Published in final edited form as:

J Mol Evol. 2009 October ; 69(4): 319–332. doi:10.1007/s00239-009-9273-y.

Phylogenetic Analysis of the NEEP21/Calcyon/P19 Family of Endocytic Proteins: Evidence for Functional Evolution in the Vertebrate CNS

Nagendran Muthusamy¹, Sanaa A. Ahmed², Brinda K. Rana³, Sammy Navarre⁴, David J. Kozlowski⁴, David A. Liberles², and Clare Bergson¹

¹Department of Pharmacology and Toxicology, Medical College of Georgia, Augusta, GA

²Department of Molecular Biology, University of Wyoming, Laramie, WY

³Department of Psychiatry, University of California at San Diego, La Jolla, CA

⁴Department of Cell Biology and Institute of Molecular Genetics, Medical College of Georgia, Augusta, GA

Abstract

Endocytosis and vesicle trafficking are required for optimal neural transmission. Yet, little is currently known about the evolution of neuronal proteins regulating these processes. Here, we report the first phylogenetic study of NEEP21, calcyon, and P19, a family of neuronal proteins implicated in synaptic receptor endocytosis and recycling, as well as in membrane protein trafficking in the somatodendritic and axonal compartments of differentiated neurons. Database searches identified orthologs for P19 and NEEP21 in bony fish, but not urochordate or invertebrate phyla. Calcyon orthologs were only retrieved from mammalian databases and distant relatives from teleost fish. *In situ* localization of the P19 zebrafish ortholog, and extant progenitor of the gene family, revealed a CNS specific expression pattern. Based on non-synonymous nucleotide substitution rates, the calcyon genes appear to be under less intense negative selective pressure. Indeed, a functional group II WW domain binding motif was detected in primate and human calcyon, but not in non-primate orthologs. Sequencing of the calcyon gene from 80 human subjects revealed a non-synonymous single nucleotide polymorphism that abrogated group II WW domain protein binding. Altogether, our data indicate the NEEP21/calcyon/P19 gene family emerged, and underwent two rounds of gene duplication relatively late in metazoan evolution (but early in vertebrate evolution at the latest). As functional studies suggest NEEP21 and calcyon play related, but distinct roles in regulating vesicle trafficking at synapses, and in neurons in general, we propose the family arose in chordates to support a more diverse range of synaptic and behavioral responses.

Correspondence: Clare Bergson, Ph.D., Department of Pharmacology and Toxicology, Medical College of Georgia, 1459 Laney Walker Blvd, Augusta, GA 30912-2300, tel: 706-721-1926; fax: 706-721-2347, cbergson@mail.mcg.edu.

Authors Contributions

NM performed molecular genetic studies on zebrafish clones, database searches and sequence alignment, initial phylogenetic analyses, and drafted the manuscript. SA performed and contributed to the interpretation of the phylogenetic analysis. SN performed the *in situ* hybridizations. DK contributed to the acquisition and interpretation of the zebrafish expression and bony fish phylogenetic data. BR contributed the SNP discovery data, and analysis. DL contributed to the conception and design of the phylogenetic analyses, and with CB wrote the final draft of the manuscript. CB designed and integrated the overall study.

Keywords

Synaptic plasticity; transcytosis; single nucleotide polymorphism; zebra fish; group II WW domain; primate innovation; Ka/Ks

Introduction

With only twice as many neurons and minor differences in the number of overall coding sequences, a major goal of evolutionary neurobiologists is to identify species differences at the genomic level that might account for the marked increase in cognitive capacity of the human brain compared to that of our closest phylogenetic relative, the chimpanzee. Structurally, the human brain is distinguished from the chimpanzee brain by a vast increase in levels of neuronal connectivity. The greater neuronal connectivity in human brain is thought to allow for more efficient communication between sensory, association, limbic and motor brain regions. However, little is known about the phylogenetic or cellular mechanisms that might drive or functionally support the increased connectivity.

Synapses are the basic unit of neuronal connectivity in the central nervous system (CNS), and are site of chemical neurotransmission between interconnected neurons. Activity-dependent regulation of synaptic transmission is thought to underlie learning and memory (Heynen et al., 2000; Teyler and Discenna, 1984; Whitlock et al., 2006). Hence, there is widespread interest in elucidating the molecular scaffold involved in regulating pre- and postsynaptic aspects of synaptic transmission. It is thought that such information could shed light on neurological and neuropsychiatric disorders in humans involving defects in connectivity such as proposed for schizophrenia, autism, Alzheimer's disease, and amyloid lateral sclerosis, among others.

Functional studies carried out in the invertebrate and vertebrate nervous system indicate that vesicle trafficking within both pre- and postsynaptic elements of the synaptic junction plays a crucial role in basal as well as activity-dependent adaptations in synaptic transmission (Bredt and Nicoll, 2003; Malinow and Malenka, 2002; Sudhof, 2004). Such 'plasticity' in synaptic transmission is proposed to underlie learning, memory and cognitive flexibility (McCormack et al., 2006; Nicholls et al., 2008; Teyler and Discenna, 1984). Little is currently known about the evolution of the neuronal proteins that regulate vesicle trafficking at synapses. However, recent studies suggest that the molecular machinery regulating these processes in mammals is substantially more elaborate than in invertebrates (Emes et al., 2008).

Here we report the first phylogenetic study of NEEP21 (Entrez Gene Name: neuron specific gene family member 1; alias, nsG1, p21), calcyon (Entrez Gene Name: calcyon neuron-specific vesicular protein; alias, Drd1p, Caly), and P19 (Entrez Gene Name: neuron specific gene family member 2; alias, nsG2, p19 and p1A75), a family of endocytic proteins recently found to play a role in regulating excitatory synaptic transmission (Saberan-Djoneidi et al., 1995; Saberan-Djoneidi et al., 1998; Steiner et al., 2002; Xiao et al., 2006). The best characterized family members, NEEP21 and calcyon appear to perform distinct roles in vesicle trafficking, consistent with a homologous relationship. For example, NEEP21

regulates the sorting of endosomes into either a recycling or degradative pathway (Debaigt et al., 2004; Steiner et al., 2002), whereas calcyon stimulates clathrin mediated endocytosis of nutrients and plasma membrane receptors into vesicles (Xiao et al., 2006). Intriguingly, recent work suggests the homologous roles of the two proteins extend to neuronal activity-driven adaptations in synaptic strength including synaptic plasticity. That is, down-regulation of NEEP21 leads to deficits in constitutive recycling of AMPA receptors and activity-dependent synaptic insertion of receptors (Alberi et al., 2005), whereas knockout of calcyon interferes with activity dependent removal of surface AMPA receptors and synaptic weakening (Davidson et al., 2009).

Recent studies suggest the calcyon/NEEP21/P19 gene family plays an important role in the polarized sorting of axonal and somatodendritic membrane components in differentiated neurons. For example, NEEP21 regulates transcytosis of molecules endocytosed in somatodendritic regions to axons (Yap et al., 2008). The role of P19 in vesicle trafficking, if any, is currently unknown (Saberan-Djoneidi et al., 1995). Nevertheless, in mammals, the expression of P19 like that of NEEP21 and calcyon is enriched in the hypothalamus, hippocampus, ventral basal ganglia, amygdala and prefrontal cortex, as well as the pituitary. The functions of these nuclei are largely elaborated in vertebrates, and quite diverse spanning homeostasis to emotion processing, social interactions, blood pressure, thirst, long term memory, reward and motivation. Using phylogenetic methods coupled with an analysis of codon and amino acid substitution rates, we present a picture of the functional evolution of the tripartite NEEP21/calcyon/P19 gene family pointing to its emergence in vertebrates.

Materials and Methods

Zebrafish strains and animal care

Breeding and staging of zebrafish embryos was performed according to standard protocols (Westerfield, 1995) and experiments were carried out with Tubingen Wild-Type or *brass* embryos. All experimental protocols using zebrafish (*Danio rerio*) were reviewed and approved by the MCG Institutional Animal Care and Use Committee.

Protein homology search

Proteins sequences corresponding to human P19 (Accession: CAG33423), NEEP21 (Accession: NP_055207) and calcyon (Accession: AAF34714) were used in CDART searches of the National Center for Biotechnology Information (NCBI) non-redundant (nr) protein database (including GenBank coding sequence translations, sequences derived from the 3-D structure Protein Data Bank, and the SWISS-PROT protein sequence database). The CDART search tool retrieves proteins with similar domain architecture. Although CDART is thought to be more sensitive, we also searched for orthologs in translated non-redundant nucleotide databases (including all GenBank, EMBL, DDBJ, and PDB sequences) using the same protein query sequences. Synthetic sequences were removed, as were sequences deemed partial alignment with ClustalW (Larkin et al., 2007). The canine calcyon sequence (XP_548818.2) was modified to terminate after 226 amino acids since the sequence codes for a protein 350 amino acids in length, and the C-terminal 124 amino acids bears no sequence similarity with the rest of the gene family. Tetraodon nigroviridis P19

(CAG11727) was used to query the *Gallus gallus* EST database, and *Xenopus tropicalis* NEEP21 (NP_001072849) to query *Danio rerio* EST databases in TBLASTN searches. Additionally, all available *C. elegans*, *C. intestinalis*, *D. melanogaster* and *S. cerevisiae* genomic and EST databases as well as Genbank were queried to investigate the presence of orthologs in invertebrates.

Multiple sequence alignment and Phylogenetic analysis

Alignment of the sequences in Table 1 was performed using ClustalW with the MegAlign tool of Lasergene 6 software (Larkin et al., 2007). At the nucleotide level using back-translated gaps from amino acid level alignments, MODELTEST (Posada and Crandall, 1998) was run with GTR+ Λ +I as the suggested model. This was used to calculate a phylogenetic tree using RAXML (Stamatakis, 2006). Branch lengths and clade bootstrap supports associated with the tree were also collected. Softparsimony (Berglund-Sonnhammer et al., 2006) was used to minimize the number of duplications and losses in generating the most parsimonious tree roots.

Ancestral sequence reconstruction coupled to counting (PBL method) was used as a non-statistical screen for positive selection (Liberles, 2001) as implemented in www.bioinfo.no. The one ratio, free ratio, neutral sites, branches and sites, and branches models were run using codeml from the PAML package (Yang, 2007) independently on the NEEP21, P19, and calcyon clades. Likelihood ratio tests were run to identify which tests were significant, with values reported in Table 2. No correction for multiple testing was applied. Diverge v1.4 (Gu and Vander, 2002) was run to compare the NEEP21, P19, and calcyon clades in a pairwise manner, with results presented in Table 3.

Cloning of Zebrafish P19

Zebrafish RNA (from embryos staged at 24, 48, and 72h post-fertilization) was prepared using the Trizol method (Invitrogen, Carlsbad, CA), and first strand cDNA by reverse transcription using the SuperScript kit (Invitrogen, Carlsbad, CA). Primers (3'AGTTGACCGCTTTCGGCT 5' and 3' GGATGAAACATTTCCACCAGAT 5') were designed based on the zebrafish LOC324971 mRNA sequence (accession number AY394975). A 1338 nucleotide cDNA was amplified by PCR and subcloned using the pGEM-Teasy Vector (Promega, Madison, WI). The identity of the putative LOC324971 cDNA clone was confirmed by DNA sequencing both strands with an ABI 3730 XL 96-capillary sequencer (Applied Biosystems) operated by the Genomics Core Lab at the Medical College of Georgia. The LOC324971 clone isolated includes 552, 45, and 741 nucleotides of predicted coding, 5', and 3' flanking sequence, respectively.

Whole-mount in situ hybridization

Whole-mount *in situ* hybridization was carried out under standard conditions (Thisse et al., 1993). Riboprobe templates were prepared by linearizing the 1338 nucleotide LOC324971 cDNA clone described above with the appropriate restriction enzymes (sense, SacII; antisense, SpeI), and transcribing with either T7 (antisense) or SP6 (sense) RNA polymerase in the presence of digoxigenin-labeled nucleotides (Roche Applied Science, Indianapolis,

IN). Sense probes were used for control hybridization reactions and resulted in only low level diffuse staining.

List of databases searched

Ciona intestinalis—<http://genome.jgi-psf.org/ciona4/ciona4.home.html>ghost.zool.kyoto-u.ac.jp/indexr1.html

Caenorhabditis elegans—<http://www.ncbi.nlm.nih.gov/genome/seq/BlastGen/BlastGen.cgi?taxid=6239>

Danio rerio—<http://www.ncbi.nlm.nih.gov/genome/seq/BlastGen/BlastGen.cgi?taxid=7955>

Drosophila melanogaster—<http://www.ncbi.nlm.nih.gov/genome/seq/BlastGen/BlastGen.cgi?taxid=7227>

Gallus gallus—<http://www.ncbi.nlm.nih.gov/genome/seq/BlastGen/BlastGen.cgi?taxid=9031>

Oryzias latipes—http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/gimain.pl?gudb=o_latipes

Tetraodon nigroviridis—<http://www.ensembl.org/Multi/blastview>

Saccharomyces cerevisiae—<http://www.ncbi.nlm.nih.gov/genome/seq/BlastGen/BlastGen.cgi?taxid=4932>

Results and Discussion

The P19/NEEP21/Calcyon Gene Family is Highly Conserved

At the amino acid level, the human (h) protein hcalcyon exhibits about 30% identity to that of hNEEP21 and hP19 which show about 50% identity to each other. Search of the NCBI non-redundant (nr) sequence database for hP19, hNEEP21 or hCalcyon orthologs using the Conserved Domain Architecture Retrieval Tool (CDART) retrieved an identical set of sequences from a variety of vertebrate species (Table 1). Each protein sequence retrieved contained a single copy of the so-called ‘calcyon’ protein domain (PF06387), a 179 amino acid sequence characterized by a single transmembrane segment in the middle. None of the hits contained any other identified protein domains. Although neither the 3-D structure nor the function of the calcyon domain is solved, this domain currently appears to be the defining characteristic of the P19/NEEP21/Calcyon gene family (Finn et al., 2007).

Apart from retrieving protein sequences from a variety of mammalian species, CDART searches identified closely related proteins in avian, piscine and amphibian species, e.g., chicken, pufferfish, frog, and zebrafish. All of the sequences appeared to belong to the P19/NEEP21/Calcyon gene family based on sequence similarity scores. Notably, the degree of similarity between family members in the same species was conserved from fish to human

(Supplementary Table 1). These results suggested that the P19/NEEP21/Calcyon family is highly conserved and is not inclusive of other genes based on their common evolutionary origin.

Absence of the P19/NEEP21/Calcyon Gene Family in Invertebrates

We also searched invertebrate sequences in Genbank as well as a set of invertebrate databases including *Caenorhabditis elegans*, *Drosophila melanogaster* and *Saccharomyces cerevisiae* with the human P19, NEEP21, and calcyon protein and nucleotide sequences, as well as those for the putative zebrafish P19 ortholog LOC324971 (NP 957167) (Table 1). However, no significant hits were identified in the protein, genomic, mRNA or EST databases of these invertebrate organisms. Similarly, no hits were obtained in BLAST searches of two separate genomic DNA databases for the lower chordate *Ciona intestinalis* using all the human sequences and the sole zebrafish ortholog (genome.jgi-psf.org/ciona4/ciona4.home.html; ghost.zool.kyoto-u.ac.jp/indexr1.html). Search of the *Ciona* 5' and 3' EST databases with the putative zebrafish P19 ortholog LOC324971 (NP 957167) sequences using TBLASTN only yielded hits with partial query sequence coverage and low bit score. The failure to identify sequences with significant bit scores either in the EST or genomic database searches suggests calcyon/P19/NEEP21 family orthologs are not present in *Ciona*, indicating that the duplication events leading to these gene families were vertebrate-specific. The pattern of duplication might be consistent with the two rounds of known whole genome duplication during the chordate-vertebrate transition (Dehal and Boore, 2005; Hughes and Liberles, 2008). It would also be consistent with pre-duplication rapid sequence divergence of the ancestral vertebrate protein from a non-vertebrate ancestor. These genes do likely have a single ancestor in non-vertebrates, but the inability to detect this gene using sequence-based methods is suggestive of a novel, perhaps duplication-mediated set of roles for the 3 vertebrate-specific genes. Gene duplication is known to enable more rapid sequence divergence (Hughes and Liberles, 2007).

Expression of P19 in the Zebrafish Developing Nervous System

Studies in mammals suggest that mRNA for calcyon/NEEP21/P19 gene family members is expressed primarily in the CNS (Saberan-Djoneidi et al., 1995; Saberan-Djoneidi et al., 1998; Steiner et al., 2002; Zelenin et al., 2002). Given the high degree of sequence identity (>70%) between human P19 and the putative P19 ortholog found in zebrafish, LOC324971 (NP 957167), we used RT-PCR and whole mount *in situ* hybridization to determine where and when this predicted calcyon/NEEP21/P19 gene family member is expressed during zebrafish embryogenesis. To determine the temporal expression of LOC324971 mRNA during embryonic development we performed RT-PCR using LOC324971-specific primers and RNA prepared from different embryonic stages (Fig. 1A, open arrow). Low levels of maternally provided LOC324971 mRNA are present at the 4-cell stage of embryogenesis and 6 hours post fertilization (hpf), but then decrease through somite stages. By 24 hpf, the expression of LOC324971 mRNA dramatically increases, continues to increase at 48 hpf, and remains high through 72 hpf (data not shown).

We determined the spatial distribution of LOC324971 mRNA during embryonic development by whole mount *in situ* hybridization (Fig. 1B–D). At 18 hpf, LOC324971

mRNA expression is detectable in the developing forebrain (Fig. 1B, open arrow), ventral midbrain (Fig. 1B, open triangles), and along the lateral edges of hindbrain rhombomeres (Fig. 1B, solid arrows). By 24 hpf LOC324971 mRNA expression is increased in the telencephalon (Fig. 1C, open arrow), diencephalon (Fig. 1C, line arrow), ventrolateral tegmentum (Fig. 1C, open triangle), hindbrain rhombomeres (Fig. 1C, solid arrows), and along the anterior spinal cord. At 48 hpf (Fig. 1D) and later stages up to 72hpf (data not shown) LOC324971 mRNA continues to be expressed exclusively in the developing telencephalon (Fig. 1D, open arrow), diencephalon (Fig. 1D, bracket), hindbrain (Fig. 1D, solid arrows), and spinal cord (Fig. 1C, bracket). The distribution of LOC324971 in developing zebrafish suggests that CNS expression is a conserved feature of this gene family even in the lowest extant common ancestor.

Phylogenetic Analysis of the Calcyon/NEEP21/P19 Gene Family

A multiple sequence alignment of the calcyon/NEEP21/P19 proteins compiled from CDART and EST searches is shown in Fig. 2. The maximum likelihood phylogenetic tree shown in Fig. 3 indicates that the refined data set corresponds to three sets of related proteins representing clusters of P19, NEEP21 and calcyon orthologs. The tree was rooted using soft parsimony to minimize the number of implied duplication and loss events. This is a robust method for rooting a tree in the absence of any known outgroup sequence, providing an assessment that is independent of sequence divergence measures. The Softparsmap rooting supports three clades of orthologous sequences that diverged from the base of the vertebrate tree. Although ambiguous, this rooting is equally parsimonious with regard to the order of the duplication events, and more parsimonious than other potential rootings. Overall, the branching topology of the three clades suggests that all three gene families emerged before the divergence of teleosts from amniotes. All three genes were identified in teleosts and in mammals, while calcyon genes were not identified in birds or amphibians, suggesting either missing data or two lineage-specific gene loss events. This rooting is also consistent with the known two rounds of whole genome duplication during the chordate-vertebrate transition (Dehal and Boore, 2005; Hughes and Liberles, 2008).

Apart from addition and deletion of genes from the genome, other evolutionary mechanisms like changes in gene expression and coding sequence changes may underlie the vast increase in computing capacity, behavioral repertoire, homeostatic mechanisms and cognitive flexibility of the human brain (Hill and Walsh, 2005). Measurement of the ratio of non-synonymous (K_a) to synonymous (K_s) nucleotide substitution is an accepted approach for identifying branches in phylogenetic trees undergoing adaptive evolution (Anisimova and Liberles, 2007). To estimate the K_a/K_s rate ratio, we used the PAML program (Yang, 2007) package to evaluate various models and their parameterization (Fig. 3). To reduce the total tree length and mis-parameterization due to saturation of synonymous sites, models were run independently on P19, NEEP21, and calcyon (see Table 2). The following strategy was undertaken. First a non-statistical ancestral sequence reconstruction and counting method (PBL) was applied to the detection of lineage-specific K_a/K_s values greater than 1 (Liberles, 2001). In parallel, model testing was used to test for support for the free ratio model to identify individual lineages undergoing positive selection. These hypothetical lineages from the non-statistical and free ratios models were individually evaluated by model testing using

the branches and branches and sites models. The free ratios model suggested three candidates for positive selection while the nonstatistical approach suggested just one (which corresponded to one of the three from the free ratios model). The one ratio model suggests generally strong negative selection on all gene families, with the most relaxed selection on calcyon. The free ratios model was significant for P19 and calcyon, but not NEEP21. Lineages with $\omega > 0.9$ were tested individually for significant evidence of positive selection using the branches and branches and sites tests. Only two lineages, one in calcyon and one in P19 had significant evidence for positive selection from the free ratio, branches, and branches and sites tests. The calcyon branch indicated positive selection with the non-statistical approach.

Possible explanations for the more constrained evolution of the NEEP21 family includes a larger set of interacting protein partners, exclusive somatodendritic distribution (whereas calcyon is also present in axons), or more strongly conserved function (Wang et al., 2007) (Steiner et al., 2002; Steiner et al., 2005). Along the lines of the latter possibility, recent studies in primary cultures of neurons suggest NEEP21 regulates the transcytosis of adhesion molecules implicated in the early phases of myelination from the somatodendritic compartment to axons (Yap et al., 2008) (White et al., 2008). Myelination of axons is a vertebrate innovation facilitating rapid conduction of nerve impulses. Interestingly, NEEP21 expression in rodents is highest early in development during periods correlated with robust myelination and synapse formation (Saberan-Djoneidi et al., 1998).

At the amino acid level, DIVERGE was run to evaluate evidence for divergence in amino acid substitution rates between clades, as the previous analysis was only able to evaluate divergence within clades. Significant evidence for such a divergence in rates was found between calcyon and NEEP21. Several sites (see Table 3) were identified as having undergone clade-specific rate shifts. The site-specific profiles indicate NEEP21/Calcyon to have a statistically detectable class of sites with differing selective pressures, including the following sites: 57–59, 67–74, 87–100, 103–114, 116–123, 125–130, 149, 152–168. Two sites, 125–130 and 152–168, situated in the predicted cytosolic domains of the proteins, are noteworthy in light of available structure-function information. For example, NEEP21 regulates GluR2 AMPA receptor recycling via interaction with GRIP1 through a domain in NEEP21 that includes residues 129–164 (Steiner et al., 2005). Similarly, calcyon appears to regulate AMPA receptor endocytosis via a clathrin light chain binding domain situated in calcyon within residues 123–155 (Davidson et al., 2009). While intriguing, the ‘Diverge’ data are only predictive, and further studies are necessary to confirm that sequence divergence in the 125–130 and 152–168 segments confers the distinct endocytic functions subserved by NEEP21 and calcyon.

Human Calcyon Polymorphism Discovery within a Primate-Specific Protein Interaction Motif

Cortical and especially the forebrain regions of cortex are greatly expanded in mammals compared to other vertebrates. While P19 and NEEP21 are also expressed in cortex, expression of the human calcyon gene is enriched in prefrontal cortex. In humans, the expression of calcyon is significantly up-regulated in patients with schizophrenia, a disease

associated with deficits in prefrontal cortical function (Bai et al., 2004; Baracskey et al., 2006; Clinton et al., 2005; Koh et al., 2003). Additionally, genome wide scans have implicated 10q26, the chromosomal region where the calcyon gene resides, in schizophrenia as well as attention deficit hyperactivity disorder, a psychiatric syndrome also associated with deficits in prefrontal function (Laurin et al., 2005; Williams et al., 2003).

To look for markers that might prove useful in association or linkage studies, we sequenced 4,866 base pairs of calcyon consisting of coding exons 2 and 3, in an ethnic diversity panel of 80 unrelated human subjects (30 African-Americans, 34 Caucasian-Americans, 8 Hispanic, 6 Filipino, 2 East Asian). We identified 9 novel SNPs (not previously reported in public databases or manuscripts). Among these was a non-synonymous SNP identified in one male African-American individual that results in the substitution of a leucine (leu, L) for a proline (pro, P) at residue 40 in the calcyon protein (Dai and Bergson, 2003) (Supplementary Fig. 1). Proline at residue 40 is also conserved among all calcyon mammalian proteins. However, only in primate lineages does Pro40 lie within a consensus (P-P-L-P) binding motif for Group II WW domain containing proteins (Bedford et al., 1997; Ilsley et al., 2002) (Fig. 4). WW domains are found in about 50 human proteins (Ingham et al., 2005), and WW domain containing proteins play a role in a wide range of cell processes including cell-cycle control, signal transduction, RNA splicing, and ubiquitin ligation (Ilsley et al., 2002). Like SH3 domains, WW domains mediate protein:protein interactions by binding proline rich motifs (Ilsley et al., 2002) and are grouped into four classes based on ligand specificity. Group I WW domain proteins like YAP65 recognize P-P-X-Y motifs; whereas the group II proteins like formin binding protein 11 (FPB11) bind P-P-L-P. The group III WW domain proteins including FEB65 bind P-P-R motifs, and the group IV proteins like Nedd4 bind (S/T)pP motifs.

Intrigued that the P40L SNP might provide structure/function information on calcyon, we obtained plasmids encoding glutathione *S* transferase (GST) fusions to prototypical group I, II and III WW domains, as well as the SH3 domain of Abl (binds a P-X-X-P motif) (Bedford et al., 1997). An S-Tag fusion protein including the Pro40 variant of calcyon (hWT(1-80)) was found to bind the group II WW domain of FBP11 in protein overlay assays (Fig. 5A, B). In contrast, binding to the group I and III WW domains of YAP65 or FEB65, or the SH3 domain of Abl was not detected (Fig. 5B). Binding of the FBP11 WW domain to either the Leu40, or the murine calcyon S-tagged fusion proteins (hP40L or mWT) was similarly not detectable (Fig. 5C). These data suggest that the P-P-L-P motif in calcyon can interact with a group II WW domain containing proteins like FBP11.

The overlay data shown in Fig. 5 argue that PPLP motif found in calcyon in primate lineages is a functional WW domain binding site since it can be disrupted by a single amino acid substitution. Further studies are required to confirm whether the PPLP motif represents a bona fide example of lineage-specific functional change. However, as discussed above, in the 160 chromosomes investigated in this region, only 1 nucleotide substitution was identified. Notably, the calcyon PPLP sequence is found exclusively in primate lineages with the suggestion of emergence through one or possibly two substitutions along the primate lineage as shown in Fig. 4. The low frequency of SNPs in this newly emergent motif is consistent with selection playing a role in its maintenance.

Previous biochemical analysis indicates calcyon acts as a clathrin adaptor or accessory protein as it facilitates clathrin coated vesicle formation via a cytosolic segment in the protein (Xiao et al., 2006). Intriguingly, the PPLP motif lies in the predicted luminal or extracellular segment of the calcyon protein. Clathrin coated vesicles assemble in a dynamic fashion to sort cargo from intracellular vesicles to the cell surface, in addition to serving as a compartment for the internalization of receptors, growth factors and nutrients from the cell surface (Deborde et al., 2008; Roth, 2008). The luminal/extracellular location of the PPLP motif in primate calcyon suggests that in neurons the interaction takes place either at the plasma membrane or within a vesicle or organelle, or perhaps both, as calcyon accumulates at cell surface in a calcium-dependent fashion (Ali and Bergson, 2003). Indeed, as the array of clathrin accessory and adaptor proteins, including tissue specific isoforms, is greatly expanded in vertebrates and mammals, one would predict an expansion in the regulation of clathrin coated vesicle formation during vertebrate evolution as well (Schmid and McMahon, 2007). Hence, it is tempting to speculate that although yet to be characterized, the interaction of the putative WW domain-containing ligand with the PPLP motif in calcyon is associated with an endocytic or exocytic specialization (cargo or regulatory step) unique to the primate CNS.

Conclusions

Phylogenetic analysis of NEEP21, calcyon, and P19 protein coding sequences points to emergence of an entirely new gene family involved in vesicle trafficking in neurons as late as early in vertebrate speciation. The tripartite family is highly conserved and bears no significant similarity with domains found in other vertebrate or invertebrate proteins. Our studies in zebrafish also indicate that predominant expression in the CNS is another conserved feature. Previous studies with mammalian orthologs of NEEP21 and calcyon implicate the family in molecular pathways important for synaptic connectivity such as the trafficking of neurotransmitter receptors at synapses, and the transcytosis of adhesion molecules to axons (von Bartheld, 2004). The phylogenetic studies reported here found evidence for positive selection in the P19 and calcyon, but not in the NEEP21 lineages. In particular, domains with significantly different rates of evolution were identified in NEEP21 and calcyon. Notably, we discovered a novel protein interaction motif present only in the primate calcyon lineages. Future studies could examine whether these domains and motifs potentially underlie their different subcellular localization or their distinct roles in receptor trafficking at synapses, or in the transcytosis of cargos from dendrites to axons.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the NIH MH063271 (C.B.). This work was also funded by an institutional NIH-INBRE award to University of Wyoming and by NSF award 0743374 (DAL). We thank Haiqin Jin for her excellent technical support.

References

- Alberi S, Boda B, Steiner P, Nikonenko I, Hirling H, Muller D. The endosomal protein NEEP21 regulates AMPA receptor-mediated synaptic transmission and plasticity in the hippocampus. *Mol Cell Neurosci*. 2005; 29:313–319. [PubMed: 15911354]
- Ali MK, Bergson C. Elevated intracellular calcium triggers recruitment of the receptor cross-talk accessory protein calcyon to the plasma membrane. *J Biol Chem*. 2003; 278:51654–51663. [PubMed: 14534309]
- Anisimova M, Liberles DA. The quest for natural selection in the age of comparative genomics. *Heredity*. 2007; 99:567–579. [PubMed: 17848974]
- Bai J, He F, Novikova SI, Undie AS, Dracheva S, Haroutunian V, Lidow MS. Abnormalities in the dopamine system in schizophrenia may lie in altered levels of dopamine receptor-interacting proteins. *Biol Psychiatry*. 2004; 56:427–440. [PubMed: 15364041]
- Barackskay KL, Haroutunian V, Meador-Woodruff JH. Dopamine receptor signaling molecules are altered in elderly schizophrenic cortex. *Synapse*. 2006; 60:271–279. [PubMed: 16786528]
- Bedford MT, Chan DC, Leder P. FBP WW domains and the Abl SH3 domain bind to a specific class of proline-rich ligands. *EMBO J*. 1997; 16:2376–2383. [PubMed: 9171351]
- Berglund-Sonnhammer AC, Steffansson P, Betts MJ, Liberles DA. Optimal gene trees from sequences and species trees using a soft interpretation of parsimony. *J Mol Evol*. 2006; 63:240–250. [PubMed: 16830091]
- Bredt DS, Nicoll RA. AMPA receptor trafficking at excitatory synapses. *Neuron*. 2003; 40:361–379. [PubMed: 14556714]
- Clinton SM, Ibrahim HM, Frey KA, Davis KL, Haroutunian V, Meador-Woodruff JH. Dopaminergic abnormalities in select thalamic nuclei in schizophrenia: involvement of the intracellular signal integrating proteins calcyon and spinophilin. *Am J Psychiatry*. 2005; 162:1859–1871. [PubMed: 16199832]
- Dai R, Bergson C. Structure and expression of the murine calcyon gene. *Gene*. 2003; 311:111–117. [PubMed: 12853145]
- Davidson HT, Xiao J, Dai R, Bergson C. Calcyon is necessary for activity-dependent AMPA receptor internalization and LTD in CA1 neurons of hippocampus 1. *Eur J Neurosci*. 2009; 29:42–54. [PubMed: 19120439]
- Debaigt C, Hirling H, Steiner P, Vincent JP, Mazella J. Crucial role of neuron-enriched endosomal protein of 21 kDa in sorting between degradation and recycling of internalized G-protein-coupled receptors. *J Biol Chem*. 2004; 279:35687–35691. [PubMed: 15187090]
- Deborde S, Perret E, Gravotta D, Deora A, Salvarezza S, Schreiner R, Rodriguez-Boulant E. Clathrin is a key regulator of basolateral polarity. *Nature*. 2008; 452:719–723. [PubMed: 18401403]
- Dehal P, Boore JL. Two rounds of whole genome duplication in the ancestral vertebrate. *PLoS Biol*. 2005; 3:e314. [PubMed: 16128622]
- Emes RD, Pocklington AJ, Anderson CN, Bayes A, Collins MO, Vickers CA, Croning MD, Malik BR, Choudhary JS, Armstrong JD, Grant SG. Evolutionary expansion and anatomical specialization of synapse proteome complexity. *Nat Neurosci*. 2008; 11:799–806. [PubMed: 18536710]
- Finn RD, Tate J, Mistry J, Coghill PC, Sammut SJ, Hotz HR, Ceric G, Forslund K, Eddy SR, Sonnhammer EL, Bateman A. The Pfam protein families database. *Nucleic Acids Res*. 2007
- Gu X, Vander VK. DIVERGE: phylogeny-based analysis for functional-structural divergence of a protein family. *Bioinformatics*. 2002; 18:500–501. [PubMed: 11934757]
- Heynen AJ, Quinlan EM, Bae DC, Bear MF. Bidirectional, activity-dependent regulation of glutamate receptors in the adult hippocampus in vivo. *Neuron*. 2000; 28:527–536. [PubMed: 11144361]
- Hill RS, Walsh CA. Molecular insights into human brain evolution. *Nature*. 2005; 437:64–67. [PubMed: 16136130]
- Hughes T, Liberles DA. The pattern of evolution of smaller-scale gene duplicates in mammalian genomes is more consistent with neo- than subfunctionalisation. *J Mol Evol*. 2007; 65:574–588. [PubMed: 17957399]

- Hughes T, Liberles DA. Whole-genome duplications in the ancestral vertebrate are detectable in the distribution of gene family sizes of tetrapod species. *J Mol Evol.* 2008; 67:343–357. [PubMed: 18815825]
- Ilsley JL, Sudol M, Winder SJ. The WW domain: linking cell signalling to the membrane cytoskeleton. *Cell Signal.* 2002; 14:183–189. [PubMed: 11812645]
- Ingham RJ, Colwill K, Howard C, Dettwiler S, Lim CS, Yu J, Hersi K, Raaijmakers J, Gish G, Mbamalu G, Taylor L, Yeung B, Vassilovski G, Amin M, Chen F, Matskova L, Winberg G, Ernberg I, Linding R, O'donnell P, Starostine A, Keller W, Metelnikov P, Stark C, Pawson T. WW domains provide a platform for the assembly of multiprotein networks. *Mol Cell Biol.* 2005; 25:7092–7106. [PubMed: 16055720]
- Koh PO, Bergson C, Undie AS, Goldman-Rakic PS, Lidow MS. Up-regulation of the D1 dopamine receptor-interacting protein, calcyon, in patients with schizophrenia. *Arch Gen Psychiatry.* 2003; 60:311–319. [PubMed: 12622665]
- Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. Clustal W and Clustal X version 2.0. *Bioinformatics.* 2007; 23:2947–2948. [PubMed: 17846036]
- Laurin N, Misener VL, Crosbie J, Ickowicz A, Pathare T, Roberts W, Malone M, Tannock R, Schachar R, Kennedy JL, Barr CL. Association of the calcyon gene (DRD11P) with attention deficit/hyperactivity disorder. *Mol Psychiatry.* 2005; 10:1117–1125. [PubMed: 16172615]
- Liberles DA. Evaluation of methods for determination of a reconstructed history of gene sequence evolution. *Mol Biol Evol.* 2001; 18:2040–2047. [PubMed: 11606700]
- Malinow R, Malenka RC. AMPA receptor trafficking and synaptic plasticity. *Annu Rev Neurosci.* 2002; 25:103–126. [PubMed: 12052905]
- McCormack SG, Stornetta RL, Zhu JJ. Synaptic AMPA receptor exchange maintains bidirectional plasticity. *Neuron.* 2006; 50:75–88. [PubMed: 16600857]
- Nicholls RE, Alarcon JM, Malleret G, Carroll RC, Grody M, Vronskaya S, Kandel ER. Transgenic mice lacking NMDAR-dependent LTD exhibit deficits in behavioral flexibility. *Neuron.* 2008; 58:104–117. [PubMed: 18400167]
- Posada D, Crandall KA. MODELTEST: testing the model of DNA substitution. *Bioinformatics.* 1998; 14:817–818. [PubMed: 9918953]
- Roth MG. Cell biology: porter and sorter. *Nature.* 2008; 452:706–707. [PubMed: 18401400]
- Saberan-Djoneidi D, Marey-Semper I, Picart R, Studler JM, Tougard C, Glowinski J, Levi-Strauss M. A 19-kDa protein belonging to a new family is expressed in the Golgi apparatus of neural cells. *J Biol Chem.* 1995; 270:1888–1893. [PubMed: 7829526]
- Saberan-Djoneidi D, Picart R, Escalier D, Gelman M, Barret A, Tougard C, Glowinski J, Levi-Strauss M. A 21-kDa polypeptide belonging to a new family of proteins is expressed in the Golgi apparatus of neural and germ cells. *J Biol Chem.* 1998; 273:3909–3914. [PubMed: 9461575]
- Schmid EM, McMahon HT. Integrating molecular and network biology to decode endocytosis. *Nature.* 2007; 448:883–888. [PubMed: 17713526]
- Stamatakis A. RAxML-VI-HP: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics.* 2006; 22:2688–2690. [PubMed: 16928733]
- Steiner P, Sarria JC, Glauser L, Magnin S, Catsicas S, Hirling H. Modulation of receptor cycling by neuron-enriched endosomal protein of 21 kD. *J Cell Biol.* 2002; 157:1197–1209. [PubMed: 12070131]
- Steiner P, Alberi S, Kulangara K, Yersin A, Sarria JC, Regulier E, Kasas S, Dietler G, Muller D, Catsicas S, Hirling H. Interactions between NEEP21, GRIP1 and GluR2 regulate sorting and recycling of the glutamate receptor subunit GluR2. *EMBO J.* 2005; 24:2873–2884. [PubMed: 16037816]
- Sudhof TC. The synaptic vesicle cycle. *Annu Rev Neurosci.* 2004; 27:509–547. [PubMed: 15217342]
- Teyler TJ, Discenna P. Long-term potentiation as a candidate mnemonic device. *Brain Res.* 1984; 319:15–28. [PubMed: 6324959]
- Thisse C, Thisse B, Schilling TF, Postlethwait JH. Structure of the zebrafish snail1 gene and its expression in wild-type, spadetail and no tail mutant embryos. *Development.* 1993; 119:1203–1215. [PubMed: 8306883]

- von Bartheld CS. Axonal transport and neuronal transcytosis of trophic factors, tracers, and pathogens 3. *J Neurobiol.* 2004; 58:295–314. [PubMed: 14704960]
- Wang HY, Chien HC, Osada N, Hashimoto K, Sugano S, Gojobori T, Chou CK, Tsai SF, Wu CI, Shen CK. Rate of evolution in brain-expressed genes in humans and other primates. *PLoS Biol.* 2007; 5:e13. [PubMed: 17194215]
- Westerfield, M. *The Zebrafish Book: a guide for the laboratory use of zebrafish (Brachydanio rerio)*. University of Oregon Press; Eugene, OR: 1995.
- White R, Gonsior C, Kramer-Albers EM, Stohr N, Huttelmaier S, Trotter J. Activation of oligodendroglial Fyn kinase enhances translation of mRNAs transported in hnRNP A2-dependent RNA granules. *J Cell Biol.* 2008; 181:579–586. [PubMed: 18490510]
- Whitlock JR, Heynen AJ, Shuler MG, Bear MF. Learning induces long-term potentiation in the hippocampus. *Science.* 2006; 313:1093–1097. [PubMed: 16931756]
- Williams NM, Norton N, Williams H, Ekholm B, Hamshire ML, Lindblom Y, Chowdari KV, Cardno AG, Zammit S, Jones LA, Murphy KC, Sanders RD, McCarthy G, Gray MY, Jones G, Holmans P, Nimgaonkar V, Adolfson R, Osby U, Terenius L, Sedvall G, O'Donovan MC, Owen MJ. A systematic genomewide linkage study in 353 sib pairs with schizophrenia. *Am J Hum Genet.* 2003; 73:1355–1367. [PubMed: 14628288]
- Xiao J, Dai R, Negyessy L, Bergson C. Calcyon, a novel partner of clathrin light chain, stimulates clathrin-mediated endocytosis. *J Biol Chem.* 2006; 281:15182–15193. [PubMed: 16595675]
- Yang Z. PAML 4: phylogenetic analysis by maximum likelihood 1. *Mol Biol Evol.* 2007; 24:1586–1591. [PubMed: 17483113]
- Yap CC, Wisco D, Kujala P, Lasiecka ZM, Cannon JT, Chang MC, Hirling H, Klumperman J, Winckler B. The somatodendritic endosomal regulator NEEP21 facilitates axonal targeting of L1/ NgCAM. *J Cell Biol.* 2008; 180:827–842. [PubMed: 18299352]
- Zelenin S, Aperia A, Diaz HR. Calcyon in the rat brain: cloning of cDNA and expression of mRNA. *J Comp Neurol.* 2002; 446:37–45. [PubMed: 11920718]

Figure 1.

Expression of LOC324971 mRNA during zebrafish (*Danio rerio*) embryogenesis. (A) Temporal expression of LOC324971 mRNA by RT-PCR analysis (open arrow). Stages are as indicated; hpf, hours post-fertilization. Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) expression is shown for comparison (solid arrow). (B–D) Detection of LOC324971 mRNA by *in situ* hybridization. (B) 18 hpf. mRNA is expressed in the developing forebrain (open arrow), midbrain (open triangles), and hindbrain (solid arrows). The otic vesicle is marked by an asterisk. (C) 24 hpf. Expression in the telencephalon (open arrow), diencephalon (line arrow), tegmentum (open triangle), hindbrain (solid arrows), and spinal cord (bracket) is marked as indicated. (D) 48 hpf. Symbols the same as in (C) except the midbrain is marked by a bracket. Panels B and D are dorsal views of flat mounted embryos and panel C is a lateral view of whole-mount embryo. Anterior is to the left in panels B–D. Scale bar in panel B is 50µm and in panels C, D is 100µm.

Figure 3.

Rooted Phylogenetic Tree. Predicted evolutionary relationship of the P19, NEEP21, and calcyon orthologs (shown in Fig. 2) constructed using RAXML (Stamatakis, 2006) with the GTR+ Λ +I model. A. The tree branch detected as undergoing positive selection using the Ka/Ks ratio from a non-statistical approach plus from the free ratios model and confirmed using branch and branch-site tests according to PAML (Yang, 2007) is shown in blue, while that from the free ratios model and confirmed by both the branch-sites and branches model is shown in brown. The branch showing positive selection by the free ratios model, but not by the other two models is labeled in pink. B. The same tree is depicted with branches represented proportional to their lengths from RAXML.

Positive selection detected with the null model, branches-sites model and ASR +PBL

Positive selection detected with the null model and the branchessites model

Positive selection detected only with the null model

Figure 4.
The PPLP motif as a primate innovation. Sequence alignment of mammalian calcyon from diverse species in which identical amino acids are boxed. The most parsimonious ancestral sequence reconstruction scenarios are shown on the tree.

Figure 5.

The prime PPLP motif mediates interaction of calcyon with WW domain proteins. **A.** Diagram of the transmembrane (TM) domain structure of calcyon. The N-terminal 80 residues of murine calcyon and both the P40 and L40 variants of hCalcyon were fused to the S-Tag for 'overlay' assays shown in **B** and **C**. **B.** Binding specificity of the calcyon WW domain. Bacteria were transformed with pGEX expressing GST fusion proteins of the indicated WW domains (FBP11, FEB65, and YAP65), or the SH3 domain of Abl. Fusion proteins were induced with isopropyl β -D-thiogalactoside (IPTG). Uninduced (U) and induced (I) bacterial extracts were separated on an SDS-PAGE gel, stained with coomassie blue (lower panel) or transferred to PVDF membrane (upper panel). After blocking, blots were incubated with 1 mg purified S-hWT (1-80) in blocking buffer overnight. Binding of S-Tag-hWT (1-80) was detected with anti-S protein conjugated to HRP, followed by ECL and exposure to autoradiography film. **C.** The FBP11 WW domain binds the calcyon hWT(P40), but not to either the murine or hP40L S-Tag fusion proteins. Overlay assay performed by incubating purified GST-FBP11 with a blot of uninduced and induced lysates of bacteria transformed with the indicated calcyon S-Tag fusion proteins. Bound GST-FBP11 was detected with HRP conjugated anti-GST antibodies as in **B**. For both **B** and **C**, similar results were obtained in at least two independent experiments.

Table 1

Refined data set of CDART and TBLASTN searches.

Proteins retrieved in CDART and TBLASTN searches with human P19, NEEP21 and calcyon.

Gene	NCBI Accession Number*	Organism
P19	NP_032767	Mus musculus
	NP_001029324	Rattus norvegicus
	BAE01858	Macaca fascicularis
	NP_001069764	Bos taurus
	NP_057064	Homo sapiens
	(BU320355.1)	Gallus gallus
	NP_001072197	Xenopus tropicalis
	NP_001090016	Xenopus laevis
	NP_957167	Danio rerio
	CAG11727	Tetraodon nigroviridis
	(BJ731590)	Oryzias latipes
NEEP21	CAG01867	Tetraodon nigroviridis
	(BJ731194)	Oryzias latipes
	NP_001072849	Xenopus tropicalis
	NP_989896	Gallus gallus
	NP_035072	Mus musculus
	NP_077042	Rattus norvegicus
	NP_001071425	Bos taurus
	NP_055207	Homo sapiens
	BAE87649	Macaca fascicularis
Calcyon	(CN013832.1)	Danio rerio
	(BJ704146)	Oryzias latipes
	NP_001069421	Bos taurus
	NP_056537	Homo sapiens
	NP_081045	Mus musculus
	NP_620270	Rattus norvegicus

Likelihood Ratio Test Results for the P19, Calcyon, and NEEP21 gene families. Branch numbers are as labeled in Figure 3A.

Parameter estimates and log likelihood scores Model									
	P19	p	l	NEEP21	p	l	Calcyon	p	l
Null Model	w = 0.039	22	-1676.77	w = 0.029	20	-2397.85	w = 0.101	20	-2263.16
	on branch 1, w = 999.0	41	-1652.77	on branch 2, w = 999.00	37	-2390.05	on branch 6, w = 999.00	37	-2239.75
Free-Ratio Model				on branch 3, w = 999.00			on branch 7, w = 999.00		
				on branch 4, w = 999.00					
Branches-sites Null Model				on branch 5, w = 23.800					
	p0 = 0.97691, w0 = 0.03811	23	-1674.02	p0 = 0.99368, w0 = 0.02747	21	-2395.48	p0 = 0.92047, w0 = 0.08631	21	-2252.47
	p1 = 0.02309, w1 = 1.00000			p1 = 0.00632, w1 = 1.00000			p1 = 0.07953, w1 = 1.00000		
	on branch 1	25	-1670.72	on branch 2			on branch 6	23	-2219.98
Branches-sites Model	p0 = 0.84028, w0 = 0.03883			p0 = 0.00000, w0 = 0.02371	23	-2394.61	p0 = 0.79419, w0 = 0.06354		
	p1 = 0.00885, w1 = 1.00000			p1 = 0.00000, w1 = 1.00000			p1 = 0.03893, w1 = 1.00000		
	p2a = 0.14930, w2a = 999.0000			p2a = 0.99354, w2a = 1.00000			p2a = 0.15908, w2a = 999.0000		
	p2b = 0.00157, w2b = 999.0000			p2b = 0.00646, w2b = 1.00000			p2b = 0.00780, w2b = 999.0000		
Branches Model				on branch 3			on branch 7	23	-2251.90
				p0 = 0.51862, w0 = 0.02754	23	-2393.52	p0 = 0.48078, w0 = 0.08255		
				p1 = 0.00361, w1 = 1.00000			p1 = 0.04474, w1 = 1.00000		
				p2a = 0.47447, w2a = 999.0000			p2a = 0.43408, w2a = 1.00000		
Branches Model				p2b = 0.00330, w2b = 999.0000			p2b = 0.04040, w2b = 1.00000		
				on branch 4					
				p0 = 0.00000, w0 = 0.02638	23	-2394.97			
				p1 = 0.00000, w1 = 1.00000					
Branches Model				p2a = 0.99389, w2a = 103.03332					
				p2b = 0.00611, w2b = 103.03332					
				on branch 5					
				p0 = 99367, w0 = 0.02747	23	-2395.48			
Branches Model				p1 = 0.00632, w1 = 1.00000					
				p2a = 0.00001, w2a = 1.00000					
				p2b = 0.00000, w2b = 1.00000					
	on branch 1, w = 999.0	23	-1674.69	on branch 2, w = 0.4652	21	-2396.97	on branch 6, w = 999.00	21	-2257.38

Parameter estimates and log likelihood scores Model	P19	<i>p</i>	<i>l</i>	NEEP21	<i>p</i>	<i>l</i>	Calcyon	<i>p</i>	<i>l</i>
				on branch 3, <i>w</i> = 999.00	21	-2396.98	on branch 7, <i>w</i> = 0.1539	21	-2263.12
				on branch 4, <i>w</i> = 999.00	21	-2397.41			
				on branch 5, <i>w</i> = 0.0001	21	-2397.85			
Likelihood ratio test statistics									
Null Model vs branches-sites null Model		2(<i>l</i>)	df	<i>P</i> value	2(<i>l</i>)	df	<i>P</i> value		
		5.492	2	0.0642	4.743	2	0.0933		
Null Model vs Free ratio Model		48.001	19	0.0003	15.602	17	0.5522		
Branches-sites Model vs Null Model	on branch 1	12.096	3	0.0071	on branch 2	6.480	3	0.0905	on branch 6
					on branch 3	8.665	3	0.0341	on branch 7
					on branch 4	5.760	3	0.1239	
					on branch 5	4.743	3	0.1916	
Branches-sites Model vs Branches-sites null Model	on branch 1	6.604	2	0.0368	on branch 2	1.737	2	0.4196	on branch 6
					on branch 3	3.923	2	0.1407	on branch 7
					on branch 4	1.017	2	0.6015	
					on branch 5	0.006	2	0.9970	
Branches Model vs Null Model	on branch 1	4.162	1	0.0413	on branch 2	1.756	1	0.1851	on branch 6
					on branch 3	1.737	1	0.1876	on branch 7
					on branch 4	0.879	1	0.3486	
					on branch 5	0.000	1	0.9980	

Table 3

Divergence in amino acid substitution rates within the P19, NEEP21, and calcyon clades.

	P19/NEEP21	P19/Calcyon	NEEP21/Calcyon
Theta ML	0.0672	0.0010	0.9992
Alpha ML	1.0129	1.8474	1.4900
SE Theta	0.3909	0.0224	0.2303
LRT Theta	0.0296	0.0000	18.8220