Extensive Changes in the Expression of the Opioid Genes between Humans and Chimpanzees

Peter Cruz-Gordillo a, Olivier Fedigo a, Gregory A. Wray a, b, Courtney C. Babbitt a

 a Department of Biology and Institute for Genome Sciences & Policy, and b Department of Evolutionary Anthropology, Duke University, Durham, N.C., USA

Key Words
Gene expression • Nociception • Opioid peptides • Pain • Primate evolution

Abstract
The various means by which the body perceives, transmits, and resolves the experiences of pain and nociception are mediated by a host of molecules, including neuropeptides within the opioid gene signaling pathway. The peptide ligands and receptors encoded by this group of genes have been linked to behavioral disorders as well as a number of psychiatric affective disorders. Our aim was to explore the recent evolutionary history of these two gene families by taking a comparative genomics approach, specifically through a comparison between humans and chimpanzees. Our analyses indicate differential expression of these genes between the two species, more than expected based on genome-wide comparisons, indicating that differential expression is pervasive among the opioid genes. Of the 8 family members, three genes showed significant expression differences (PENK, PNOC, and OPRL1), with two others marginally significant (OPRM1 and OPRD1). Accelerated substitution rates along human and chimpanzee lineages within the putative regulatory regions of OPRM1, POMC, and PDYN between the human and chimpanzee branches are consistent with positive selection. Collectively, these results suggest that there may have been a selective advantage to modulating the expression of the opioid genes in humans compared with our closest living relatives. Information about the cognitive roles mediated by these genes in humans may help to elucidate the trait consequences of these putatively adaptive expression changes.

Introduction
The opioid family of genes is an ancient network of genes coding for the endogenous neuropeptide ligands and their cell surface receptors. The genes encoding the precursors for the opioid ligands are prodynorphin (PDYN), proopiomelanocortin (POMC), proenkephalin (PENK), and pronociceptin (PNOC). Their receptors are opioid receptor \( \kappa_1 \) (OPRK1), opioid receptor \( \mu_1 \) (OPRM1), opioid receptor \( \delta_1 \) (OPRD1), and opioid receptor-like 1 (OPRL1). The primary and secondary binding affinities of these ligands and receptors are illustrated in figure 1 [Chen et al., 1993a; Meng et al., 1996; Gong et al., 1998; Zaveri et al., 2001; Snook et al., 2008; Liu et al., 2009]. Most investigations surrounding this gene family have focused on the role its members play in pain and nocicep-
tion [reviewed in Stein and Zollner, 2009], stress response [Sonetti et al., 2005], behavior [reviewed in Bodnar, 2009], substance abuse [Kreek, 1996; Kreek et al., 2005; Drakenberg et al., 2006; Xuei et al., 2006, 2007; Huang et al., 2008; Nikoshkov et al., 2008], and some psychiatric affective disorders [reviewed in Ogden et al., 2004; Kennedy et al., 2006; Bodnar, 2009].

The opioid ligand and receptor families evolved through gene duplication events [Barrallo et al., 1998; Dorés et al., 2002; Bradford et al., 2005; Khalap et al., 2005; Bradford et al., 2006; Pinal-Seoane et al., 2006; Stevens et al., 2007; Dreborg et al., 2008; Sundstrom et al., 2010]. Both the receptor and ligand paralogs appear to be mainly due to the genome duplications in early vertebrate evolution [Dores et al., 2002; Dreborg et al., 2008; Sundstrom et al., 2010]. Most gene duplicates resulting from the whole-genome duplications that occurred early in vertebrate evolution [reviewed in Panopoulou and Poustka, 2005] were subsequently lost. In contrast, all of the genes encoding the opioid ligands and receptors have been retained, suggesting that these paralogs were kept due to some selective advantage [Kondrashov et al., 2002]. The fact that these genes have been conserved throughout vertebrate evolution provides evidence that the members of this gene family each have a distinct physiological role.

To date, most studies on the evolutionary history of opioid signaling have focused on comparisons between deeply diverged vertebrate species, through vertebrate evolution. Overall, opioid gene family sequence identities between vertebrate species are quite high, suggesting broadly conserved functions [Stevens et al., 2007; Stevens, 2009]. As with many protein-coding regions, the coding regions of the opioid genes are ~99% conserved between humans and our closest primate relatives, chimpanzees. As opioid signaling has been implicated in a number of behavioral and disease susceptibilities, human-specific cognitive traits may, instead, result from cis-regulatory functional divergence.

An evolutionary analysis of the cis-regulatory region of PDYN provides evidence for this claim [Rockman et al., 2005]. This study focused on a functional 68-bp repeat within the cis-regulatory region of PDYN. Based on the rate of substitutions within and nearby this repeat, Rockman et al. [2005] determined that this regulatory region had been subjected to positive selection during human origins. They also found that different variants are being selectively maintained among distinct human populations, a trend consistent with balancing selection-maintaining unique segregating variants. A later study showed that, beyond the repeat region, multiple segregating regulatory polymorphisms modulate PDYN expression, in a cell-type and sex-specific manner [Babbitt et al., 2010b]. In addition, a coding variant segregating in OPRM1 is known to affect transcript levels, and a putative regulatory polymorphism of OPRM1 has been associated with changes in nociception [Shabalina et al., 2009] and substance abuse [Drakenberg et al., 2006]. The complexity of opioid regulation is just beginning to be understood, but there is evidence that, in the case of PDYN, changes in expression may have been due to adaptive processes.

As reviewed above, there is evidence for rapid adaptive evolution (positive selection) in the PDYN cis-regulatory region, affected by positive selection during human origins [Rockman et al., 2005]. The goal of this study was to search for similar evidence throughout the opioid gene network by comparison of humans and chimpanzees. We utilized two sources of evidence. First, we measured differential gene expression for each ligand and receptor in one brain region of humans and chimpanzees, the frontal cortex. We then scanned for signatures of positive selection characterized by a significant acceleration of substitution rate within putative regulatory and coding regions.
of each gene. Finally, we inquired as to whether these regions showed any evidence of adaptive changes. Understanding the interspecific expression differences may provide insight into interspecific changes, variations, and functional differences across this network of genes.

Materials and Methods

Sample Preparation

The frontal cortex samples used in this study were acquired from 4 human and 4 chimpanzee individuals. All samples were obtained through opportunistic sampling; hence, no primates were sacrificed for the purposes of this research. Samples were obtained from two sources: BioChain Institute Incorporated (Homo sapiens), and Southwest National Primate Research Center (Pan troglodytes; online suppl. table 1; for all online supplementary material, see www.karger.com/doi/10.1159/000320968). Postmortem tissue samples were collected within 12 h of the time of death. All samples were collected from adult males. Total RNA was isolated with an RNeasy kit (Qiagen); RNA concentration and quality were determined using a NanoDrop spectrophotometer (Thermo Scientific) and an Experion system (BioRad), respectively. Only total RNA samples with high-quality 18S and 28S ribosomal bands, lacking obvious contamination but containing adequate 28S rRNA ratios, were used. Total RNA was converted to cDNA using a high-capacity cDNA reverse transcription kit (Applied Biosystems). All samples were checked by PCR to ensure absence of genomic DNA contamination using a SDHA (Stainucate dehydrogenase complex, subunit A) primer pair designed to jump across an intronic region (online suppl. table 2).

Primer Design

Gene sequences were downloaded from the Ensembl genome browser (http://www.ensembl.org/) using the H. sapiens 36.3 and P. troglodytes 2.1 genome builds. PCR primers (Sigma-Aldrich) were designed within completely conserved exonic regions among all transcript isoforms and species based on current University of California Santa Cruz (UCSC) annotations [Karolchik et al., 2003] (online suppl. table 2). Primers were selected using Primer3 Input v0.4.0 [Rozas et al., 2003]. The primer sequences were blasted to all the human and chimpanzee sequences using the Ensembl BLAST tool and a test PCR was performed on human cDNA to ensure that only one product existed for each primer pair in each species.

Quantitative RT-PCR

Quantitative RT-PCR (qPCR) measurements were conducted on an ABI PRISM 7000 (Applied Biosystems). Each reaction consisted of: 15 μl of 2x ABGene Absolute qPCR SYBR® Green Mix, 0.75 μl of each primer (10 μM), 1 μl of cDNA template, and PCR quality water to reach a total volume of 30 μl. The following PCR program was used for all reactions: 95°C for 5 min, 40 cycles at 95°C for 15 s and 58°C for 30 s, followed by a melt curve from 60 to 95°C. Ct values were determined using the CalQplex setting. For each primer pair, a standard curve was set up on human brain or skeletal muscle cDNA over a 12-point, factor-of-two dilution series to determine the efficiency and working Ct range. All primer sets had an efficiency between 94 and 100% and R² >0.99 (online suppl. table 2).

We ran each individual for each gene in each species in triplicate wells. Control samples were run in technical duplicates. Only measurements with standard deviations <0.4 Ct across replicates were used. Within plates, expression was normalized with two control genes TBP (TATA box binding protein) and EIF2B2 (Eukaryotic translation initiation factor 2B, subunit 2 beta; online suppl. table 2) [Fedrigo et al., 2010]. These genes show even expression among human tissues in the Novartis Gene Expression Atlas (http://biogps.gnf.org/), no statistically significant differences in expression between humans and chimpanzees [Fedrigo et al., 2010], and a similar expression level to the genes of interest within the frontal cortex [Vandesompele et al., 2002]. Between plates, an internrun calibration was conducted by running the control gene, EEF2 (Eukaryotic translation elongation factor 2), on IMR-32 cell cDNA in technical triplicate [Hellemans et al., 2007]. To convert the raw Ct expression into normalized relative expression, we utilized a modified ΔΔCt method [Vandesompele et al., 2002; Hellemans et al., 2007]. Our code is available at: http://www.duke.edu/~ofedrigo/Olivier_Fedrigo/PYthonScripts.html.

Assessing Significant Changes in Gene Expression

To assess significant differences in gene expression between species, we used Student's t test. To correct for multiple comparisons, p values were corrected using the method of Benjamini and Hochberg [1995] in R [R Development Core Team, 2005]. p <0.05 was considered significantly different, whereas values of 0.10 > p > 0.05 are listed as marginally significant. The subsampling of the data from Babbitt et al. [2010a] was performed in R [R Development Core Team, 2005].

Detecting Signatures of Positive Selection

Scans for positive selection in the opioid genes and flanking regions were performed in three steps:

Sequence Data and Gene Compartment Annotation. For the eight genes in the opioid gene family, we used customized scripts to extract coding, 5’ and 3’ untranslated regions (UTRs) and 5’ flanking putative promoter regions upstream of the 5’ UTR. The 5’ flanking and 5’ and 3’ UTR regions are collectively called non-coding regions in the following. We downloaded the human (H. sapiens; assembly hg19), chimpanzee (P. troglodytes; assembly panTro2), orangutan (Pongo pygmaeus abelii; assembly panAbe2), and macaque (Macaca mulatta; assembly rMac2) sequences from the UCSC Genome Bioinformatics website (http://genome.ucsc.edu/). 5’ and 3’ UTRs and exons of members of the opioid gene family were defined from the UCSC and RefSeq annotations. We assigned a compartment identity to a sequence by overlapping all the known transcript isoforms and using their intersection (e.g. segments identified as 5’ UTRs are always 5’ UTRs for all transcripts). This conservative approach is required to eliminate the confounding signal caused by multiple functions (e.g. UTRs that are also coding exons) [Haygood et al., 2007]. When segments of the transcripts did not overlap, we considered the union, instead of the intersection, allowing us to check for the absence of any signatures of positive selection, regardless of their function.

We defined the putative promoter sequence by a 5-kb region 5’ flanking upstream of the most 5’ transcription start site since most of the core promoter is believed to reside within few kilobases from the transcription start site [Wray et al., 2003; Blanch-
et al., 2006; Crawford et al., 2006]. We aligned all sequences using MUSCLE [Edgar, 2004]. We then masked nucleotide sites that contain Ns or indels for any of the four species, and visually checked alignments for quality.

Tests for Positive Selection. To test for lineage-specific signatures of positive selection, we used the modified branch site model [Zhang et al., 2005a]. Essentially, this method aims to detect a lineage-specific accelerated nucleotide substitution rate. This rate is estimated relative to a neutral rate in the form of a substitution rate ratio (sequence of interest/neutral sequence). It is key to find the most appropriate genomic region to use as a neutral proxy (see next paragraph). This method compares a null model with no positive selection but that accounts for relaxed constraint, and an alternative model with positive selection on the branch of interest. The two models were contrasted with a likelihood ratio test. We assessed the significance of the likelihood ratio test using a $\chi^2$ with one degree of freedom. A significant p value is suggestive of signatures for positive selection. To avoid local optima while model fitting, we kept the best fit of 10 replicates. We performed these tests with customized and available scripts in the HyPhy software [Pond et al., 2005]. We analyzed exonic sequences and noncoding sequences using similar methods [Wong and Nielsen, 2004; Zhang et al., 2005a; Haygood et al., 2007] and the same quartet of species. We performed these tests on both the human and chimpanzee branches.

Estimating the Neutral Proxy. For the coding analyses, we used the synonymous substitution rate for neutral proxy as it is common usage [Zhang et al., 2005a]. For the noncoding analyses, a similar neutral proxy was applied in order to gain statistical power [Haygood et al., 2007]. Because it has been shown that introns are the least constrained sequence of the genome [Hellmann et al., 2003; Keightley et al., 2005], we collected all introns from genes in 100-kb regions [Chuang and Li, 2004; Gaffney and Keightley, 2005] centered around the gene of interest with customized scripts [Haygood et al., 2007]. An important statistical concern arises when introns contain regulatory elements that can be more slowly evolving than neutral sites. Our strategy was to eliminate any putatively functional regions that might be conserved between species and that would artificially inflate the substitution rate ratio and can lead to erroneous detection of positive selection. We excluded 100 bp at each extremity of the introns, with the goal of eliminating splicing signal sites [Sorek and Ast, 2003]. We also excluded first introns because they are known to often contain regulatory elements, and we included a maximum of 2,500 bp from any one intron, drawn from the edges, since some long introns have been shown to contain regulatory elements at their center [Blanchette et al., 2006]. We used the remaining intronic sequences as neutral proxy for detecting positive selection in non-coding regions (noncoding as defined as the 5' flanking, 5' and 3' UTR regions described above). Because functional elements may still be present in this selected subset of introns, we constructed 100 bootstrap replicates from the intronic data, performed the test for selection with each of the bootstrap replicates, and considered the median p value as an indicator of positive selection.

Results

Expression Level Comparisons of the Opioid Genes

We used qPCR to detect the relative expression levels of the opioid genes in chimpanzee and human cortical tissue, respectively. Three genes were significantly differentially expressed after a correction for multiple comparisons, and two others marginally so.

$\text{PNNOC}$ and $\text{OPRL1}$ were both significantly higher in human cortices relative to chimpanzee cortices ($p = 0.0125$ and $p = 0.0052$, respectively; fig. 2). $\text{PENK}$ had the most significantly different expression, and is higher in
human cortices relative to chimpanzee cortices (p = 0.0008), while OPRD1 was marginally significant (p = 0.0573). The reverse pattern was observed between the POMC and OPRM1 ligand-receptor pair. OPRM1 was significantly higher in human cortices relative to chimpanzee cortices (p = 0.0230), whereas POMC was not differentially expressed (p = 0.1538). Interestingly, neither PDYN nor OPRK1 were found to be significantly different in human cortices relative to chimpanzee cortices (p = 0.1395 and p = 0.2621, respectively).

For a background comparison, we randomly subsampled groups of eight genes from a genome-wide expression comparison between human and chimpanzee brains [Babbitt et al., 2010a]. Our finding of five genes out of the 8 total family members is in the 99th percentile of that sampling distribution at either a corrected significance level of p < 0.05 or p < 0.10. Although the data from Babbitt et al. [2010a] were measured on a different platform, some of the individuals and the region measured are the same, and so those are currently the best data available to provide at least a rough background estimate of the expected amount of change.

The magnitude of these changes in expression varies manifold depending on the gene. Comparing the mean expression differences between the species samples, we see that OPRM1, PNOC, and OPRL1 expression in humans is approximately 2–3 times greater than the mean expression in chimpanzees, whereas it is approximately 10 times higher for PENK (fig. 2).

**Evidence for Positive Selection in Different Gene Regions**

The protein-coding regions of the four opioid ligands and four receptors are highly conserved between humans and chimpanzees, with a range of 98.8–100% nucleotide sequence identity. Therefore, the patterns of differential expression reported above between species may be due to changes within regulatory sequences. It is very challenging to locate the specific changes that are responsible for expression differences, so we attempted to look for important nucleotide substitutions by scanning possible regulatory regions for evidence of positive selection. Positively selected regulatory regions would appear as an overabundance of substitutions in the regulatory sequence as compared to nearby intronic regions, which are assumed to be evolving in a neutral fashion. We scanned for signatures of positive selection on the human and chimpanzee lineages using a four-species tree of known phylogenetic relationships (human, chimpanzee, orangutan, and rhesus macaque).

All of the statistically significant change appears to be concentrated within 5’ regulatory regions (partitioned between the 5’ flanking regions and the 5’ UTR). We found three genes with a signature of positive selection in a putative regulatory region: the 5’ UTR of POMC (p = 0.0026 and 0.009 on the human and chimpanzee branches, respectively), the 5,000-bp 5’ flanking region of OPRM1 (p = 0.00011 for the chimpanzee branch), and the 5’ flanking region of PDYN (p = 0.0671) on the human branch (table 1; online suppl. table 3). The signals for
POMC and OPRM1 remain significant after a q-value correction (online suppl. table 3). These are more changes than expected by chance, given results from genome-wide scans for selection within these regions [Haygood et al., 2007]. It is important to note that the type of scan for selection we employed requires multiple sequence changes, accumulated in a defined region, for an inference of positive selection. Although our scans only surveyed possible regulatory regions located close to the genes, some important regulatory elements may be located at more distal regions that we did not survey. This point, and the generally underpowered nature of tests for selection, indicates that a negative result does not rule out the possibility that positive selection operated on regulatory sequences.

Discussion

Opioid gene signaling is a critical link in a number of human behavioral responses [reviewed in Sonetti et al., 2005; Bodnar, 2009; Stein and Zollner, 2009], as well as in many human disease susceptibilities [Kreek, 1996a; Ogden et al., 2004; Kreek et al., 2005; Drakenberg et al., 2006; Kennedy et al., 2006; Xuei et al., 2006, 2007; Huang et al., 2008; Nikoshkov et al., 2008; Bodnar, 2009]. Several studies have presented evidence of functional changes that have occurred in the coding sequences of these genes, both across vertebrate evolution [Dores et al., 2002; Stevens et al., 2007; Dreborg et al., 2008] and within human populations (e.g. one known functional variant within the coding region of OPRM1 in certain human populations [Zhang et al., 2005b]). However, presented with the indication that at least one of the opioids has been a target of natural selection in a cis-regulatory region during human evolution [Rockman et al., 2005], we attempted to investigate changes throughout this gene family.

We examined the gene expression levels of all eight opioid genes in a single brain region as one index of differential function. Recent studies comparing gene expression between humans and chimpanzees in brain tissue have found that ~12–18% of genes are differentially expressed, depending on the tissue measured, the platform used, and the thresholds employed [Khaitovich et al., 2004; Uddin et al., 2004; Khaitovich et al., 2005; Babbitt et al., 2010a]. Our finding that ~37% of the opioid genes are significantly differentially expressed (~62% if OPRM1 and OPRD1 are also included) exceeds the global genome-wide average of differentially expressed genes within the brain between these two species (fig. 2) and is clearly at the tail of the distribution when compared to a background model of change.

Some important caveats to our results are that we only measured gene expression from a single brain region, not in all the brain regions in which opioid signaling is known to be important [Harris, 1959; Harris and Roos, 1959; Comb et al., 1982; Noda et al., 1982; Chen et al., 1993b; Mollereau et al., 1994; Simonin et al., 1994, 1995; Mollereau et al., 1996; Telkov et al., 1998; Cowley et al., 2001; Hurd, 2002; Nikoshkov et al., 2005; Drakenberg et al., 2006]. Expression of opioids is also known to be environmentally variable [reviewed by Stein and Zollner, 2009], and we have no information on any previous exposures to opioid compounds for these individuals. We attempted to control for these issues here by measuring multiple individuals. However, larger sample sizes may assist in identifying other species-specific expression differences – particularly if future studies are able to assay other brain regions and more individuals or species. Another layer of complexity that needs to be explored is the role of alternative splicing in opioid signaling in the brain. The expression levels measured here are from exons conserved between known splice forms, but some of the opioid genes are known to be alternatively spliced, and different splice forms have different binding affinities [Pan, 2005]; therefore, this may be an important method of regulation between brain regions or between species. A further challenge will be to understand the polymorphisms in the regulatory regions that have become fixed between species that might be driving differences in expression. To date, several studies have investigated the functional impact of regulatory variants of the PDYN gene [Zimprich et al., 2000; Nikoshkov et al., 2005, 2008; Yuferov et al., 2009; Babbitt et al., 2010b] and OPRM1 [Shabalina et al., 2009] within human populations. There are also OPRM1 coding and 5’ UTR variants found in macaques that have similar effects on in vitro expression [Vallender et al., 2008] and in behavioral associations [Miller et al., 2004; Barr et al., 2008] as the well-studied A118G variant identified in humans. This suggests that the expression level of OPRM1 at least has an impact on behavioral phenotypes across species. The results presented here provide motivation for identifying the genetic basis for additional human-specific features of opioid gene regulation.

In addition to the expression assays, we also searched for evidence of positive selection [Haygood et al., 2007] in all eight of the genes in the opioid family. The previous evidence of positive selection in humans within a 68-bp
repeat region ~3 kb upstream of PDYN [Rockman et al., 2005] suggested that we should conduct an analogous test for selection in the other opioid ligands and receptors. We found evidence of positive selection within the 5’ putative regulatory regions of three opioid genes. The 68-bp repeat region of PDYN did not appear to be significant in our scan here, most likely because repeat regions were excluded from our very conservative analysis pipeline. Other motifs regulating the seven other opioid genes, like the repeat region, may have been the substrate for selection, but locating these regions will require more sequence information from additional species.

The opioid genes exhibiting expression differences and evidence of positive selection are not particularly congruent. Nonetheless, both tests show a much higher than expected amount of change within these two gene families between humans and chimpanzees. The lack of overlap between the scans for selection and the expression data could be due to a number of reasons. The first is that signatures of selection may occur as a result of adaptive advantages in any tissue and at any developmental time, whereas we only measured expression within one brain region in adults. The second possible reason is that our scans for selection were tremendously conservative and would, therefore, miss regions where there were few changes (though possibly of strong effect) or regulatory motifs that occur in small or large repeat regions. This could also be due to differences in ligand-receptor affinities and/or post-translational processes, rather than mRNA expression levels, functioning as the predominate regulatory mechanism. Future functional studies are needed to establish the underlying sequence changes driving differences in expression and possibly in protein interactions, as well as their history during human evolution. The opioid signaling system has critical behavioral roles and pathological significance. The striking expression changes observed in our comparison between humans and chimpanzees suggest that differential regulation of the opioid pathway may underlie unique behaviors and pathologies that have evolved between humans and our closest relatives.

Acknowledgments

We would like to thank the members of the Wray lab for advice and support, especially David Garfield and Lisa Warner, Adam Pfefferle and William Nielsen for assistance with qPCR, and Julie Horvath, Jera Pecotte and Mary Jo Aivaliotis for their help in obtaining the biological materials from the Southwest National Primate Research Center. This project was funded by National Science Foundation Grant NSF-BCS-08-27552 (HOMINID) and the Institute for Genome Sciences & Policy at Duke University. P.C.-G. was funded by a summer undergraduate training fellowship from the Institute for Genome Sciences & Policy at Duke University.

References


