Published in final edited form as:
Neuropeptides. 2010 February ; 44(1): 31. doi:10.1016/j.npep.2009.11.002.

Genome-wide Census and Expression Profiling of Chicken Neuropeptide and Prohormone Convertase Genes

K. R. Delfino, B. Sc.[*,†], **B. R. Southey, Ph. D.**[*,†], **J. V. Sweedler, Ph. D.**[*], and **S. L. Rodriguez-Zas, Ph. D.**[†,§]

K. R. Delfino: delfino2@illinois.edu; B. R. Southey: southey@illinois.edu; J. V. Sweedler: jsweedle@illinois.edu; S. L. Rodriguez-Zas: rodrgzzs@illinois.edu

[*] Department of Chemistry, University of Illinois, Urbana IL, USA

[†] Department of Animal Sciences, University of Illinois, Urbana IL, USA

Abstract

Neuropeptides regulate cell-cell signaling and influence many biological processes in vertebrates, including development, growth, and reproduction. The complex processing of neuropeptides from prohormone proteins by prohormone convertases, combined with the evolutionary distance between the chicken and mammalian species that have experienced extensive neuropeptide research, has led to the empirical confirmation of only 18 chicken prohormone proteins. To expand our knowledge of the neuropeptide and prohormone convertase gene complement, we performed an exhaustive survey of the chicken genomic, EST, and proteomic databases using a list of 95 neuropeptide and 7 prohormone convertase genes known in other species. Analysis of the EST resources and 22 microarray studies offered a comprehensive portrait of gene expression across multiple conditions. Five neuropeptide genes (apelin, cocaine-and amphetamine-regulated transcript protein, insulin-like 5, neuropeptide S, and neuropeptide B) previously unknown in chicken were identified and 62 genes were confirmed. Although most neuropeptide gene families known in human are present in chicken, there are several gene not present in the chicken. Conversely, several chicken neuropeptide genes are absent from mammalian species, including C-RF amide, c-type natriuretic peptide 1 precursor, and renal natriuretic peptide. The prohormone convertases, with one exception, were found in the chicken genome. Bioinformatic models used to predict prohormone cleavages confirm that the processing of prohormone proteins into neuropeptides is similar between species. Neuropeptide genes are most frequently expressed in the brain and head, followed by the ovary and small intestine. Microarray analyses revealed that the expression of adrenomedullin, chromogranin-A, augurin, neuromedin-U, platelet-derived growth factor A and D, proenkephalin, relaxin-3, prepronociceptin, and insulin-like growth factor I was most susceptible (P-value < 0.001) to changes in developmental stage, gender, and genetic line among other conditions studied. Our complete survey and characterization facilitates understanding of neuropeptides genes in the chicken, an animal of importance to biomedical and agricultural research.

§Corresponding author: rodrgzzs@illinois.edu, 1207 W Gregory Dr, Urbana, IL 61801, Phone 217-333-8810 Fax: 217-333-8286.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

Keywords

neuropeptide; prohormone; convertase; chicken genome; microarray experiment

INTRODUCTION

Neuropeptides encompass a wide range of small signaling peptides, such as neurotransmitters and peptide hormones that regulate many biological processes, including reproduction, development, growth, memory, feeding, and behavior (Hook, 2008). These important intercellular messengers derive from larger prohormone proteins via a complex series of post-translational cleavages, spearheaded by prohormone convertases (PCs) and other post-translational modifications, which challenge their detection solely based on sequence homology to other more extensively studied species (Fricker, 2005; Hook, 2008). The chicken was the first avian genome sequenced (International Chicken Genome Consortium, 2004) and thus lacks of closely related species with neuropeptide sequence information, although the song bird is currently being sequenced and annotated. The availability of the genome sequence allows one to uncover genes with limited or no empirical confirmation using bioinformatics tools, and the growing number of gene expression microarray experiments supports the functional annotation of these genes (Cogburn et al., 2003).

Although approximately 95 neuropeptide genes that code for prohormones have been identified in human and mammalian model organisms, only 65 of these genes have been reported or predicted from the chicken genome and, in addition, prohormone peptide YY has only been reported at the protein level. The incomplete status of the chicken neuropeptide and PC gene complement is a notable deficiency considering the well-recognized status of the chicken as a model organism in biomedical and agricultural research (Stern, 2005). The role and expression patterns of a small percentage of these neuropeptides have been explored in chicken. Insulin-like growth factor 1 (**IGF1**) has a role in chicken fetal growth, as well as axonal growth and myelination (Duclos, 2005); Bennett et al. (2006) identified polymorphisms in IGF1 and insulin (**INS**) associated with weight at 5 weeks and 55 weeks in a layer-broiler cross in chickens; Zhou et al. (2005) reported significant associations between IGF1 and bone size and strength in 8-week old female and male chickens. Vasoactive intestinal peptide (**VIP**) relaxes the smooth muscle of trachea, stomach, and gall bladder. Jozsa et al. (2006) demonstrated that the brain levels of VIP and pituitary adenylate cyclase-activating polypeptide (**PACA**) change in chicken and rats after food deprivation and concluded that the 2 peptides are differentially involved in feeding.

The public Gene Expression Omnibus (GEO) database contains multiple chicken microarray gene expression platforms, including many with more than 10,000 probes (e.g. GPL1731, GPL1461, GPL1836, GPL2719, GPL2863, GPL3213, GPL4993, GPL5618, GPL6049). Although some of these platforms include neuropeptide and PC gene probes, the incomplete knowledge of the chicken neuropeptide and PC gene complement has challenged the profiling of these genes. In addition, the ability of mass spectrometry experiments to detect and characterize neuropeptides is aided by the availability of accurate prohormone gene identification and annotation (Li and Sweedler, 2008).

The objective of this study was to obtain the first genome-wide census and functional annotation of the chicken neuropeptide and PC genes. First, an exhaustive master list of known neuropeptide and PC genes in the human and chicken was constructed. Second, the master list was searched against various complementary chicken genome databases. Third, neuropeptide and PC gene expressions were profiled using a database of approximate expression patterns inferred from EST sources and a set of 22 chicken microarray experiments. Lastly, cleavage

sites on the prohormone protein sequences were predicted and compared to known neuropeptide sequences and associated cleavages.

METHODS

Detection of Chicken Neuropeptide and Convertase Genes

A search for neuropeptide and PC genes across the chicken genome (1.1 Mb, including 30 microchromosomes and 9 macrosomes) was undertaken. A master list of candidate genes was generated based on known human and chicken prohormone gene sequences available in public databases and a literature review (Amare et al, 2006; Southey et al., 2008; Southey et al., 2009). The human sequences offer a good representation of the mammalian genes (Tegge et al., 2008) and were complemented with already known chicken sequences not detected in mammalian species. The candidates were first searched for among the chicken sequences already available in the GenBank (release 173.0, August 15 2009) and UniProt databases (release 15.8, September 22, 2009). To uncover chicken genes not previously reported or with different nomenclature from that of the master list, the human prohormone gene sequences were aligned against three resources stemming from the chicken genome build 2.1, the genome (**Genome**), the expressed sequence tag (**EST**), and the high throughput genome sequence (**HTGS**) databases available in NCBI. Sequence searches were implemented using the chicken NCBI BLAST website (http://www.ncbi.nlm.nih.gov/genome/seq/BlastGen/BlastGen.cgi?taxid=9031) with default parameters (BLOSUM62 scoring matrix and maximum E-value of 10) and no filtering of low complexity regions. To augment the likelihood of identifying functionally conserved homologues, the protein sequence was used as a query. The matches were screened based on the alignment E-value and distribution of the alignment identities, close matches, mismatches, and gaps along the sequence. The matches were also screened for alignments to related genes in the same neuropeptide family.

Characterization of the Neuropeptide and Convertase Gene Expression Profiles

The expression patterns of neuropeptide and PC genes were obtained from two resources. One resource was the UniGene database (build #41) which includes the expression of chicken neuropeptide EST across tissues and maturation stages. The other resource was the GEO database which encompasses gene expression experiments that used chicken microarray platforms and included probes for neuropeptide and PC genes. Affymetrix Chicken Genome Array GPL3213 (http://www.affymetrix.com/support/technical/byproduct.affx?product=chicken) was selected among the chicken microarray platforms because it had the highest number of relevant probes, including 53 neuropeptide and 5 PC genes transcripts. In addition, the GPL3213 platform had the highest number of gene expression experiments (22) of all chicken platforms. This unique abundance enabled a comprehensive analysis of all the experiments and the identification neuropeptide and PC gene expression patterns across a wide range of conditions. The studies were grouped into 6 classes: retina, heart and breast muscle, brain and head, liver and duodenum, oocyte and gonad, and other tissues; the number of studies (and GEO series identification) within each class was 4 (GSE6543, GSE7176, GSE11439, and GSE15382), 4 (GSE6843, GSE8693, GSE9251, and GSE15413), 4 (GSE6844, GSE6868, GSE8693, GSE12268), 4 (GSE6856, GSE8483, GSE15413-liver, GSE15413-duodenum), 2 (GSE8693, GSE10231), and 5 (GSE8010, GSE8016, GSE8018, GSE8483, GSE9884), respectively. Experiments GSE8693 (Ellegren et al., 2007) and GSE15413 included comparisons of conditions across multiple tissues, and the samples corresponding to each tissue were analyzed separately to facilitate the interpretation of results.

Pre-processing and normalization of the microarray data was done using the Affy R package (Irizarry et al., 2009) and included the $\log_2$ transformation of the intensities and GC-robust multichip average normalization of expression measurements. The expression measurements of all the probes in the platform were analyzed, and the statistical significance of the differential expression was adjusted for multiple testing across all probes using the false discovery rate approach (Benjamini and Hochberg, 1995). The microarray analyses were done using Beehive (http://stagbeetle.animal.uiuc.edu/Beehive).

Prediction of Cleavage Sites

Several models have been proposed to predict the cleavage of prohormone proteins coded by neuropeptide genes (Hummon et al., 2005; Southey et al., 2006b; Tegge et al., 2008; Southey et al. 2008; Southey et al. 2009). However, no cleavage model has been trained on avian species. The accuracy to predict avian cleavage sites of the "known motif" model (Southey et al., 2006b) and the logistic regression model trained on human sequences (Tegge et al., 2008) was evaluated using the 25 chicken prohormone sequences that have peptide information (and in most cases with signal peptide information) available in UniProt. Both models are available at NeuroPred (http://neuroproteomics.scs.uiuc.edu/neuropred.html; Southey et al., 2006a).

RESULTS AND DISCUSSION

Chicken Neuropeptide Genes

A master list of 95 neuropeptide genes and 7 PC genes were identified from the literature review and the Gene, UniGene, and UniProt databases. Table 1 summarizes the distribution of the genes on the master list across the 3 databases used to compile already known chicken genes (Gene, UniProt, UniGene) and the 3 databases used to uncover previously unreported chicken genes (Genome, HTGS and EST databases).

A total of 62 chicken neuropeptide genes were present in the Gene database and among them, 49 had the corresponding complete or partial prohormone sequence in UniProt. This count includes augurin or esophageal cancer-related gene 4 protein (**ECRG4**); although not currently present in the Gene database, a region in a genomic contig on chromosome 1 (ref|NW_001471545.1|Gga1_WGA43_2) had been assigned as being similar to ECRG4. Further, this count excludes the peptide YY-like (**PYY**-like) gene because although this peptide is reported in UniProt (P29203), no evidence for the corresponding gene sequence was found in any of the NCBI databases. The inability to confirm the UniProt PYY-like entry in other databases prompted us to remove this peptide from known chicken peptides. The proportion of prohormones coded by neuropeptide genes in UniProt that had empirical evidence at the protein level was 0.34 (17/50) and the remaining prohormones had evidence at the transcript level, or were based on sequence similarity or predictions (Table 1). Of the 62 neuropeptide genes in the Gene database, 59 had a corresponding record in UniGene. The absence of UniGene entries for motilin (**MOTI**), oxytocin (**NEU1**), and orexigenic neuropeptide QRFP (OX26) reflects the lack of ESTs reported for these neuropeptide genes.

Of the 62 neuropeptide gene sequences in UniGene, 8 were not located in the chicken Genome, HTGS, or EST databases. Gastrin (**GAST**) and PACA were not located in Genome, HTGS, and EST databases, meanwhile C-type natriuretic peptide (**ANFC**), chromogranin (**CMGA**), prolactin-releasing peptide (**PRRP**), parathyroid hormone-related protein (**PTHR**), neuropeptide VF precursor (**RFRP**), and secretogranin-1 (**SCG1**) were not located in the HTGS database. UniProt does not have records forGAST, CMGA, and SCG1that would support the corresponding UniGene records. A likely explanation for the neuropeptide genes not located in either of the three databases is that the Genome assembly and EST libraries may be incomplete at the locations of these genes.

Chicken Prohormone Convertase Genes

Of the 7 PC enzymes in the master list, only PC 4 (**PCSK4**) was not reported in the Gene database and only 2 PC genes were reported in the UniProt database (Table 1). The 6 chicken PC were confirmed in the chicken Genome, HTGS, and EST databases. Five PC genes in the Gene database had a corresponding record in UniGene. In addition to the UniGene partial record for PC 1 (**PCSK1**, Gga.9357), UniGene has a record (Gga.31439) predicted from the genome that is annotated to be similar to PC 1, although there is no corresponding Gene record for this UniGene record. The alignment between these two sequences has an *E-value* of 9.-× 10^{-154} and 94% identity. Our survey will support the currently limited work on PC in chicken. Ling et al. (2004) detected PC1 and PC2 mRNA in multiple chicken tissues including heart, lung, gizzard, pancreas, spleen, bursa of Fabricius, kidney, adipose tissue, skeletal muscle, pituitary gland, cerebrum, mid-brain and cerebullum and Richards and McMurtry (2008) reported that PC2 mRNA was mostly expressed in pancreas and proventriculus, whereas PC1 mRNA was more expressed in duodenum and brain of chicken.

There was no suitable match for PCSK4 in the chicken genome suggesting that PCSK4 may have evolved after the split between chickens and mammals. The absence of evidence for PCSK4 is noteworthy because this convertase plays an essential role in the process of fertilization and is located in testicular germ cells in mice (Gyamera-Acheampong et al., 2006). The differences in the reproductive biology of mammals and chicken and the overlap of processing of some convertases (Baea et al., 2008) bar the conclusion that the absence of PCSK4 may hinder any particular prohormone cleavage or presence of a particular neuropeptide in chicken.

Previously Unreported Genes Detected in Chicken

The bioinformatics approach uncovered evidence for 5 neuropeptide genes that have not been previously reported. Confirmatory evidence in the Genome, HTGS, and EST databases that was further validated using complementary resources supports the discovery of evidence for the chicken homologues to the neuropeptide genes apelin (**APEL**), cocaine- and amphetamine-regulated transcript protein (**CART**), insulin-like 5 (**INSL5**), neuropeptide S (**NPS**), and neuropeptide B (**NPB**). The multi-step strategy varied across sequences and depended on the strength of the evidence, quality of the genome sequence, and availability of homologue sequences to confirm these findings. In the first step, 2 criteria were used to rule the finding of a previously unreported gene in the 3 databases: a low *E-value* of the sequence alignment (encompassing a high percentage of identities and similarities with a minimum percentage of mismatches and gaps), and conservation of the region encompassing the gene product known in human were required. The additional confirmatory steps unique to each gene and a brief review of the implications of our findings in the understanding of chicken physiology, production, and health follows. Supplementary materials Table S1 presents a detailed description of sequence alignments for each of the 5 neuropeptide genes from the 3 databases, and underlined is the region corresponding to the functional neuropeptide known in humans.

APEL—The region of the chicken genome that matched the human APEL sequence (NP_059109) contained many gaps that prevented complete identification of chicken APEL from the genome. An EST (BU323997) that have a good match (*E-value* = 0.034) to the human sequence was identified but the resulting BLAST alignment between the human sequence and BU323997 had 2 sections indicating a probably frame-shift in the EST sequence. The alignment of BU323997 against a region of the chicken genome on chromosome 4 indicates that there is an extra nucleotide ‘G’ in the EST that is not present in the genome sequence (Table S1a). After removing this extra ‘G’ and non-coding sequence from the chicken EST, a chicken APEL sequence was predicted. The trace archives (NCBI Trace-Other database) were searched using the putative nucleic APEL sequence to overcome limitations of the genome assembly. There

are only 2 matches to this putative sequence which correspond to 2 exons as expected from the genomic structure of the human APEL gene. The detection of APEL in the chicken is significant because this neuropeptide has been implicated in a variety of roles, including cardiac function, drinking behavior, regulation of adiposity, lipid and energy metabolism, and gastric cell proliferation in humans (Higuchi et al., 2007).

CART—There was no clear and prolonged match to the human CART sequence in the chicken genome, suggesting that the CART gene region was not fully assembled in the chicken. However, following the strategy used for APEL, an EST (BM490862) with a good match (*Evalue* = 7×10^{-32}) to the human CART sequence was identified. Considering that the human sequence matched the third reading frame of the EST and that the first 2 nucleotides of the EST were 'TG', we hypothesized that the EST sequence could be missing the initial 'A' nucleotide that would code for the first methionine amino acid in the CART protein. The start of the sequence was confirmed using the chicken trace archives (Trace-Others) and visualization of the matches showed that the middle of CART is missing in the genome sequences. The chicken EST mapped to a genomic contig yet to be located on the chicken genome (NW_001476554.1|GgaUn_WGA14361_2). The significance of the discovery of CART in the chicken is highlighted by its known role in reward, feeding, changes in body weight and fat mass, and stress (Asnicar et al., 2001; Kuhar et al., 2002).

INSL5—The search for human INSL5 in the chicken genome results in a match (*E-value* = 0.11) on chromosome 8 that was no attributable to members of the relaxin or insulin gene families. The NCBI annotation to this genomic region is "similar to WD repeat domain 78" (WDR78; Gene identifier LOC429114). The Map View feature in NCBI indicated that the chicken LOC424701 and TCTEX1D1 Tctex1 domain containing 1 genes are adjacent to this gene which is remarkable because human INSL5 is located on the negative strand between these 2 genes. Although the human INSL5 match was insufficient to identify a chicken INSL5 gene, a putative sequence was obtained using 3 fish sequences in the relaxin gene family; *Danio rerio* (Zebrafish, B1AAQ6_DANRE), *Fugu rubripes* (Japanese pufferfish, B1AAR5_FUGRU), and *Tetraodon nigroviridis* (Green puffer, B1AAR6_TETNG). The *Danio rerio* sequence is shorter than the other 2 but the overlap is extensive. The alignment of the longer sequences to the chicken genome produced 2 very good alignments (*E-values* = 5×10^{-20} and 2×10^{-16}, respectively) that are 1210 bp apart (start of both chicken genome matches are 17631700 and 17630490, respectively). Both matching regions are annotated in the genome region as WDR78,further confirming that the inaccurate annotation of the chicken WDR78 gene prevented the annotation of the chicken INSL5 gene. The identification of INSL5 in chicken is notable because it has been postulated that INSL5 plays a role in gut contractility, remodeling and repair of the gastrointestinal tract and neuroendocrine signaling (Conklin et al., 1999; Dun et al., 2006; Haugaard-Jönsson et al., 2009).

NPS—A match (*E-value* = 2×10^{-15}) to human NPS was identified on chicken chromosome 6 (Table S1b). The chicken genomic region that matched the human NPS sequence (± 5000 bp) was extracted, and the Wise2 version 2.1.20 software (http://www.ebi.ac.uk/Tools/Wise2/index.html; Birney et al., 2004) was used to predict the coded protein on the extracted region. The predicted chicken protein matched the complete human and other mammalian NPS sequences present in UniProt (Table S1b). The uncovering of NPS in the chicken genome is noteworthy because of its role in behavior (e.g. anxiolytic action, hyperlocomotion, wakefulness, altered sleep behavior, panic disorder), and intake (Castro et al., 2009; Pape et al., 2009)

NPB—The search for human NPB in the chicken genome matched a "hypothetical protein" (Gene database identifier 769277 LOC769277) on chromosome 18 (*E-value* = 2 ×

10^{-6}). The genome region corresponding to actual bioactive NPB peptide was conserved in the alignment. Furthermore, the UniGene entry associated with this hypothetical protein clusters it with "anaphase promoting complex subunit 11" representatives from other species including human, mouse, and zebra fish. The parsing of the chicken nucleic genomic region of the human match using Wise2 uncovered one gene with 2 exons. Exon 1 includes the neuropeptide and is well conserved with the human NPB gene that also has 2 exons. Similar analyses using the known NPB sequences of the zebra fish and salmon offered results consistent with those obtained using the human sequence. The alignment of the NPB prohormone sequences predicted from the chicken Trace-Other archives includes gaps indicating that incomplete genome assembly prevented the annotation of the chicken NPB gene. Our discovery of NPB in chicken is significant because the NPB/NPW neuropeptide system regulates energy homeostasis, pain, and emotion, and NPB exerts strong synergistic anorectic effects in mice when co-administered with CRF (Hondo et al., 2008; Aikawa et al., 2008)

Neuropeptide Genes Not Located In Either Chicken or Mammals

There was insufficient evidence to locate 26 neuropeptide genes from the master list in the chicken genome: intermedin (**ADM2**), natriuretic peptide B (**ANFB**), calcitonin-related polypeptide beta (**CALCB**), cortistatin (**CORT**), galanin-like peptide (**GALP**), hepcidin (**HEPC**), insulin-like 3 and 6 (**INSL3** and **INSL6**, respectively), metastasis-suppressor KiSS-1 (**KISS1**), neuromedin S (**NMS**), neuropeptide FF (**NPFF**), neuropeptide W (**NPW**), proprotein convertase subtilisin/kexin type 1 inhibitor (**PCSK1N**), proenkephalin B (**PDYN**), putative peptide YY-2 (**PYY2**), pro-relaxin 1 and 2 **(REL1** and **REL2**, respectively), regulated endocrine-specific protein 18 (**RES18**), spexin (**SPXN**), tachykinin 4 (**TAC4**), parathyroid hormone 2 (**TIP39**), tachykinin 3 (**TKNK**), torsin family 2 member A isoform prosalusin (**TOR2X**), urocortin (**UCN1**), and urocortin 2 (**UCN2**). Although there is a Gene and UniGene entry for chicken torsin family 2 member A (XP_415507, Gga.5228), the chicken sequence corresponds to a human torsin alternative splicing isoform (**TOR2A**) that does not code for the neuropeptide salusin and thus was considered not located in the chicken genome. In contrast, 3 neuropeptide genes in the master list (C-RF amide, c-type natriuretic peptide 1 precursor, and renal natriuretic peptide) were present in chicken but were not located in mammalian species.

A remarkable finding is that the vast majority of the mammalian genes not located in the chicken genome have at least 1 neuropeptide gene in the same family present in the chicken genome (Table 1). For example UCN1 and UCN2 are absent from the chicken genome, while urocortin 3 (**UCN3**) is present in the chicken genome. The exceptions to the presence of at least 1 member of the neuropeptide family in the chicken genome are CORT, HEPC, KISS1, NPFF, NPW, RES18, and SPXN. Burt (2007) noted the low number of genes identified in the chicken genome relative to the human genome and hypothesized that there were more duplication events in the mammalian linage of some genes and more losses of other genes in the avian lineage. Our results reinforce this hypothesis that some neuropeptide genes have undergone substantially lower gene duplication in the chicken compared to mammals.

Neuropeptide Gene Expression Across Tissues and Developmental Stages

The distribution of the expression of most neuropeptide and PC genes available in UniGene was used to gain an initial understanding of the expression profiles across tissues and developmental stages. All PC and neuropeptide genes in UniGene, with the exceptions of c-type natriuretic peptide 1 (**CNP1**), CRF, GAST, progonadoliberin 1 (**GON1**), pancreatic polypeptide (**PAHO**), parathyroid hormone (**PTHY**), prothyroliberin (**TRH**), thyroid stimulating hormone subunit beta (**TSHB**), and urotensin 2 (**UTS2**), had expression information. For PCSK1, the corresponding UniGene record (Gga.9357) did not have expression information, but another UniGene record (Gga.31439), annotated as "similar to

PC1", was used as proxy because of the availability of expression information and similarity to the PCSK1 sequence. Table 2 provides a summary of the expression profile of 51 neuropeptide and 6 PC genes across the 5 tissues and 2 stages with most frequent neuropeptide gene expression out of 19 tissues or body parts and 4 development or maturation stages. Supplementary materials Table S2 presents the distribution of expression across all 19 tissues and 4 stages.

The tissue or body part with highest number of neuropeptide gene expression reports (expressed in absolute number and percentage) was the brain (33, 64.71%), followed by head (21, 41.18%), ovary (18, 35.29%), small intestine (16, 31.37%), and heart (13, 25.49%). A similar distribution was observed for the PCs, with the brain and small intestine being the body parts with highest frequency of gene expression. These results are consistent with the role of neuropeptides in physiology, health, and behavior (Hook, 2008). The developmental-maturation stage with the most reports of neuropeptide gene expression was adult (42, 82.35%), followed by embryo (39, 76.47%). The neuropeptide genes with the highest number of reports of expression across tissues or body parts were platelet-derived growth factor alpha polypeptide (**PDGFA**, 11, 57.89%), SCG1 (11, 57.89%), and ECRG4 (9, 47.37%). These results are consistent with neuropeptide research across species. For example, PDGFA is expressed in the seminiferous epithelium and interstitial mesenchymal cells, and studies with mice show it may play a role in cell proliferation, migration in osteoblastic cells, and in production of Leydig cells (Yang et al., 2008). Likewise, the tyrosine-sulfated secretory protein SCG1, found in a wide variety of peptidergic endocrine cells in mice, may play a role in the early phase of neoplastic progression (Lukinius et al., 2003). Also, the expression of ECRG4 in multiple tissues including the heart, brain, placenta, lung, liver, skeletal muscle, kidney, and pancreas suggests a role in the modulation of salt and energy homeostasis, cardiovascular function, and cerebral spinal fluid composition (Mirabeau et al., 2007; Mori et al., 2007).

Although the distribution of expression reports can be influenced by the imbalanced distribution of EST libraries and experimental interest across tissues, developmental stages, and neuropeptide genes, all chicken tissues and developmental stages had at least 1 neuropeptide gene with a UniGene expression report. This confirms the importance of neuropeptides on all aspects of chicken physiology, growth, reproduction, and health.

Expression Profiling Based On 22 Microarray Experiments

Although the information in UniGene offers a broad picture of the expression of neuropeptide and PC genes across a wide range of tissues and stages, additional conditions can influence the expression profile. To fully investigate the variation in neuropeptide and PC gene expression across a wide range of conditions and augment the understanding of the impact of these genes on reproduction, health, growth, and other traits of importance to biomedical research and agricultural production, the information from a large number of microarray gene expression experiments investigating numerous conditions was mined.

We present results from the first simultaneous analysis of 22 microarray experiments to characterize the expression of neuropeptide genes and PC across a wide range of conditions. The in-situ synthesized microarray platform selected has the highest representation of chicken neuropeptide and PC genes available in GEO and is most widely used. A total of 73 probes representing 53 neuropeptide and 5 PC genes was available in the platform. The experiments were broadly grouped into retina, heart and breast muscle, brain and hypothalamus, liver and duodenum, gonad and oocyte, chicken-quail comparisons, and other conditions. To facilitate the interpretation of the results, a summary of the experiments and their features (e.g., tissues, treatments, age, gender, genetic line) is presented in Table 3. More detailed descriptions of the experiments are available in supplementary material Table S3 and in the GEO database. Due to the multiple probes analyzed, a minimum false discovery rate multiple-test adjusted *P*-

value < 0.05 threshold (corresponds to an approximate unadjusted *P-value* < 0.005) and a minimum fold-change equal to 1.25 was used to identify differentially expressed genes. Table 4 summarizes the number of probes with differential expression across experimental group and probes corresponding to the same gene. Supplementary materials Table S4 presents the detailed distribution of the differential expression level of each probe and experiment. A summary of the main findings by tissue group are described below.

Retina—The chicken is a well-established model for the human eye and retinal development and degeneration (Burt et al., 2007). Two independent microarray experiments GSE6543 (McGlinn et al., 2007) and GSE11439 (Schippert et al., 2008) investigated the effect of myopia and the lens in the retina in chicken, respectively. The neurotensin (**NEUT**) gene was significantly over-expressed in the treated samples relative to the control samples in both studies. In addition, glucagon (**GLUC**) and NEUT were over-expressed in the treated samples relative to the control in the GSE11439 experiment, and VIP was under-expressed in the treated samples relative to the control in the GSE6543 experiment. These findings confirm the important role of neuropeptides in vision in the chicken (Schwippert et al., 1998; Chapman and Debski, 1995). Likewise, from experiment GSE7176 (Rizzolo et al., 2007) that studied the chicken retina across embryo developmental stages, we uncovered that the expression of gene adrenomedullin (**ADML**) was significantly lower in embryonic day 7, or E7 (*P-value* < 1×10^{-6} and 0.29 average fold change) relative to more advanced ages (E10, E14, and E18). Both the ADML peptide and its mRNA have been detected in embryonic mice in the outer neuroblastic layer of the retina (Montuenga et al., 1997) and in the human retinal pigment epithelial cells, suggesting an important physiological role for ADML in eye development (Udono et al., 2000). On the other hand, the variation in the expression of VEGFC across developmental stages was significant (*P-value* < 3.0×10^{-4}) but higher at E7 relative to E10, E14, and E18 (2.24 average fold change). Neuropeptide VEGFC is expressed in the retinal astrocytes and promotes both endothelial cell proliferation and migration (Alon et al., 1995; Stone et al., 1995; Pierce et al., 1996; Provis et al., 1997). Likewise, the expression had a significant fluctuation across stages (*P-value* < 1.0×10^{-16}) with the expression at E7 higher relative to E10, E14, and E18 (3.44 average fold change). This profile is consistent with studies in mice that have shown that the PDGFA receptor, which is activated mainly by PDGFA located in retinal neurons, is expressed at all stages of maturation and is important for retinal astrocyte proliferation and migration (Mudhar et al., 1993; Fruttiger et al., 1996). Over-expression of PDGFA in transgenic mice causes a significant increase in retinal astrocytes, resulting in proportional overgrowth of the retinal vasculature (Fruttiger et al., 1996).

Experiment GSE15382 (Kubo et al., 2009) aimed to investigate the potential impact of c-hairy 1, a gene that inhibits neuronal differentiation on gene expression. Our analysis uncovered that samples with this gene exhibited over-expression of neuropeptide Y (**NPY**) (*P-value* < 3.0×10^{-2}, 11.31 fold change) and to a lesser extent, secretogranin 2 (**SCG2**, *P-value* < 2.0×10^{-2}, 1.13 fold change), relative to control samples. In addition, the differential expression between c-hairy1 and control was similar to that between c-hairy1 and Delta and Wnt2b, 2 other genes expected to also inhibit neuronal differentiation. Thus, c-hairy1 had a strong association with the expression of 2 neuropeptide genes that is not observed with the other 2 potential inhibitor genes. These results are consistent with reports that NPY, along with its receptors, are present in the retina of both mammalian and non-mammalian species (D'Angelo et al., 2004) and that this neuropeptide may modulate the development of retinal circuitry in rats (Bagnoli et al., 2003). Recent experiments with rats have shown that NPY produces a 2-fold increase in retinal neural cell proliferation and promotes the proliferation of committed neural immature cells (Álvaro et al., 2008).

Heart and Breast Muscle—Neuropeptides can contract muscles, and the action of a neuropeptide is of significance in the control of antagonistic contractions (Cho et al. 1996).

Experiments GSE6843 (Itoh et al., 2007) and GSE8693 investigated gene expression in embryonic heart tissue, and our analyses did not detect differential expression among the neuropeptide genes studied between females and males, suggesting that the role of the neuropeptides may be equally important in both genders. On the other hand, neuropeptide gene expression can exhibit significant variation across breast muscle at different development stages. For example, the analysis of experiment GSE15413 uncovered numerous neuropeptide genes that were differentially expressed in the breast muscle between 7-d-old and just hatched (0-d-old) chickens. Specifically, prepronociceptin (**PNOC**, 2 probes), PDGFA, platelet-derived growth factor beta polypeptide (**PDGFB**), platelet derived growth factor D (**PDGFD**), ADML, and ECRG4 were over expressed in 0-d-old relative to 7-d-old chickens (Table 4, Supplementary materials Table S4). These findings are consistent with studies that report platelet-derived growth factors are important in avian embryonic development (Van Den Akker et al., 2005). Conversely, proenkephalin (**PENK**) and IGF1 were over-expressed in 7-d-old relative to 0-d-old chickens. No neuropeptide gene was differentially expressed between male and female embryo heart samples at 18-d-old in GSE8693 or between male and female embryos at late stages of development in GSE6843.

Experiment GSE9251 (Zheng et al., 2009) profiled the expression of genes in the breast muscle of broiler and layer genetic lines at different ages. Our analysis found that PNOC appears to have a quadratic expression pattern, regardless of genetic line, because it is under-expressed in both broiler and layer at young (1-d-old) and old (6-wk-old to 8-wk-old) ages relative to intermediate ages (2-wk-old to 4-wk-old). Both lines had similar levels of fluctuation across ages (approximate *P-value* < 0.0005 and approximate maximum fold change 2.22). This result is consistent with the role of nociceptin in stimulating locomotion (Florin et al., 1997). The PNOC gene is highly conserved within the mouse, rat, and human, and studies have shown that it is broadly expressed in the nervous system, primarily in the brain and spinal cord (Mollereau et al., 1996). For PRRP, SCG2, ANFC, musclin (**OSTN**), neuromedin-U (**NMU**), TSHB, TRH, somatostatin (**SMS**), islet amyloid polypeptide (**IAAP**), gastric inhibitory polypeptide (**GIP**), and tachykinin, precursor 1 (**TKN1**), the expression at 1-d-old was lower than at older ages in both genetic lines, and the fold change did not differ significantly between lines (maximum $P\text{-}value = 8.4 \times 10^{-3}$). The ECRG4 gene was highly differentially expressed ($P\text{-}value < 1.0 \times 10^{-16}$) and had the highest expression in 1-d-old broiler and layers with both lines showing similar fold changes. Gene ADML exhibited differential expression ($P\text{-}value < 1.0 \times 10^{-16}$) with the level at 1-d-old being higher than at older ages in broilers; meanwhile for layers, ADML is under-expressed in 1-d-old chickens relative to older ages, and the level is significantly different between the lines in 1 d-old chickens. The difference in expression between genetic lines may be associated with the angiogenic role of ADML, albeit weaker in chicken than in human and mouse (Martínez et al., 2006). The expression of IGF1 varied significantly across ages and lines ($P\text{-}value < 2.4 \times 10^{-13}$) with 2-wk-old chickens having the highest expression among layers, with expression not varying across ages within broilers not showing a clear pattern significant trend or deviation at a particular time point, and the maximum difference between lines (2.15 fold change over-expression in layers relative to broilers) was observed at 2-wk of age. For the IGF2 probe set, the minimum expression for both genetic lines was in 1-d-old chickens ($P\text{-}value < 3.9 \times 10^{-3}$) and the lines did not differ in the level of expression at that age. This result does not support the hypothesis postulated by Wang et al. (2005) that IGF2 can be a candidate gene influencing growth and carcass traits, although their work only used broilers. The differential expression ($P\text{-}value < 8.4 \times 10^{-5}$) of the vascular endothelial growth factor C (**VEGFC**) gene exhibited the maximum at 2-wk-old across lines. The differential expression ($P\text{-}value < 1.0 \times 10^{-16}$) of vascular endothelial growth factor D (**VEGFD**) encompassed the minimum and maximum expression in 1-d-old and 2-wk-old broilers, respectively. The level of expression in broilers is significantly lower than layers at 1-d-old ($P\text{-}value < 1.0 \times 10^{-16}$, 2.05 fold change) but similar by 2-wk-old. This result is consistent with work that demonstrated the presence of lymphatic capillaries throughout

VEGFC and VEGFD in the muscles of humans and mice (Kivelä et al., 2007) and the role of VEGFC lymphatic regeneration in tissue repair of the intestinal muscle coat (Shimoda and Kato, 2006). For the PENK gene, a linear trend of expression can be identified, with lower ages having significantly higher expression (*P-value* < 1.0×10^{-6}) than advanced ages for the broiler line and at lower significance levels for the layer line. The level of expression in broilers was significantly higher than layers at 1-d-old (*P-value* < 4.0×10^{-4}, 1.61 fold change) but similar by 2-wk-old. This result is consistent with reports that during development, PENK mRNA is abundant in skeletal muscle, bone, and intestine and that the levels of PENK mRNA tend to decrease as mice mature, with the exception of the brain, gut, lungs, and heart (Prasad et al., 2008).

Brain and Hypothalamus—The limited number of neuropeptide genes differentially expressed in the brains of chickens under different conditions compared to other tissues and body parts was an unexpected finding. This result may be because the conditions compared within these studies did not allow the detection of major and consistent differential expression in the samples available. Study GSE12268 compared the brain of male and female embryos (6.5-d of incubation), and the present analysis determined that the pro-melanin-concentrating hormone (**MCH**) gene was over-expressed; meanwhile, ECRG4 and GAST were borderline over-expressed in males (*P-value* < 2.5×10^{-2}, *P-value* < 9.5×10^{-2} and *P-value* < 9.5×10^{-2}, respectively). These results support the hypothesis that MCH acts as a neuromodulator involved in a wide variety of physiological and behavioral adaptations (arousal) with regard to feeding, drinking, and reproduction in birds (Cardot et al., 1999).

Results from the analysis of the other study (GSE6844; Itoh et al., 2007) of embryo female and male brains identified REL3 as being differentially over-expressed in females (*P-value* < 9.7×10^{-3}); meanwhile SCG1 and VIP were over-expressed in males (*P-value* < 3.3×10^{-2} and *P-value* < 2.1×10^{-2}). Also, study GSE8693 identified REL3 as being over-expressed in the brain of females (*P-value* < 9.7×10^{-3}) relative to males. The profile of REL3 was expected because the relaxin hormone is renowned for its function in pregnancy, parturition, and other aspects of female reproduction (Agoulink, 2007). Relaxin-3 is a hypothalamic neuropeptide expressed in the nucleus incertus of the brainstem and plays a role in energy homeostasis (Tanaka et al., 2005; McGowan et al., 2009). Lastly, analysis of study GSE6868 (Rosenquist et al., 2007) that compared treated (homocysteine congenital defect cell culture) versus control neural crest samples did not uncover differentially expressed neuropeptide genes. The lack of differential expression of the ADML gene in the brain is consistent with a report that the levels of ADML protein are almost undetectable in the chicken brain (Zudaire et al., 2005).

Liver and Duodenum—In the analysis of gene expression obtained in the study GSE15413, which compared the liver of chicken at two ages, only IGF1 was found over-expressed in 7-d-old relative to 0-d-old chickens (*P-value* < 2×10^{-4}, fold change 12.47) followed by prokineticin 2 (**PROK2**), which had the same profile but was only moderately differentially expressed. These results are consistent with reports that in avian species, IGF1 mRNA is found in the liver, muscle, kidney, testes, heart, ovary, brain, and intestine and that the metabolic effects of IGF include increased amino acid and glucose uptake (Amills et al., 2003). Also, Wang et al. (2007) reported that the hepatic expression of IGF1 was altered by hypothyroidism in chicken. As with heart and muscle, no neuropeptide gene differential expression was observed between the liver of female and male chicken embryos (study GSE6856; Itoh et al., 2007). Analysis of the GSE8016 study (Nakao et al., 2008) of liver from chicken and quail samples uncovered that PNOC, ADML, REL3, NMU, NPY, and ECRG4 were significantly under-expressed in the liver of the chicken relative to the quail (maximum *P-value* < 7.3×10^{-3}).

From the analysis of study GSE15413, multiple neuropeptide genes were differentially expressed in the duodenum of newly hatched (0-d-old) relative to 7-d-old broiler chickens. The genes PENK and INS were significantly over-expressed in 0-d-old chickens (*P-value* < 9.0×10^{-4}, 6.23 fold change, and *P-value* < 3.5×10^{-3}, 14.72 fold change, respectively), and VEGFD was borderline over-expressed in 0-d-old chickens (*P-value* < 7.8×10^{-2}). The former finding is in agreement with studies showing that PENK mRNA is expressed in the human esophagus, gastrointestinal tract, pancreas, and gallbladder (Monstein et al., 2006) and enkephalins, such as PENK, have potent effects on gastrointestinal function, such as motility (Edin et al., 1980; Bitar and Makhlouf, 1982; Reynolds et al., 1984), intestinal secretion (Dobbins et al., 1980; McKay et al., 1981; Powell, 1981), and gastric acid secretion (Konturek et al., 1980; Feldman et al., 1980).

Oocyte and Gonads—The analysis of gene expression in oocytes across and within genetic lines from study GSE10231 uncovered that UTS2 was over-expressed in the genetic line with the long fertile period (DPF+) compared to the short fertile period (DPF−), high-growth (HG+), and non-high-growth (HG−) genetic lines. This outcome is consistent with the role of UTS2 in reproduction that is associated with its spasmogenic activity and in humans, UTS2 has been linked to preeclampisa-eclampsia (Balat et al., 2005). Also, and as expected, there were numerous neuropeptide genes differentially expressed between female and male gonads based on the profiles obtained from study GSE8693. Neuropeptide genes C-RF, SCG2, TKN1, PDGFA, PAHO, IGF1, and PDGFD were over-expressed in males relative to females (maximum *P-value* < 4.0×10^{-2} and 1.29 average fold change), and NEU2 was borderline under-expressed (*P-value* < 6.0×10^{-2}). The neuropeptide genes with over-expression in females relative to males were ADML (*P-value* < 3.5×10^{-2} and 2.98 fold change) and GLUC (*P-value* < 1.3×10^{-2} and 2.84 fold change).

Other Tissues—The analysis of the gene expression information from study GSE8018 (Nakao et al., 2008), which investigated the effect of day length on quail using the chicken microarray platform, identified numerous differentially expressed neuropeptides with unique profiles. The expression of REL3 did not vary within the long-day cycle but was higher at 18 hr in the short-day cycle relative to the other time points in the same cycle (overall time-by-day cycle interaction *P-value* < 2.8×10^{-2}). The expression of TSHB and cholecystokinin (**CCKN**) did not vary across the day within the day cycle but each was higher in the long-day cycle relative to the short-day cycle at every sampled time point (overall time-by-day cycle interaction *P-value* < 5.8×10^{-6} and *P-value* < 1.4×10^{-4}, respectively). There was no differential expression among the neuropeptide genes considered in the studies that compared neural crest cells treated with homocysteine versus control (study GSE6868), adipose tissue from lean and fat genetic lines (study GSE8010, Wang et al., 2007), circulating red and non-red blood cells (study GSE9884, McIntyre et al., 2008), and histone H1 variants (study GSE8483, Takami et al., 1997).

Prohormone Cleavage Prediction

An outcome of the comprehensive survey of neuropeptide gene sequences is the ability to predict previously unidentified biologically active neuropeptides that can used in high throughput experiments such as proteomic mass spectrometry experiments. We undertook the first prediction of cleavage sites in chicken prohormone sequences to gain insight into neuropeptide processing in avian species. The cleavage sites of all 24 chicken prohormone sequences, with empirically confirmed sequences and known or predicted neuropeptides available in UniProt (summarized in Table 1), were predicted using the empirically derived known motif and the human logistic regression cleavage models. The predictions were evaluated against the neuropeptides reported in UniProt. The comparison of the cleavage

prediction models allows to assess the performance of the human cleavage model to predict avian cleavage sites.

The number of true positives (correctly predicted cleaved sites), true negatives (correctly predicted non-cleaved sites), false positives (incorrectly predicted cleaved sites), and false negatives (incorrectly predicted non-cleaved sites) obtained by the known motif and human models, respectively, were 36, 2811, 75, 13 and 35, 2851, 35, 14. The calculation of sensitivity, specificity, and correct classification rate obtained by the known motif and human models, respectively, were 73.5, 97.4, 97.0 and 71.4, 98.8, and 98.3%. Model performance by individual neuropeptide prohormone is presented in Table 5. Overall, the sensitivity and the specificity of both models to predict cleavage was high, especially considering that neither model was trained using chicken sequences. The highest number of correctly predicted cleavage sites was identified by the known motif model, and the highest number of correctly predicted non-cleavage sites was identified by the human model. Thus, the sensitivity (percentage of all cleaved sites that were correctly predicted) is higher in the known motif model and the specificity (percentage of all non-cleaved sites that were correctly predicted) is higher in the human model. Although the human model correctly predicted 40 additional non-cleaved sites and one less cleaved site, the overall difference in correct classification rate was minor (1.3%); this is because of the relatively higher number of non-cleaved sites than cleaved sites, which results in more weight for the correctly predicted cleaved sites relative to the correctly predicted non-cleaved sites. The previous results confirm that the processing of prohormone proteins into neuropeptides is similar between chicken and human species. Until there are more empirically confirmed neuropeptides to train and validate an avian model, the known motif and human models offer a good solution for predicting prohormone cleavages and determining the resulting neuropeptides in the chicken.

CONCLUSIONS

The role of neuropeptides on reproduction, development, growth, and health has been widely recognized. However, a comprehensive study of the representation and expression of neuropeptide genes in chicken has never been undertaken. In this study, the first survey of neuropeptide genes, prohormone sequences, and prohormone convertase enzyme genes in chicken was completed. The integration of multiple bioinformatic resources allowed us to uncover evidence supporting 5 new neuropeptide genes, in addition to the 62 previously reported in the chicken genome. Among these chicken neuropeptide genes, 3 genes are not present in Eutherian mammals. There was insufficient evidence to detect in the chicken genome, 26 neuropeptide genes that are known in mammalian species. A remarkable finding was that for most of the missing genes, another gene in the same neuropeptide family has been identified in the chicken genome. This finding suggests that neuropeptide genes have undergone less duplication or more gene loss, or both processes in the chicken than in mammalian species. The high correct prediction of cleavage and non-cleavage sites in prohormones obtained with a model trained in human sequences indicates that the processing of prohormones into neuropeptides does not differ substantially between chicken and human species.

To gain a broad picture of the incidence of neuropeptide genes, we built a panel of expression across tissues and developmental stages. This panel will be of great value in streamlining neuropeptide research by helping to identify the tissues and developmental stages most likely to exhibit differential neuropeptide gene expression and subsequently, neuropeptide activity. Noteworthy findings include identifying the regions with highest number of neuropeptide gene expression reports (brain, head, small intestine, and heart) and the most frequently reported expressed genes (PDGFA, SCG1, and ECRG4). To further understand the role of neuropeptides in reproduction, growth, and health, we analyzed the expression of neuropeptide

genes across 22 microarray experiments that evaluated a wide range of ages, genders, tissues, genetic lines, and other conditions. Notable findings include various neuropeptide genes differentially expressed between the brain of male and female; these include MCH, ECRG4, GAST, REL3, and SCG1. Also, numerous neuropeptide genes, including PNOC, PDGFA, PDGFB, PDGFD, ADML, ECRG4, PENK, and IGF1, were differentially expressed in the breast muscle between 7- and 0-d-old (just hatched) chickens. Lastly, the expression profiles of the neuropeptide genes ADML, IGF1, VEGFC, VEGFD, and PENK across age differed significantly between broiler and layer genetic lines.

The chicken is a fundamental model for avian species and more insight into the neuropeptide complement of this species can be expected from proteomic mass spectrometry studies in the chicken and also from the sequencing of the zebra finch, an avian model system used to study brain development, learning, and memory (The Zebra Finch Genome Consortium, 2005). The list of chicken neuropeptide genes will also support the annotation of homolog genes in avian species with genomes that are in the process of being sequenced and annotated or that do not have sequenced genomes. The series of bioinformatics steps used in this study is applicable to surveying neuropeptide or other gene sets in organisms with similar bioinformatics resources. The chicken neuropeptide gene sequences and prohormone cleavage prediction approaches are available at http://neuroproteomics.scs.uiuc.edu/neuropred.html. The expression panel developed here will facilitate neuropeptide research by aiding with identification of the tissues and developmental stages most likely to exhibit neuropeptide gene expression and subsequently, neuropeptide activity in avian species.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The project described was supported by Award Number P30DA018310 and Award Number 1R21DA027548 from the National Institute On Drug Abuse (NIDA), Award Number 1R01GM068946 from the National Institute of General Medical Science (NIGMS), and by Award No. ILLU-538-311 from the United States Department of Agriculture (USDA), Cooperative State Research, Education and Extension Service. The content is solely the responsibility of the authors and does not necessarily represent the official views of NIDA, NIGMS, the National Institutes of Health or the USDA. Stephanie Baker is gratefully acknowledged for her excellent technical assistance.

Sources of support:

NIH/NIDA Award Numbers P30DA018310 and 1R21DA027548

NIH/NCI Award Number 5R03CA1439751

USDA CSREES Award Number ILLU-538-311

References

Agoulnik AI. Relaxin and related peptides in male reproduction. Adv Exp Med Biol 2007;612:49–64. [PubMed: 18161481]

Aikawa S, Ishii M, Yanagisawa M, Sakakibara Y, Sakurai T. Effect of neuropeptide B on feeding behavior is influenced by endogenous corticotropin-releasing factor activities. Regul Pept 2008;151:147–152. [PubMed: 18760311]

Alon T, Hemo I, Itin A, Pe'er J, Stone J, Keshet E. Vascular endothelial growth factor acts as a survival factor for newly formed retinal vessels and has implications for retinopathy of prematurity. Nat Med 1995;1:1024–1028. [PubMed: 7489357]

Álvaro A, Martins J, Araújo I, Rosmaninho-Salgado J, Ambrósio A, Cavadas C. Neuropeptide Y stimulates retinal neural cell proliferation involvement of nitric oxide. J Neurochem 2008;105:2501–2510.

Amare A, Hummon AB, Southey BR, Zimmerman TA, Rodriguez-Zas SL, Sweedler JV. Bridging neuropeptidomics and genomics with bioinformatics, Prediction of mammalian neuropeptide prohormone processing. J Proteome Res 2006;5:1162–1167. [PubMed: 16674105]

Amills M, Jimenez N, Villalba D, Tor M, Molina E, Cubilo D, Marcos C, Francesch A, Sanchez A, Estany J. Identification of three single nucleotide polymorphisms in the chicken insulin-like growth factor 1 and 2 genes and their associations with growth and feeding traits. Poult Sci 2003;82:1485–1493. [PubMed: 14601723]

Asnicar MA, Smith DP, Yang DD, et al. Absence of cocaine-and amphetamine- regulated transcript results in obesity in mice fed a high caloric diet. Endocrinology 2001;142:4394–4400. [PubMed: 11564703]

Baea JA, Parka HJ, Seoa YM, Rohb J, Hsuehc AJW, Chun SY. Hormonal regulation of proprotein convertase subtilisin/kexin type 5 expression during ovarian follicle development in the rat. Mol Cell Endocrinol 2008;289:29–37. [PubMed: 18502031]

Bagnoli P, Dal Monte M, Casini G. Expression of neuropeptides and their receptors in the developing retina of mammals. Histol Histopathol 2003;18:1219–1242. [PubMed: 12973690]

Balat O, Aksoy F, Kutlar I, Ugur MG, Iyikosker H, Balat A, Anarat R. Increased plasma levels of Urotensin-II in preeclampsia-eclampsia: a new mediator in pathogenesis? Eur J Obstet Gynecol Reprod Biol 2005;120:33–38. [PubMed: 15866083]

Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. J R Stat Soc B 1995;57:289–300.

Bennett AK, Hester PY, Spurlock DE. Polymorphisms in vitamin D receptor, osteopontin, insulin-like growth factor 1 and insulin, and their associations with bone, egg and growth traits in a layer--broiler cross in chickens. Anim Genet 2006;37:283–286. [PubMed: 16734694]

Birney E, Clamp M, Durbin R. GeneWise and Genomewise. Genome Res 2004;14:988–995. [PubMed: 15123596]

Bitar KN, Makhlouf GM. Specific opiate receptors on isolated mammalian gastric smooth muscle cells. Nature 1982;297:72–74. [PubMed: 6122163]

Burt DW. Emergence of the chicken as a model organism: implications for agriculture and biology. Poult Sci 2007;86:1460–1471. [PubMed: 17575197]

Cardot J, Griffond B, Risold PY, Blähser S, Fellmann D. Melanin-concentrating hormone-producing neurons in birds. J Comp Neurol 1999;411:239–256. [PubMed: 10404250]

Castro AA, Casagrande TS, Moretti M, Constantino L, Petronilho F, Guerra GC, Calo' G, Guerrini R, Dal-Pizzol F, Quevedo J, Gavioli EC. Lithium attenuates behavioral and biochemical effects of neuropeptide S in mice. Peptides. 2009

Chapman AM, Debski EA. Neuropeptide Y immunoreactivity of a projection from the lateral thalamic nucleus to the optic tectum of the leopard frog. Vis Neurosci 1995;12:1–9. [PubMed: 7718491]

Cho K, McFarlane ID. Physiological actions of the neuropeptide Antho-RNamide on antagonistic muscle systems in sea anemones. Neurosci Lett 1996;219:171–174. [PubMed: 8971807]

Cogburn LA, Wang X, Carre W, Rejto L, Porter TE, Aggrey SE, Simon J. Systems-wide chicken DNA microarrays, gene expression profiling, and discovery of functional genes. Poult Sci 2003;82:939–951. [PubMed: 12817449]

Conklin D, Lofton-Day CE, Haldeman BA, Ching A, Whitmore TE, Lok S, Jaspers S. Identification of INSL5, a new member of the insulin superfamily. Genomics 1999;60:50–66. [PubMed: 10458910]

D'Angelo I, Brecha NC. Y2 receptor expression and inhibition of voltage-dependent Ca(2+) influx into rod bipolar cell terminals. Neuroscience 2004;125:1039–1049. [PubMed: 15120863]

Dobbins J, Racusen L, Binder HJ. Effect of d-alanine methionine enkephalin amide on ion transport in rabbit ileum. J Clin Invest 1980;66:19–28. [PubMed: 6249844]

Duclos MJ. Insulin-like growth factor-I (IGF-1) mRNA levels and chicken muscle growth. J Physiol Pharmacol 2005;56:25–35. [PubMed: 16077194]

Dun SL, Brailoiu E, Wang Y, Brailoiu GC, Liu-Chen LY, Yang J, Chang JK, Dun NJ. Insulin-like peptide 5:expression in the mouse brain and mobilization of calcium. Endocrinology 2006;147:3243–3248. [PubMed: 16601133]

Edin R, Lundberg J, Terenius L, Dahlstrom A, Hokfelt T, Kewenter J, Ahlman H. Evidence for vagal enkephalinergic neural control of the feline pylorus and stomach. Gastroenterology 1980;78:492–497. [PubMed: 6985879]

Ellegren H, Hultin-Rosenberg L, Brunström B, Dencker L, Kultima K, Scholz B. Faced with inequality: chicken do not have a general dosage compensation of sex-linked genes. BMC Biol 2007;5:40. [PubMed: 17883843]

Feldman M, Walsh JH, Taylor IL. Effect of naloxone and morphine on gastric acid secretion and on serum gastrin and pancreatic polypeptide concentrations in humans. Gastroenterology 1980;79:294–298. [PubMed: 7399233]

Florin S, Suaudeau C, Meunier JC, Costentin J. Orphan neuropeptide NocII, a putative pronociceptin maturation product, stimulates locomotion in mice. Neuroreport 1997;8:705–707. [PubMed: 9106751]

Fricker LD. Neuropeptide-processing enzymes: applications for drug discovery. AAPS J 2005;7:E449–E455. [PubMed: 16353923]

Fruttiger M, Calver AR, Krüger WH, Mudhar HS, Michalovich D, Takakura N, Nishikawa S, Richardson WD. PDGF mediates a neuron-astrocyte interaction in the developing retina. Neuron 1996;17:1117–1131. [PubMed: 8982160]

Gyamera-Acheampong C, Tantibhedhyangkul J, Weerachatyanukul W, Tadros H, Xu H, van de Loo JW, Pelletier RM, Tanphaichitr N, Mbikay M. Sperm from mice genetically deficient for the PCSK4 proteinase exhibit accelerated capacitation, precocious acrosome reaction, reduced binding to egg zona pellucida, and impaired fertilizing ability. Biol Reprod 2006;74:666–673. [PubMed: 16371590]

Haugaard-Jönsso LM, Hossain MA, Daly NL, Craik DJ, Wade JD, Rosengren KJ. Structure of human insulin-like peptide 5 and characterization of conserved hydrogen bonds and electrostatic interactions within the relaxin framework. Biochem J 2009;419:619–627. [PubMed: 19178384]

Higuchi K, Masaki T, Gotoh K, Chiba S, Katsuragi I, Tanaka K, Kakuma T, Yoshimatsu H. Apelin, an APJ Receptor Ligand, Regulates Body Adiposity and Favors the Messenger Ribonucleic Acid Expression of Uncoupling Proteins in Mice. Endocrinology 2007;148:2690–2697. [PubMed: 17347313]

Hondo M, Ishii M, Sakurai T. The NPB/NPW neuropeptide system and its role in regulating energy homeostasis, pain, and emotion. Results Probl Cell Differ 2008;46:239–256. [PubMed: 18204824]

Hook V, Funkelstein L, Lu D, Bark S, Wegrzyn J, Hwang SR. Proteases for processing proneuropeptides into peptide neurotransmitters and hormones. Annu Rev Pharmacol Toxicol 2008;48:393–423. [PubMed: 18184105]

Hummon AB, Hummon NP, Corbin RW, Li L, Vilim FS, Weiss KR, Sweedler JV. From precursor to final peptides: a statistical sequence-based approach to predicting prohormone processing. J Proteome Res 2003;2:650–656. [PubMed: 14692459]

International Chicken Genome Sequencing Consortium. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature 2004;432:695–716. [PubMed: 15592404]

Irizarry, RA.; Gautier, L.; Bolstad, BM.; Miller, C. Methods for Affymetrix Oligonucleotide Arrays. < http://bioconductor.org/packages/2.5/bioc/html/affy.html >. 2009 [November 10, 2009].

Itoh Y, Melamed E, Yang X, Kampf K, Wang S, Yehya N, Van Nas A, Replogle K, Band MR, Clayton DF, Schadt EE, Lusis AJ, Arnold AP. Dosage compensation is less effective in birds than in mammals. J Biol 2007;6:2. [PubMed: 17352797]

Jozsa R, Nemeth J, Tamas A, Hollosy T, Lubics A, Jakab B, Olah A, Lengvari I, Arimura A, Reglodi D. Short-term fasting differentially alters PACAP and VIP levels in the brains of rat and chicken. Ann NY Acad Sci 2006;1070:354–358. [PubMed: 16888191]

Kivelä R, Havas E, Vihko V. Localisation of lymphatic vessels and vascular endothelial growth factors-C and -D in human and mouse skeletal muscle with immunohistochemistry. Histochem Cell Biol 2007;127:31–40. [PubMed: 16924525]

Konturek SJ, Tasler J, Cieszkowski M, Mikos E, Coy DH, Schally AV. Comparison of methionine-enkephalin and morphine in the stimulation of gastric acid secretion in the dog. Gastroenterology 1980;78:294–300. [PubMed: 6243110]

Kubo F, Nakagawa S. Hairy1 acts as a node downstream of Wnt signaling to maintain retinal stem cell-like progenitor cells in the chick ciliary marginal zone. Development 2009;136:1823–1833. [PubMed: 19386663]

Kuhar MJ, Adams S, Dominguez G, Jaworski J, Balkan B. CART peptides. Neuropeptides 2002;36:1–8. [PubMed: 12147208]

Li L, Sweedler JV. Peptides in the Brain: Mass Spectrometry-Based Measurement Approaches and Challenges. Annual Review of Analytical Chemistry 2008;1:451–483.

Ling MK, Hotta E, Kilianova Z, Haitina T, Ringholm A, Johansson L, Gallo-Payet N, Takeuchi S, Schiöth HB. The melanocortin receptor subtypes in chicken have high preference to ACTH-derived peptides. Br J Pharmacol 2004;143:626–637. [PubMed: 15466451]

Lukinius A, Stridsberg M, Wilander E. Cellular expression and specific intragranular localization of chromogranin A, chromogranin B, and synaptophysin during ontogeny of pancreatic islet cells:an ultrastructural study. Pancreas 2003;27:38–46. [PubMed: 12826904]

Martínez A, Bengoechea JA, Cuttitta F. Molecular evolution of proadrenomedullin N-terminal 20 peptide (PAMP): evidence for gene co-option. Endocrinology 2006;147:3457–3461. [PubMed: 16574790]

McGlinn AM, Baldwin DA, Tobias JW, Budak MT, Khurana TS, Stone RA. Form-deprivation myopia in chick induces limited changes in retinal gene expression. Invest Ophthalmol Vis Sci 2007;48:3430–3436. [PubMed: 17652709]

McGowan BM, Stanley SA, Ghatei MA, Bloom SR. Relaxin-3 and its role in neuroendocrine function. Ann NY Acad Sci 2009;1160:250–255. [PubMed: 19416197]

McIntyre BA, Alev C, Tarui H, Jakt LM, Sheng G. Expression profiling of circulating non-red blood cells in embryonic blood. BMC Dev Biol 2008;8:21. [PubMed: 18302797]

McKay JS, Linaker BD, Turnberg LA. Influence of opiates on ion transport across rabbit ileal mucosa. Gastroenterology 1981;80:279–284. [PubMed: 6256253]

Mirabeau O, Perlas E, Severini C, Audero E, Gascuel O, Possenti R, Birney E, Rosenthal N, Gross C. Identification of novel peptide hormones in the human proteome by hidden Markov model screening. Genome Res 2007;17:320–327. [PubMed: 17284679]

Mollereau C, Simons MJ, Soularue P, Liners F, Vassart G, Meunier JC, Parmentier M. Structure, tissue distribution, and chromosomal localization of the prepronociceptin gene. Proc Natl Acad Sci USA 1996;93:8666–8670. [PubMed: 8710928]

Monstein HJ, Grahn N, Ohlsson B. Proenkephalin-A mRNA Is Widely Expressed in Tissues of the Human Gastrointestinal Tract. Eur Surg Res 2006;38:464–468. [PubMed: 16940731]

Montuenga LM, Martinez A, Miller MJ, Unsworth EJ, Cuttitta F. Expression of adrenomedullin and its receptor during embryogenesis suggests autocrine or paracrine modes of action. Endocrinology 1997;138:440–451. [PubMed: 8977434]

Mori Y, Ishiguro H, Kuwabara Y, Kimura M, Mitsui A, Kurehara H, Mori R, Tomado K, Ogawa R, Katada T, Harata K, Fujii Y. Expression of ECRG4 is an independent prognostic factor for poor survival in patients with esophageal squamous cell carcinoma. Oncol Rep 2007;18:981–985. [PubMed: 17786363]

Mudhar HS, Pollock RA, Wang C, Stiles CD, Richardson WD. PDGF and its receptors in the developing rodent retina and optic nerve. Development 1993;118:539–552. [PubMed: 8223278]

Nakao N, Ono H, Yamamura T, Anraku T, Takagi T, Higashi K, Yasuo S, Katou Y, Kageyama S, Uno Y, Kasukawa T, Iigo M, Sharp PJ, Iwasawa A, Suzuki Y, Sugano S, Niimi T, Mizutani M, Namikawa T, Ebihara S, Ueda HR, Yoshimura T. Thyrotrophin in the pars tuberalis triggers photoperiodic response. Nature 2008;452:317–322. [PubMed: 18354476]

Pape HC, Jüngling K, Seidenbecher T, Lesting J, Reinscheid RK. Neuropeptide S: A transmitter system in the brain regulating fear and anxiety. Neuropharmacology. 2009 June 10;

Pierce EA, Foley ED, Smith LE. Regulation of vascular endothelial growth factor by oxygen in a model of retinopathy of prematurity. Arch Ophthalmol 1996;114:1219–1228. [PubMed: 8859081]

Powell DW. Muscle or mucosa:the site of action of antidiarrheal opiates? Gastroenterology 1981;80:406–408. [PubMed: 7450430]

Prasad SK, Clerk A, Cullingford TE, Chen AW, Kemp TJ, Cannell TM, Cowie MR, Petrou M. Gene expression profiling of human hibernating myocardium:increased expression of B-type natriuretic peptide and proenkephalin in hypocontracile vs normally-contracting regions of the heart. Eur J Heart Fail 2008;10:1177–1180. [PubMed: 18974021]

Provis JM, Leech J, Diaz CM, Penfold PL, Stone J, Keshet E. Development of the human retinal vasculature:cellular relations and VEGF expression. Exp Eye Res 1997;65:555–568. [PubMed: 9464188]

Reynolds JC, Ouyang A, Cohen S. Evidence for an opiate-mediated pyloric sphincter reflex. Am J Physiol 1984;246:G130–136. [PubMed: 6696109]

Richards MP, McMurtry JP. Expression of proglucagon and proglucagon-derived peptide hormone receptor genes in the chicken. Gen Comp Endocrinol 2008;156:323–338. [PubMed: 18299131]

Rizzolo LJ, Chen X, Weitzman M, Sun R, Zhang H. Analysis of the RPE transcriptome reveals dynamic changes during the development of the outer blood-retinal barrier. Mol Vis 2007;13:1259–1273. [PubMed: 17679949]

Rosenquist TH, Bennett GD, Brauer PR, Stewart ML, Chaudoin TR, Finnell RH. Microarray analysis of homocysteine-responsive genes in cardiac neural crest cells in vitro. Dev Dyn 2007;236:1044–1054. [PubMed: 17326132]

Schippert R, Schaeffel F, Feldkaemper MP. Microarray analysis of retinal gene expression in chicks during imposed myopic defocus. Mol Vis 2008;14:1589–1599. [PubMed: 18769560]

Schwippert WW, Röttgen A, Ewert JP. Neuropeptide Y (NPY) or fragment NPY 13–36, but not NPY 18–36, inhibit retinotectal transfer in cane toads Bufo marinus. Neurosci Lett 1998;253:33–36. [PubMed: 9754798]

Shimoda H, Kato S. A model for lymphatic regeneration in tissue repair of the intestinal muscle coat. Int Rev Cytol 2006;250:73–108. [PubMed: 16861064]

Southey BR, Amare A, Zimmerman TA, Rodriguez-Zas SL, Sweedler JV. NeuroPred: a tool to predict cleavage sites in neuropeptide precursors and provide the masses of the resulting peptides. Nucleic Acids Res 2006;34:W267–W272. [PubMed: 16845008]

Southey BR, Rodriguez-Zas SL, Sweedler JV. Prediction of neuropeptide prohormone cleavages with application to RFamides. Peptides 2006;27:1087–1098. [PubMed: 16494967]

Southey BR, Sweedler JV, Rodriguez-Zas SL. A python analytical pipeline to identify prohormone precursors and predict prohormone cleavage sites. Front Neuroinformatics 2008;2:7. [PubMed: 19169350]

Southey BR, Rodriguez-Zas SL, Sweedler JV. Characterization of the prohormone complement in cattle using genomic libraries and cleavage prediction approaches. BMC Genomics 2009;10:228. [PubMed: 19445702]

Stern CD. The chick; a great model system becomes even greater. Dev Cell 2005;8:9–17. [PubMed: 15621526]

Stone J, Itin A, Alon T, Pe'er J, Gnessin H, Chan-Ling T, Keshet E. Development of retinal vasculature is mediated by hypoxia-induced vascular endothelial growth factor (VEGF) expression by neuroglia. J Neurosci 1995;15:4738–4747. [PubMed: 7623107]

Takami Y, Nakayama T. A single copy of linker H1 genes is enough for proliferation of the DT40 chicken B cell line, and linker H1 variants participate in regulation of gene expression. Genes Cells 1997;2:711–723. [PubMed: 9491804]

Tanaka M, Iijima N, Miyamoto Y, et al. Neurons expressing relaxin 3/INSL 7 in the nucleus incertus respond to stress. Eur J Neurosci 2005;21:1659–1670. [PubMed: 15845093]

Tegge AN, Southey BR, Sweedler JV, Rodriguez-Zas SL. Comparative analysis of neuropeptide cleavage sites in human, mouse, rat, and cattle. Mamm Genome 2008;19:106–120. [PubMed: 18213482]

Udono T, Takahashi K, Nakayama M, Murakami O, Durlu YK, Tamai M, Shibahara S. Adrenomedullin in cultured human retinal pigment epithelial cells. Invest Ophthalmol Vis Sci 2000;41:1962–1970. [PubMed: 10845623]

Van Den Akker NM, Lie-Venema H, Maas S, Eralp I, DeRuiter MC, Poelmann RE, Gittenberger-De Groot AC. Platelet-derived growth factors in the developing avian heart and maturating coronary vasculature. Dev Dyn 2005;233:1579–1588. [PubMed: 15973731]

Wang G, Yan B, Deng X, Li C, Hu X, Li N. Insulin-like growth factor 2 as a candidate gene influencing growth and carcass traits and its bialleleic expression in chicken. Sci China C Life Sci 2005;48:187–194. [PubMed: 15986892]

Wang HB, Li H, Wang QG, Zhang XY, Wang SZ, Wang YX, Wang XP. Profiling of chicken adipose tissue gene expression by genome array. BMC Genomics 2007;8:193. [PubMed: 17594506]

Yang X, Chrisman H, Weijer CJ. PDGF signalling controls the migration of mesoderm cells during chick gastrulation by regulating N-cadherin expression. Development 2008;135:3521–3530. [PubMed: 18832396]

Zebra Finch Genome Consortium. Proposal to Sequence the Genome of the Zebra Finch (Taeniopygia guttata). < http://www.songbirdgenome.org/pdfs/ZebraFinchGenomeNHGRIjuly05a.pdf >. 2005 [November 10, 2009].

Zheng Q, Zhang Y, Chen Y, Yang N, Wang X, Zhu D. Systematic identification of genes involved in divergent skeletal muscle growth rates of broiler and layer chickens. BMC Genomics 2009;10:87. [PubMed: 19232135]

Zhou H, Mitchell AD, McMurtry JP, Ashwell CM, Lamont SJ. Insulin-like growth factor-I gene polymorphism associations with growth, body composition, skeleton integrity, and metabolic traits in chickens. Poult Sci 2005;84:212–219. [PubMed: 15742956]

Zudaire E, Cuesta N, Martínez A, Cuttitta F. Characterization of adrenomedullin in birds. Gen Comp Endocrinol 2005;143:10–20. [PubMed: 15993100]

Table 1

Neuropeptide and convertase gene and protein master list

Neuropeptide Prohormone						
Abbreviated Name	**Name**	**UniProtID[1]**	**GeneID[2]**	**UniGeneID[3]**	**Evidence in Chicken[4]**	**EST/Genome/HTGS**
ADM2	Intermedin	NA[5]	NA	NA	NA	NA/NA/NA
ADML	Adrenomedullin	NA	423042	Gga.12006	NA	1[6]/1/1
ANF	Atrial natriuretic factor	P18908	395765	Gga.5157	protein	1/1/1
ANFB	Natriuretic peptides B	NA	NA	NA	NA	NA/NA/NA
ANFC	C-type natriuretic peptide	A9CDT6	419487	Gga.12392	transcript	1/1/0
APEL	Apelin	NA	NA	NA	NA	NA/NA/NA
C-RF AMIDE	C-RF amide peptide	B0LF68	420716	Gga.3202	predicted	1/1/1
CALC/CALCA	Calcitonin/Calcitonin gene-related peptide 1	P07660, P10286	396256	Gga.4991	transcript/protein	1/1/1
CALCB	Calcitonin gene-related peptide 2	NA	NA	NA	NA	NA/NA/NA
CART	Cocaine- and amphetamine-regulated transcript protein	NA	NA	NA	NA	NA/NA/NA
CCKN	Cholecystokinin	Q9PU41	414884	Gga.2441	protein	1/1/1
CMGA	Chromogranin-A	NA	423420	Gga.19002	NA	1/1/0
CNP1	C-type natriuretic peptide 1	A9CDT5	NA	Gga.47230	transcript	NA/NA/NA
COLI	Pro-opiomelanocortin	Q9YI93	422011	Gga.6271	predicted	1/1/1
CORT	Cortistatin	NA	NA	NA	NA	NA/NA/NA
CRF	Corticoliberin	Q703P0	404297	Gga.11323	transcript	1/1/1
ECRG4	Augurin (Esophageal cancer-related gene 4 protein)	NA	771055	Gga.8435	NA	1/1/1
EDN1	Endothelin-1	NA	420854	Gga.25090	NA	1/1/1
EDN2	Endothelin-2	NA	419559	Gga.8238	NA	1/1/1
EDN3	Endothelin-3	Q3MU75	768509	Gga.22840	transcript	1/1/1
GALA	Galanin	P30802	423117	Gga.12649	protein	1/1/1
GALP	Galanin-like peptide	NA	NA	NA	NA	NA/NA/NA
GAST	Gastrin	P09859	396365	Gga.782	protein	0/0/0
GHRL	Obestatin	Q8AV73, Q7T2V1	408185	Gga.16	homology	1/1/1
GIP	Gastric inhibitory polypeptide	A1DPK0	419989	Gga.7981	transcript	1/1/1

Neuropeptide Prohormone						
Abbreviated Name	**Name**	**UniProtID[1]**	**GeneID[2]**	**UniGeneID[3]**	**Evidence in Chicken[4]**	**EST/Genome/HTGS**
GLUC	Glucagon	P68259	396196	Gga.704	protein	1/1/1
GON1	Progonadoliberin-1	P37042	770134	Gga.41802	protein	1/1/1
GRP	Gastrin-releasing peptide	P01295	425213	Gga.43422	protein	1/1/1
HEPC	Hepcidin	NA	NA	NA	NA	NA/NA/NA
IAPP	Islet amyloid polypeptide	Q90743	396362	Gga.780	transcript	1/1/1
IGF1	Insulin-like growth factor I	P18254	418090	Gga.850	protein	1/1/1
IGF2	Insulin-like growth factor 2 (somatomedin A)	P33717	395097	Gga.8511	protein	1/1/1
INS	Insulin	P67970	396145	Gga.673	protein	1/1/1
INSL3	Insulin-like 3	NA	NA	NA	NA	NA/NA/NA
INSL5	Insulin-like 5	NA	NA	NA	NA	NA/NA/NA
INSL6	Insulin-like 6	NA	NA	NA	NA	NA/NA/NA
KISS1	Metastasis-suppressor KiSS-1	NA	NA	NA	NA	NA/NA/NA
MCH	Pro-melanin-concentrating hormone	NA	418091	Gga.14659	NA	1/1/1
MOTI	Motilin	Q9PRP6	768422	NA	protein	1/1/0
NEU1	Oxytocin	Q2ACD0	768516	NA	predicted	1/1/1
NEU2	Neurophysin-II	P24787	396101	Gga.652	transcript	1/1/1
NEUT	Neurotensin	P13724	417883	Gga.10167	protein	1/1/1
NMB	Neuromedin-B	A0MAR5	415333	Gga.8071	transcript	1/1/1
NMS	Neuromedin-S	NA	NA	NA	NA	NA/NA/NA
NMU	Neuromedin-U	P34963	422748	Gga.18392	protein	1/1/1
NPB	Neuropeptide B	NA	NA	NA	NA	NA/NA/NA
NPFF	Neuropeptide FF	NA	NA	NA	NA	NA/NA/NA
NPS	Neuropeptide S	NA	NA	NA	NA	NA/NA/NA
NPW	Neuropeptide W	NA	NA	NA	NA	NA/NA/NA
NPY	Neuropeptide Y	P28673	396464	Gga.837	homology	1/1/1
OREX	Orexin	Q8AV17	374005	Gga.11	transcript	1/1/1
OSTN	Osteocrin (Musclin)	A5JNH0	424907	Ggta.13448	transcript	1/1/1
OX26	Orexigenic neuropeptide QRFP	B2CL09	771867	NA	transcript	1/1/1

Neuropeptide Prohormone						
Abbreviated Name	**Name**	**UniProtID[1]**	**GeneID[2]**	**UniGeneID[3]**	**Evidence in Chicken[4]**	**EST/Genome/HTGS**
PACA	Pituitary adenylate cyclase-activating polypeptide	P41534	408251	Gga.616	protein	0/0/0
PAHO	Pancreatic polypeptide	P68248	395564	Gga.308	protein	1/1/1
PCSK1N	Proprotein convertase subtilisin/kexin type 1 inhibitor	NA	NA	NA	NA	NA/NA/NA
PDGFA	Platelet-derived growth factor alpha polypeptide	Q90WK2, Q9PUF7	374196	Gga.3899	transcript	1/1/1
PDGFB	Platelet-derived growth factor beta polypeptide	Q90W23	374128	Gga.71	transcript	1/1/1
PDGFD	Platelet derived growth factor D	O57658	418978	Gga.43662	transcript	1/1/1
PDYN	Proenkephalin-B	NA	NA	NA	NA	NA/NA/NA
PENK	Proenkephalin	NA	421131	Gga.11430	NA	1/1/1
PNOC	Prepronociceptin	NA	422019	Gga.10041	NA	1/1/1
PROK2	Prokineticin 2	NA	771674	Gga.10528	NA	1/1/1
PRRP	Prolactin-releasing peptide	A3RJ26	424018	Gga.10552	predicted	1/1/0
PTHR	Parathyroid hormone-related protein	P17251	396281	Gga.2626	protein	1/1/0
PTHY	Parathyroid hormone	P15743	396436	Gga.78	homology	1/1/1
PYY	Peptide YY	P29203	NA	NA	protein	NA/NA/NA
PYY2	Putative peptide YY-2	NA	NA	NA	NA	NA/NA/NA
REL1	pro-relaxin 1	NA	NA	NA	NA	NA/NA/NA
REL2	pro-relaxin 2	NA	NA	NA	NA	NA/NA/NA
REL3	Relaxin-3	B1AC67	427223	Gga.37019	transcript	1/1/1
RES18	Regulated endocrine-specific protein 18	NA	NA	NA	NA	NA/NA/NA
RFRP	Neuropeptide VF precursor	Q6T2D1, Q75XU6	378785	Gga.9285	transcript	1/1/0
RNP	Renal natriuretic peptide	A9CDT7	NA	NA	transcript	NA/NA/NA
SCG1	Secretogranin-1	NA	421312	Gga.10025	NA	1/1/0
SCG2	Secretogranin-2	NA	424808	Gga.11999	NA	1/1/1
SECR	Secretin	P01280	423015	Gga.14227	protein	1/1/1
SLIB	Somatoliberin	Q1KNA8, Q1KNA7	419178	Gga.11231	transcript	1/1/1
SMS	Somatostatin	P33094	396279	Gga.742	homology	1/1/1
SPXN	Spexin	NA	NA	NA	NA	NA/NA/NA

Abbreviated Name	Name	UniProtID[1]	GeneID[2]	UniGeneID[3]	Evidence in Chicken[4]	EST/Genome/HTGS
Neuropeptide Prohormone						
TAC4/TKN4	Tachykinin-4	NA	NA	NA	NA	NA/NA/NA
TIP39	Parathyroid hormone 2	NA	NA	NA	NA	NA/NA/NA
TKN1	Tachykinin, precursor 1	NA	420573	Gga.12286	NA	1/1/1
TKNK	Tachykinin 3	NA	NA	NA	NA	NA/NA/NA
TOR2X	Torsin family 2, member A	NA	NA	NA	NA	NA/NA/NA
TRH	Prothyroliberin	Q6ZXC3	414344	Gga.19489	transcript	1/1/1
TSHB	Thyroid-stimulating hormone subunit beta	O57340	395937	Gga.551	transcript	1/1/1
UCN1	Urocortin	NA	NA	NA	NA	NA/NA/NA
UCN2	Urocortin 2	NA	NA	NA	NA	NA/NA/NA
UCN3	Urocortin 3	NA	769274	Gga.11141	NA	1/1/1
UTS2	Urotensin 2	Q6Q2J6	404535	Gga.14388	transcript	1/1/1
UTS2D	Urotensin II-related peptide	Q6Q273	404534	Gga.9482	transcript	1/1/1
VEGFC	Vascular endothelial growth factor C	NA	422573	Gga.12347	NA	1/1/1
VEGFD	Vascular endothelial growth factor D	Q8QGD7	395255	Gga.3219	transcript	1/1/1
VIP	Vasoactive intestinal peptides	P48143	396323	Gga.666	protein	1/1/1
Prohormone Convertase Enzyme						
PCSK1	Proprotein convertase subtilisin/kexin type 1	NA	395137	Gga.9357	NA	1/1/1
PCSK2	Proprotein convertase subtilisin/kexin type 2	NA	395136	Gga.9404	NA	1/1/1
PCSK3	Furin	Q91000	395457	Gga.1751	transcript	1/1/1
PCSK4	Proprotein convertase subtilisin/kexin type 4	NA	NA	NA	NA	NA/NA/NA
PCSK5	Proprotein convertase subtilisin/kexin type 5	NA	395456	Gga.12660	NA	1/1/1
PCSK6	Proprotein convertase subtilisin/kexin type 6	NA	395454	Gga.21090	NA	1/1/1
PCSK7	Proprotein convertase subtilisin/kexin type 7	Q5ZKB5	395455	Gga.5311	transcript	1/1/1

[1]: UniProt identifier

[2]: Gene database identifier

[3]: UniGene database identifier

[4]: Evidence in chicken in UniProt at the protein or transcript level, inferred from homology or predicted

[5]: NA = Not Available

[6]: 0 denotes absent and 1 denotes present in the corresponding database

Table 2

Abbreviated distribution of neuropeptide and convertase gene EST across tissues and stages

Neuropeptide Prohormones	Unigene ID[1]	Brain	Head	Ovary	Small Intestine	Embryo Stage	Adult Stage
ADML	Gga.12006	1[2]	0	0	0	1	1
ANF	Gga.5157	1	0	0	0	1	1
ANFC	Gga.12392	1	0	0	0	0	1
C-RF AMIDE	Gga.3202	1	1	1	0	1	1
CALCA	Gga.4991	1	1	0	0	1	0
CCKN	Gga.2441	1	0	1	1	1	1
CMGA	Gga.19002	1	1	1	1	1	1
COLI	Gga.6271	1	0	0	0	0	0
ECRG4	Gga.8435	1	1	1	1	1	1
EDN1	Gga.25090	0	0	0	0	1	1
EDN2	Gga.8238	1	0	1	1	1	1
EDN3	Gga.22840	0	0	1	0	1	1
GALA	Gga.12649	0	0	0	1	0	1
GHRL	Gga.16	0	0	1	0	1	1
GIP	Gga.7981	0	0	0	1	0	1
GLUC	Gga.704	0	1	0	1	1	1
GRP	Gga.43422	0	0	0	0	0	0
IAPP	Gga.780	1	1	0	0	1	1
IGF1	Gga.850	0	0	1	0	0	1
IGF2	Gga.8511	0	0	0	0	0	1
INS	Gga.673	0	0	1	0	0	1
MCH	Gga.14659	1	0	0	0	1	0
NEU2	Gga.652	1	0	0	0	1	1
NEUT	Gga.10167	1	1	0	1	1	1
NMB	Gga.8071	1	0	0	0	1	1
NMU	Gga.18392	1	0	0	1	1	1
NPY	Gga.837	1	0	1	0	1	1
OREX	Gga.11	0	0	0	0	1	0
OSTN	Gga.13448	0	1	0	0	1	1

Neuropeptide Prohormones	Unigene ID[1]	Brain	Head	Ovary	Small Intestine	Embryo Stage	Adult Stage
PACA	Gga.616	1	1	0	0	1	1
PDGFA	Gga.3899	1	1	1	1	1	1
PDGFB	Gga.71	1	1	1	0	1	1
PDGFD	Gga.43662	1	1	0	0	1	1
PENK	Gga.11430	1	1	1	0	1	1
PNOC	Gga.10041	1	0	0	0	1	0
PROK2	Gga.10528	1	0	1	0	1	1
PRRP	Gga.10552	0	0	1	0	0	1
PTHR	Gga.2626	0	0	0	0	1	0
REL3	Gga.37019	1	0	1	0	1	1
RFRP	Gga.9285	1	1	0	0	1	1
SCG1	Gga.10025	1	1	1	0	1	1
SCG2	Gga.11999	1	1	0	0	1	1
SECR	Gga.14227	0	0	0	1	0	1
SLIB	Gga.11231	1	1	0	0	1	0
SMS	Gga.742	1	0	0	1	1	1
TKN1	Gga.12286	1	0	0	1	0	1
UCN3	Gga.11141	0	0	0	1	0	1
UTS2D	Gga.9482	0	1	0	0	1	0
VEGFC	Gga.12347	0	1	0	0	1	1
VEGFD	Gga.3219	1	1	1	1	1	1
VIP	Gga.666	1	1	0	1	1	1
Total		**33**	**21**	**18**	**16**	**39**	**42**
Prohormone Convertases							
PCSK1	Gga.9357	N/A	N/A	N/A	N/A	N/A	N/A
PCSK1 similar	Gga.31439	1	0	0	1	1	1
PCSK2	Gga.9404	1	1	0	1	1	1
PCSK3	Gga.1751	1	0	0	0	0	1
PCSK5	Gga.12660	1	0	1	1	1	1

Neuropeptide Prohormones	Unigene ID[1]	Brain	Head	Ovary	Small Intestine	Embryo Stage	Adult Stage
PCSK6	Gga.21090	0	0	1	1	1	1
PCSK7	Gga.5311	1	1	1	1	1	1
Total		**4**	**2**	**3**	**4**	**4**	**5**

[1]: UniGene database identifier

[2]: 1 denotes presence and 0 denotes absence

Table 3

Abbreviated description of the 22 chicken microarray experiments analyzed

Exp. ID[1]	Tissue	Gender	Age	Reference
GSE6543	Retina	Female	1-wk	McGlinn et al.. (2007)
GSE7176	Retinal epithelium	NA[2]	7-d-old embryo	Rizzolo et al. (2007)
GSE11439	Retina	Male	9-d-old	Schippert et al. (2008)
GSE15382	Retina	NA	Embryo	Kubo et al. (2009)
GSE6843	Embryonic heart	M/F[3]	Late stage embryo	Itoh et al. (2007)
GSE8693	Embryonic heart	M/F	18-d-old embryo	Ellegren et al. (2007)
GSE9251	Pectoralis muscles	Female	1-d-old to 8-wk-old	Zheng et al. (2009)
GSE6844	Embryonic brain	M/F	Late stage embryo	Itoh et al. (2007)
GSE6868	Neural tube explants	NA	Stage 9+ embryo	Rosenquist et al. (2007)
GSE8018	Hypothalamus	NA	NA	Nakao et al. (2008)
GSE8693	Embryonic brain	M/F	18-d-old embryo	Ellegren et al. (2007)
GSE12268	Brain	M/F	Stage 29 embryo	NA
GSE15413	Brain	NA	Newly hatched/7-d-old	NA
GSE6856	Embryonic liver	M/F	Late stage embryo	Itoh et al. (2007)
GSE8016	Liver	Female	NA	Nakao et al. (2008)
GSE15413	Liver	NA	Newly hatched/7-d-old	NA
GSE15413	Duodenum	NA	Newly hatched/7-d-old	NA
GSE8693	Embryonic gonad	M/F	18-d-old embryo	Ellegren et al. (2007)
GSE10231	F1 oocyte stage	Female	NA	NA
GSE6868	Neural tube explants	NA	Stage 9+embryo	Rosenquist et al. (2007)
GSE8010	Adipose Tissue	Female	7-wk	Wang et al. (2007)
GSE8483	DT40 cells	NA	NA	Takami et al. (1997)
GSE9884	Embryonic heart blood	NA	Embryo	McIntyre et al. (2008)

[1]: ID = Gene Expression Omnibus GEO Series identifier

[2]: NA = Not Available

[3]: M/F = Male and Female

Table 4

Number of differentially expressed neuropeptide and convertase genes (*P-value* < 0.005) across 22 microarray studies grouped by tissue type

Prohormone	UniGene Probe[1]	Retina	Heart-breast	Brain-head	Liver-duod.[2]	Oocyte-gonad	Others	Total by Probe
ADML	**Gga.12006.1.S1_at**	1	2	0	0	1	1	5
ANF	**Gga.5157.1.S1_at**	0	0	0	0	0	0	0
ANFC	**Gga.12392.1.S1_a_at**	0	1	0	0	0	1	2
CALCA	**Gga.4991.2.S1_a_at**	1	0	0	0	0	0	1
CCKN	**GgaAffx.21834.1.S1_s_at**	0	0	0	0	0	1	1
CCKN	**Gga.2441.1.S1_at**	0	0	0	0	0	0	0
CMGA	**GgaAffx.21576.1.S1_s_at**	1	1	0	1	0	1	4
CMGA	**Gga.12437.2.S1_at**	0	1	0	0	0	0	1
COLI	**Gga.6271.1.S1_at**	0	1	0	0	0	1	2
CRF	**Gga.11323.1.S1_at**	0	0	0	0	1	0	1
ECRG4	**Gga.8435.1.S1_at**	0	2	0	0	0	1	3
ECRG4	**Gga.11232.1.A1_at**	0	1	0	0	0	0	1
GALA	**Gga.12649.1.S1_at**	0	0	0	0	0	0	0
GAST	**Gga.782.1.S1_at**	0	1	0	0	0	0	1
GHRL	**Gga.16.1.S1_at**	0	0	0	0	0	0	0
GIP	**Gga.7981.1.S1_at**	0	1	0	0	0	0	1
GLUC	**GgaAffx.21780.2.S1_s_at**	0	0	0	0	1	0	1
IAPP	**Gga.780.1.S1_at**	0	1	0	0	0	0	1
IGF1	**Gga.850.1.S1_at**	0	2	0	1	1	0	4
IGF2	**Gga.8511.1.S1_at**	0	1	0	0	0	1	2
INS	**Gga.673.1.S1_at**	0	0	0	1	0	1	2
MCH	**Gga.14659.1.S1_at**	0	0	1	0	0	0	1
NEU2	**Gga.652.1.S1_at**	0	0	0	0	0	0	0
NEUT	**Gga.10167.1.S1_at**	1	0	0	0	0	0	1
NMB	**Gga.8071.1.S1_a_at**	0	0	0	0	0	0	0
NMU	**Gga.18392.1.S1_at**	0	1	0	1	0	1	3
NPY	**Gga.837.1.S1_a_at**	0	0	0	0	0	1	1
OREX	**Gga.11.1.S1_at**	0	0	0	0	0	0	0
OSTN	**Gga.13448.1.S1_at**	0	1	0	0	0	0	1

Prohormone	UniGene Probe[1]	Retina	Heart-breast	Brain-head	Liver-duod.[2]	Oocyte-gonad	Others	Total by Probe
PACA	Gga.11409.1.S1_at	0	1	0	0	0	0	1
PACA	Gga.616.1.S1_s_at	0	0	0	0	0	0	0
PAHO	Gga.308.1.S1_at	0	0	0	0	1	0	1
PDGFA	Gga.3899.3.S1_a_at	1	1	0	0	1	0	3
PDGFB	Gga.71.1.S1_at	0	1	0	0	0	0	1
PDGFD	Gga.9675.1.S1_at	0	2	0	0	1	0	3
PENK	Gga.11430.1.S1_at	0	2	0	1	0	0	3
PNOC	Gga.10041.1.S1_a_at	0	2	0	0	0	0	2
PNOC	Gga.10041.2.A1_at	0	1	0	0	0	1	2
PNOC	GgaAffx.20191.1.S1_s_at	0	0	0	0	0	0	0
PROK2	Gga.10528.1.S1_a_at	0	0	0	1	0	0	1
PROK2	Gga.10528.2.A1_at	0	0	0	0	0	0	0
PRRP	Gga.10552.1.S1_at	0	1	0	0	0	0	1
PTHR	Gga.2626.1.S1_at	0	1	0	0	0	1	2
PTHY	Gga.78.1.A1_at	0	0	0	0	0	0	0
REL3	Gga.12454.1.S1_at	0	0	1	0	0	2	3
RFRP	Gga.9285.1.S1_at	0	0	0	0	0	1	1
SCG1	Gga.10025.1.S1_at	0	0	1	0	0	0	1
SCG2	Gga.11999.1.S1_at	0	0	0	0	0	0	0
SCG2	Gga.11999.1.A1_s_at	0	1	0	0	1	0	2
SCG2	Gga.11999.1.A1_at	0	0	0	0	0	0	0
SECR	Gga.14227.1.S1_at	0	0	0	0	0	0	0
SLIB	Gga.11231.1.S1_at	0	0	0	0	0	0	0
SMS	Gga.742.1.S1_at	0	1	0	0	0	0	1
TKN1	Gga.12286.1.S1_at	0	1	0	0	1	0	2
TRH	Gga.19489.1.A1_at	0	1	0	0	0	0	1
TRH	Gga.19489.1.S1_at	0	0	0	0	0	0	0
TSHB	Gga.551.1.S1_at	0	1	0	0	0	1	2
UCN3	Gga.11141.1.S1_at	0	0	0	0	0	0	0
UTS2	Gga.14388.1.S1_at	0	0	0	0	0	0	0
UTS2D	Gga.9482.1.S1_at	2	0	0	0	0	0	2
VEGFC	Gga.12347.1.S1_at	1	1	0	0	0	0	2

Prohormone	UniGene Probe[1]	Retina	Heart-breast	Brain-head	Liver-duod.[2]	Oocyte-gonad	Others	Total by Probe
VEGFD	**Gga.3219.1.S1_at**	1	1	0	0	0	0	2
VIP	**Gga.666.1.S1_a_at**	1	0	1	0	0	0	2
Total by Tissue		10	36	4	6	9	16	81
Prohormone Convertases								
PCSK2	**Gga.2786.1.S1_at**	0	1	0	0	0	0	1
PCSK2	**Gga.9404.1.S1_at**	0	0	0	0	0	0	0
PCSK3	**Gga.1751.1.S1_at**	0	0	0	0	1	0	1
PCSK5	**Gga.247.1.S1_at**	1	0	0	0	0	0	1
PCSK5	**Gga.12660.2.S1_a_at**	0	0	0	0	0	1	1
PCSK6	**Gga.20041.1.S1_at**	0	1	0	1	1	0	3
PCSK6	**GgaAffx.20832.1.S1_s_at**	0	0	0	0	1	0	1
PCSK6	**Gga.246.1.S1_at**	0	0	0	0	0	0	0
PCSK7	**GgaAffx.12272.1.S1_s_at**	0	0	0	0	0	0	0
PCSK7	**Gga.17539.1.S1_s_at**	0	0	0	0	1	0	1
Total by Tissue		1	2	0	1	4	1	9

[1]: UniGene identifier of microarray gene probe

[2]: Duodenum

Table 5

Evaluation of the prediction of cleavage sites in chicken prohormone sequences

	Known Motif Model				Human Model			
Neuropeptide Prohormone	**TP**	**TN**	**FP**	**FN**	**TP**	**TN**	**FP**	**FN**
ANF	0	107	4	1	0	107	4	1
CALC	2	105	2	0	2	107	0	0
CALCA	2	91	3	0	1	93	1	1
CCKN	2	102	0	2	2	102	0	2
GALA	2	88	0	0	1	88	0	1
GLUC	5	171	2	2	5	173	0	2
GON1	1	63	1	0	1	64	0	0
GRP	1	104	0	1	2	103	1	0
IGF1	0	96	4	1	0	100	0	1
IGF2	0	152	7	1	0	156	3	1
INS	2	77	0	0	2	77	0	0
MOTI	1	85	0	0	1	85	0	0
NEU2	1	135	2	0	1	136	1	0
NEUT	2	137	2	0	2	137	2	0
NMU	2	125	4	1	3	129	0	0
NPY	1	62	1	0	1	63	0	0
PACA	3	139	5	1	3	142	2	1
PAHO	1	49	1	0	1	49	1	0
PTHR	2	133	12	0	2	142	3	0
PTHY	1	86	3	0	1	88	1	0
SECR	2	126	1	0	1	127	0	1
SMS	1	86	0	1	1	86	0	1
TRH	0	218	14	0	0	220	12	0
TSHB	0	112	2	0	0	113	1	0
VIP	2	162	5	2	2	164	3	2
Total	36	2811	75	13	35	2851	35	14

	Known Motif Model				Human Model			
Neuropeptide Prohormone	**TP**	**TN**	**FP**	**FN**	**TP**	**TN**	**FP**	**FN**
Sensitivity		73.5%				71.4%		
Specificity		97.4%				98.8%		
CCR[2]		97.0%				98.3%		

[1]: TP: true positives; TN: true negatives; FP: false positives; FN: false negatives; positives=cleavage sites; negatives=non-cleavage sites

[2]: Correct classification rate.