

**Reference genome and annotation.** We used the chromosome-scale Hd-rR medaka reference genome (Ensembl release 94) for alignments. In preliminary analyses, we noticed that many reads for expected pituitary-specific transcripts were mapped near, but outside the 3′ boundary of gene annotations. Therefore, in order to improve transcript quantification accuracy, we manually inspected the read alignment to these and other highly expressed genes (using Samtools v 1.10 and the Integrative Genomics Viewer v 2.3.3), and adjusted 3′ UTR annotations where necessary (Table 1). The resulting custom gene transfer file (GTF) was then supplied using the 10x Genomics Cell Ranger v 3.0.2 (https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/what-is-cell-ranger) mkref command.

**Initial data processing.** We used the standard 10x Genomics Cell Ranger (v 3.0.2) pipeline with default parameters for preliminary analyses. It includes quality control of the FASTQ files, alignment of FASTQ files to the customized medaka reference using STAR (v 2.5.1b), demultiplexing of cellular barcodes, and quantification of gene expression. Table 2 summarizes the Cell Ranger quality control and output.
Cell Ranger also performs automated cell clustering based on the similarity of gene expression profiles. We used the Loupe Cell Browser v 3.1.1 ([https://support.10xgenomics.com/single-cell-gene-expression/software/visualization/latest/what-is-loupe-cell-browser]) to visualize these results. Unsupervised graph-based clustering suggested nine and eleven cell populations in the female and male medaka pituitary, respectively (see Code Availability). For a more detailed analysis, we used the Seurat27 R toolkit (v 3.1.5) and dropEst28 (v 0.8.5) for accurate estimation of molecular counts per cell (Table 3). We used the .bam files produced by Cell Ranger as the input for dropEst.

Quality control (QC). dropEst by default reports any barcode with more than 100 UMIs as cellular (filtered cells in Table 3). Presumably, this relaxed threshold still includes many non-cellular and debris barcodes. Therefore, we explored these data using Seurat, and derived additional selection criteria for cellular barcodes based on the UMIs counts, mitochondrial fraction and globin expression. After imposing these thresholds, 2644 and 3921 cells remained for the female and male pituitary, respectively.

| Gene   | Ensembl ID          | Feature       | Original            | Updated            |
|--------|---------------------|---------------|---------------------|--------------------|
| pomca  | ENSORLG00000025908  | 3’ exon       | 2:10971088–10973672 | 2:10971088–10971837 |
| prl    | ENSORLG00000016928  | 3’ exon       | 8:23591803–23592288 | 8:23591803–23592634 |
| ENSLG00000013460   | 3’ exon       | 13:26485894–26486532 | 13:26485894–26486136 (–) |
| gh     | ENSORLG00000019556  | 3’ exon       | 8:75404–76751 (+)   | 8:75404–75568 (+)   |
| thiba  | ENSORLG0000029251   | 3’ exon       | 5:13105690–13108242 | 5:13107640–13108242 (–) |
| Jdh    | ENSORLG00000029237  | 3’ exon       | 3:9888850–9890682 (–) | 3:9890408–9890682 (–) |
| Hbh    | ENSORLG00000030553  | 3’ exon       | 15:14339670–14339965 (–) | 15:14339670–14340309 (+) |
| nr5a1  | ENSORLG0000013196   | 3’ exon       | 12:24979868–24980234 (–) | 12:24979864–24980234 (–) |
| Cgsa   | ENSORLG00000022598  | 3’ exon       | 22:10122990–10126745 (–) | 22:10126430–10126745 (–) |
| 3’ exon | 22:10124168–10144948 (–) | 22:10124168–10144948 (–) |
| mdkb   | ENSORLG00000014169  | 3’ exon       | 6:26806299–26806315 (–) | 6:26805916–26806315 (–) |
| eyf1a  | ENSORLG00000007614  | 3’ exon       | 11:17258988–17259109 (–) | 11:17258988–17258678 (–) |
| fth1a  | ENSORLG00000005872  | 3’ exon       | 3:17888694–17888835 (–) | 3:17888418–17888835 (–) |
| exon   | 3:17888677–17888835 (–) | 3:17888418–17888835 (–) |
| mcee   | ENSORLG00000022009  | 3’ exon       | 3:17885682–17890772 (+) | 3:17885682–17886310 (+) |

Table 1. Genome annotation adjustments.

| Gene   | Ensembl ID          | Feature       | Female   | Male    |
|--------|---------------------|---------------|----------|---------|
| pomca  | ENSORLG00000025908  | 3’ exon       | 165818560 | 142738097 |
| prl    | ENSORLG00000016928  | 3’ exon       | 97.5%    | 97.4%   |
| ENSLG00000013460   | 3’ exon       | 79.7%    | 80.3%   |
| gh     | ENSORLG00000019556  | 3’ exon       | 83.4%    | 81.5%   |
| thiba  | ENSORLG0000029251   | 3’ exon       | 84.1%    | 84.1%   |
| Jdh    | ENSORLG00000029237  | 3’ exon       | 59.9%    | 62.5%   |
| Hbh    | ENSORLG00000030553  | 3’ exon       | 82.4%    | 85.8%   |
| nr5a1  | ENSORLG0000013196   | 3’ exon       | 2890     | 4321    |
| Cgsa   | ENSORLG00000022598  | 3’ exon       | 3:17888677–17888835 (–) | 3:17888418–17888835 (–) |
| 3’ exon | 3:17888505–17888070 (–) | 3:17885682–17889077 (–) |

Table 2. Sequencing and Cell Ranger statistics.

| Gene   | Ensembl ID          | Feature       | Female   | Male    |
|--------|---------------------|---------------|----------|---------|
| pomca  | ENSORLG00000025908  | 3’ exon       | 74338    | 77520   |
| prl    | ENSORLG00000016928  | 3’ exon       | 4945     | 7334    |
| ENSLG00000013460   | 3’ exon       | 18786    | 19043    |
| gh     | ENSORLG00000019556  | 3’ exon       | 17284    | 17882   |
| thiba  | ENSORLG0000029251   | 3’ exon       | 15036    | 15598   |
| Jdh    | ENSORLG00000029237  | 3’ exon       | 2644     | 3921    |
| Hbh    | ENSORLG00000030553  | 3’ exon       | 2592     | 3804    |

Table 3. Final cell quantification statistics.
Since single-cell capture is sensitive to the formation of doublets, we used the R package DoubletFinder\textsuperscript{29} (v 2.0.3) to detect hybrid expression profiles in our data. The number of expected real doublets depends on the number of cells captured. We selected a 2% doublet rate for the female sample and a 3% doublet rate for the male sample, based on the 10x Genomics specifications (https://support.10xgenomics.com/single-cell-gene-expression/index/doc/user-guide-chromium-single-cell-3-reagent-kits-user-guide-v2-chemistry). We removed 52 and 117 putative doublets from the female and male pituitary data, respectively.

**Clustering.** After QC, we normalized the data with the LogNormalize method in Seurat and identified highly variable genes by using the vst method. Subsequently, we used these variable genes (n = 2000) to identify significant principal components (PCs) based on the jackStraw function. We used ten informative PC dimensions in both samples as the input for uniform manifold approximation and projection (UMAP). After the dimensionality reduction, we clustered the cells using the FindClusters function with a resolution of 0.5 and 0.9 for the female and male data sets, respectively, and initially characterized 11 different clusters in both the female and male pituitary.

**Cell type assignment.** We used differentially expressed genes expressed per cluster (FindAllMarkers with min. pct = 0.25) to help assign tentative cell type identities to these clusters. Marker genes (Table 4) reported in previous studies\textsuperscript{2,14-16,21,30-32} were used for cell type assignment.

**Cluster refinement.** During initial data exploration, we observed several clusters sharing many common markers. Therefore, we further inspected their heterogeneity iterating the clustering function on selected subsets of the data. Subclusters were detected across multiple clustering resolutions (0.2 and 0.5) of the FindClusters function in Seurat. We subsequently used the FindAllMarkers function to find differentially expressed genes between each of these subclusters to highlight the differences and finally decided on the cell type boundaries for these ‘zoomed-in’ sub-clusters.

**Data integration.** We combined the female and male datasets using the Seurat FindIntegrationAnchors and IntegrateData functions, using default settings. For subsequent visualization, we randomly downsampled the male data to 2592 cells.

**Bulk RNA-seq analyses.** We quantified the Cell Ranger .bam files without cellular barcode demultiplexing using htsq-count\textsuperscript{30} (v 0.11.2) with the intersection-nonempty setting. In addition, we quantified gene expression from comparable (adult) samples of a separate bulk RNA-seq study\textsuperscript{14,15} on the medaka pituitary gland. Sequencing libraries for these samples (females 122–124, 249–277 days post-fertilization; and males 137–139, 96–129 days post-fertilization) were previously prepared using the Smart-SEQ HT kit (Takara Bio), sequenced at 2 × 151 nt on an Illumina NovaSeq 6000, and aligned to the modified medaka reference genome using STAR. For both studies, 23635 annotated protein-coding genes were quantified in counts per million reads (CPM).

**Table 4.** Cells per cluster in scRNA-seq data.

| Cell types          | Hormone (gene) | Female cells | Male cells | Marker gene | References |
|---------------------|----------------|--------------|------------|-------------|------------|
| Melanotropes        | α-Msh (cga)    | 54           | 53         | cga, oacyl  |            |
| Corticotropes       | Acth (c1qb)    | 25           | 55         | c1qb, oacyl|            |
| Lactotropes         | Prl (c1qc, c1qb, c1qc) | 123 | 175 | prl |            |
| Lactotropes         | Prl (c1qc, c1qb, c1qc) | 69 | 222 | prl |            |
| Somatolactotropes   | Sl (tshba)     | 15           | 38         | tshba       |            |
| Somatotropes        | Gh (gb)        | 52           | 127        | gb          |            |
| Thyrotropes         | Tsh (gb)       | 69           | 46         | gb          |            |
| Fish- gonadotropes  | Fsh (fshb)     | 72           | 198        | fshb        |            |
| Lh-gonadotropes     | Lh             | 535          | 895        | Lhb         |            |
| Gonadotrope-like    | Lh             | 597          | 444        | cga, m5a1b  |            |
| Red blood cells     |                | 371          | 559        | gb          |            |
| Macrophages         | 15             | 15           | 0.6        | c1qa, c1qb, c1qc | 14,15 |
| Uncharacterized     | 277            | 10.7         | 410        | 10.8        |            |
| Uncharacterized     | 198            | 7.6          | 386        | 10.2        |            |
| cga-expressing cells| 120            | 4.6          | 160        | 4.2         | cga, prl   |            |
| Uncharacterized     |                | 21           | 0.6        |             |            |

**Data Records**

We present a characterization of the cell types of the medaka pituitary gland, as a basis for more in-dephts analyses of the regulation of and inter-communication between different hormones and cell types.

Our data set on the medaka pituitary consists of gene expression values for 2592 and 3804 adult female and male cells, respectively. These can be assigned to 15 or 16 distinct cell populations, respectively, one of which appears to be unique to male pituitaries. We assigned biological identities to 12 populations based on the expression of known hormone-encoding genes. Figure 1 and Table 4 provide an overview of the cell types of the medaka pituitary.
The main data record consists of barcodes.tsv, features.tsv and matrix.mtx files, listing raw UMI counts for each gene (feature) in each cell (barcode) in a sparse matrix format. In addition, we provide a table of cell type assignments and UMAP projections for each individual cell (barcode), as well as a summary of the top differentially expressed genes per cell type. Raw sequencing data are available in FASTQ format. Finally, the data record contains matrix files on exonic and intronic expression, which can be used for Velocyto gene expression dynamics analyses. These data can be accessed through the project accession number GSE162787 at the NCBI Gene Expression Omnibus.

The bulk RNA-seq data used for validation can be accessed through GSE179598 at the NCBI Gene Expression Omnibus, and will be described in more detail in a separate publication.

Technical Validation
In order to validate the quality of our data, we investigated their reproducibility, the influence of technical variables, and the biological perspective.

Reproducibility. At the level of cell type clustering (Fig. 1a–b), both datasets appear superficially similar. We further investigated whether the cells in the female and male medaka pituitary are comparable by integrating both datasets in a single analysis. The resulting UMAP visualization (Fig. 1c) suggests this is indeed the case, with all major cell clusters remaining intact in this projection. In general, the numbers of cells per cluster are
similar between the two samples (Fig. 1d), with minor differences that are consistent with previously established sex-specific patterns.

This analysis also suggests both datasets are highly reproducible. We further established this by comparing the compound expression of protein-coding genes in the data (Fig. 2a), i.e. without deconvolution into constituent cells. At the level of read counts per million (CPM), female and male data are again highly similar (Spearman rank correlation 0.98).

As one of our intended uses for this scRNA-seq dataset is to interpret patterns in additional RNA-seq studies, we also compared these counts to CPM derived from bulk RNA-seq on biologically similar samples (mean of three samples). (d) Idem for male scRNA-seq and three comparable male RNA-seq samples.

---

**Fig. 2** Sequencing reproducibility. (a) Stacked barplots of read mapping to genomic fractions for both samples. (b) Scatterplot of gene expression in both scRNA-seq samples processed as bulk samples, showing good correspondence between females and males. Yellow dots indicate the major protein hormone-encoding genes. CPM: counts per million. (c) Scatterplot of compound gene expression (CPM) in female scRNA-seq compared to bulk RNA-seq (CPM) on comparable samples (mean of three samples). (d) Idem for male scRNA-seq and three comparable male RNA-seq samples.
Technical quality control. Interpretation of single-cell transcriptomics data is highly sensitive to technical artifacts. For example, expression profiles can derive from low quality cells, cellular debris, or ambient RNA. We therefore determined empirical thresholds distinguishing genuine cellular expression profiles from noise (Fig. 3). We defined selection criteria for cellular barcodes based on UMI counts (Fig. 3a), expression of mitochondrial genes (Fig. 3b, cyan dots), expression of genes specific to red blood cells (Fig. 3c, red dots), and a probability score for cellular doublets (Fig. 3a–c, orange dots). Based on the distribution of UMI counts per barcode, we defined
as cellular barcodes those with ≥1500 UMIs in both female and male samples. High levels of expression of mitochondrial genes have been interpreted as an indication of damaged cells\textsuperscript{40,41}. We therefore excluded barcodes in which mitochondrial gene expression contributes more than 5% of the total (for both females and males, Fig. 3b).

In initial clustering analyses, we noticed that these criteria exclude both technical artifacts (data not shown) and a single cluster of cells with low overall UMI counts, but a very clearly defined expression profile. Based on their expression profile we interpret these barcodes to represent red blood cells. Although they are developmentally only very distantly related to the other cells in the pituitary, our data show that they are an intrinsic (if transient) component. Therefore, including them in our dataset will allow for more precise interpretation of patterns in bulk RNA-seq studies (e.g. Figure 2c–d).

With the exception of these red blood cells, all other defined cell types show a high and consistent number of genes expressed per individual cell (Fig. 4a). In addition, a saturation analysis (Fig. 4b), contrasting the number of genes expressed per cell and the sequencing depth per cell, suggests the latter is sufficient to capture most of the transcriptomic complexity for each cell type.

**Biological quality control.** In order to validate the biological relevance of the cell type identities shown in Fig. 1, we have compared their gene expression patterns with known expression profiles from both teleosts and mammals. Initially, we identified eleven different populations based on optimized Seurat clustering. However, there were several clusters that shared high expression of known, specific marker genes. We therefore...
evaluated expression heterogeneity within each cluster and manually defined final clusters (see Methods). We defined two subclusters for the initial cluster expressing *pomca* (ENSORLG00000025908), which encodes pro-opiomelanocortin, the protein precursor for α-Msh and Acth. The first subcluster has high expression of *pomca*, *oacyl* (ENSORLG00000005993), *crhbp* (ENSORLG00000002228), *pcsk2* (ENSORLG00000006472) and *pax7* (ENSORLG00000004269), which are markers for the α-Msh-secreting melanotropes; the second subcluster was identified as Acth-secreting corticotropes due to lower expression of *pax7* and *pomca*.

Next, a large compound cluster on the right side of Fig. 1 (both females and males) shares three important hormone-encoding genes: *tshb*, *gh* and *smtla*. We therefore divided it into three subclusters using Seurat, gene expression patterns and differentially expressed (DE) genes (see available code for details). The first subcluster has strong expression of *tshba* (ENSORLG00000029251, encoding the β subunit of Tsh), which is a robust marker for thyrotropes (orange color in Fig. 1). The second subcluster expresses *gh* (ENSORLG00000019556), which demarcates the somatotropes (yellow). The last subcluster (light green) has expression of *smtla* (ENSORLG00000013460,
which encodes Sl, a hormone produced by the teleost, but not the mammalian, pituitary gland)\textsuperscript{16,21,30,42}. In contrast to the zebrafish pituitary\textsuperscript{18}, we did not observe the expression of pouf1f (ENSORLG00000015870) in somatolactotropes.

Finally, Seurat initially divided the big cluster of gonadotrope cells (expressing the gonadotropin α-subunit cga, ENSORLG00000022598) at the bottom of Fig. 1 into two subclusters for the males, but not for the females. However, based on the heterogeneity of the expression pattern in the female cluster, and its apparent homology to the two male clusters, we defined two subclusters: one is enriched for tihb (ENSORLG00000035553, encoding the β subunit of Lh) and gnrhr2 (ENSORLG00000012659, also known as gnrhr2a\textsuperscript{43}) expression, both of which are exclusively associated with Lh production in teleosts\textsuperscript{43,44}; the other expresses nr5a1b (ENSORLG00000013196), encoding a nuclear receptor involved in gonadotropin regulation in the mammalian pituitary and teleosts\textsuperscript{16,32}.

Three remaining cell clusters we directly assigned to specific hormone-producing cell types, based on the very high expression of hormone-encoding genes: two distinct populations of lactotropes expressing prl (ENSORLG00000016928) and fshb-expressing gonadotropes (ENSORLG00000029237, encoding the β subunit of Fsh).

After final clustering, we define 15 and 16 distinct expression patterns in female and male data, respectively, which we interpret as distinct cell types (see Fig. 1 and Table 4). Based on their expression patterns, we could assign unambiguous cell type identities to twelve of these (Fig. 5). Ten clusters show high expression of genes encoding protein hormones (Fig. 5a and Table 4). In addition, one cluster expresses cga, but no gene encoding a β subunit. Two clusters appear to be red blood cells and macrophages (Fig. 5b and Table 4). Based on detailed expression patterns (Fig. 5b) we could not unequivocally define the elusive folliculostellate cell population of the pituitary. In the mammalian pituitary, in which cells are often responsible for producing and releasing multiple hormones\textsuperscript{14}.

Overall, our assignments of biological function largely confirm a ‘one cell type, one hormone’ (Fig. 5a) division of labour, in which major hormones are produced by a single, dedicated cell type. This is in apparent contrast to the mammalian pituitary, in which cells are often responsible for producing and releasing multiple hormones\textsuperscript{14}.

**Code availability**

The R code used in the analysis of the scRNA-seq data is available on GitHub (https://github.com/sikh09/Medaka-pituitary-scRNA-seq).

Received: 5 January 2021; Accepted: 9 September 2021; Published online: 28 October 2021

**References**

1. Kelberman, D., Rizzoti, K., Lovell-Badge, R., Robinson, I. C. & Dattani, M. T. Genetic regulation of pituitary gland development in human and mouse. Endocr. Rev. 30, 790–829 (2009).
2. Weltzien, F. A., Andersson, E., Andersen, O., Shalchian-Tabrizi, K. & Norberg, B. The brain—pituitary—gonad axis in male teleosts, with special emphasis on flatfish (Pleurocephaliformes). Comp. Biochem. Physiol. A Mol. Integr. Physiol. 137, 447–477 (2004).
3. Ooi, G. T., Tawadros, N. & Escalona, R. M. Pituitary cell lines and their endocrine applications. Mol. Cell. Endocrinol. 228, 1–21 (2004).
4. Pogoda, H. M. & Hammerschmidt, M. Molecular genetics of pituitary development in zebrafish. Semin. Cell. Dev. Biol. 18, 543–558 (2007).
5. Perez-Castro, C., Renner, U., Haedo, M. R., Stall, G. K. & Arzt, E. Cellular and molecular specificity of pituitary gland physiology. Physiol. Rev. 92, 1–38 (2012).
6. Golan, M., Martin, A. O., Mollard, P. & Levavi-Sivan, B. Anatomical and functional gonadotrope networks in the teleost pituitary. Sci. Rep. 6, 23777 (2016).
7. Yaron, Z. et al. Regulation of fish gonadotropins. Int. Rev. Cytol. 225, 131–185 (2003).
8. Fontaine, R. et al. Gonadotrope plasticity at cellular, population and structural levels: A comparison between fishes and mammals. Gen. Comp. Endocrinol. 287, 113344 (2020).
9. He, W., Dai, X., Chen, X., He, J. & Yin, Z. Zebrafish pituitary gene expression before and after sexual maturation. J. Endocrinol. 221, 429–440 (2014).
10. Ager-Wick, E. et al. The pituitary gland of the European eel reveals massive expression of genes involved in the melanocortin system. PLoS One 8, e77396 (2013).
11. Ager-Wick, E., Henkel, C. V., Haug, T. M. & Weltzien, F. A. Using normalization to resolve RNA-Seq biases caused by amplification from minimal input. Physiol. Genomics 46, 808–820 (2014).
12. Kolodziejczyk, A. A., Kim, J. K., Svensson, V., Marioni, J. C. & Teichmann, S. A. The technology and biology of single-cell RNA sequencing. Nat. Rev. Genet. 16, 547–551 (2017).
13. Lineage analysis reveals an endodermal contribution to the vertebrate pituitary. Science 370, 463–467 (2020).
14. Fshb, a hormone produced by the teleost, but not the mammalian, pituitary gland. Salmo salar model organism. From the far East. Nat. Rev. Genet. 3, 53–64 (2002).
15. Ishikawa, Y. Medaka as a model system for vertebrate developmental genetics. Bioessays 22, 487–495 (2000).
16. Phyllogenetic classification of bony fishes. BMC Evol. Biol. 17, 162 (2017).
17. Developmental tracing of luteinizing hormone beta-subunit gene expression using green fluorescent protein transgenic medaka (Oryzias latipes) reveals a putative novel developmental function. Dev. Dyn. 241, 1665–1677 (2012).
18. Fontaine, R., Ager-Wick, E., Hodne, K. & Weltzien, F. A. Plasticity of Lh cells caused by cell proliferation and recruitment of existing cells. J. Endocrinol. 240, 361–377 (2019).
19. Ager-Wick, E. et al. Preparation of a high-quality primary cell culture from fish pituitaries. J. Vis. Exp. 138, e58159 (2018).
20. Li, H. et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079 (2009).
25. Thorvaldsdóttir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief. Bioinform. 14, 178–192 (2013).
26. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21 (2013).
27. Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene expression data. Nat. Biotechnol. 33, 495–502 (2015).
28. Petukhov, V. et al. dropEst: pipeline for accurate estimation of molecular counts in droplet-based single-cell RNA-seq experiments. Genome Biol. 19, 78 (2018).
29. McGinnis, C. S., Murrow, L. M. & Gartner, Z. J. DoubletFinder: Doublet detection in single-cell RNA sequencing data using artificial nearest neighbors. Cell Syst. 8, 329–337 (2019).
30. Weltzien, F. A., Hildahl, J., Hodne, K., Okubo, K. & Haug, T. M. Embryonic development of gonadotrope cells and gonadotropin hormones—lessons from model fish. Mol. Cell. Endocrinol. 385, 18–27 (2014).
31. Kazeto, Y. et al. Japanese eel follicle-stimulating hormone (Fsh) and luteinizing hormone (Lh): production of biologically active recombinant Fsh and Lh by Drosophila S2 cells and their differential actions on the reproductive biology. Biol. Reprod. 79, 938–946 (2008).
32. Fletcher, P. A. et al. Cell type- and sex-dependent transcriptome profiles of rat anterior pituitary cells. Front. Endocrinol. 10, 623 (2019).
33. Anders, S., Pyl, P. T. & Huber, W. HTSeq – a Python framework to work with high-throughput sequencing data. Bioinformatics 31, 166–169 (2015).
34. Ager-Wick, E. et al. Ectopic expression of lipid homeostasis genes in rare cells of the teleost pituitary gland. Preprint at BioRxiv, https://doi.org/10.1101/2021.06.11.448009 (2021).
35. Gene Expression Omnibus https://identifiers.org/geo:GSE179598 (2021).
36. Gene Expression Omnibus https://identifiers.org/geo:GSE162787 (2020).
37. NCBI Sequence Read Archive https://identifiers.org/insdc.sra:SRP296792 (2020).
38. La Manno, G. et al. RNA velocity of single cells. Nature 560, 494–498 (2018).
39. Royan, M. R. et al. 3D atlas of the pituitary gland of the model fish medaka (Oryzias latipes). Front. Endocrinol. 12, 719843 (2021).
40. Ilicic, T. et al. Classification of low quality cells from single-cell RNA-seq data. Genome Biol. 17, 29 (2016).
41. Luckeen, M. D. & Theis, F. J. Current best practices in single-cell RNA-seq analysis: a tutorial. Mol. Syst. Biol. 15, e8746 (2019).
42. Herkenhoff, M. E. et al. Expression profiles of growth-related genes in two Nile tilapia strains and their crossbred provide insights into intraspecific breeding effects. Anim. Genet. 51, 611–616 (2020).
43. Ciani, E. et al. Gnrh receptor gnrhr2bbalpha is expressed exclusively in lhb-expressing cells in Atlantic salmon male parr. Gen. Comp. Endocrinol. 285, 11293 (2020).
44. Hodne, K., Fontaine, R., Ager-Wick, E. & Weltzien, F. A. Gnrh1-induced responses are indirect in female medaka Fsh cells, generated through cellular networks. Endocrinology 160, 3018–3032 (2019).
45. Lloyd, R. V. & Mailloux, J. Analysis of S-100 protein positive folliculo-stellate cells in rat pituitary tissues. Am. J. Pathol. 133, 338–346.
46. Iakura, E. et al. Generation of transgenic rats expressing green fluorescent protein in S-100beta-producing pituitary folliculo-stellate cells and brain astrocytes. Endocrinology 148, 1538–1523 (2007).

Acknowledgements

We are thankful to Susanne Lorenz for sequencing and Lourdes Carreon G Tan for fish facility maintenance. This work was supported by NMBU and the Norwegian Research Council grants no. 251307, 255601, and 248828.

Author contributions

Eirill Ager-Wick, Finn-Arne Weltzien and Christiaan Henkel conceived and designed the project. Eirill Ager-Wick isolated the cells. Khadeeja Siddique, Romain Fontaine and Christiaan Henkel analyzed the data and wrote the paper with input from all other authors.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to F.-A.W. or C.V.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

The Creative Commons Public Domain Dedication waiver http://creativecommons.org/publicdomain/zero/1.0/ applies to the metadata files associated with this article.

© The Author(s) 2021