Beyond Neuroanatomy: Novel Approaches to Studying Brain Evolution

Alexander A. Pollen\textsuperscript{a}  Hans A. Hofmann\textsuperscript{b}

\textsuperscript{a}Program in Neuroscience, Stanford University, Stanford, Calif., and \textsuperscript{b}Section of Integrative Biology, Institute for Molecular and Cellular Biology, Institute for Neuroscience, The University of Texas at Austin, Austin, Tex., USA

**Introduction**

Remarkable differences in brain and behavior can be found even among closely related species. By studying these differences, the field of brain evolution can highlight relationships between structure and function, reveal constraints and selective pressures, and address questions about the evolution of our own brains. However, the field has been criticized for focusing too heavily on size measures of the whole brain or of functionally heterogeneous structures. A recent review argues that more measurements and correlations involving brain structure size will not further our understanding of the function or evolution of the nervous system [Healy and Rowe, 2007]. Thus a major challenge for the field is to develop experimental approaches within a comparative framework that allow the functional analysis of neural phenotypes in the context of development, physiology and behavior.

In the following, we present four comparative approaches for studying variability beyond neuroanatomy (see fig. 1 for a conceptual overview), and we discuss examples of the functional insights – from neurobiology and other areas – yielded by these approaches (table 1). First, we emphasize the crucial, yet often under-appreciated, role behavioral testing has to play in any comparative analysis. Second, in line with much recent attention given to the burgeoning field of ‘evo-devo’, we highlight how careful analyses of neural development and physiology across species, given recent molecular advances, can provide important and often unexpected insights into...
functional relationships as well as constraints. Third, exploiting modern techniques in genetics and genomics allows us to identify genetic and molecular factors underlying species differences in behavior and/or brain structure or function. Fourth, the advent of genome-scale sequence repositories enables us to identify regions of the genome that have experienced increased selective pressure and might contribute to adaptations of the nervous system. For all of these approaches, we underscore the vital importance of strong phylogenetic hypotheses; only if the evolutionary relationships between the species under study are taken into account, can meaningful insights emerge. Exemplary systems such as Microtus voles and electric fish have been analyzed by multiple approaches, and we return to these systems in several sections (table 1). By integrating these approaches, comparative studies can describe additional levels of variability across species and examine the functional role of these differences.

We do not mean to minimize the importance of additional fine-scale neuroanatomical approaches. Indeed,
many studies at the anatomical level have focused on specific structures that are likely relevant to ecological correlates measured [see Shumway, 2008], and have highlighted functional relationships, including between the avian high vocal center and song complexity [DeVoogd et al., 1993; Spencer et al., 2005], the avian auditory midbrain nucleus and auditory localization [Iwaniuk et al., 2006], the hippocampus and spatial learning [Krebs et al., 1989; Clayton and Krebs, 1995; Reboreda et al., 1996], and many others [e.g., Barton, 1998, 2004, 2006; Rilling et al., 2008]. Additionally, anatomical studies, coupled with phylogenetic analysis, can use gene expression, neurochemistry, and hodology to identify homologous circuits and fields across distantly related species [Maler and Hincke, 1999; Reiner et al., 2005; Castro et al., 2006; Harvey-Girard et al., 2007; Northcutt, 2008]. However, this review focuses on approaches beyond neuroanatomy that can contribute to the study of vertebrate brain evolution [see Holland and Short, 2008 for a discussion of the transition from the invertebrate to the vertebrate brain in chordate evolution; and Farris, 2008 for an in-depth review of the evolution of higher brain centers at the protostome-deuterostome boundary], and we argue that comparisons of behavior and of the underlying developmental, physiological and genetic factors are now ripe for highlighting mechanistic and selective factors underlying brain evolution.

**Approach 1: Behavioral Assays**

The astonishing diversity of brain structures across species stands in direct relationship to the equally impressive diversity of animal behavior. It has long been assumed that social behavior, foraging behavior, anti-predator behavior and cognition (to name just a few behavioral traits) underlie much of the diversity in brain structure and function [Roth and Wullimann, 2000]. Although most studies that have pursued these hypotheses have focused on comparing neuroanatomical features in a quantitative and standardized fashion, very few have employed standardized behavioral assays across species using the comparative approach and related the results to brain structure and function. It is often difficult to design an experimental paradigm that is ‘fair’, i.e., not biased towards any one of the species under investigation, as species differences that are not directly relevant to the behavior under study might interfere in non-obvious ways. For example, in studies on spatial learning using a food reward, one species might simply be more motivated by the food reward used, and yet would appear to be superior at spatial learning [Odling-Smee and Braithwaite, 2003]. It is therefore one of the greatest challenges in understanding brain evolution to devise ways of examining social, cognitive and other behaviors in a way that is efficient, robust and relevant given the differences in natural history encountered by many species.

Studies of food-storing birds provide a classic example of the comparative method as applied to neuroanatomy and behavior. These studies have been reviewed extensively elsewhere [e.g., Clayton, 1998; Bolhuis and Macphail, 2001; Shettleworth, 2003; Emery, 2006], but below we discuss the insights and controversies that emerge from these studies as examples of opportunities and challenges for the comparative study of behavior. Early studies showed that hippocampal size is greater in food-storing than in non-storing species [Krebs et al., 1989; Sherry et al., 1989], and that hippocampal size within the corvid and parid families correlates with the intensity of food-storing behavior [Healy and Krebs, 1992, 1996; Hampton et al., 1995]. These initial comparative studies of neuroanatomy separated food-storing ecology into only a few categories, but subsequent laboratory studies of behavior have found some evidence that spatial learning performance on other tasks also correlates with food-storing and neuroanatomy [e.g., Basil et al., 1996; Biegler et al., 2001]. The insights from these studies are functional – that the hippocampus likely plays a role in spatial memory – and evolutionary – that selective pressures for improved spatial learning might have driven a size increase in the hippocampus. The functional prediction complements experiments on the function of the hippocampus in standard model systems, and the selective hypothesis has been explored by analyzing hippocampus size in species that may have faced other ecological pressures, such as dispersal, brood parasitism, and migration, for improved spatial memory [reviewed in Clayton, 1998].

Although the correlation between hippocampus size and food-storing ecology appears robust [Lucas et al., 2004], the correlation of these traits with spatial learning abilities is more controversial. Macphail and Bolhuis [2001] review over thirty laboratory studies and note that even though there is a trend for superior performance in species with higher food-storing demands, a few studies show the opposite result and many studies do not show significant differences. These imperfect correlations are likely due in part to limitations of the comparative method. For example, regressions examining functional hypotheses about brain structure size (e.g., correlations of...
forebrain and social group size) also cannot explain all data points, even when the relationship is highly significant [Shettleworth, 2003]. Nonetheless, studies of spatial learning compare fewer species than studies of neuroanatomy, making counter-examples more difficult to interpret and highlighting the need for efficient behavioral assays that can be repeated many times. Conflicting results could also stem from the use of different experimental paradigms across studies, which might involve different behavioral capacities (e.g., memory retention over short vs. long timescales) and different contextual variables (e.g., size and type of reward, type of stimuli used, history of animals), and could thus be more suitable or biased towards some species. One way to address this issue of fairness, discussed by Shettleworth [2003], has been to ask a slightly different question: does food-storing ecology correlate with an increased preference for spatial cues relative to color cues in memory tasks? In a sense, measuring this within-species preference normalizes for contextual variables in the paradigm that may contribute to absolute differences across species, and results suggest that spatial cues are in fact more important to food-storing species [reviewed in Shettleworth, 2003].

In another review, Bolhuis and Macphail [2001] advance an even more fundamental critique of comparative studies. They argue that comparative studies address only levels of ultimate causation – the functional role of a phenotype and the selective forces that favor the phenotype – and cannot provide insights into the mechanistic basis of behavior. However, comparative studies in several systems have been extremely successful at revealing mechanistic factors underlying behavioral evolution. For example, the fast-start escape behavior of fishes is an excellent comparative model system for analyzing the neural circuitry and musculoskeletal function underlying this vital behavior. In an elegant series of experiments Hale et al. [2002] carefully described the electromyogram (EMG) features and kinematics of rapid escape behavior in four fish species at key positions in the vertebrate phylogeny and showed that several of the control features associated with this behavior exhibit a mosaic pattern of ancestral and derived traits. The mechanistic insights from these studies depended on the development of robust, ‘fair’ (i.e., unbiased), and efficient behavioral assays. A similar case can be made for Microtus voles, where species comparisons beyond neuroanatomy have come to include developmental, physiological and genetic differences (which we discuss in subsequent sections).

Microtus voles are monogamous and biparental, whereas others are polygynous with maternal care. By housing monogamous prairie vole males with a female in the laboratory, it is possible to induce a partner preference, which can be quantitatively measured as ‘social affiliation’ in the partner association test. Using a two-choice paradigm, males are placed between the familiar female and an unfamiliar female, and the amount of time spent near each is measured. Importantly, this assay can also be applied to polygynous species and thus overcomes the issue of ‘fairness’ (or species bias). At the same time, it is sufficiently robust to characterize subtle within-species differences [Hammock and Young, 2005], and efficient enough to achieve reasonable sample sizes.

The partner association test, in combination with comparisons beyond neuroanatomy, has made it possible to examine which neurobiological differences in vole species contribute to divergence in social behavior. The neuropeptide arginine vasopressin was implicated as a candidate for behavioral differences based on studies of vasopressin in other systems and on differences in the distribution of the neuropeptide and its receptor across vole species. By injecting the arginine vasopressin or a receptor antagonist, the partner association test demonstrated that vasopressin is necessary and sufficient for the formation of partner preference [Winslow et al., 1993]. This example and many subsequent experiments in Microtus [reviewed in Lim and Young, 2006] illustrate the potential for comparative studies of behavior to provide mechanistic insights. In this case, early behavioral comparisons categorized Microtus species as monogamous or non-monogamous, and the subsequent development of an unbiased, robust, and efficient assay in prairie voles made it possible to examine candidate proximate factors derived from other comparative studies.

Although less complex than affiliative behavior, feeding behavior can often be objectively compared across species by using the underlying morphology as a proxy for behavior. Two well-known examples that illustrate this point particularly well are the association between beak shape and trophic niche in Darwin’s finches [Grant and Grant, 2006], and the movement of mouthparts in relation to feeding specialization in cichlid fishes [Liem, 1979; Wainwright et al., 2001]. However, no attempts have been made to correlate these morphological structures with brain structure or function.

An exceptionally elegant system for comparing behavior across species in an unbiased (‘fair’) manner – and one that comes with much neurobiological insight – is available for weakly electric fishes. These animals can
sense changes to the electric field around them and produce electrical emissions for communication and electrolocation, which can vary dramatically in terms of duration and shape [see, for example, Hopkins, 1999]. Because active electro-sensing and communication is conducted already in the currency of the nervous system – i.e., changes in the membrane potential of neurons, receptors and electrocytes – a recording electrode is all that is needed for acquiring comparable behavior data. Few other systems exist that allow such elegant comparisons, but there is an increasing need to overcome this obstacle independent of taxonomic group and the type of behavior under investigation, especially given the opportunities that are arising from other approaches. In the next three sections, we discuss these comparative approaches beyond neuroanatomy that can be used to identify proximate factors that may contribute to brain and behavioral evolution.

**Approach 2: Development and Physiology as Functionally Relevant Traits**

Comparative approaches to development and physiology have yielded mechanistic insights about factors underlying differences in brain structure and behavior, and there is great potential for future work. Because the neo-cortex is the anatomical location underlying many higher human cognitive functions, there is great interest in understanding how and why the human cortex expanded. Many studies have compared the size of the cortex across mammals and evaluated whether ecological factors or developmental constraints are correlated with changes in size [e.g., Harvey and Krebs, 1990; Finlay and Darlington, 1995; Reader and Laland, 2002]. However, by combining the comparative method with advances in our understanding of the molecular and developmental basis of cortical development, recent studies have highlighted mechanistic factors that may contribute to cortical expansion.

Two general models have been advanced to explain the evolutionary expansion of the cortex. Both models are based on modifications to one of the three phases of cell division that produce cortical excitatory neurons. In phase 1, progenitors along the neuroepithelium divide symmetrically and ultimately produce radial glia cells. The Radial Unit Hypothesis [Caviness et al., 1995; Rakic, 1995] predicts that simply altering phase 1 can scale the size of the cortex. During phase 2, radial glia cells divide asymmetrically producing a single daughter neuron and a radial glia cell. In phase 3, radial glia cells produce intermediate progenitor cells, which divide symmetrically in the subventricular zone producing at least two daughter neurons. A radial glia cell that produces an intermediate progenitor cell will thus ultimately produce at least twice the number of neurons per a cell division, compared with directly producing a single neuron. The Intermediate Progenitor Hypothesis [Kriegstein et al., 2006] posits that proportionately more neurogenesis occurs during phase 3 in the evolutionary expansion of the cortex. These models would be impossible to reconcile by focusing experiments on a single model system, such as mouse. In fact, mutations in mice can increase the size of the cortex according to either model. Over-expression of a stabilized beta-catenin transgene increases founder cell division in phase 1 [Chenn and Walsh, 2003], whereas knocking out the gene Cux2 increases the proportion of neurogenesis in phase 3 while leaving phase 1 unchanged [Cubelos et al., 2007].

Nonetheless, by applying the comparative method to developmental processes, several studies have evaluated the predictions of each of these models. The Radial Unit Hypothesis predicts that the adult cortex size should correlate with the size of the embryonic ventricular zone. The Intermediate Progenitor Hypothesis, on the other hand, predicts that the size of the adult cortex should correlate with the size of the subventricular zone (SVZ), which contains intermediate progenitor cells. Additionally, because SVZ cells produce predominantly upper layer neurons [Wu et al., 2005], the proportion of upper layer neurons should be greater in animals with increased relative cortex size. At this point, the Intermediate Progenitor Hypothesis appears to have stronger support, as it has been shown that SVZ size correlates with cortex size in turtle, rats, ferrets, macaques and humans [Smart et al., 2002; Martínez-Cerdeno et al., 2006; Bayatti et al., 2008]. Furthermore, primates and humans in particular have a greater proportion and diversity of upper-layer neurons than do other mammals [Cajal, 1909; Hill and Walsh, 2005].

Despite these insights, there is still an urgent need for more comparative experiments. For example, the Intermediate Progenitor Hypothesis predicts that local changes in SVZ thickness might contribute to gyri and sulci formation, but this has only been observed indirectly in human and macaque [Kriegstein et al., 2006]. Do changes in SVZ thickness predict gyri and sulci formation in orders that independently evolved folded brains? Similarly, does the proportion of upper-layer neurons also increase in non-primate orders with an expanded cortex? At a mo-
lecular level, do changes in the regulation or sequence of genes, such as Cux2, that control the division of intermediate progenitors, correlate with changes in neocortex size or gyriﬁcation? Birds and – to an even larger extent – teleost ﬁshes show a remarkable diversity in forebrain sizes just as seen in mammals [Huber et al., 1997; Iwaniuk and Hurd, 2005; Pollen et al., 2007; Lefebvre and Sol, 2008]. Do changes in patterns of proliferation also underlie these differences? Recent developmental studies in chick suggest that a SVZ is also present in the avian striatum and dorsal ventricular ridge, opening the door to comparisons of SVZ size across avian species [Striedter and Keefer, 2000; Charvet et al., 2007; Cheung et al., 2007]. Meanwhile, recent work describes a much more prominent role for adult neurogenesis in the forebrain of ﬁshes than in mammals, indicating a possible mecha-

nism in addition to embryonic neurogenesis by which patterns of proliferation could affect forebrain structure [Ekström et al., 2001; Zupanc et al., 2006].

Evolutionary modiﬁcations of developmental processes can have obvious effects on brain structure, but evolutionary changes in physiological processes can also affect brain structure and behavior [see also Wang, 2008]. As mentioned previously, morphological changes that affect physiology such as in the cichlid jaw apparatus or the beak of Darwin’s ﬁnches have been related to feeding behavior, and might also correlate with changes in brain structure. With respect to neurophysiology, studies of communication in weakly electric ﬁsh provide an excellent example of insights into mechanistic neurobiology that physiological comparisons combined with other approaches can produce. Below, we discuss studies of electric organ physiology in electric ﬁsh that incorporate direct sequence comparisons and signatures of adaptive selection (further discussed in Approach 3: Genetic and Genomic Analyses, and Approach 4: Signatures of Selection in DNA Sequence, respectively).

Weakly electric ﬁsh use electrical emissions for communication and electrolocation, yet the duration of signals can vary 100-fold across species, and the waveform is also variable [Hopkins, 1999]. Although communication is generally diﬃcult to study at the genetic level, ion channels are known to be a key component in electrical discharges. Na+ channels were previously implicated in the discharges of one species of electric ﬁsh [Ferrari et al., 1995], and mutations at functional sites in Na+ channels are known to aﬀect muscle and neuron ﬁring rate in many human clinical syndromes [George, 2005], making Na+ channels good candidate genes for a role in electric communication. However, identifying whether divergence in the physiology of Na+ channels of electric ﬁsh underlies species diﬀerences requires comparative studies.

Electric communication has evolved independently in two groups of teleosts, the mormyrids in Africa and the gymnotiforms in South America. Zakon et al. [2006, 2008] compared the expression and coding sequence of two candidate genes encoding sodium channels, NaV1.4a and NaV1.4b, in three gymnotiforms and one mormyrid with four related non-electric ﬁsh. They observed that NaV1.4a expression had been independently lost from muscle and gained in the electric organ in gymnotiforms and mormyrids. By integrating their physiological comparison with additional approaches, the authors asked whether changes in NaV1.4a were functional (see Approach 3: Genetic and Genomic Analyses) and adaptive (see Approach 4: Signatures of Selection in DNA Sequence). Direct sequence comparisons revealed that in both gymnotiforms and mormyrids, many highly conserved amino acids have been replaced, particularly in the domain of the channel responsible for the ﬁnal steps of fast inactivation. In fact, some of these amino acid substitutions overlap with mutations underlying human channelopathies, and are consistent with the rapid electric pulses observed in these species, whereas replacements at other sites may provide new insights into natural mechanisms for regulating the function of Na+ channels. The loss of muscle expression and gain of electric organ expression also coincided with an increased rate of change in the coding sequence of NaV1.4a in both lineages suggesting that changes in the protein were in fact adaptive. Thus comparative analysis of the physiology underlying electric emissions, combined with additional approaches, not only illuminates the neurobiological mechanism, but it also suggests that the number of ways to produce rapid electric pulses might be limited, as changes in the same ion channel gene appear to be relevant in convergent examples.

Again, comparative studies can address many more questions. In this case, only four species of electric ﬁsh were compared, and correlations between the structure of Na+ channels and the properties of the electric discharge were not explicitly examined. Future studies that include species spanning the great diversity of signal forms in electric ﬁsh may reveal further unknown properties of Na+ channels and additional convergent events. The quantitative nature of electric organ discharges also lends itself to behavioral comparisons. Moreover, Na+ channels probably act in concert with other ion channels that can also be examined. In fact, the regulation of K+...
channels has also been implicated within species in sex and life history differences in electric organ discharge patterns [Stoddard et al., 2006]. More generally, the comparative study of physiology can be used to examine the functional basis for many species-specific phenotypic differences, to identify examples of convergent evolution, and in some cases to infer the sequence of evolutionary changes [e.g., Berenbrink et al., 2005; Bridgham et al., 2006]. Thus comparative studies of developmental and physiological processes underlying natural diversity can contribute to a mechanistic and evolutionary understanding of the nervous system.

**Approach 3: Genetic and Genomic Analyses**

The ability to compare genomic variability across populations and species complements the mechanistic comparisons of Approach 2 (Development and Physiology). In many cases, developmental and physiological pathways underlying a derived trait are either poorly understood or involve too many genes to narrow down top candidates underlying evolutionary change. By contrast, genome-wide comparisons can provide unbiased measures of how well genomic regions, expression patterns, and sequence variations associate with derived traits. Two major techniques, genome-wide linkage mapping and transcriptional profiling, have been used to study the genetic basis for some morphological differences in a diverse range of species. A third technique of direct sequence comparisons can also be used to highlight genetic changes that may be functional. These techniques, particularly when combined with functional assays, are powerful ways to determine how derived traits are produced, and are also likely to illuminate novel roles for genes in producing phenotypes of interest.

Genome-wide linkage mapping involves hybridizing phenotypically divergent individuals from related populations or species, and allows for the identification of chromosomal markers that segregate with traits of interest. The power of this technique is that the actual genomic regions responsible for evolved traits are revealed, and their effect sizes and level of dominance can be estimated. Indeed, quantitative trait loci have been identified that control substantial variation in fruitfly mating behavior [Moehring and Mackay, 2004; Gleason and Ritchie, 2004], jaw and tooth specializations in cichlid fish [Streelman et al., 2003; Albertson et al., 2005; Streelman and Albertson, 2006], the reduction of pelvic spines and armor plates in stickleback fish [Shapiro et al., 2004; Colosimo et al., 2004], albinism in cavefish populations [Protas et al., 2006], and morphological traits in many other vertebrates.

These studies provide insights into fundamental evolutionary questions about the type and magnitude of genetic changes underlying recent phenotypic diversification [Kocher, 2004; Peichel, 2005]. However, the application of genome-wide linkage mapping to brain and behavior remains a challenge for several reasons. First, analysis often requires over 100 individuals, and efficient and standardized comparisons of behavior can be difficult. Second, the plasticity characterizing many neural phenotypes makes mapping difficult because phenotypes might not reliably correspond to genotypes. Finally, mapping studies identify chromosomal windows, but often lack the resolution to identify individual genes. Therefore, identifying the key gene within a large window of sequence could be easier for a trait such as pigmentation, whose genetic basis is well understood [e.g., Protas et al., 2006; Miller et al., 2007], than for behavioral traits, where the genetic basis is much less clear. Despite these challenges, neural phenotypes are amenable to genome-wide linkage mapping. Indeed, recent linkage studies have identified loci controlling anxiety in mice strains [Talbot et al., 1999; Yalcin et al., 2004] and anti-predator behavior in zebrafish [Wright et al., 2006]. Similarly, loci affecting brain structure size and neuron number in inbred mouse strains have been identified [Williams et al., 1998; Dong et al., 2007], and endophenotypes such as neural gene expression, and levels of hormones and other metabolites can be measured efficiently [e.g., Fraiser et al., 2007].

Transcriptional profiling can also be used to highlight the genetic basis for phenotypic changes [Hofmann, 2003]. This approach involves comparing gene expression in specific tissue types and developmental stages thought to underlie phenotypic differences. For a targeted set of genes, in situ hybridization can achieve the histological specificity to detect qualitative changes in expression. To reveal quantitative changes, transcript abundance can be compared across the genome using microarrays or massively parallel sequencing technologies [Renn et al., 2004; Hoheisel, 2006; Renn et al., 2008; Torres et al., 2008; Vera et al., 2008]. Compared to genome-wide linkage mapping, transcriptional profiling has some limitations for identifying the specific genetic changes responsible for phenotypic changes: Regulatory changes do not necessarily underlie phenotypic differences; cis and trans factors controlling gene expression are difficult to distinguish [Osada et al., 2006; Genissel et al., 2008]; and further experiments are required to determine...
whether expression differences are causes or effects of phenotypic differences. However, transcriptional profiling also has clear advantages over genome-wide linkage mapping: it requires fewer individuals, can be applied across species in which genetic crosses are impossible, and can provide more global information about changes in networks of gene expression.

In fact, transcriptional profiling of candidate genes has been applied to a canonical example of adaptive specialization, the beaks of Darwin’s finches, to identify a regulatory change contributing to morphological evolution. By comparing beak development in six species of Darwin’s finches and the chick, Abzhanov et al. [2004] identified the developing mesenchyme of the beak prominence as the tissue in which phenotypes likely diverge. In a targeted in situ screen of growth factor genes, the authors identified striking changes of Bmp4 expression localized to this tissue, and then applied functional tests (discussed below) to confirm a role for Bmp4 in beak morphology.

Transcriptional profiling has also been applied to compare human and chimpanzee gene expression in adult post-mortem tissues, including many brain regions [Cáceres et al., 2003; Khaitovich et al., 2004, 2006]. These initial findings provide a broad overview of differences between species and also between brain regions, but are difficult to interpret because of hybridization differences to human arrays, different cellular compositions, and the challenge of distinguishing neutral and adaptive gene dosage changes. Nonetheless, follow-up studies have confirmed cortical gene expression changes in synaptogenic thrombospondins [Cáceres et al., 2007], and multi-species arrays [Gilad et al., 2005], ‘next generation’ sequencing technologies [Torres et al., 2008; Vera et al., 2008], and new methods of analysis [Khaitovich et al., 2005] can overcome many of these limitations.

Finally, direct sequence comparisons have been used to find specific nucleotide changes that may contribute to novel phenotypes. Most frequently, this approach has been applied to candidate genes. As mentioned earlier, a number of studies in Microtus voles demonstrated that changes in the pattern of arginine vasopressin receptor expression may underlie species differences in pair-bonding behaviors [for review see Lim and Young, 2006]. Direct sequence comparisons of promoter regions have helped to elucidate the genetic basis for changes in gene expression by suggesting that an expanded microsatellite repeat in the 5’ promoter region of the monogamous but not the polygynous vole might be responsible for the observed regulatory changes [Lim et al., 2004]. However, it should be noted that the presence or absence of this expanded microsatellite repeat does not correlate with the mating systems in other Microtus vole species [Fink et al., 2006]. In addition, highly conserved non-coding elements acting as long-range enhancers can also play a major role in gene regulation [Pennachio et al., 2006]. Aiming at identifying such enhancers, Sasaki et al. [2008] studied the distribution of the AmnSINE1 retrotransposon family across mammalian genomes and found that members of the AmnSINE1 family are highly conserved within mammals, suggesting that the expansion of this retrotransposon family early in the evolution of mammals might have affected the expression of many genes, including those associated with mammalian-specific forebrain development. Of course, hypotheses based on direct sequence comparisons must be tested in functional assays.

Functional analyses are extremely valuable for evaluating the results of genome-wide linkage mapping, transcriptional profiling and especially direct sequence comparisons. For example, the expanded microsatellite version of the prairie vole arginine vasopressin promoter has been shown to modulate expression of a reporter gene in cell culture [Hammock and Young, 2004], and the overall promoter region is sufficient to drive expression of a transgene in a prairie vole-like pattern in the mouse forebrain [Young et al., 1999]. Similarly, two of the duplicated AmnSINE1 sequences, near the gene Fgf8 [which affects cortical patterning: Fukuchi-Shimogori and Grove, 2001] and near the gene Satb2 [which specifies upper-layer neuron identity: Alcamo et al., 2008; Britanova et al., 2008], were shown to drive expression of a reporter gene in forebrain domains matching components of the endogenous expression pattern of these genes [Sasaki et al., 2008]. The fact that these elements act as forebrain enhancers suggests that other members of the AmnSINE1 family might also drive neural gene expression. These enhancer elements could be particularly relevant to mammal brain evolution, because the expansion and subsequent conservation of this retrotransposon family corresponds to the expansion of the dorsal cortex from three to six layers and other structural changes in the forebrain [Sasaki et al., 2008]. Thus direct sequence comparisons coupled with expression assays can highlight sequence changes affecting gene regulation across many phylogenetic levels.

In addition to expression assays, functional tests can also use transgenic and knock-in approaches to determine whether identified genes are sufficient to drive phenotypic changes. For example, genome-wide linkage mapping in stickleback fish suggested that an allele for
ectodysplasin (EDA) was responsible for loss of armor plates. Transgenic analysis in freshwater stickleback fish showing extreme plate reduction confirmed that a mouse ortholog of the gene was sufficient to drive the formation of additional plates [Colosimo et al., 2005]. Similarly, based on the striking changes in BMP4 expression in Darwin’s finches, Abzhanov et al. [2004] used a viral vector to drive BMP4 in the mesenchyme of the chick beak prominence, producing deep and broad beaks resembling those of ground finches. In polygynous montane voles, viral mediated expression of arginine vasopressin receptor in the ventral pallidum, but not the caudate, was shown to increase affiliative behaviors reminiscent of pairbonding [Pitkow et al., 2001]. Knock-in experiments are a more precise form of analysis in which the sequence of a particular locus is actually replaced with another allele. This technique most closely recapitulates evolutionary changes, and will be extremely useful for evaluating the function of evolving protein domains and enhancer elements. For example, replacing a forelimb enhancer in mouse with a bat version that had accumulated sequence changes, increased expression of the nearby developmental gene Prx1 in the forelimb, and quantitatively increased mouse forelimb length during development [Cretkos et al., 2008]. These morphological and neurobiological examples highlight the potential of genetic and genomic analyses to generate testable hypotheses about the mechanistic basis of phenotypic evolution in the nervous system, and to complement approaches that compare developmental and physiological processes across species.

**Approach 4: Signatures of Selection in DNA Sequence**

Analyses of molecular evolution can provide a shortcut to identifying genetic changes that contribute to adaptation. Whereas direct sequence comparisons can be used to highlight putative functional changes (Approach 3: Genetic and Genomic Analyses), sequence comparisons in the context of molecular evolution models might also show that adaptive selection is the best explanation for a disproportionate number of sequence changes being fixed in a given region. One general signature of adaptive evolution useful for interspecies comparisons is to identify conserved regions of the genome that have accumulated an unusually large number of functional nucleotide changes in a particular lineage. Although most studies have focused on protein-coding genes, new techniques are being developed to study adaptive changes in non-coding regions of the genome. By correlating signatures of adaptive selection in specific genes or elements with phenotypic changes that these genes might affect, studies of molecular evolution, as in other genetic and genomic analyses, can generate testable hypotheses about the genetic basis for novel phenotypes.

How can we estimate the strength and nature of selection acting on genes? In protein-coding sequences, signatures of selection may be examined by comparing the ratio of amino acid replacement substitutions (Ka) to synonymous substitutions (Ks). Nucleotide substitutions that replace amino acids comprise putative functional changes that are subject to the forces of natural selection. In contrast, nucleotide substitutions at degenerate positions of a codon or in nearby non-functional sequences approximate neutral changes that are fixed only at a rate determined by genetic drift. If amino acid replacements are deleterious, purifying selection will reduce the Ka/Ks ratio. Indeed, in most proteins, the Ka/Ks ratio is low (~0.15), highlighting constraints on protein structure. In principle, the Ka/Ks ratio should exceed one only when adaptive selection fixes replacement substitutions more quickly than drift can fix synonymous substitutions.

By comparing the sequence of genes affecting neural development across mammals, several studies have identified correlations between Ka/Ks ratios above one and changes in brain size. For example, Evans et al. [2004a, b] analyzed the sequences of ASPM and MCPH1, two genes underlying human primary microcephaly. In both genes, very high Ka/Ks values were observed at historical branches of the primate phylogeny, prior to the divergence of great apes, and also in the lineage leading to humans, and low values were observed in non-primate mammalian lineages. Additional comparative studies have identified other correlations between Ka/Ks values in neurodevelopmental genes and mammal brain size [see Vallender, 2008]. Although it is tempting to speculate that adaptive changes in these genes might affect brain size, genes involved in neural development are frequently expressed in many other tissues, and selection could be acting on non-brain phenotypes [Ponting, 2006]. Nonetheless, these mechanistic hypotheses about the role of amino acid substitutions in brain development are testable in cell culture and in standard model systems. Additionally, evidence for historical episodes of adaptive selection, as observed for ASPM and MCPH1 in early primates, may provide a window into understanding selective pressures and constraints facing ancestral species [Messier and Stewart, 1997].
Because Ka/Ks comparisons apply only to changes in protein sequences, it has been difficult to study the evolution of non-coding sequences across species. However, the majority of functional DNA in the genome, as measured by sequence conservation, appears to be non-coding [Mouse Genome Sequencing Consortium, Waterston et al., 2002], and it has been suggested that regulatory and not coding changes underlie many phenotypic differences that arise in recent evolution [Britten and Davidson, 1971; King and Wilson, 1975; Carroll et al., 2001; but see Hoekstra and Coyne, 2007]. Thus, the ability to detect not just protein coding, but also non-coding loci that have been evolving under adaptive selection will be extremely valuable to the study of brain evolution.

Several recent studies have analyzed the human genome for signatures of selection in non-coding sequences [Pollard et al., 2006; Prabhakar et al., 2006]. As in studies applying the Ka/Ks ratio, these studies identify genomic regions with a high proportion of potentially function-altering nucleotide substitutions. However, function-altering substitutions are harder to identify in non-coding sequence than in protein-coding sequence, where the amino acid code is well understood. Nonetheless, substitutions in non-coding sequence can be considered potentially function-altering if they occur in highly conserved sequence. This is because it appears likely that highly conserved non-coding sequences have been under purifying selection for functional reasons to maintain the same sequence characteristics over great phylogenetic distances. In fact, many non-coding elements that are highly conserved between humans and chick or mouse act as transcriptional enhancers for nearby genes [Pennacchio et al., 2006] or encode RNA genes and could contribute to the evolution of gene regulation.

To study non-coding changes in the lineage leading to humans, Pollard et al. [2006] identified elements that are highly conserved between rodents and chimpanzee, but that have undergone rapid sequence changes, significant against a neutral model, in the human lineage. In the most exceptional case, the authors identified HAR1F, a novel RNA gene with an element that shares 116 out of 118 bases between chimpanzee and chicken, but has undergone 18 substitutions on the human lineage. Expression studies in macaque and human indicate that HAR1F is expressed in Cajal-Retzius neurons of the subpial granular layer during stages of cortical development in midgestation, as well as several other brain regions and tissues. Given that the subpial granular layer is particularly prominent in humans [Sidman and Rakic, 1973; Zecevic and Rakic, 2001], HAR1F sequence changes might contribute to unique aspects of the human cortex. Again, future laboratory work can test this mechanistic hypothesis.

These examples illustrate the potential for molecular evolution approaches to highlight genetic changes that could contribute to adaptive phenotypes in a given lineage. In some cases, studies can suggest adaptive roles and functionally important domains of candidate genes, such as ASPM and MCPH1 in primate brain evolution. In other cases, studies can identify novel elements or genes, such as HAR1F. In all cases, functional tests are required to evaluate whether genetic changes on a lineage contribute to phenotypic changes on the lineage. Nonetheless, results of molecular evolution studies are strengthened when they intersect with other approaches, such as a high Ka/Ks in physiologically relevant NaV1.4a channels of electric fish (Approach 2: Development and Physiology), or results from other genome scale comparisons (Approach 3: Genetic and Genomic Analyses). Given improved techniques for detecting selection in non-coding sequences, rapidly increasing sequence data from diverse species, and improved techniques for functional tests, we expect studies of molecular evolution to contribute to our understanding of the mechanistic basis of brain evolution.

**Overcoming ‘Just-So Stories’ and Phylogenetic Confounds**

Evolutionary change in the structure of coding regions and regulatory elements often alters developmental and physiological processes and may ultimately result in neuroanatomical and behavioral differences across species. Although correlation studies on brain evolution have been useful for determining large scale patterns of evolutionary change, there are serious shortcomings. Chief among them, of course, is the causation problem – correlations with ecology or other factors do not prove those factors drove the change [see Shumway, 2008]. The danger is therefore, that correlative results are explained with plausible yet unproven ‘just-so stories’ and that other (including non-adaptive) alternative explanations are ignored [see Gould and Lewontin, 1979]. Another problem for interpreting these kinds of datasets lies in the complexity of life histories animals display and the habitats in which they live. The common focus on associations between brain structure and single behavioral or ecological traits may be inconclusive, as properties of the environment or social system often correlate with
Each other [see also Healy and Rowe, 2007]. For example, in birds and primates, correlations exist between the size of forebrain regions and social group size [Reader and Laland, 2002; Burish et al., 2004; Lefebvre and Sol, 2008], yet the size of forebrain regions has also been associated with many other ecological factors [see Shumway, 2008] as well as developmental constraints with the rest of the brain [Finlay and Darlington, 1995; Finlay et al., 2001].

Until recently, genomic and mechanistic layers of brain evolution were inaccessible, and the field focused on correlating structural differences between taxa within an evolutionary context, often unaware of phylogenetic relationships. However, due to shared ancestry, comparative data sets at any phenotypic level often violate statistical assumptions of independence [Felsenstein, 1985; Harvey and Pagel, 1991]. Thus, in order to draw conclusions from the covariation of traits across taxa, one needs to take into account this phylogenetic non-independence. Assuming the phylogenetic relationships between the species studied are known, one generally accepted method to overcome the effect of shared evolutionary history is to calculate differences in (extant and ancestral) trait values between sister taxa [Felsenstein, 1985; Garland et al., 1992]. Two traits are considered evolutionarily correlated (i.e., change in one trait has been accompanied by change in the other) if these (standardized) differences – or phylogenetically independent contrasts – in one trait significantly covary with contrasts in the other trait. This approach has become very common, but the more basic problem is that a well-resolved phylogeny often does not exist. Even in groups that have been relatively well studied, there are often alternative hypotheses for the phylogenetic relationships. It is therefore important to conduct comparative analyses for the different phylogenetic hypotheses if a consensus has not yet been reached. Thus comparative studies involving the four approaches presented here must continue to avoid ‘just-so’ stories and correct for phylogenetic confounds.

**Conclusion**

The novel experimental approaches discussed here allow the integrative study of behavioral, physiological, genetic and genomic changes underlying brain function within an evolutionary framework. By utilizing these approaches together with robust phylogenies, the field of brain evolution can re-invent itself to study all levels of biological organization and ultimately uncover the driving forces, constraining factors and proximate mechanisms that have resulted in the diversity of brains and behaviors as we find them everywhere in the natural world.

**Acknowledgements**

We thank Susan Renn and Caroly Shumway for many fruitful discussions over the years and Abraham Bassan, Kim Hoke, David Kingsley, Ron Oldfield, Phil Reno, Anna Sessa, and an anonymous reviewer for critically reading earlier versions of the manuscript. Supported by NSF grant IBN-021795 and the Institute for Cellular and Molecular Biology (H.A.H.) and an NDSEG fellowship (A.A.P.).

References


Britanova O, de Juan Romero C, Cheung A, Kwan KY, Schwark M, Gyorgy A, Vogel T,


