General and Comparative Endocrinology 195 (2014) 107–115

Contents lists available at ScienceDirect

General and Comparative Endocrinology

journal homepage: www.elsevier.com/locate/ygcen

# Complex structural and regulatory evolution of the pro-opiomelanocortin gene family

Rayna M. Harris [a,b], Peter D. Dijkstra [a], Hans A. Hofmann [a,b,c,*]

[a] Department of Integrative Biology, The University of Texas at Austin, Austin, TX 78712, United States
[b] Institute for Cell and Molecular Biology, The University of Texas at Austin, Austin, TX 78712, United States
[c] Institute for Neuroscience, The University of Texas at Austin, Austin, TX 78712, United States

ARTICLE INFO

*Article history:*
Received 20 February 2013
Revised 8 October 2013
Accepted 10 October 2013
Available online 1 November 2013

*Keywords:*
*pomc*
ε-MSH
Melanocortin system
Gene duplication
Genome duplication
Cichlid

ABSTRACT

The melanocortin system is a neuroendocrine machinery that has been associated with phenotypic diversification in a number of vertebrate lineages. Central to the highly pleiotropic melanocortin system is the pro-opiomelanocortin (*pomc*) gene family, a family of pre-prohormones that each give rise to melanocyte stimulating hormone (MSH), adrenocorticotropic releasing hormone (ACTH), β-lipotropin hormone, and β-endorphin. Here we examine the structure, tissue expression profile, and pattern of *cis* transcriptional regulation of the three *pomc* paralogs (α1, α2, and β) in the model cichlid fish *Astatotilapia burtoni* and other cichlids, teleosts, and mammals. We found that the hormone-encoding regions of *pomc* α1, *pomc* α2 and *pomc* β are highly conserved, with a few notable exceptions. Surprisingly, the *pomc* β gene of cichlids and pomacentrids (damselfish) encodes a novel melanocortin peptide, ε-MSH, as a result of a tandem duplication of the segment encoding ACTH. All three genes are expressed in the brain and peripheral tissues, but *pomc* α1 and α2 show a more spatially restricted expression profile than *pomc* β. In addition, the promoters of each *pomc* gene have diverged in nucleotide sequence, which may have facilitated the diverse tissue-specific expression profiles of these paralogs across species. Increased understanding of the mechanisms regulating *pomc* gene expression will be invaluable to the study of *pomc* in the context of phenotypic evolution.



## 1. Introduction

Decades of research have demonstrated that genomic changes that alter the structure of proteins and/or their regulation can have profound effects on phenotype (Carroll, 2008; Li and Johnson, 2010; Tirosh et al., 2008; Van de Peer et al., 2009; Wray, 2007; Wray et al., 2003; Yokoyama et al., 2010). Much of our knowledge regarding the influence of structural and regulatory contribution to phenotypic diversity comes from studies examining these two mechanisms in isolation. For example, a mutation in the coding region of the melanocortin-1 receptor (*mc1r*) contributes to adaptive variation in pigmentation patterns between two subspecies of oldfield mice (*Peromyscus polionotus*) such that a mainland subspecies has a cryptic dark brown dorsal coat, while a beach-dwelling subspecies has evolved a lighter coat that is well camouflaged on pale coastal sand dunes (Steiner et al., 2007). The patterning of body armor plates of stickleback fish provides a compelling example of regulatory evolution as the differences between marine and freshwater populations are due to changes in the regulatory logic of the gene encoding the secreted signaling molecule ectodysplasin (EDA) (Knecht et al., 2007). Gene and genome duplication and deletion events have long been associated with phenotypic diversity (Buchanan and Scherer, 2008; Girirajan and Eichler, 2010; Meyer and Schartl, 1999; Ohno, 1970). Once two copies of a gene exist, immediate or subsequent alteration of the coding sequence (non-synonymous substitutions, insertions, deletions) of one paralog can give rise to new proteins (with novel functional properties or a lack of functionality), and changes in the regulatory sequence can facilitate sub- or neo-functionalization and specialization. Despite these important insights, the consequences of concurrent structural and regulatory changes have rarely been examined in the same system.

The melanocortin system is a complex neuroendocrine machinery that has been associated with phenotypic diversification in a number of vertebrate lineages. Within this system, the pro-opiomelanocortin (*pomc*) gene family is an ideal candidate for three reasons: its evolutionary past is replete with a variety of duplication events, it is central to the highly pleiotropic melanocortin system, and it displays remarkable tissue specificity (Dores and Lecaude, 2005; Ducrest et al., 2008). Finally, recent insights into *pomc cis*-regulatory regions have shed light on the

* Corresponding author at: Department of Integrative Biology, The University of Texas at Austin, 2401 Speedway – C0990, Austin, TX 78712, United States. Fax: +1 512 471 3878.
*E-mail address:* hans@utexas.edu (H.A. Hofmann).

0016-6480/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ygcen.2013.10.007

mechanisms that give rise to tissue specific gene expression (Kobayashi et al., 2008).

It is generally accepted that *pomc* arose from an opioid-encoding gene following the 1R genome duplication event in the early chordate ancestor. The α- and β-MSH and β-END emerged with the subsequent radiation of jawless vertebrates (Dores and Baron, 2011). The 2R genome duplication event gave rise to a γ-MSH peptide, but γ-MSH was subsequently lost in ray-finned fishes (Dores and Baron, 2011; Takahashi and Kawauchi, 2006). Notably, the teleost ancestor underwent a third genome duplication that gave rise to *pomc* gene duplicates (Braasch and Salzburger, 2009). Among ray-finned fishes, *pomc* paralogs have been identified in salmoniformes (Okuta et al., 1996), cypriniformes (Arends et al., 1998; Gonzalez-Nunez et al., 2003), acipenseriformes (Alrubaian et al., 1999), perciformes (Cardoso et al., 2011), smegamorpha (Wunderink et al., 2012), tetraodontiformes (De Souza et al., 2005a), and pleuronectiformes (Kim et al., 2009; Takahashi et al., 2005).

*pomc* genes consists of three exons and two introns. Exon 1 and intron 1 are untranslated, but exon 2 and 3 encode a preprohormone (POMC). Exon 2 encodes a signal peptide and the N-terminal peptide, while exon 3 gives rise to the hormonal segments (De Souza et al., 2005a; Gonzalez-Nunez et al., 2003). POMC is cleaved into numerous functional peptides, including (but not limited to): melanocyte stimulating hormone (MSH), adrenocorticotropic releasing hormone (ACTH), β-lipotropin (β-LPH), and β-endorphin (β-END). Post-translational processing of POMC is tissue specific. For example, in the mammalian anterior pituitary, proprotein convertase (PC) 1/3 endoproteolytically cleaves POMC at paired basic amino acid sites (KK, KR, RK, and RR) to yield ACTH and β-LPH; however, in the mammalian intermediate pituitary and the hypothalamus PC1/3 and PC2 cleave POMC into α-, β-, γ-MSH, and β-END (Chen et al., 2011). Following endoproteolytic cleavage, some hormonal segments undergo further tissue- and species-specific modifications such as C-terminal amidation and N-acetylation (Dores and Baron, 2011).

POMC peptides affect a multitude of physiological processes such as energy homeostasis (Coll and Loraine Tung, 2009), food intake and glucose balance (Xu et al., 2011), stress response (Cerda-Reverter et al., 2011), pigmentation (Cerda-Reverter et al., 2011), and behavior (Ducrest et al., 2008; Mountjoy, 2010). These functions are mediated by several melanocortin receptor subtypes (MC1R, MC2R, MC3R, MC4R, and MC5R). It has long been known that the HFRW motif found in all MSH peptides is necessary for binding to the MC1R, MC3R, MC4R, and MC5R (Cone, 2006; Dores, 2013), and the opioid receptor binding motif in the region encoding β-END is YGGF (Dores, 2013). Increasing evidence implicates the KKRRP motif for MC2R (aka ACTH-R) binding (Costa et al., 2004; Liang et al., 2013). The affinity of each peptide for a given receptor has been characterized for a few species (Cone, 2006).

It is becoming increasingly clear that teleost *pomc* paralogs show tissue- and/or behavior-specific expression patterns and thus have likely been sub-functionalized. For example, in barfin flounder, *Verasper moseri*, *pomc A* is found only in the pituitary; *pomc B* is expressed in the brain, pituitary, and a few peripheral tissues: and *pomc C* is widely expressed and is the only pomc expressed in the skin (Takahashi et al., 2005). In the gilthead sea bream, *Sparus arautus*, *pomc α1* and *pomc α2* are expressed in brain, pituitary and periphery but not the skin (Cardoso et al., 2011). No *pomc β* has been isolated in this species. In Tetraodon, *pomc α* is expressed in the nucleus lateralis tuberis (NLT) of the hypothalamus and the pars intermedia (PI) of the pituitary, while *pomc β* is expressed in the pre-optic area (POA) and PI of the pituitary (De Souza et al., 2005a). Leder and Silverstein (2006) showed in the rainbow trout, *Oncorhynchus mykiss*, that all three *pomc* paralogs (including a splice variant) are widely expressed throughout the brain and periphery, yet only *pomc A1* appears to be involved in the feeding response. Finally, in the Senegalese sole, *Solea senegalensis*, *pomc* paralogs appear to serve different physiological roles as *pomc*-a expression in the pituitary is down-regulated in chronically stressed juveniles, whereas *pomc*-b expression levels remain unaffected (Wunderink et al., 2012).

A few studies have examined *pomc* promoter (the region upstream of the transcription start site) sequence to identify *cis*-regulatory sequences that regulate tissue-specific expression patterns. Analyses that compared the *pomc* promoter of tetrapods (which possess only one *pomc* gene) with the *pomc α* gene of *Tetraodon*, fugu, and zebrafish have identified a conserved 10 base pair (bp) motif that appears to be involved in pituitary- and hypothalamus-specific expression (Bumaschny et al., 2007; De Souza et al., 2005b). Kobayashi et al. (2008) examined the promoter the three *pomc* paralogs of the barfin flounder, *V. moseri*, and identified a number of ubiquitous as well as paralog-specific *cis*-regulatory elements (Kobayashi et al., 2008).

In the present study, we obtained sequence data from whole genome and candidate gene sequencing efforts and analyzed the *cis*-regulatory and protein coding sequence evolution of the *pomc* gene family with a focus on the radiation of cichlid fishes in East Africa. It seems likely that simultaneous structural and regulatory evolution of *pomc* during the radiation of cichlid fishes has contributed to their remarkable behavioral, physiological, and phenotypic diversity (Hofmann, 2003; Kocher, 2004). Because the genomes and tissue-specific transcriptomes of five cichlid fishes have recently been sequenced, we have the exceptional opportunity to examine the structural and *cis*-regulatory evolution of the *pomc* family in this group. Furthermore, the model cichlid *Astatotilapia burtoni* has become a powerful study system in social neuroscience to examine the neural, molecular, and genomic bases of social plasticity, and provides an excellent model system in which to examine these effects (Hofmann, 2003; Maruska et al., 2013; Robinson et al., 2008). In *A. burtoni*, body coloration, social behavior, and physiology are often correlated in males, suggestions underlying pleiotropic mechanisms. To test the extent to which cichlid *pomc* paralogs have undergone structural and evolution, we reconstructed the phylogenetic relationships of the coding sequences of *pomc* genes and POMC peptides. Then, we examined mRNA expression patterns for sub-functionalization and specialization. Finally, we examined *cis*-regulatory motifs in the *pomc* promoters using sequences to look for evidence for regulatory evolution.

## 2. Materials and methods

### 2.1. Animals

Males of the haplochromine cichlid species *A. burtoni* were obtained from lab-reared stocks and housed in 110 L tanks in naturalistic communities under conditions mimicking their native Lake Tanganyika (pH 8.5 ± 0.2; 27 ± 0.2 °C; 12:12 h. photoperiod). Adult neotropical convict cichlids, *Amatitlania nigrofasciata*, were also obtained from a laboratory-reared stock and housed in single-sex groups under the same conditions as above. An adult three-spot domino damselfish, *Dascyllus trimaculatus,* was obtained directly from the pet trade. All experiments were carried out in accordance with the Institutional Animal Care and Use Committee at The University of Texas at Austin.

### 2.2. Characterization of pomc

The complete *A. burtoni pomc α2* mRNA sequence was obtained from GenBank (accession number 89322545). To obtain *A. burtoni pomc* α1 and β, *A. nigrofasciata pomc* β, and *D. trimaculatus pomc* β sequences, we designed primers against the predicted *Oreochromis*

*niloticus* POMC α1 and β sequences using Primer3 (http://frodo.wi.mit.edu/primer3/; Rozen and Skaletsky, 2000). See Supplementary Table 1 for primer details. RNA was extracted from whole brain using Trizol (Life Technologies), and cDNA was made using Superscript II reverse transcriptase, oligo(dT), and random hexamers (Life Technologies) following the manufacturer's instructions. Using whole brain cDNA as template, we employed a touchdown PCR strategy that began with an annealing temperature of 60 °C and subsequently reduced the temperature by 1 °C per cycle for 8 cycles. We then continued for 32 more cycles with an annealing temperature of 52 °C. After confirmation of correct fragment size by electrophoresis on a 1% agarose gel, the PCR products were purified and sequenced. For *A. burtoni pomc* α1 and β and *A. nigrofasciata pomc* β, the entire coding region along with the 5′ and 3′-untranslated regions were amplified using the SMARTer RACE kit (CloneTech) according to the manufacturer's instructions using gene-specific nested primers with annealing temperatures of 70 °C. See Supplementary Table 1 for primer details.

Using the empirically determined *A. burtoni pomc* sequences as query sequences, we accessed BouillaBase.org and employed BLAST search to obtain the *pomc* DNA sequences for all sequenced cichlid genome assemblies (March 2011 release of Tilapia_broad_ V1, Metriaclima_zebra_broad_v0, Astatotilapia_burtoni_broad_v1, Neolamprologus_brichardi_broad_v1, and Pundamillia_nyererei_broad_v1). We then repeated this process with the combined transcriptomes available for the same species (March 2012 release of Tilapia_broad_combined_transcriptomes, Metriaclima_zebra_broad_combined_transcriptomes, Astatotilapia_ burtoni_broad_combined_transcriptomes, Neolamprologus_brichardi_broad_combined_transcriptomes, and Pundamillia_nyererei_broad_combined_transcriptomes) to confirm exon/intron boundaries, transcription start sites (defined as the most 5′ transcribed nucleotide processed reads for each paralog), and the length of the 3′UTRs. In July 2012, we accessed the publically available *Xiphophorus maculatus* genome using the NCBI's whole genome blast server to predict the *X. maculatus pomc* genes.

### 2.3. Phylogenetic and evolutionary analyses of melanocortin peptides

To examine the evolutionary history of POMC preprohormones, we generated bootstrapped maximum likelihood trees based on the Whelan and Goldman model (Whelan and Goldman, 2001). Phylogenetic analyses were conducted using *MEGA* version 5 (Tamura et al., 2011). We then predicted which peptides would be produced by the canonical dibasic cleavage sites (KR, RR, RK, and KK) for cleavage enzymes PC1/3 and PC2 (Tanaka, 2003). Alignments of POMC β pre-prohormone sequences and of α-, β-, γ-, δ-, and ε-MSH sequences were created using Clustal Omega (Goujon et al., 2010; Sievers et al., 2011). Accession numbers for sequences used in the phylogenetic trees and sequence alignments are listed in Supplementary Table 2.

### 2.4. Tissue-specific gene expression analysis

We conducted an *in silico* analysis of *pomc* genes expression by querying (via BLAST) the publicly available transcriptomes of five cichlid species (BouillaBase.org) using either the complete *pomc* coding sequences or a single exon of *A. burtoni*. The available transcriptomes were made from pooled RNA the following tissues: blood, whole brain, eye, heart, kidney, liver, muscle, skin, ovary, testis, and early and late stage embryos. All transcripts were reciprocally blasted back to the genome to confirm sequence specificity.

In order to validate these patterns we conducted RT-PCR in *A. burtoni* on all these tissues as well as grossly dissected brain regions for a more detailed spatial resolution. *A. burtoni* males were used to collect grossly dissected brains while male and females were used for collecting peripheral tissues and whole brains. Tissues were rapidly dissected and stored in RNALater (Ambion), except blood was immediately frozen in Trizol (Life Technologies). RNA was extracted using Trizol, DNase-treated (TURBO DNA-free, Life Technologies), and pooled. cDNA was prepared using Superscript II reverse transcriptase, oligo(dT), and random hexamers (Life Technologies) following the manufacturer's instructions. Gene-specific primers were designed using Primer3 (Rozen and Skaletsky, 2000). See Supplementary Table 1 for primer details. We performed PCR with a 60 °C annealing temperature for 40 cycles using FastStart PCR Master Mix (Roche Applied Science). Amplification products were visualized by gel electrophoresis on a 1% agarose gel using ethidium bromide staining.

### 2.5. Comparative promoter analysis

To examine *pomc* regulatory evolution across cichlids, teleosts, and mammals, we first identified the transcription start site (the first nucleotide of exon 1) of all of the available sequences. (In a few cases the Ensembl record contained little or no sequence from exon 1, so we obtained this sequence from a GenBank submission, or vice versa). For within-cichlid comparisons, we compared 1500 bp promoter sequence, which has previously been shown to allow for meaningful analysis of the proximal and distal promoter (O'Leary et al., 1996; Zhou et al., 2010). For comparisons across vertebrates, we used 500 bp promoter sequences, because longer sequences were not available for most species. We used MatInspector (Cartharius et al., 2005) to predict transcription factor binding sites (TFBS) in the promoter in order to include the proximal and more distal promoter sequences. Data were imported into Microsoft Excel for further analysis. We focused on "matrix families" (i.e., groups of similar TFBS) (Cartharius et al., 2005) to generate a broad view of classes of transcription factors likely to bind to any given promoter (see Supplementary Table 4 for abbreviations and descriptions of matrix families identified in this study). We created a list of all the predicted matrix families and used the online tool Wordle (http://www.wordle.net/) to visualize the relative abundance of putative TFBSs as word clouds. Gene structure figures were made using Adobe Illustrator CS2.1. Correlations between the abundance of each transcription factor binding motif (matrix family) were calculated using SPSS.

## 3. Results

### 3.1. Description and phylogenetic analysis of pomc

In *A. burtoni*, we identified three *pomc* genes, which we designated *pomc α1*, *pomc α2*, and *pomc β*. All three *pomc* paralogs consist of two coding exons and one untranslated exon in the 5′UTR (Fig. 1), which is consistent with most vertebrate *pomc* genes (De Souza et al., 2005b). Our phylogenetic analysis suggests that *pomc α* and *β* arose during the teleost-specific whole genome duplication event (Fig. 2). *pomc α* has undergone numerous independent duplication events during teleost evolution such that *A. burtoni pomc α1* and *pomc α2* are co-orthologous to the *pomc* paralogs found in acipenseriformes (sturgeon fishes), cypriniformes (carp), salmoniformes (trout and salmon), and smegmamorpha (swordtail) (Fig. 2). We cannot determine whether *pomc α* was duplicated in the perciform lineage immediately before or after pleuronectiformes (flounder) split from perciformes (such as cichlids), as sequence divergence following the duplication event resulted in poor resolution of the species in this lineage. No *pomc α* duplication has been observed in medaka, fugu or Tetraodon, and no duplicate *pomc β* gene has been identified in any teleost to date.

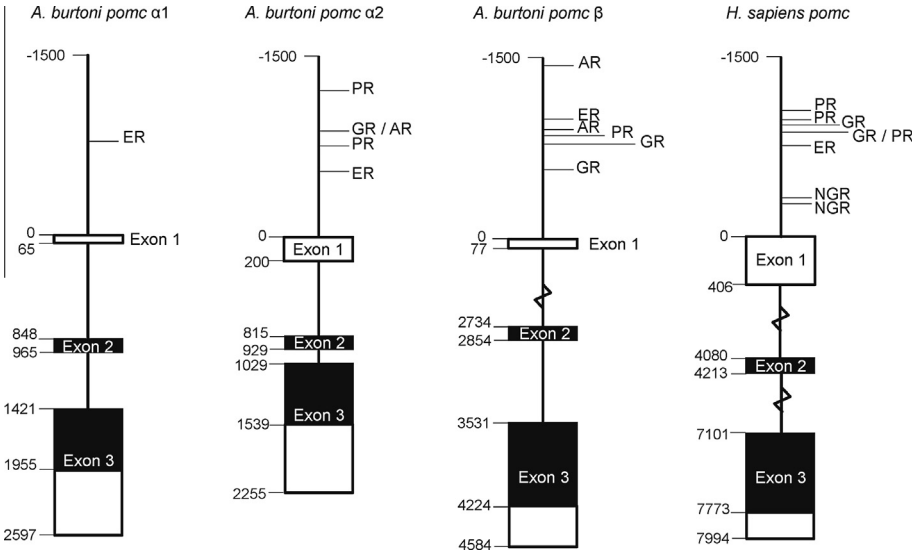


**Fig. 1.** *pomc* structure with predicted transcription factor binding sites. *A. burtoni pomc* α1 (far left), α2 (middle left), and β (middle right) and human *pomc* (far right) genes are depicted. White boxes represent the 5′ and 3′ untranslated regions, black boxes represent exons, and black lines represent genomic regions. Androgen (AR), Estrogen (ER), Progesterone (PR), and Glucocorticoid (GR or NGR) receptor binding sites are shown.

### 3.2. Tandem duplication of a novel peptide-encoding sequence within the pomc β gene

We predicted that *A. burtoni* POMC α1 and POMC α2 can be cleaved into α-MSH, β-MSH, ACTH, and β-END and that POMC β would be cleaved into α-MSH, ACTH, and β-MSH. However, when characterizing the *A. burtoni*, *Metriaclima zebra*, *Neolamprologus brichardi*, *O. niloticus and Pundamilia nyererei pomc* β genes, we discovered that each contained a tandem duplication (Fig. 3). The source of the duplication was a 90 bp (30 amino acid) fragment that encompassed part of the N-terminal fragment, all of α-MSH, and part of ACTH. By isolating the *pomc β* gene from *A. nigrofasciata* (a representative of the neotropical cichlids, which diverged from old-world cichlids ca. 65–85 million years ago (Hulsey et al., 2010) and the three-spot damselfish, *D. trimaculatus* (a representative of the pomacentridae that are closely related to cichlidae (Wainwright et al., 2012) we confirmed that the tandem duplication arose before the radiation of cichlids but sometime after the radiation of the Labroidea (Fig. 3).

This duplication likely encodes a novel peptide, which we have named ε-MSH, as α-, β-, γ-, and δ-MSH have previously been described (for review see Dores and Baron, 2011). ε-MSH contains the canonical melanocortin receptor binding motif (HFRW), which is critical for melanotropic activity, surrounded by the canonical peptide cleavage enzyme binding sites (RR and RR), yet it lacks the ACTH receptor binding motif (KKRRP) (Figs. 3 and 4). The cleavage product from this duplication is either 28 amino acids (long) or 18 amino acids (short) in length depending on whether monobasic cleavage enzymes can process a short peptide (Fig. 4). An alignment shows that it is striking similar to α-MSH but is increasingly less similar to γ-, β-, and δ-MSH, respectively.

### 3.3. Tissue-specific gene expression in African cichlid fishes

The persistence of three *pomc* genes in many teleost lineages suggests that they may have acquired specialized functions. In order to test this hypothesis we examined tissue-specific expression profiles in the five species of cichlids for which large-scale transcriptome data sets are available. *pomc* transcripts showed both remarkable tissue specificity and variable expression profiles across species (Supplementary Table 2). The brain and embryo were found to express all three paralogs in at least one species. In all species, only *pomc β* was expressed in the skin. Most tissues do not show a conserved expression pattern across all species. In the haplochromine cichlids *A. burtoni*, *P. nyererei*, and *M. zebra*, *pomc β* is the most widely expressed paralog.

We then used RT-PCR on the same set of tissues used to generate the *A. burtoni* transcriptomes. We find a high degree of concordance when the gene appears to be highly expressed but less concordance when the gene is expressed at lower levels. Specifically, the RT-PCR confirmed that *pomc β* is the most widely expressed paralog, that all three paralogs are expressed in the brain, and that only *pomc β* is expressed in the skin. In three tissues (eye, heart, and embryo) the transcriptome analysis appeared more sensitive than the RT-PCR at detecting gene expression, whereas in eye and kidney RT-PCR detected gene expression that was missed by the transcriptome analysis (Fig. 5).

Next, we examined expression patterns of *pomc* paralogs in more detail in five grossly dissected brain regions, pituitary, skin, and testes of *A. burtoni* males (Fig. 4). As with the peripheral tissues, pomc *α1* shows a more restricted expression pattern while *pomc α2* and *pomc β* are more widely expressed. Specifically, *pomc α1* is expressed only in the hypothalamus, cerebellum, and the pituitary. *pomc α2* is expressed in the telencephalon, hypothalamus, cerebellum, hindbrain, and pituitary. Finally, *pomc β* is expressed in the telencephalon, hypothalamus, optic tectum, cerebellum, medulla, and pituitary.

### 3.4. Comparative analysis of pomc promoters

We hypothesized that the tandem duplication in *pomc β* might be associated with novel regulatory motifs indicative of a possible new function for this gene. We specifically asked where along the promoter steroid hormone receptor binding sites are predicted (Fig. 1). We found estrogen receptor (ER) binding sites in all *A. burtoni pomc* paralogs as well as the human *pomc*. Glucocorticoid (GR) and progesterone (PR) receptor binding sites were predicted in *A. burtoni pomc α2* and *β* and human *pomc*. Androgen receptor (AR) binding sites were only predicted in *A. burtoni pomc α2* and *β* while negative GR binding sites (NGR) were only predicted in human *pomc*.

Next, we looked at all the predicted *cis*-regulatory motifs and found that while many were predicted in all *pomc* sequences, their abundance varied strikingly across species and paralogs (Suppl.

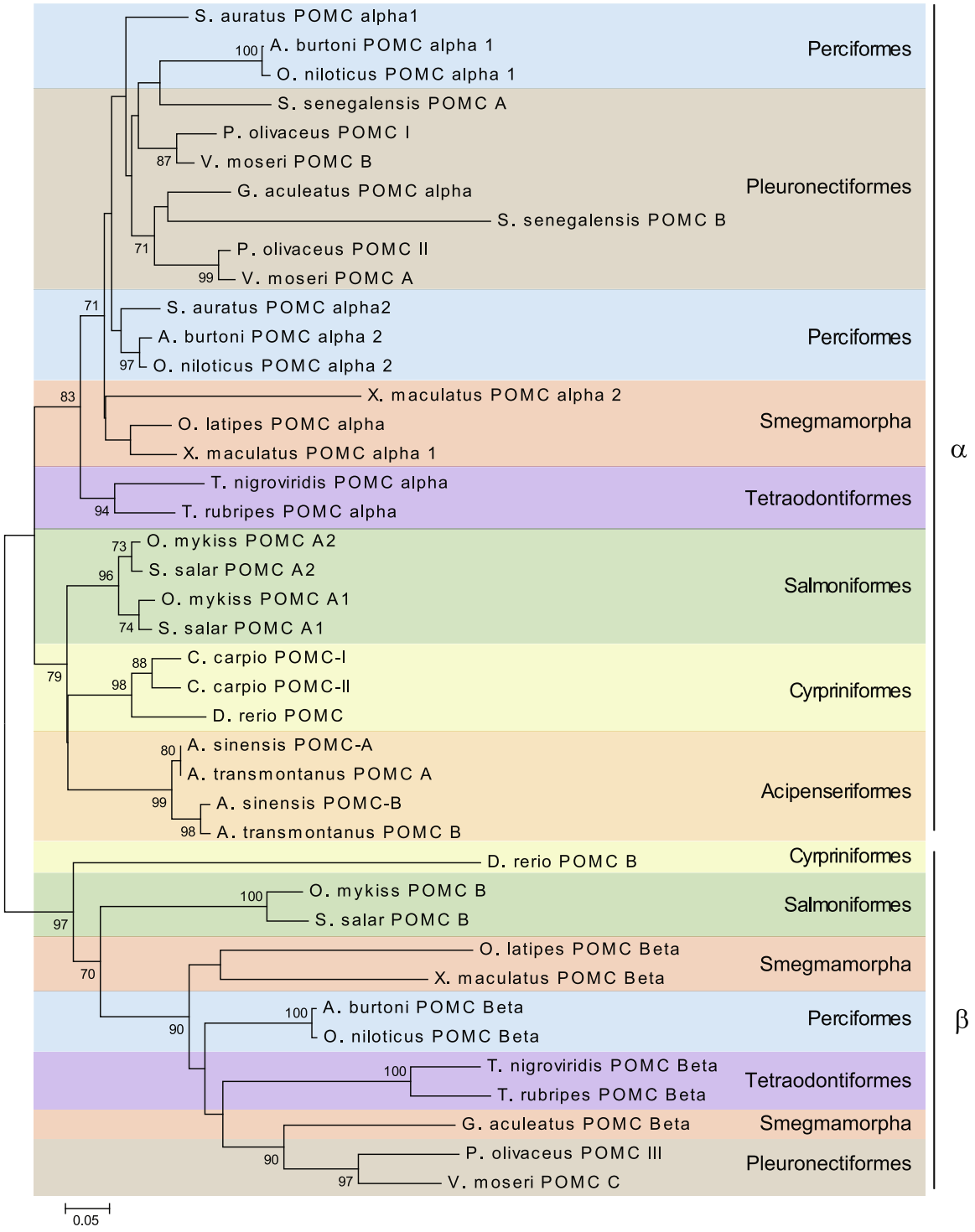


**Fig. 2.** Phylogenetic comparison of ray-finned fishes, POMC sequences. A maximum likelihood tree shows that the evolution of POMC in ray-finned fishes is characterized by multiple gene duplication events. The POMC α clade is characterized by several independent duplication events, while the *pomc* β clade likely arose only once during the teleost whole genome duplication. Genbank accession Nos. and species information are given in Supplementary Table 2.

Fig. 1A). We then asked whether the abundance of predicted transcription factor binding sites is correlated across all paralogs and orthologs examined. Looking only at cichlids (Suppl. Fig. 1B, green shading), we found the greatest similarity for the *pomc β* promoter ($r > 0.94$), followed by *pomc α2* ($r > 0.70$) then *pomc α1* ($r > 0.69$). We also found that *pomc β* promoters were more similar to those of *pomc α1* ($0.58 > r > 0.78$) than *pomc α2* ($0.49 > r > 0.65$). When comparing across teleost (Suppl. Fig. 1B, purple shading) and across vertebrates (Suppl. Fig. 1B, pink shading) we find that there is still considerable similarity in overall abundance of regulatory motifs.

## 4. Discussion

### 4.1. pomc coding sequences

Our comparison of the *pomc* protein coding sequence revealed a complex pattern of evolution. *pomc α* duplication events have

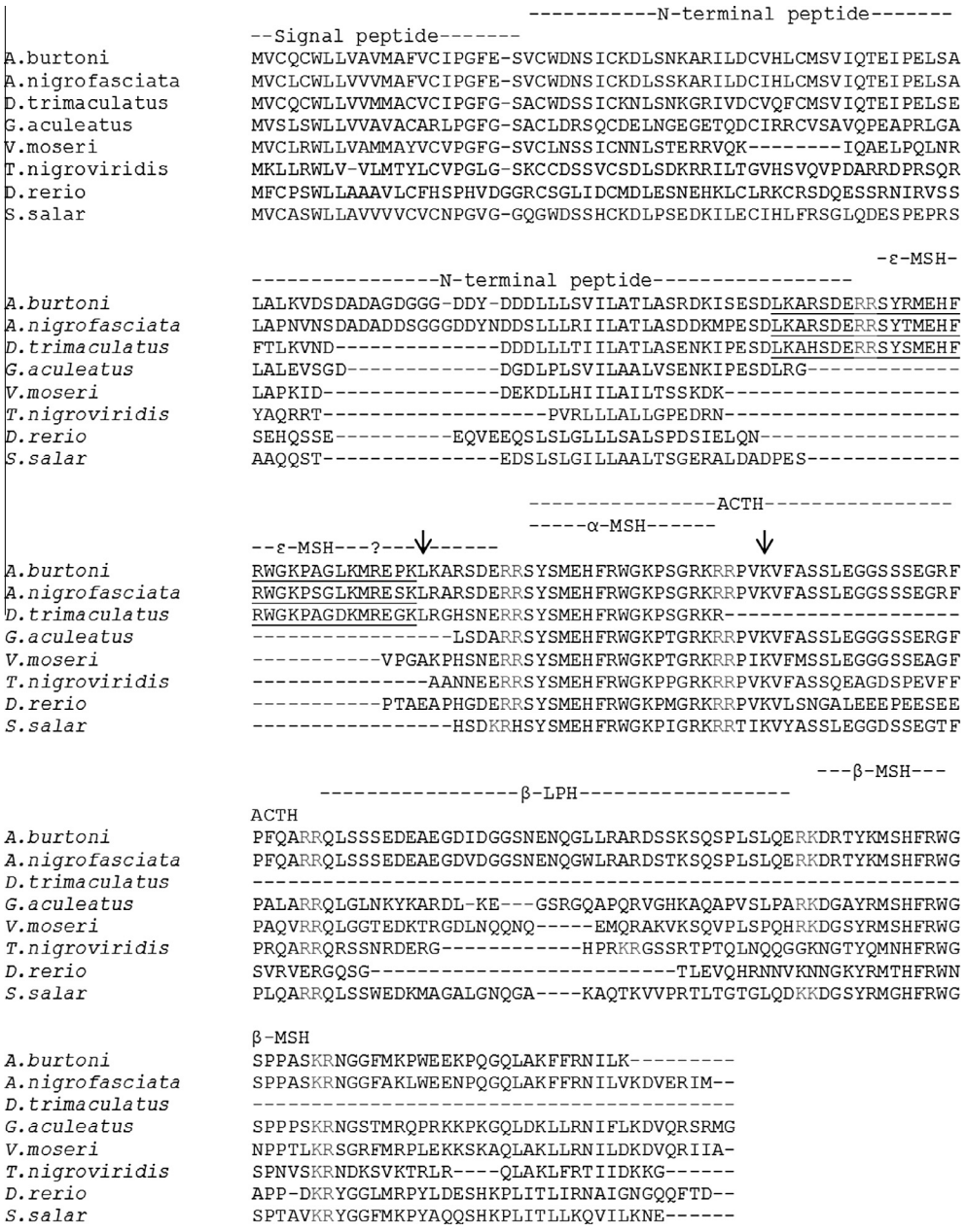


**Fig. 3.** Alignment of POMC β sequences for representative species from major teleost lineages, demonstrating the presence of a novel ε-MSH in cichlids and pomacentrids. The predicted cleavage products for the signal, N-terminal, and hormonal peptides for, *A. burtoni* (African cichlid), *A. nigrofasciata* (Central American cichlid), and *D. trimaculateus* (damselfish) are indicated above the aligned amino acids. The underlined segment marks the novel duplicated region while the arrow point to the origin of the duplication. Dibasic cleavage sites are indicated by grey letters; a possible monobasic cleavage site that would give rise to a short form of ε-MSH is indicated with a question mark. GenBank accession numbers can be found in Supplementary Table 2.

occurred multiple times in the ray-finned fishes, such that the acipenseriformes *pomc* α sequences are co-orthologs to the *pomc* α genes of salmoniformes and perciformes, for example. Despite the large amount of sequence data available for *pomc* α sequences from species in the teleost lineage Percomorpha, some of these alpha paralogs have quickly diverged in structure and coding sequence (*Sparus auratus*: Cardoso et al., 2011; *V. moseri*: Takahashi et al., 2005; *S. sole*: Wunderink et al., 2012), reducing our ability to identify orthologs or determine timing of duplication within the Percomorpha lineage. The *pomc* β gene arose from the Teleost Whole Genome Duplication (Braasch and Salzburger, 2009). Subsequently, the opioid receptor binding motif in the region encoding β-END has degenerated (from YGGF to NGGF) and lost is function in all known teleost *pomc β* sequences.

Unexpectedly, the cichlid and pomacentrid *pomc β* has further specialized in melanocortin function. In this gene, a 90 bp (30 amino acid) tandem duplication resulted in the gain of a novel melanocortin peptide product, ε-MSH, immediately upstream of the segment encoding α-MSH and ACTH. This duplication includes a PC1/3 cleave site (RR) followed by the canonical MC1R, MC4R, and MC5R binding motif (HFRW) (Dores and Baron, 2011), yet lacks the mc2r binding motif (KKRRP) (Dores, 2013; Liang et al., 2013). It is possible that cleavage at a monobasic arginine base could yield a peptide that is only 1 amino acid longer than α-MSH. Regardless, it is clear that this sequence is more similar to α-MSH then to ACTH, γ-, β-, and δ-MSH, and therefore likely functions similar to α-MSH. The presence of a tandem duplication in the coding region of the *pomc β* gene in both the Cichlidae and

```
H portusjacksoni pomc gMSH      ---LRNYVMGHFRWNKFGKK----------------------------
P annectens pomc gMSH           ----RKYMMTHFRWDKFGRR----------------------------
H sapiens pomc gMSH             ----RKYVMGHFRWDRFGRR----------------------------
A burtoni pomc B eMSH long      ---RRSYRMEHFRWGKPAGLKMREPKLKARSDERR-------------
A burtoni pomc B eMSH short     ---RRSYRMEHFRWGKPAGLKMR-------------------------
A burtoni pomc B ACTH           ---RRSYSMEHFRWGKPSGRKRRPVKVFASSLEGGSSSEGRFPFQARR
A burtoni pomc B aMSH           ---RRSYSMEHFRWGKPSGRK---------------------------
T rubripes beta ACTH            ---RRSYSMEHFRWGKPPGRKRRPVKVFSSSQEGGDSFEAAFPRQTRR
T rubripes beta aMSH            ---RRSYSMEHFRWGKPPGRK---------------------------
H portusjacksoni pomc ACTH      ---KRSYSMEHFRWGKPMGMKRKPVKVYPNNFEDGSVEN--MGPRLKR
H portusjacksoni pomc aMSH      ---KRSYSMEHFRWGKPMGMKR--------------------------
P annectens pomc ACTH           ---RRSYSMEHFRWGKPVGKKRRPIKVYPNGMEDESAES--YSPDLRR
P annectens pomc aMSH           ---RRSYSMEHFRWGKPVGKK---------------------------
H sapiens pomc ACTH             ---KRSYSMEHFRWGKPVGKKRRPVKVYPNGADDESAEA--FPLEFKR
H sapiens pomc aMSH             ---KRSYSMEHFRWGKPVGKK---------------------------
H portusjacksoni pomc dMSH      KKDGKIYKMTHFRWGRG-------------------------------
H portusjacksoni pomc bMSH      KKDGEDYKFGHFRWSVPLKDKR--------------------------
T rubripes pomc B bMSH          -GRNGTYQMSHFRWGSPNVSERSKK-----------------------
A burtoni pomc B bMSH           -RKDRTYKMSHFRWGSPPASKR--------------------------
P annectens pomc dMSH           -KKDRNYRMQHFRWNSPPKDKR--------------------------
P annectens pomc bMSH           -KKDRNYRMQHFRWNSPPKDKR--------------------------
H sapiens pomc bMSH             KKDEGPYRMEHFRWGSPPKDKR--------------------------
                                    * : ****
```

**Fig. 4.** Alignment (using Clustal Omega) of ε-MSH with α-, β-, γ-, and δ-MSH from representative lineages: α-, β-, γ-, and δ-MSH from *H. protusjacksoni*, an acipenseriform; α-, β-, and γ-, from *H. sapiens* and *P. annectens*; α- and β-MSH from *T. nigroviridus pomc* β ; α-, β-, and ε-MSH *A. burtoni pomc* β. Clustering confirms that delta δ-MSH duplicated from β-MSH and that ε-MSH recently duplicated from a segment encoding a fraction of the α-MSH/ACTH encoding region of *pomc* and inserted itself 21 nucleotides (7 amino acids with single underline) upstream of the leading dibasic cleavage site for α-MSH (double underline). Boxes surround the MC2R binding motif (KKRRP) as confirmed by (Liang et al., 2013). Dashed boxes indicate putative MC2R motifs whose binding affinity is not known. If monobasic cleavage is possible (at an R rather than RR), the ε-MSH could be cleaved within the MC2R binding site (referred to as "short form" in Fig. 4). GenBank accession numbers for all sequences are provided in Supplementary Table 3.

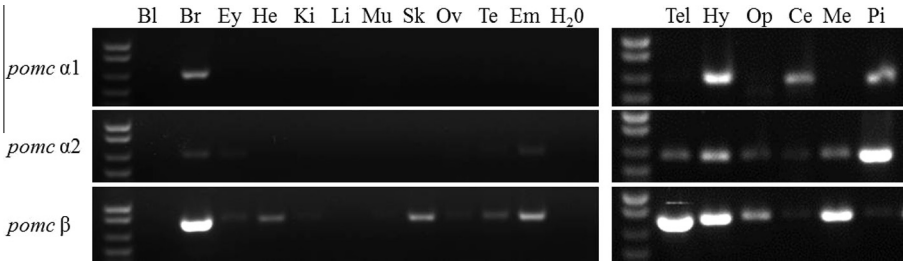


**Fig. 5.** Expression profiles of *A. burtoni pomc* α1 in central and peripheral tissues. Bl: blood, Br: brain, Ey: eye, He: heart, Ki: kidney, Li: liver, Mu: muscle, Sk: skin, Ov: ovary, Te: testis, Em: pooled early and late stage embryo, Tel: telencephalon, Hy: hypothalamus, Op: optic tectum, Ce: cerebellum, Me: medulla, Pi: pituitary.

Pomacentridae families is consistent with the current view that cichlids are more closely related to the pomacentrids than to the labroids (Wainwright et al., 2012; Westneat and Alfaro, 2005). While the discovery of this novel melanocortin peptide was unexpected, it is not unprecedented, and it reflects a common pattern in the MSH subtype evolution by tandem duplication (Dores and Baron, 2011). For example, in the lineage leading to modern sharks and rays, β-MSH duplicated and gave rise to δ-MSH (Amemiya et al., 1999a,b; Dores et al., 2003).

### 4.2. Spatial patterns of pomc paralog expression

We found that most tissues express only one or two paralogs, which supports the hypothesis that the pomc paralogs have undergone sub- or neofuncitonalization. Interestingly, across all five cichlid species, *pomc* β was the only paralog expressed in skin tissue, suggesting an important role in skin specific processes such as pigmentation. Only the whole brain and embryonic tissue express all three paralogs, which is not surprising given that these tissues are quite heterogeneous in nature.

The tissue-specific expression results for the brain, testis, and skin were very concordant between our PCR-based analysis and the BROAD transcriptomes, with few exceptions. It is possible that the discrepancy is due to low levels of expression being hard to reliably detect by either method or that gene expression in this tissue is highly plastic. Nevertheless, it is clear that some paralogs are more highly expressed in the tissues analyzed than others. Many studies have found low levels of expression of pomc expression in a wide variety of species (Karsi et al., 2004; Leder and Silverstein, 2006; Takahashi et al., 2005).

The results from RT-PCR on grossly dissected brains show that *pomc α1* is selectively expressed centrally while *pomc α2* and *β* producing neurons are more widely distributed throughout the brain. *pomc β* appears to have acquired a more specific endocrine function in the skin and gonads, while the *pomc α2* is most abundant in the pituitary. Similar to the studies on barfin flounder (Takahashi et al., 2005) and and gilthead sea bream (Cardoso et al., 2011), we found all three paralogs to be widely expressed, but some paralogs appear to have a more prominent role in the skin or pituitary. More studies examining the spatial distribution of the *pomc* expressing neurons (like De Souza et al., 2005a) are needed.

### 4.3. Regulatory evolution of pomc promoters

Following gene duplication, mutations can arise in each *pomc* sequence that may lead to loss of ancestral function and/or gain of a novel function. By examining both the tissue distribution patterns and the *cis*-regulatory motifs of each paralog, we can gain insight into which tissues may be pleiotropically regulated across a few or many cell and tissue types. The divergent expression

patterns indicate that, in addition to coding sequence evolution, the regulatory logic of these paralogs has also evolved. We found that the steroid hormone binding profile was distinct across *A. burtoni pomc* paralogs, and that none had a profile similar to human (Fig. 1). Steroid hormones activate intracellular receptors that act as transcription factors to regulate transcription of many genes (Mangelsdorf et al., 1995), including many neuropeptides and their receptors. Drouin et al. (1989) identified a *cis*-regulatory element, the negative glucocorticoid response element (NGRE) that mediates glucocorticoid dependent repression of *pomc* transcription in the rat. The human *pomc* promoter also contains a NGRE (Fig 1.), but none of the cichlid *pomc* promoters contain a NGRE. Our analysis did not find any NGRE *cis*-regulatory elements in *A. burtoni pomc* genes (Fig. 1). ERα colocalized with POMC expression neurons in the mammalian arcuate nucleus and has been shown to bind to a hypothalamic *pomc* enhancer (nPE2) (de Souza et al., 2011). While the *A. burtoni pomc α2* possesses a putative ER binding site, an ER antagonist did not alter *pomc α2* gene expression in dominant or subordinate *A. burtoni* males (O'Connell and Hofmann, 2012). Other studies suggest that ER has a stimulatory effect on *pomc* levels in females. It would be interesting to test whether these transcription factors indeed preferentially bind specific *pomc* paralogs under different physiological and environmental conditions.

*cis*-Regulatory motif prediction is an emerging field for examining evolution of regulatory sequences, yet only a subset of these motifs have been functionally confirmed. Nevertheless, given that non-coding regions cannot be analyzed in terms of synonymous and non-synonymous substitutions, the analysis of gain/loss of a *cis*-regulatory motifs provides important initial insights. We found that while many canonical motifs were predicted in all *pomc* sequences, their abundance varied strikingly across species and paralogs, suggesting that these sequences are undergoing evolution, although the source of the variation (point mutations, indels, etc.) is unknown. Furthermore, by examining correlations between the abundance of all predicted motifs across paralogs, we found evidence that different paralogs might be evolving at different rates. Interestingly, despite the dramatic change in the coding region of *pomc* β (the gain of ε-MSH), there has been little evolutionary change in the *cis*-regulatory sequence of this gene as compared to other teleosts that lack the novel ε-MSH. While future work will examine the functional consequences of gain and loss of *cis*-regulatory motifs in more detail, this global view of motif patterns already provides unique insight into the evolution of non-coding regions.

## 5. Conclusion

In this study, we analyzed both coding and non-coding sequence evolution of the pro-opiomelanocortin (*pomc*) gene family, which is central to the highly pleiotropic and complex neuroendocrine machinery known as the melanocortin system. Notably, we found that cichlids – along with other members of the Labroidea clade – possess a tandem duplication in *pomc β* that appears to encode a novel melanotropic peptide, ε-MSH. We found that, overall, there is remarkable conservation at both the regulatory and structural level. Our results provide new insights into the evolution of the *pomc* gene family following duplication events. Future studies will need to functionally test the expression of *pomc* paralogs across tissues and how they are regulated under different physiological and environmental conditions.

## Acknowledgments

We thank Carly Kenkel, Ben Liebeskind, Da-Jiang Zheng, and four anonymous reviewers for comments on earlier versions of the manuscript; Michelle Faria, Sean Maguire, Eva Wu, and Da-Jiang Zheng for technical assistance; members of the Hofmann Lab for discussion. This work was supported by the EU International Outgoing Marie Curie fellowship to PDD and a NSF Grant 0843712, a Dwight W. and Blanche Faye Reeder Centennial Fellowship in Systematic and Evolutionary Biology and an Institute for Cellular and Molecular Biology Fellowship to HAH.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ygcen.2013.10.007.

## References

Alrubaian, J., Danielson, P., Fitzpatrick, M., Schreck, C., Dores, R.M., 1999. Cloning of a second proopiomelanocortin cDNA from the pituitary of the sturgeon, Acipenser transmontanus. Peptides 20, 431–436.
Amemiya, Y., Takahashi, A., Meguro, H., Kawauchi, H., 1999a. Molecular cloning of lungfish proopiomelanocortin cDNA. Gen. Comp. Endocrinol. 115, 415–421.
Amemiya, Y., Takahashi, A., Suzuki, N., Sasayama, Y., Kawauchi, H., 1999b. A newly characterized melanotropin in proopiomelanocortin in pituitaries of an elasmobranch, *Squalus acanthias*. Gen. Comp. Endocrinol. 114, 387–395.
Arends, R.J., Vermeer, H., Martens, G.J., Leunissen, J.A., Wendelaar Bonga, S.E., Flik, G., 1998. Cloning and expression of two proopiomelanocortin mRNAs in the common carp (*Cyprinus carpio* L.). Mol. Cell. Endocrinol. 143, 23–31.
Braasch, I., Salzburger, W., 2009. In ovo omnia: diversification by duplication in fish and other vertebrates. J. Biol. 8, 25.
Buchanan, J.A., Scherer, S.W., 2008. Contemplating effects of genomic structural variation. Genet. Med. 10, 639–647.
Bumaschny, V.F., de Souza, F.S., Lopez Leal, R.A., Santangelo, A.M., Baetscher, M., Levi, D.H., Low, M.J., Rubinstein, M., 2007. Transcriptional regulation of pituitary POMC is conserved at the vertebrate extremes despite great promoter sequence divergence. Mol. Endocrinol. 21, 2738–2749.
Cardoso, J.C., Laiz-Carrion, R., Louro, B., Silva, N., Canario, A.V., Mancera, J.M., Power, D.M., 2011. Divergence of duplicate POMC genes in gilthead sea bream *Sparus auratus*. Gen. Comp. Endocrinol. 173, 396–404.
Carroll, S.B., 2008. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. Cell 134, 25–36.
Cartharius, K., Frech, K., Grote, K., Klocke, B., Haltmeier, M., Klingenhoff, A., Frisch, M., Bayerlein, M., Werner, T., 2005. MatInspector and beyond: promoter analysis based on transcription factor binding sites. Bioinformatics 21, 2933–2942.
Cerda-Reverter, J.M., Agulleiro, M.J., R, R.G., Sanchez, E., Ceinos, R., Rotllant, J., 2011. Fish melanocortin system. Eur. J. Pharmacol. 660, 53–60.
Chen, Q., Layton, K., Veo, K., Angleson, J.K., Dores, R.M., 2011. Alpha-TC1.9 cells – a model system for analyzing the endoproteolytic processing of POMC. Gen. Comp. Endocrinol. 172, 96–106.
Coll, A.P., Loraine Tung, Y.C., 2009. Pro-opiomelanocortin (POMC)-derived peptides and the regulation of energy homeostasis. Mol. Cell. Endocrinol. 300, 147–151.
Cone, R.D., 2006. Studies on the physiological functions of the melanocortin system. Endocr. Rev. 27, 736–749.
Costa, J.L., Bui, S., Reed, P., Dores, R.M., Brennan, M.B., Hochgeschwender, U., 2004. Mutational analysis of evolutionarily conserved ACTH residues. Gen. Comp. Endocrinol. 136, 12–16.
De Souza, F.S., Bumaschny, V.F., Low, M.J., Rubinstein, M., 2005a. Subfunctionalization of expression and peptide domains following the ancient duplication of the proopiomelanocortin gene in teleost fishes. Mol. Biol. Evol. 22, 2417–2427.
De Souza, F.S., Santangelo, A.M., Bumaschny, V., Avale, M.E., Smart, J.L., Low, M.J., Rubinstein, M., 2005b. Identification of neuronal enhancers of the proopiomelanocortin gene by transgenic mouse analysis and phylogenetic footprinting. Mol. Cell. Biol. 25, 3076–3086.
De Souza, F.S., Nasif, S., Lopez-Leal, R., Levi, D.H., Low, M.J., Rubinsten, M., 2011. The estrogen receptor alpha colocalizes with proopiomelanocortin in hypothalamic neurons and binds to a conserved motif present in the neuron-specific enhancer nPE2. Eur. J. Pharmacol. 660, 181–187.
Dores, R.M., 2013. Observations on the evolution of the melanocortin receptor gene family: distinctive features of the melanocortin-2 receptor. Front. Neurosci. 7, 28.
Dores, R.M., Baron, A.J., 2011. Evolution of POMC: origin, phylogeny, posttranslational processing, and the melanocortins. Ann. N.Y. Acad. Sci. 1220, 34–48.
Dores, R.M., Lecaude, S., 2005. Trends in the evolution of the proopiomelanocortin gene. Gen. Comp. Endocrinol. 142, 81–93.
Dores, R.M., Cameron, E., Lecaude, S., Danielson, P.B., 2003. Presence of the delta-MSH sequence in a proopiomelanocortin cDNA cloned from the pituitary of the galeoid shark, *Heterodontus portusjacksoni*. Gen. Comp. Endocrinol. 133, 71–79.
Drouin, J., Trifiro, M.A., Plante, R.K., Nemer, M., Eriksson, P., Wrange, O., 1989. Glucocorticoid receptor binding to a specific DNA sequence is required for

hormone-dependent repression of pro-opiomelanocortin gene transcription. Mol. Cell. Biol. 9, 5305–5314.
Ducrest, A.L., Keller, L., Roulin, A., 2008. Pleiotropy in the melanocortin system, coloration and behavioural syndromes. Trends Ecol. Evol. 23, 502–510.
Girirajan, S., Eichler, E.E., 2010. Phenotypic variability and genetic susceptibility to genomic disorders. Hum. Mol. Genet. 19, R176–R187.
Gonzalez-Nunez, V., Gonzalez-Sarmiento, R., Rodriguez, R.E., 2003. Identification of two proopiomelanocortin genes in zebrafish (*Danio rerio*). Brain Res. Mol. Brain Res. 120, 1–8.
Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., Lopez, R., 2010. A new bioinformatics analysis tools framework at EMBL–EBI. Nucleic Acids Res. 38, W695–W699.
Hofmann, H.A., 2003. Functional genomics of neural and behavioral plasticity. J. Neurobiol. 54, 272–282.
Hulsey, C.D., Hollingsworth Jr., P.R., Fordyce, J.A., 2010. Temporal diversification of Central American cichlids. BMC Evol. Biol. 10, 279.
Karsi, A., Waldbieser, G.C., Small, B.C., Liu, Z., Wolters, W.R., 2004. Molecular cloning of proopiomelanocortin cDNA and multi-tissue mRNA expression in channel catfish. Gen. Comp. Endocrinol. 137, 312–321.
Kim, K.S., Kim, H.W., Chen, T.T., Kim, Y.T., 2009. Molecular cloning, tissue distribution and quantitative analysis of two proopiomelanocortin mRNAs in Japanese flounder (*Paralichthys olivaceus*). BMB Rep. 42, 206–211.
Knecht, A.K., Hosemann, K.E., Kingsley, D.M., 2007. Constraints on utilization of the EDA-signaling pathway in threespine stickleback evolution. Evol. Dev. 9, 141–154.
Kobayashi, Y., Chiba, H., Amiya, N., Yamanome, T., Mizusawa, K., Amano, M., Takahashi, A., 2008. Transcription elements and functional expression of proopiomelanocortin genes in the pituitary gland of the barfin flounder. Gen. Comp. Endocrinol. 158, 259–267.
Kocher, T.D., 2004. Adaptive evolution and explosive speciation: the cichlid fish model. Nat. Rev. Genet. 5, 288–298.
Leder, E.H., Silverstein, J.T., 2006. The pro-opiomelanocortin genes in rainbow trout (*Oncorhynchus mykiss*): duplications, splice variants, and differential expression. J. Endocrinol. 188, 355–363.
Li, H., Johnson, A.D., 2010. Evolution of transcription networks – lessons from yeasts. Curr. Biol. 20, R746–R753.
Liang, L., Angleson, J.K., Dores, R.M., 2013. Using the human melanocortin-2 receptor as a model for analyzing hormone/receptor interactions between a mammalian MC2 receptor and ACTH(1–24). Gen. Comp. Endocrinol. 181, 203–210.
Mangelsdorf, D.J., Thummel, C., Beato, M., Herrlich, P., Schütz, G., Umesono, K., Blumberg, B., Kastner, P., Mark, M., Chambon, P., Evans, R.M., 1995. The nuclear receptor superfamily: the second decade. Cell 83, 835–839.
Maruska, K.P., Zhang, A., Neboori, A., Fernald, R.D., 2013. Social opportunity causes rapid transcriptional changes in the social behaviour network of the brain in an African cichlid fish. J. Neuroendocrinol. 25, 145–157.
Meyer, A., Schartl, M., 1999. Gene and genome duplications in vertebrates: the one-to-four (-to-eight in fish) rule and the evolution of novel gene functions. Curr. Opin. Cell Biol. 11, 699–704.
Mountjoy, K.G., 2010. Distribution and function of melanocortin receptors within the brain. Adv. Exp. Med. Biol. 681, 29–48.
O'Connell, L.A., Hofmann, H.A., 2012. Social status predicts how sex steroid receptors regulate complex behavior across levels of biological organization. Endocrinology 153, 1341–1351.
Ohno, S., 1970. Evolution by Gene Duplication. Springer-Verlag, Berlin, New York.
Okuta, A., Ando, H., Ueda, H., Urano, A., 1996. Two types of cDNAs encoding proopiomelanocortin of sockeye salmon, *Oncorhynchus nerka*. Zool. Sci. 13, 421–427.
O'Leary, K.A., McQuiddy, P., Kasper, C.B., 1996. Transcriptional regulation of the TATA-less NADPH cytochrome P-450 oxidoreductase gene. Arch. Biochem. Biophys. 330, 271–280.
Robinson, G.E., Fernald, R.D., Clayton, D.F., 2008. Genes and social behavior. Science 322, 896–900.
Rozen, S., Skaletsky, H., 2000. Primer3 on the WWW for general users and for biologist programmers. Methods Mol. Biol. 132, 365–386.
Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J.D., Higgins, D.G., 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. 7, 539.
Steiner, C.C., Weber, J.N., Hoekstra, H.E., 2007. Adaptive variation in beach mice produced by two interacting pigmentation genes. PLoS Biol. 5, e219.
Takahashi, A., Kawauchi, H., 2006. Evolution of melanocortin systems in fish. Gen. Comp. Endocrinol. 148, 85–94.
Takahashi, A., Amano, M., Itoh, T., Yasuda, A., Yamanome, T., Amemiya, Y., Sasaki, K., Sakai, M., Yamamori, K., Kawauchi, H., 2005. Nucleotide sequence and expression of three subtypes of proopiomelanocortin mRNA in barfin flounder. Gen. Comp. Endocrinol. 141, 291–303.
Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28, 2731–2739.
Tanaka, S., 2003. Comparative aspects of intracellular proteolytic processing of peptide hormone precursors: studies of proopiomelanocortin processing. Zoolog. Sci. 20, 1183–1198.
Tirosh, I., Weinberger, A., Bezalel, D., Kaganovich, M., Barkai, N., 2008. On the relation between promoter divergence and gene expression evolution. Mol. Syst. Biol. 4, 159.
Van de Peer, Y., Maere, S., Meyer, A., 2009. The evolutionary significance of ancient genome duplications. Nat. Rev. Genet. 10, 725–732.
Wainwright, P.C., Smith, W.L., Price, S.A., Tang, K.L., Sparks, J.S., Ferry, L.A., Kuhn, K.L., Eytan, R.I., Near, T.J., 2012. The evolution of pharyngognathy: a phylogenetic and functional appraisal of the pharyngeal jaw key innovation in labroid fishes and beyond. Syst. Biol. 61, 1001–1027.
Westneat, M.W., Alfaro, M.E., 2005. Phylogenetic relationships and evolutionary history of the reef fish family Labridae. Mol. Phylogenet. Evol. 36, 370–390.
Whelan, S., Goldman, N., 2001. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol. Biol. Evol. 18, 691–699.
Wray, G.A., 2007. The evolutionary significance of *cis*-regulatory mutations. Nat. Rev. Genet. 8, 206–216.
Wray, G.A., Hahn, M.W., Abouheif, E., Balhoff, J.P., Pizer, M., Rockman, M.V., Romano, L.A., 2003. The evolution of transcriptional regulation in eukaryotes. Mol. Biol. Evol. 20, 1377–1419.
Wunderink, Y.S., de Vrieze, E., Metz, J.R., Halm, S., Martinez-Rodriguez, G., Flik, G., Klaren, P.H., Mancera, J.M., 2012. Subfunctionalization of POMC paralogues in Senegalese sole (*Solea senegalensis*). Gen. Comp. Endocrinol. 175, 407–415.
Xu, Y., Elmquist, J.K., Fukuda, M., 2011. Central nervous control of energy and glucose balance: focus on the central melanocortin system. Ann. N.Y. Acad. Sci. 1243, 1–14.
Yokoyama, K.D., Thorne, J.L., Wray, G.A., 2010. Coordinated genome-wide modifications within proximal promoter *cis*-regulatory elements during vertebrate evolution. Genome Biol. Evol. 3, 66–74.
Zhou, X., Sumazin, P., Rajbhandari, P., Califano, A., 2010. A systems biology approach to transcription factor binding site prediction. PLoS One 5, e9878.