3.05 Cortical Evolution as the Expression of a Program for Disproportionate Growth and the Proliferation of Areas

B L Finlay and P Brodsky, Cornell University, Ithaca, NY, USA

Glossary

allometry	The study of alterations in form causally correlated with changes in size.
cortical area	A proposed unit of the cortex, typically containing a topographically mapped representation of a sensory, motor, or computed surface, and a characteristic set of thalamic inputs and subcortical outputs.
module	In computer science and cognitive science, a functionally encapsulated unit performing a particular computation. In neurobiology, often used to refer to any repeating unit of structure.
pleiotropy	The case where a single gene contributes to the execution of many functions, preventing its optimization for any single function.
prosomere	Hypothesized embryonic segmental structure of the forebrain.
radial unit hypothesis	The hypothesis that the essential unit of mature cortical function, the cortical column, and also the unit of evolutionary selection and replication is the assembly of neuroblasts migrating to the cortical plate on a single radial glial guide in development.
sensory exploitation	The idea that in the evolution of communication systems, the sender evolves to maximally activate the generic sensory system of the receiver.
subnet	An assembly of neurons performing a logical unit of computation.
symmetric and asymmetric cell division	Symmetric cell divisions give rise to identical daughter cells, both capable of further division; asymmetric division produces a cell that will become a differentiated mature form.
topographic map	The feature of retention of nearest-neighbor relationships when one array of neurons projects upon another.

3.05.1 General Introduction to Developmental Structure in Brain Evolution

Evolved structures are the result of successful adaptation to the environment, and evolution occurs by the variation and selection of genetic programs, as they are expressed in the development of the organism and in mature phenotype. In this article, we will discuss the evolution of the cortex in a developmental context, focusing on how the various versions of cortex we see in different mammalian radiations are expressions of a generally conserved developmental program, produced over variable lengths of time and variable scales. In classical evolutionary biology (Gould, 1977, 1980), conserved developmental programs are often viewed as constraints, limiting the range of variation offered for selection (see Principles of Brain Scaling). Possible reasons for stabilization of developmental programs are multiple. Very early events in development may become fixed, as they may contribute to the structure of so many mature systems no single system can be caused to vary independently. Or, independent of development, single genes may contribute to disparate functions (termed pleiotropy) so that variation in the gene cannot be linked to the optimization of a single system. More recently, however, the extent to which fundamental developmental programs are conserved across phyla has proved to be quite breathtaking and caused a substantive rethinking of the significance of conservation, away from the negative-to-neutral interpretations of constraints just described. Rather, the explanatory focus has shifted to the structures of genetic systems that allow robustness and stability to coexist with variability (Gerhart and Kirschner, 1997; Radman *et al.*, 1999; Wilkins, 2001).

Conserved features include the polarity, symmetry, and segmentation of the fundamental invertebrate and vertebrate body plan (Duboule and Dollé, 1989; Graham *et al.*, 1989); the designation of areas for special senses and limbs (Callaerts *et al.*, 1997); the control of the cell cycle and symmetry-breaking events that control cell cycle entry and exit; and other features of cell specification (Gerhart and Kirschner, 1997). Many particular mechanisms central to nervous system construction are similarly stable across taxa, such as mechanisms for axon extension and inhibition (Dickson, 2002), mechanisms for approaching and crossing the neural midline (Stein and Tessier-Lavigne, 2001), and activity-dependent stabilization of synaptic connections (Greenough and Bailey, 1988; see Scaling the Brain and Its Connections). Though highly conserved, such a set of fundamental mechanisms seems ill-described as constraints. If these fundamental mechanisms represent optimal or near-optimal solutions to repeatedly encountered developmental problems, the sense of the word 'constraint' as limit is poorly applied to them.

A longer view of the definition of 'adaptation' and 'environment' than each particular animal's interaction with its immediate physical setting will provide some of the appropriate context for understanding why developmental programs are likely to be conserved along with the conserved neural circuitry that the developmental programs will eventually represent. In addition, a more sophisticated understanding of how complex behavioral functions might be distributed over the essentially conserved architecture but widely varying size of the cortical surface is absolutely essential. We will discuss the evolution of complex sensory and behavioral functions first.

An example from current neuroethology of a phenomenon called sensory exploitation throws better light on what kinds of possible adaptations are good solutions to conflicting adaptive demands (Ryan, 1998). In the first days of neuroethology, the observation that an anolis lizard might have a bright orange dewlap that it extends for aggressive displays, or a male frog a particular croak with acoustic features that attracts females of its own species were hypothesized to be roads to insight into neural coding in both sensory and motor systems (and they were, but not in the way anticipated). Researchers imagined isomorphic specializations in the nervous systems of both the senders and the receivers, from the sensory periphery on into the central nervous system.

While there must be some committed circuitry for conspecific recognition, even after years of work neuroethologists generally failed to find specific adaptations in receivers, especially in the sensory periphery, when comparing closely related species using different signaling systems – that is, there were no orange dewlap detectors in the retina, or anywhere else in early visual processing. Instead, they found that the signaler had evolved to produce a maximally perceivable signal for the receiver's sensory system or a maximally contrastive signal for the immediate environment, thus exploiting the generic visual or auditory system of the receiver (Ryan and Rand, 1995; Persons *et al.*, 1999). This makes sense when the multifunctional nature of sensory systems is understood: the same visual system that must respond to aggressive social signals must also recognize food, recognize threat, navigate terrain, and respond to a wide range of other social signals,

which simply cannot be realized computationally with any efficiency as a collection of committed detectors (Field, 1994).

The style of the initial steps of neuroethology has unfortunately traveled unmodified into some current theorizing about cortical organization, sometimes replacing detectors with modules. One example is the discussion of the evolution of primate vision and the neural representation of trichromacy in the cortex (Barton, 1998). In this case, the differentiation of a photopigment capable of improving discrimination in the red–green end of the visual spectrum is attributed to the single function of fruit detection with associated central nervous system alterations rather than a neatly placed transducer employing generic contrastive processing, which can measurably improve not only the discrimination of particular fruits, but their ripeness, leaf maturity, other edible prey, assignment of boundaries and edges to improve navigation through cluttered forest environments, social signaling, and so forth (Moller and Hurlbert, 1996; Nickle and Heymann, 1996; Regan *et al.*, 2001; Dominy and Lucas, 2001; Finlay *et al.*, 2005b). Views of how behavioral adaptations are translated into neural specializations will be important when we examine hypotheses about how cortical areas might proliferate.

What environment are organisms adapted to? In classic adaptation scenarios, individuals compete for reproductive success in a stable environment, improving their perceptual capacities, signaling of reproductive quality, and physiological and behavioral abilities in general, producing an adaptive walk through a universe of potential adaptive states responsive both to environmental pressures and the nature of the particular competition they have engaged (Dawkins, 1976, 1986). This type of scenario often generates greater and greater specialization, particularly arms races in the means of sexual selection. However, most individuals in their own lifetimes or in their immediate ancestors' lifetimes have faced environments of great disturbance: climatic shifts and ecological catastrophes both local and global (Alvarez *et al.*, 1980; Albritton, 1989). Those individuals that survive have the robust and stable genomes and nervous systems suitable for both classes of environments, the stable and the catastrophic. Our genomes and nervous systems contain histories both of adaptive specialization in stable environments and successful survival through massive environmental changes (Gerhart and Kirschner, 1997).

Finally, as our sophistication grows both in understanding nervous systems (e.g., Schüz and Miller, 2002) and generalized computer and network structures (Watts and Strogatz, 1998; Nolfi and Floreano, 2002; Newman, 2003), we are beginning to get a better idea of what kinds of modifications might be necessary to scale up brains to larger bodies, adapt brains to new behavioral niches, and produce intelligent behavior. Quite analogously to the first-pass detector guess for how to evolve sensory systems, often the first guess about how to generate a new behavior is to propose a new committed module for the brain (Chomsky, 1975; Barkow *et al.*, 1992; Fodor, 1992), which has rarely proved to be the actual case; for many of the same types of reasons, committed detectors are poor solutions to sensory problems. In addition, it proves that aspects of modular construction (to be discussed) pose difficult challenges for nervous system scaling.

The allometry of scaling of brains with bodies (Jerison, 1973; Schmidt-Nielsen, 1984) and the coscaling of sensory systems are phenomena that have never been satisfactorily understood (Finlay *et al.*, 2001). Why brains should scale regularly in size with bodies at all is unclear, as we know that in many nonbiological systems there is no need for the size of a control system to match the physical size of the entity controlled. Why visual acuity should scale roughly with body size (Kiltie, 2000) and why the ratio of brain size to body size is somewhat better correlated with behavioral complexity than absolute brain size alone (Jerison, 1973; but see Gibson, 2002) are similar puzzles (see Encephalization: Comparative Studies of Brain Size and Structure Volume in Mammals). Our available explanatory schemes have not coped well with these questions, but the application of new work in the properties of network architectures may begin to help.

In the following pages we will review some of the basic generalizations of how mammalian nervous systems tend to scale and adapt, and how those are linked to conserved developmental programs. When possible, we will try to evaluate whether we are looking at conserved programs that exist because of some constraint limiting the range of evolutionary solutions, or conserved programs that represent selected optimal solutions. Finally, we will propose and evaluate some exploratory models of network scaling in this evolutionary and developmental context.

3.05.1.1 A Conserved Order of Neurogenesis across Mammals and Its Relationship to Brain Allometry

In 1995, in an initial analysis of the relationship of variations in the order and relative duration of

neurogenesis to variations in the size of mammalian brain regions, Darlington and Finlay compared two large sets of data on both phenomena and found an unusually close and predictable relationship (Finlay and Darlington, 1995). Using the information collected for primates, insectivores, and bats for comparable brain regions by Stephan *et al.* (1981), a data set that has been the subject of numerous analyses, we first reiterated a finding that had been known before, but not usually highlighted: approximately 97% of the variance in the sizes of brain parts was predicted by the size of the whole brain, and 99% if a second limbic or olfactory factor was added (Gould, 1975; Jolicoeur *et al.*, 1984; Barton *et al.*, 1995; Figure 1a). This was an unusual emphasis, because most investigators, interested in mapping the 'differences' in size of brain parts to 'differences' in animal's behavior and niche, disposed of the shared variance and examined the residual variance, using various statistical approaches (Barton and Harvey, 2000; Clark *et al.*, 2001; de Winter and Oxnard, 2001). Since we were interested in the relationship of absolute volumes of brain or neuron numbers to neurogenesis, it was necessary, as well as fortuitous, to focus our attention on shared variance.

In addition to the predictability of brain component scaling from brain size, a second important feature was disproportionality – different brain components enlarge with markedly different slopes, such that the mammalian brain comes to be dominated in volume by different structures as it enlarges, notably the cortex. The large human cortex is just the size it should be for a primate brain our size (Hofman, 1989) and for absolutely larger brains, like the elephant's, the cortex is an even greater proportion of total brain volume.

The proximate cause of the disproportionality in the enlargement of brain parts can be understood by looking at neurogenesis, how neurons are generated in early development across mammalian species (Figure 1b). The ordinal position of the peak day that neurogenesis ceases for each cell group and structure in the brain is very highly conserved in the mammals studied (this end of neurogenesis is called the cell group or structure's birthday), although the total duration of neurogenesis varies from approximately 10 days in the mouse to over 100 days in monkeys. A two-factor equation can be written that captures 99% of the variance in this species/structure matrix (Clancy *et al.*, 1999, 2001).

Underlying the conservation of ordinal position, however, is a very strong nonlinearity: the numbers of cells generated do not increase linearly with extended developmental time, but exponentially over time, reflecting the doubling and redoubling

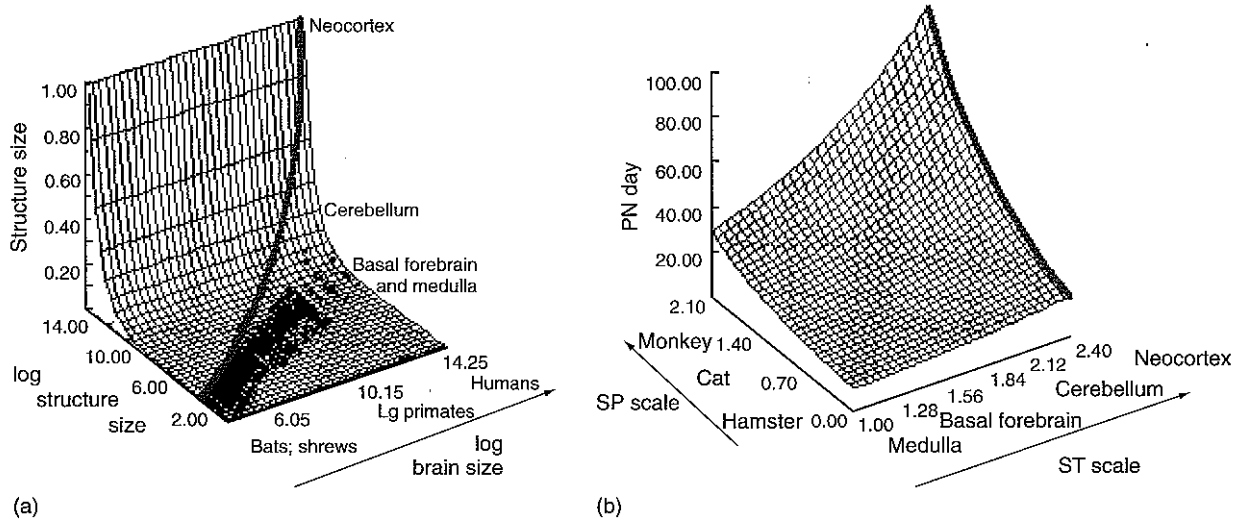

Figure 1 a, A combined log (*y*-axis) and nonlog (*z*-axis) plot of brain structure volumes versus log brain volume (*x*-axis) for the 131 primates, bats, and insectivores of the Stephan data set. This style of graphing is chosen to highlight both the predictability and disproportionality of those structures scaling at the steepest slopes with respect to brain size, particularly the neocortex. b, Model of the predictability of the birth date of a structure in a species given the ordinal position of each structure's birth date across animals (structure scale, ST) and the relative duration of neurogenesis in a species compared to others (species scale, SP). The seven species modeled in this analysis are hamsters, mouse, rat, spiny mouse, possum, cat, and monkey; 51 neural structures are modeled, from motor nuclei of the medulla to cortical layers. Those structures with high ST values are the latest-generated ones and are the same that become disproportionately large as brain volume enlarges. Lg primates, large primates such as rhesus macaque and the anthropoid apes; PN day, number of days postconception. a, Redrawn from Finlay, B. L. and Darlington, R. B. 1995. Linked regularities in the development and evolution of mammalian brains. *Science* 268, 1578–1584.

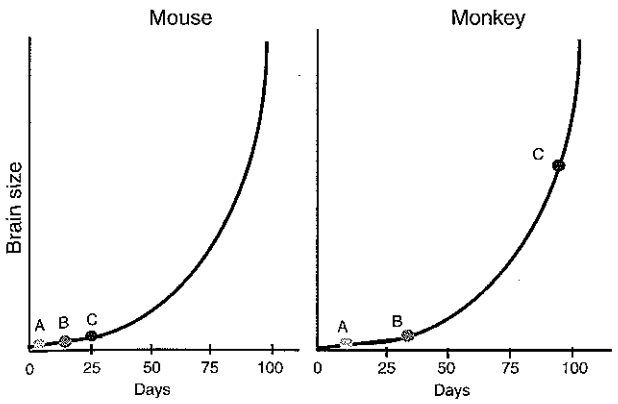

Figure 2 Late equals large. A schematic of the consequences for eventual size of a structure generated early (A), intermediate (B), or late (C) in the order of neurogenesis for a species with a short period of neurogenesis and a small brain, like a mouse, versus one with a long period of neurogenesis and a large brain. In the long-development species, the precursor pool for late-generated structures has a longer time to multiply and becomes disproportionately large.

nature of the symmetric phase of cell division in early embryogenesis (Figure 2). The consequences of exponential growth for lengthening the period of neurogenesis by roughly a factor of 10, the ratio difference from mouse to monkey, are quite different for the end neuron number in structures with early birth dates (like the medulla), middle birth dates (like the midbrain), and late ones (like the cortex). Our shorthand term for this relationship is late equals large.

Thus, particular parts of the brain increase disproportionately by a 'developmental rule'. This was quite a disturbing finding in that most previous accounts of disproportionate enlargement of some parts of the brain were cast as special adaptations due to the virtues of those parts. Particularly, some special organization or adaptive advantage was usually ascribed to the cortex to produce its disproportionate size in primates and ourselves: its efficient layering, the columnar structure, its hierarchical/parallel associative connectivity. However, the developmental rule that will produce a disproportionately large cortex in a large animal is already present in the much smaller stem primates with their small cortices, who presumably had no particular plans of their own for generating a useful structure to house a language cortex or any other kind of elaborated cognition. In addition, after a decision to allocate more tissue volume to olfactory-limbic structures or to cortex, the cortex scales disproportionately in all mammalian radiations we have studied (Reep *et al.*, 2007).

Before we go on to discuss this perplexing observation and the relationship of brain structure and function in evolution, we will go into a little more detail about just what the developmental rule is. We have published a number of further analyses and reviews to which we refer the reader for detail about such subjects as differences in neurogenesis in eutherian and noneutherian mammals (Darlington *et al.*, 1999), closer structural analyses including scaling of the thalamus (Finlay *et al.*, 1998, 2001), the special case of the limbic system concentrating on olfactory bulb and hippocampus (Kaskan and Finlay, 2001; Reep *et al.*, 2007), the scaling of the visual system (Kaskan *et al.*, 2005; Finlay *et al.*, 2005b), and the relationship of other developmental events to neurogenesis (Clancy *et al.*, 1999, 2001).

3.05.1.2 Prosomeres Are the Developmental Units Organizing the Duration of Neuron Proliferation in the Forebrain

Mammals differ from most other vertebrates by confining most of their neurogenesis to early development, rather than generating brain throughout life (there are exceptions to this generalization, of much current interest; Scharff, 2000). Does the conserved order of neurogenesis we see reflect a random pattern that happened to crystallize at the time of divergence of mammals, or some more fundamental organization? It proves that the conserved pattern of early neurogenesis we see, as well as where ongoing neurogenesis can be found in adult mammalian brains, can be explained by reference to an organizational scheme defined by patterns of expression of regulatory genes and transcription factors in early neurogenesis, the prosomere model (Rubenstein *et al.*, 1994). The basic axes that define this structure are common to the entire brain, which begins as an extended plate, the neural plate, which subsequently rounds up and connects its lateral-most edges to become the neural tube. The neural tube, whose original form is most obviously visible in the spinal cord, consists of repeating segments of similar fundamental structure with local variations. The part of the plate (and later tube) near the midline is called basal for its embryonic position, and in the spinal cord, this basal plate gives rise to motor neurons. The lateral part is called alar, Latin for wing, again for embryonic position, and in the spinal cord will produce secondary sensory neurons. The topology of the embryonic tube is maintained in the adult brain, with generative areas that initially neighbored each other, producing neighboring but translocated adult neurons. The embryonic relationships of cell groups in the adult brain can be recovered using classical neuroanatomical methods combined with painstaking developmental observations up to

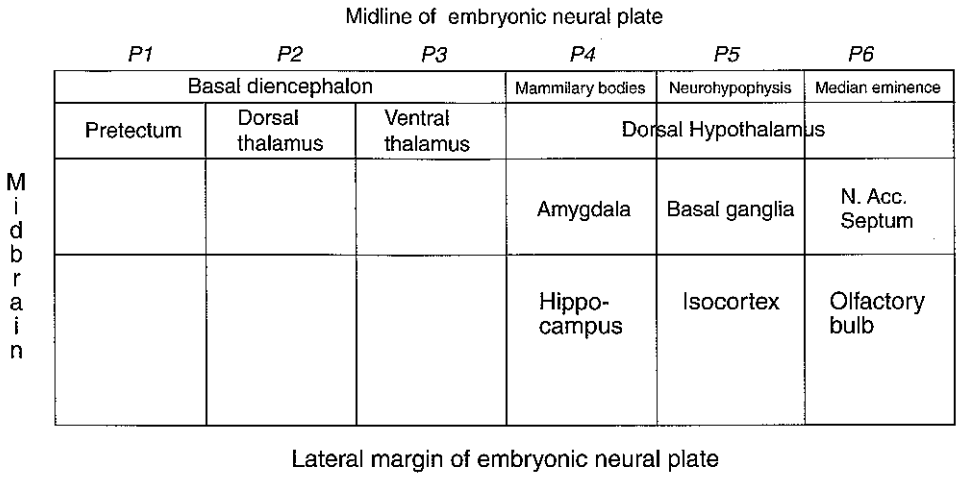

Figure 3 Components of prosomeres, the embryonic segments of the telencephalon, described by Rubenstein *et al.* (1994). The lateral-most part of the early neural plate is the part that undergoes the most extended cell division and becomes disproportionately large in large brains. N. Acc., nucleus accumbens.

about the level of the midbrain. The extended and convoluted pattern of neurogenesis in the forebrain, however, makes it impossible to track unlabeled cell groups from their place of origin. The ability to visualize gene expression gradients was required in order to trace each cell group back to its position of origin on the two axes of the embryonic brain, anterior–posterior and basal–alar.

An assignment of the traditionally named brain parts to this axial system of 'prosomeres' is given in Figure 3. For the most part (with some exceptions), as adult brain divisions reflect embryonic neural tube positions, these assignments can be made unambiguously. Hypothalamic and some basal forebrain structures are in the basal prosomeres; the large cellular masses of the forebrain, for example, the basal ganglia, are intermediate, and the cortical structures – olfactory bulb, hippocampus – most lateral or alar. If embryonic axial position is correlated with birth date (Finlay *et al.*, 1998), both axes contribute to the solution, but the alar–basal axis predominates, with late birth dates for cell groups associated with alar positions, early birth dates with basal. Our shorthand for the relationship of timing of neurogenesis to brain part size 'late equals large' can now be extended to a spatial axis of gene expression 'lateral equals late equals large'. Note that for two of the most anterior structures in the most lateral–alar position, the olfactory bulb and hippocampus, there is in fact no terminal birth date and neurogenesis continues throughout life, even in mammals (Bayer, 1980, 1983). For the cortex, at the same alar position, neurogenesis does appear to stop (Rakic, 2002; there is debate on the issue; Gould *et al.*, 1999).

Therefore, the conserved pattern of neurogenesis is not the crystallization of an arbitrary order that happened at some point in mammalian evolution, but is an expression of an axial pattern that at least in part is common to all vertebrates. The fact that the telencephalon is a likely division (but not the only one) for disproportionate enlargement across all vertebrates, and the cortex specifically enlarges in mammals appears to find its roots in this aspect of embryonic structure.

3.05.1.3 Brain Adaptations, Specializations, and Residual Variance

It is important to understand the actual physical characteristics of brain evolution to understand the nature of the shared and residual variance. In the Stephan data set (Stephan *et al.*, 1981), brain weights vary from a fraction of a gram to over a kilogram, a factor of about 20 000. At any particular brain weight, the residual variance of individual structures is approximately 2.5; that is, two species similar on the two factors (whole brain and limbic) might commonly have individual structures varying by over a factor of 2, occasional pairs considerably larger, which would be very conspicuous to an investigator looking for individual or species differences in the sizes of brain components. It also proves that the distribution of variance in volume across structures is quite uneven (Glendenning and Masterton, 1998).

This residual variation is an interesting aspect of species variation, and we do not discount it as an important window into brain structure, but we have set our job to understanding the significance of the factor of 2 in the context of the factor of 20 000, not the factor of 2 alone. Should we expect all adaptation-relevant increases in brain size to be carved out of the residual variation? The answer is

no, but it requires that our assumption of structure–function identity be loosened somewhat, particularly in the case of potentially multimodal regions like the cortex. There is no doubt that the sensory periphery is usually committed to its particular tasks – olfactory bulbs are for chemoreception, the retina for vision, and later on we will list large differences in the sensory periphery for animals in particular environmental niches that require different types of competence. A useful analogy for understanding brain evolution is the two-hit model of cancer initiation (Knudsen, 2001). In order for tumor genesis to begin, both proliferation 'and' mutation must occur. Rather than selecting for increased size for a function bound to a particular structure, a structure that has accessible variation in its size (first hit) may be well placed to acquire new functions, either by genetic specification (second hit) or through the epigenetic route of experience.

3.05.2 The Particular Case of Proliferation of the Cortex

The proliferation of the cortex, considering either total volume or general structure, is not identical across all mammals, and we would note a few of the interesting differences before concentrating on some commonalities. First, different radiations of mammals appear to allocate proliferation preferentially between limbic forebrain (olfactory bulb and hippocampus) and iso- or neocortex, with carnivores and primates showing more neocortical proliferation and insectivores, rodents, and anteaters greater olfactory bulb and hippocampus proliferation (Gould, 1975; Jolicoeur *et al.*, 1984; Reep *et al.*, 2007). As is well known, the cortex increases in area as it enlarges, but not exclusively – the cortex also increases in depth, as measured in number of neurons and particularly those concentrated in the upper cortical layers – in the smallest brains, a differentiated layer 4, the thalamic input layer, cannot often be detected, while in primate brains, often layer 4 has obvious sublamination (Valverde, 1990). Interestingly, the cetacean brain does not show an increase in depth with increasing size, suggesting an altered pattern of proliferation (Hof *et al.*, 2000). Finally, there is a great amount of local variation in such features as periodic expression of cytochrome oxidase, and various neurotransmitter receptors and neuromodulators both within the cortex, which have not been assessed systematically across species. Here we will discuss the proliferation of cortical areas, the size of cortical areas in relationship to niche, and some aspects of connectivity.

3.05.2.1 What Is a Unit of Cortex in Developmental Terms?

3.05.2.1.1 Radial units and cortical areas These two concepts, the first a developmental process and the second a characteristic of the mature cortex, are the best-known hypotheses about the structure of evolutionary variability in the cortex. Rakic has proposed that the radial unit, which is a region in the embryonic neural tube containing the precursor cells that will give rise to almost all of the cell types in an adult cortical column, named after the radial glia that provide the highway from the ventricular surface, is the fundamental unit of cortical proliferation (Rakic, 1990). By extending development, as measured in cell cycles producing the precursor cells in the cortical ventricular zone, more radial units are generated, generating more cortical columns and more cortical areas (Takahashi *et al.*, 1997). Rakic has characterized the generation of new cortical areas as a developmental negotiation between the cortex and the thalamus between their prespecified arrays of regions, initial protomaps (Rakic, 1988).

Cortical areas were first described as features of adult cortical organization: within the general uniformity of cortical layering, discontinuities in cell size, density, and total number and differences in axon distribution were noted (Brodmann, 1909). These regions of cytoarchitectonic discontinuities correspond to areas receiving particular thalamic input and thus correspond to the representation of modalities, which are usually topographically mapped, and a host of other features (Kaas, 1987; Krubitzer, 1995). At least the primary sensory and motor areas (V1, A1, and S/M1) are specified in their relative locations in the embryonic cortex by early polarizing events (Ragsdale and Grove, 2001). Thus, a cortical area could be an addressable unit of cortex variation and evolution: perhaps the whole set of instructions for producing a new area could be duplicated, as Kaas has suggested, much as the number of segments in a vertebrate whole body plan may increase or decrease under selection. In addition, if a cortical area is a unit analogous to a body plan segment, secondary modifications might be able to be attached to it, just as the segments of an insect body or vertebrate spinal cord have pronounced local specializations. Thus, new features could be attached to specific regions, like ultrahigh temporal resolution in cortical areas used for echolocation in bats (Suga *et al.*, 1981; see Somatosensory

Adaptations of Flying Mammals). Similarly, cortical areas could reasonably be increased or reduced in size, depending upon the relationship of their modality to each species' particular adaptive needs.

3.05.2.1.2 Multidimensional models The physical nature of scaling of biological tissue over the range of cortical surface areas that go from millimeters to meters requires some amplification of the first two hypotheses. A large cortex, or cortical area, appears to have much more internal detail than a smaller cortex. By this we mean that a large cortex contains not only more cortical columns and identifiable areas, but also more anatomical features such as stripes, puffs, and blobs in transmitter expression, and aspects of activity, interleaved ordered thalamic, intracortical, and callosal connectivity, as well as the elaborated functional maps correlated with the anatomical differences. They neither appear to be scaled-up versions of a stem insectivore cortex, nor the stem cortex replicated over and over across the cortical surface. All mammalian cortices, with the exception of those of monotremes, which differ in topology, appear to share three fundamental sensory and motor regions, and elaborate themselves along the same general lines, but with substantial variation in the details of all the features of cortical organization described above (Krubitzer, 1995).

Various hybrid models have been proposed, therefore, beginning from the assumption of a cortex that has specification of the rules of connectivity and general organization for three primary sensory areas, and then also for elaborating the cortical structure as it expands in size and functionality. One aspect of several of these models is the idea that cortical areas and other features arise in a combinatorial fashion by the interplay of genetically specified and environmental sources of variation. Because of some unusual aspects of gene expression in cortical neurons, cortical neurons themselves may generate an unusual variety of phenotypes (Kaushal *et al.*, 2003). Maturational gradients in cell generation and innervation in combination with intrinsic periodicities in the expression of various neurotransmitters and neuromodulators and the core set of specified cortical regions may also cause emerging arealization and segmentation, realized at different scales (Kingsbury and Finlay, 2001). Finally, intrinsically produced and environmentally specified activity is a powerful organizer of cortical areas in experimentally rewired cortices, and presumably must also be in development when it is unperturbed (Pallas, 2001).

Other than the very general feature of the allometric scaling of the cortex and some of its areas previously described (for example, Frahm *et al.*, 1984), and some systematic features of changes in the pattern of cortical cytoarchitecture with cortical size (Valverde, 1990), systematic comparisons of cortical organization at different scales that might allow us to better discriminate hypotheses about the process of cortical evolution have not been done. We undertook to quantify the manner in which the cortex proliferates, in the number of areas, and in the size of identifiable cortical areas, using primarily the extensive cortical mapping work in a variety of species of Kaas and Krubitzer as our sources for these analyses (Finlay *et al.*, 2005a). Both of our investigations using measurements taken from published studies have as their fundamental question, what the units of cortical expansion and differentiation might be and their developmental mechanisms.

3.05.2.2 Proliferation of the Number of Cortical Areas

The number of cortical areas increases generally with brain size, but the regularity of the increase was not known. We examined the proliferation of the number of cortical areas with respect to brain size in 24 mammals representing six orders, comparing visual, somatosensory, and total areal proliferation (Finlay *et al.*, 2005a). For each species, we ascertained or measured overall brain weight and overall cortical surface area; for the details of the studies employed, how a cortical area should be defined, the method of measurement and statistical analysis, we refer the reader to the original study. However, we will briefly describe our central choices and tactics.

Not all researchers agree on the definition or the number of cortical areas. Some make significantly fewer subdivisions (for example, Zilles, 1985), while others argue for the existence of a large number of smaller areas even in small brains (for example, Olavarria and Montero, 1990). Essentially, we chose to remain agnostic on the true definition of a cortical area and to rely instead on the pragmatic consideration of which explicit criteria allow us to examine the most species. The arguments of Kaas and Krubitzer on what constitutes an area are, however, compelling (for example, Kaas, 1987; Krubitzer, 1995). Their criteria for identification of an area are multidimensional, and include the presence of a fully mapped visuotopic, somatotopic, or other computed dimension, internally consistent patterns of thalamic, intracortical, and callosal input and output, and in some cases identification of the features of cortical cytoarchitecture or neurotransmitter or modulator expression.

The number of cortical areas might be best predicted by one of several independent variables,

including cortical surface area, or the weight of the whole brain, or one of the various formulas for encephalization, the ratio of brain to body weight. There was (and still is) good reason to suspect that in the case of two animals with equal absolute brain sizes, the more encephalized one might have a greater number of cortical areas. Unfortunately, in this data set overall, the largest-bodied animals and most encephalized are all primates, and the smallest, and least encephalized are all insectivores. Thus, the two measures of cortical area and encephalization are highly correlated ($r = 0.98$, $R^2 = 0.96$, $n = 19$). In addition, brain weight also correlates highly with cortical surface area, ($r = 0.95$, $R^2 = 0.91$, $n = 19$). Since total cortical surface area is the most proximate variable to the dependent variables we measured (number of cortical areas, ocular dominance column width, and axonal spread in the cortex), we have done our statistical analyses with respect to total cortical surface area, but in explanation of these data, the co-variation of cortical area with other brain measures should not be forgotten. Finally, because species may share traits through common descent rather than through independent adaptation, we employed the method of the comparison of independent contrasts (Purvis and Rambaut, 1995) in order to correct for the effects of phylogenetic relatedness.

Figure 4 shows the predictability of the 'number' of cortical areas overall, somatomotor areas only, and visual areas only from total surface cortical area (note we compare the number of areas, as a name of a unit of cortex, to the surface area, a measure of the cortex, an unfortunate ambiguity arising from their normal terminology). Though all three relationships were regular and highly statistically significant, surface area captured less of the variance when predicting the number of visual areas. The unusual scaling of striate cortex, which particularly in the primate lineage does not divide into subareas as the somatomotor regions do, may account for some of this variation.

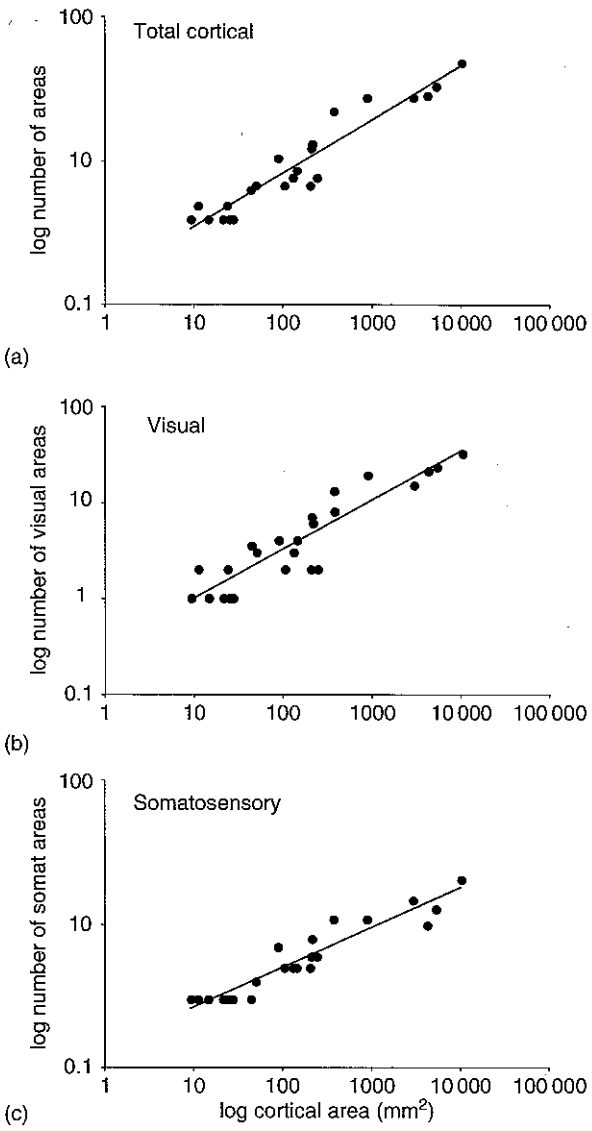

Figure 4 Simple regression of a, log total cortical area proliferation on log cortex surface area (log total areas = 0.172 + 0.379 log area, $R^2 = 0.89$, $n = 23$); b, visual area proliferation (log visual areas = –0.529 + 0.509 log area, $R^2 = 0.82$, $n = 24$); and c, somatomotor area proliferation (log somatomotor areas = 0.154 + 0.287 log area, $R^2 = 0.87$, $n = 23$).

The observation that the log of cortical area strongly predicts the log of the number of cortical areas shows that the relationship is predictable, but a logarithmic graph is not directly instructive about the kind or number of developmental mechanisms underlying this pattern of proliferation. As most biological developmental mechanisms bear some relationship to cell size, they operate over finite physical distances and not their ratios. To better visualize the change in cortex size for comparison to developmental mechanisms, the number of cortical areas is plotted against cortical surface area 'without' logarithmic transformation in Figure 5. For increases from the smallest cortices (from 10 to 400 mm^2), cortical area proliferation is rapid. Thereafter, however, only massive increases in cortical area produce new cortical areas. It is also important to understand the absolute range of cortical surface expansion: in the approximately 500-fold range of the cortical areas graphed, the entire cortex of the least shrew (*Cryptotis parva*) could fit comfortably within a small fraction of the striate cortex alone of the rhesus monkey (see Finlay *et al.*, 2005a for the data employed and list of sources).

With the exception of the primary cortical areas, whose approximate position and boundaries appear to be fixed by early genetic specification, the

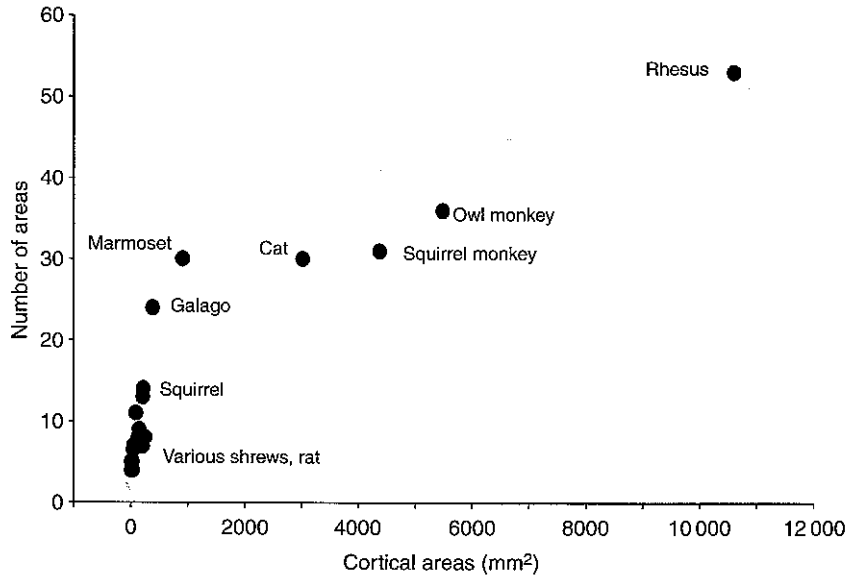

Figure 5 The number of cortical areas plotted as a function of cortex surface area, without logarithmic transformation.

mechanism by which other cortical areas emerge is not understood, but the size ranges over which the rate of proliferation changes, suggests two separate mechanisms. As cortical area increases from approximately 20 (the smallest shrews) to 400 mm^2 (galago) the number of cortical areas increases rapidly, from 4 to 24. Thereafter, another 400–4000 mm^2 of cortical area nets only another six cortical areas (those animals with 30 cortical areas enumerated are the cat, the marmoset, and the squirrel monkey). However, it is at the brain size of cats and small monkeys that the well-studied substructure of visual cortical areas begins to emerge, the ocular dominance columns, puffs, and blobs of primary visual cortex (Hubel and Wiesel, 1962, 1968), and we suggest that the cortical areas of small brains and the morphological specializations such as ocular dominance columns within cortical areas of large brains may both be manifestations of the same underlying developmental mechanism, activity-dependent axonal (and dendritic) sorting. The spatial extents of single axon arbors, dendritic trees, and ocular dominance columns in the cortex essentially do not scale with brain size (Kaas, 2000; Finlay *et al.*, 2005a; see also Manger *et al.*, 1998). Most models of stripe-in-topographic-map formation in brain tissue essentially pit axon–axon affinities against axon–substrate affinities (for example, Swindale, 1980). We hypothesize that as brains get bigger, more specific aspects of sensory stimuli may provide the correlational structure necessary (that is, the increased axon affinities) to allow the segregation of new, functionally specific cortical areas once additional volume of cortical tissue (diluting axon–substrate affinities) is made available. We have no model, as yet, for the increase in cortical areas in the largest brains.

Therefore, though the nature of peripheral sensory specializations may have a direct effect on the nature of the maps formed in the cortex, we are proposing that cortical areas also proliferate by a developmental rule, at least over the smaller ranges of brain size. This hypothesis is distinct from the idea of a cortical area as a specially selected, modular processing region. While a cortical area might be heavily involved in the processing of a particular submodality important to a species, we propose it emerges due to developmental rules present in all species (axon–substrate affinities and Hebbian fire-together, wire-together) influenced by its peripheral specialization and not the selection of new mapping and processing rules.

3.05.2.3 The Size of Cortical Areas, Sensory Specializations, and Behavioral Niche

The relative size of cortical areas in animals specialized for different niches can also inform us whether it is possible to select for the size of a cortical area and allocate resources to a particular modality differentially by changing developmental rules. Using essentially the same data sources and methods of the prior analysis, we quantified instead the surface area of primary sensory cortices (visual, somatomotor, and auditory), and the total cortical surface area devoted to the same modalities in either nocturnal or diurnal animals (Kaskan *et al.*, 2005). The nocturnal animals of this analysis all make use of their eyes for nocturnal vision, such as the rat, possum, or

owl monkey, and are not fossorial animals whose eyes have degenerated (for this contrast, see Cooper *et al.*, 1995). The fundamental assumption to be tested is that nocturnal animals should allocate fewer resources to vision and more to somesthesis and audition, and thus, across taxa, nocturnal animals should show a grade shift with a smaller primary, or total visual cortex with respect to total cortical area, and relatively larger somatosensory and auditory regions.

Prior to comparing animals in different niches, we first examined the scaling of cortical areas independent of niche and replicated a finding that has already been described in various studies with completely different species – that the volumes of identifiable cortical areas or regions each have specific, statistically discriminable allometric scaling with total cortical area, with primary visual cortex scaling at a steeper slope than other primary sensory regions (Frahm *et al.*, 1984; Stevens, 2001). The observation that frontal cortex scales at a steeper slope than other cortical regions and that humans have a larger but predictable volume of frontal cortex (Jerison, 1997; Semendeferi *et al.*, 2002) is analogous to this analysis.

Figure 6 plots the regression slopes of V1 (a) and total visual cortex, including all secondary visual areas of the parietal cortex (b) for 20 nocturnal and 8 diurnal species against total neocortex. Again, as measured by the method of independent contrasts, no significant difference appears between nocturnal and diurnal mammals in their visual cortical scaling – either in V1 or in all visual cortical areas. The regression plots themselves are notable

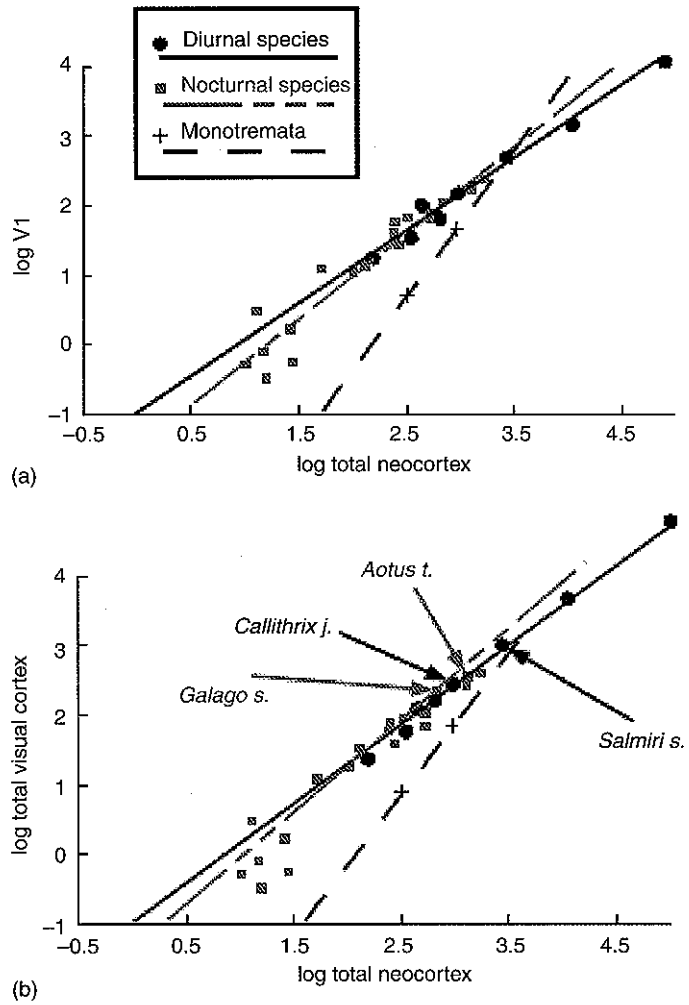

Figure 6 The volume of primary visual cortex (a) or all mapped visual cortical areas (b) versus total cortex volume in nocturnal (orange points) and diurnal animals (blue points). There is no statistically significant difference between these groups for visual cortex or motor or somatosensory cortex (not shown). Reproduced from Kaskan, P., Franco, C., Yamada, E., Silveira, L. C. L., Darlington, R., and Finlay, B. L. 2005. Peripheral variability and central constancy in mammalian visual system evolution. *Proc. R. Soc. Biol. Sci.* 272, 91–100.

for their complete overlap. If diurnal animals devoted more cortex to vision, we would expect to see a grade shift in these plots, with the diurnal regression line displaced upward but parallel to the nocturnal. Neither did nocturnal animals possess more somatosensory or auditory cortex – primary nor total – as might be expected if they could designate more anatomically defined cortical areas to modalities important in the dark. The retinas of these animals are quite different, as would be expected, with nocturnal animals possessing many more rods. Since we were able to capture statistically significant differences in the allometry of V1 and total visual cortex compared to other cortical areas independent of niche, our confidence that we could detect a difference between nocturnal and diurnal species in cortical organization is reasonable.

Should a nocturnal visual system require less neural volume than a diurnal visual system? If the nocturnal computational problem of vision could be demonstrated to be harder, it could be argued that it should require more. Our central observation is that it appears to require neither less nor more (as defined by cortical area volume) in the face of greatly different photoreceptor distributions. Nevertheless, the interpretation that greater usage should have a direct correlation with the size of corresponding neural systems is prevalent in the literature (Barton *et al.*, 1995; Barton, 1998; Barton and Harvey, 2000). These observations range from large tactical allocations of brain mass between the limbic system and neocortex, attributed to preference for the visual modality, to very specific differences in cell types associated with fugivory versus folivory. In the latter case, it is perplexing that subtle differences in visual requirements for frugivory versus folivory should be associated with differences in visual system organization, while major differences to accommodate nocturnality and diurnality should not. Niche, brain size, and cortex size are difficult to dissociate in primates. It may be, therefore, that the observations are confounded by conserved scaling phenomena unrelated to function. Alternatively, frugivory and a certain visual cortical organization might be necessarily related, but with selection on gross brain size as the only means of inducing the required detail in brain organization. For example, in a study of dexterity in animals whose dexterity ranged from hooves to hands, brain area devoted to control of the forelimbs was highly positively correlated with dexterity, but the area devoted to the forelimbs was in turn accounted for entirely by total neocortex size (Nudo and Masterton, 1990).

The hypothesis that visual cortex differs between nocturnal and diurnal mammals may yet be true, in two different ways: first, other aspects of cortex than size must surely be important, and second, our measurement of size may inaccurately assess true functional allocation. Evolution might well act on variables such as dendritic structure or density, receptive field size, axonal arborization, myelination or basic cell physiology; for example, consider the intriguing case of dyslexia, now associated in some of its forms with difficulties in detecting rapid transients in more than one modality (Merzenich *et al.*, 1993). Differential expression of some feature of synaptic transmission in the cortex may allow some individuals to represent information in rapid stimulus transients, producing ordered phonemic representations (Petersen and Fiez, 1992), while others without it do not. Details of basic cortical processing could well be different in animals of different niches.

Considering the second point, how well do the anatomically defined cortical areas reflect actual functional allocation in the cortex? A number of new imaging and other functional studies suggest that we may have been overimpressed by the major thalamic input to an area when cortical regions were named, and we thus assess true functional allocation inaccurately. Functions may play more freely over the cortical matrix specified early in development than we have imagined, with the most likely substrate through long-range intracortical connectivity. We find violations of modality specificity of cortical areas in evolution, where, for example, the visual cortex comes to respond to auditory stimulation in the blind mole rat (Bronchti *et al.*, 2002), in essentially normal function, where the auditory cortex responds to the visual stimulation of lip-reading (Calvert *et al.*, 1997), in experimental rewiring of visual information into auditory areas (Gao and Pallas, 1999), and in cases of sensory loss in humans where visual cortex becomes critical to tactile Braille reading in the early and late blind (Sadato *et al.*, 1996). In the adult, our understanding of corticocortical connectivity is limited, but recent work shows that connections may be widespread and fail to conform to traditional hierarchies and notions of connectivity (Falchier *et al.*, 2001; Rockland, 2001). Thalamocortical connections also show a distributed nature with a matrix of superficially projecting cells not confined to the intralaminar nuclei, which may serve to bind sensory experiences by connecting multiple cortical and thalamic areas (Jones, 1998). On the whole, the findings that such broad structure–function matches in the cortex exist imply that the neocortex is not a piecemeal collection

of areas, each with its own discrete function, but a generalized processing device.

3.05.2.4 Integration of Individual Variability and Developmental Plasticity into an Evolutionary Account

A final perplexing feature in the understanding of brain size regulation is the report of remarkable individual variability in the size of cortical areas, which would seem to be in stark contradiction to the very regular scaling we have just described in allometric studies. We will assume that published allometric studies have managed to determine representative mean structure volumes for scaling work and that the scaling results are accurate at the species level for which they are intended. Here we address the question of how we are to understand the importance of structure sizes if individual members of a species may occasionally but substantially differ from one another in the relative sizes of brain parts.

What kinds of variations are reported at the individual level, within species? The best information comes from a number of studies of the primate visual system, particularly the rhesus macaque. Van Essen *et al.* (1984) have found individual animals whose primary visual cortex differed by a factor of 2 or more. Similarly, the variability of the human visual cortex exceeds substantially the variability of the entire cortex (Gilissen and Zilles, 1995). There are only a few studies, to our knowledge, of the variability at the individual level of the number and arrangement of cortical areas (Qi and Kaas, 2004; Airey *et al.*, 2005), but comprehensive imaging of individuals may soon allow this kind of comparison to be made. Few of these observations have as yet been tracked onto individual variation in visual capacity, and it would be interesting to do so. However, there is reason to believe that with the exception of variations in cell density in the visual periphery that directly affect acuity, the basic processing of the visual system is robust to wide variations in number of neurons in interconnecting populations, due to the equilibrating effect of processes such as activity-dependent stabilization in early development and compensatory norming in adulthood, producing the remappings described in the previous section.

Preferential allocation of space to various kinds of sensory specializations is commonplace 'within' particular cortical areas (Suga *et al.*, 1981; Silveira *et al.*, 1989; Catania and Kaas, 1997). In addition, complete loss of a sensory system either phylogenetically (Cooper *et al.*, 1995) or very early in development reduces the volume and area of a cortical area through the epigenetic route of degeneration of the intermediate thalamic nucleus (Finlay and Pallas, 1989; Rakic *et al.*, 1991; Kahn and Krubitzer, 2002). Both of these kinds of alterations, within-modality specializations, and sensory system loss, seem likely to use the epigenetic pathway of mapping a new thalamic organization onto the cortex rather than alteration of cortical specification.

All evidence reviewed so far – regular allometric scaling of the entire cortex (Finlay *et al.*, 1998, 2001), of particular cortical areas (Jerison, 1997; Stevens, 2001; Kaskan *et al.*, 2005), and of the number of cortical areas (Finlay *et al.*, 2005a); independence of specific cortical area volume from niche (Kaskan *et al.*, 2005); individual variability (Van Essen *et al.*, 1984); and developmental plasticity (Pallas, 2001) – converges on the same interpretation. The regularity of major components of neural allometric scaling, best predicted by cross-mammalian developmental constraints, suggests that mismatches of neural ratios or of typical structure/function allocations must be a regular, compensated phenomenon in mammalian evolution (Xiong and Finlay, 1996). The independence of the relative size of primary sensory areas from niche (Kaskan *et al.*, 2005) and the coupling of dexterity to whole cortex size rather than the relative size of the somatomotor cortex (Nudo and Masterton, 1990) suggest that the relative size of a cortical area with respect to the whole cortex is either very difficult to change or unimportant to function. Large individual differences in the sizes of particular brain areas unaccompanied by flagrant disabilities tell the same story about individual development, as do innumerable instances of developmental plasticity. Thus, particularly for intrinsically cross-modal structures like the cortex, structure and function may not be uniquely linked at neurogenesis, and neural resources may be allocated to new functions as necessary. The fact that we have named a structure visual cortex (because that is typically what it does) does not prevent it from becoming Braille cortex when circumstance permits (Sadato *et al.*, 1996).

3.05.3 Computational Considerations in Cortex Proliferation

The description of the evolution of brain size, cortex size, and the size and number of cortical areas is qualitatively and quantitatively consistent, but fundamentally unsatisfying. With the exception of very broad, tactical allocations of space to different classes of neural processing (such as cortex vs. hippocampus; McClelland *et al.*, 1995) either for

whole vertebrate radiations (Gould, 1975; Finlay *et al.*, 1998) or for very broad behavioral niche (Clark *et al.*, 2001; de Winter and Oxnard, 2001), we have not found much to tie features of brain organization to environmental niche, nor to any aspect of behavior. All we have done is to speak negatively about premature allocation of function to structure and speak vaguely and positively about the cortex as a general-purpose processing device. Yet the brain, a very expensive tissue to maintain, increases with body size at an extremely predictable rate for parts that have a direct relationship to body size, and for parts that do not. Sensory systems, particularly in what appears to be their central allocation of neural volume, do not economically scale their size to their apparent usage. Cortical areas proliferate in a fairly direct relationship to overall cortical area, and there is really no direct evidence of any kind as yet linking any explicit feature of behavioral adaptation to genetically produced changes in cortical architecture. We will argue now that the missing component of our understanding may be the nature of network architecture in the brain, particularly the cortex, and an understanding of how networks scale.

We now approach the question of brain scaling and resource allocation in a different manner, considering candidates for the class of network architecture that the cortex might possess (Milo *et al.*, 2004), with a brief review of what is known about their scaling properties. We will then lay out one toy model of a network whose properties we can perturb to see the nature of its scaling. It is important to understand here that we are making no claim that the cortex necessarily resembles the model that we are exploring here. Rather, we employ models to make hypotheses about how networks might scale in an explicit and testable manner, and in so doing, perhaps recognize in the behavior of the models some aspects of cortical scaling and brain–body allometry that have been determined empirically.

3.05.3.1 What Are the Dynamics of Brain Scaling?

As brains get bigger, it is important that they do so efficiently. Brains face the problem of fitting into necessarily limited spatial and metabolic budgets (for example, Aiello and Wheeler, 1995), placing a premium on brains that minimize the amount of tissue not contributing to the organism's fitness. While pruning unnecessary cells is a relatively simple matter, decreasing average axon length is a considerably more complex problem. A longer axon is just as good as a shorter one computationally, but the longer one squanders the brain's spatial and metabolic budgets (Swindale, 2001). The ratio of somas to axons and dendritic trees decreases with scale, requiring new wiring strategies (Zhang and Sejnowski, 2000). Moreover, brain size is limited to the amount of space the head can provide, given that the head has more functions than enclosing the brain, and by the sheer expense of hauling a heavier body part. In mammals, head size itself is under some constraint by the female reproductive organs (Leutenegger, 1982).

Given the brain's limited spatial and metabolic budgets, a wide range of fundamental questions in brain network architecture must be addressed. Does the same network architecture apply to the scaling of large and small brains? What are the properties of various classes of proposed modules in the brain, from the sense of repeating, open units such as a cortical column, to encapsulated, closed processing units in the sense of Fodor (1983, 2000)? What happens when new functions are introduced into new networks, or the size of components is altered? In the following section, we outline a few of the issues arising in our first explorations of these computational questions.

There are several immediate consequences of body scaling. Different body types and sizes require different motor programs to maintain a reasonable level of energy efficiency. Additionally, changes in body type and size also change the relative muscle mass and pH buffering capacity, which in turn directly affects what motor programs will maximize the aerobic to anaerobic ratio. Typically, the larger and more complex the body type, the greater the required level of motor sophistication and physics modeling (Keimel and Roth, 1992).

While receptive field sizes and higher-level convergence vary not only across species but even across the body of a single animal, there appear to be minimum tactile acuity levels beyond which larger body size requires a larger somatosensory region in the brain. In specialized systems of various well-studied mammals, from the star-nosed mole, to the echolocation systems of bats, to the primate fovea, an increase in sensory organ acuity is closely related to the relative volume of primary sensory cortex devoted to the high-acuity region (Suga *et al.*, 1981; Silveira *et al.*, 1989; Catania and Kaas, 1997). For instance, in the acoustic fovea of bats, a hypertrophied sensitivity to a particular frequency in the cochlea is paralleled by a highly disproportionate amount of auditory cortex devoted to that frequency. Since such sensory foveas are not simply increases in receptor density, but also interrelated with unusual motor and perceptual processing demand, the subsequent within-area cortical

gerrymandering may be the result not only of an increase in afferents but of a need for more computational power. Any necessary relationship between the number of afferent units and the amount of cortical area necessary to process their signals has not, to our knowledge, been documented. Thus, the dynamics that couple afferent numbers with total brain size is an unexplored question. We will argue that a critical component of understanding brain–body allometry lies in the governing dynamics of wiring strategies employed by the brain.

3.05.3.2 Classes of Wiring Strategies

A wiring strategy is a plan by which neurons, or groups of neurons, interconnect efficiently (that is, minimize metabolic and spatial expense) while taking into account the settings of several parameters, perhaps the most important of which is overall brain size. For the purposes of this paper, we will focus our discussion on synaptic interconnections at the level of cortical columns, nuclei, or lobes; or more generally, groups of neurons, which we refer to as subnets. Ultimately, to be efficient, a wiring strategy must attain good performance on two axes: reuse and intersubnet communication.

1. *Interconnecting subnets for the purpose of reuse.* To increase efficiency, subnets should combine and produce behavior not implemented by any one subnet. The recombination of subnets results in a reuse of existing resources, conserving the number of neurons needed to accomplish any given set of tasks. Such reuse of cortical resources is very commonly observed in current imaging work. Increased executive demand and task difficulty, for instance, causes activation of the frontal cortex for a variety of different tasks across different modalities (Duncan and Owen, 2000). How subnet reusability might be achieved is a question we will address later.
2. *Interconnecting subnets for the purpose of communication.* The fundamental purpose of the brain is to control the body. No matter how elaborate the internal architecture, on output, the brain must act as a whole and cannot execute contradictory commands. Different subnets, participating in the planning and execution of any given task, must maintain some level of communication to ensure that eventually subnets work in concert with each other. In order to avoid the kind of dysfunctional situations seen following spinal cord or corpus callosum transection, which essentially produces two brains in one body (Sauerwein and Lassonde, 1994), every neuron in the brain must be connected to every other neuron in the brain either directly or indirectly.

Both subnet reusability and subnet interconnectivity can be regarded as two sides of the same problem. The average distance – measured by the number of synapses, or the number of hops, any neuron is from any other – directly dictates:

1. the extent to which that neuron can be reused in different neural circuits;
2. the recombinatorial power of the subnets, to which that neuron belongs; and
3. the ability of those subnets to communicate with other subnets.

To maximize 1, 2, and 3, a connection strategy must minimize average hop count. There are various ways this can be accomplished, depending upon computational goals and constraints.

3.05.3.2.1 Small worlds The problem of minimizing the average hop count has been extensively studied and modeled (Watts, 1999). Consider a daisy-chained group of nodes arranged in a circle. The maximum (worst case) internode distance is the number of nodes in the chain minus one; the average is the number of nodes in the circle divided by two. Provided there are two or more nodes in the circle, the constant formula for the average hop count remains fixed irrespective of the number of nodes. However, with just one projection, spanning the diameter of the circle, both the worst case as well as the average internode distance is dramatically reduced. By adding just a few long-distance projections, the average distance between neurons could be decreased in just the same way.

However, small worlds has been primarily used to model connections between computers on the Internet or the connections between people in a social network. While it has been argued (Manev and Manev, 2004) that a small world network could be a reasonable description of the brain, there are important aspects to small world networks that make them unsuitable as a model for synaptic connectivity. Unlike computers, humans, or mail hubs, neurons are not routers that are capable of directing a message to its appropriate destination. They have only one axon and any message a neuron sends reaches every dendrite on which the axon terminates. Moreover, individual neurons are virtually never mere communication relays: they process and mutate incoming data in the process of transmission.

Furthermore, single-neuron-to-single-neuron communication is not the essential wiring challenge. In most mammalian neural systems, an individual

neuron has very little computational weight. Rather, parcels of processed information are sent in parallel from one subnet on to other subnets. Neural subnets of even very modest complexity would be unlikely to communicate with each other via a single synapse; the process of decoding a binary serial stream requires many more resources than sending the data in parallel. It is important to recognize that maintaining a certain number of degrees of separation between neurons (or average hop distance) is not a wiring strategy in and of itself, but the result of a wiring strategy. Random projections, for example, could easily reduce the degrees of separation between subnets and yet be useless in actually transmitting the relevant information.

3.05.3.2.2 Nearest neighbor If the small world model does not seem to be a good model for the mammalian brain, what other models are there? Arguably, an alternative approach to efficiency would involve physically moving related subnets closer to each other, decreasing the average synapse length, although the average node distance (hops) would remain unchanged. If, for instance, two subnets are both responsible for processing information from the same modality, it might make sense to move them as close to each other as possible. The principle of minimizing axon length might be driving sensory systems to stay together in the cortex and may also be behind the ubiquitous feature of topographic maps (Kaas, 1997). On the other hand, intramodal processing of various kinds is central to cortical processing at every stage, which will necessarily disrupt wiring strategies dependent on single-modality topography. The nature of the topographic dimension to be represented often varies with the processing stage, for example, moving from cochleotopic to spatiotopic representations within the auditory system. Overall, the Hebbian fire-together, wire-together strategy cannot produce a single nearest-neighbor solution based on modality, though it may contribute to aspects of the wiring strategy.

3.05.3.2.3 Modules So far this paper has not attempted to define subnets in any other way than as simply a group of neurons working in tandem. Sometimes, such constellations of neurons have been called modules: self-contained components that are used in combination with other modules. In terms of neural circuitry, a module is a cluster of interconnected neurons that have a set of input neurons and a set of output neurons – collectively called an interface – with any number of interconnected neurons in between. Additionally, a module provides a particular functionality to other modules without requiring other modules to know how the functionality is implemented.

Modules, as a concept borrowed from computer science, strive for opacity and encapsulation, such that their implementation is hidden behind an interface that defines the scope and nature of the set of inputs and outputs the module can accept and produce (Parnas, 1972). Software-like modules have been argued to exist in the brain (Chomsky, 1975; Fodor, 1983). There are several reasons why software-like modules might make sense in a neural context. Modules are extensible – that is, modules place no requirements on the internal wiring of other modules nor connect to internal components, and thus can be added *ad lib*. Moreover, modularity could provide a level of insulation from conflicts and cross-talk that can arise in complex multifunctional systems, especially ones that have been extended by accretion (like evolution). Often, a degree of modularity is required to prevent clusters of neurons from being reused in different circuits. For example, fail-safe critical systems such as those controlling heart rate and respiration should function as autonomous (and potentially redundant) modules, but leave open a command interface that can accept high-level input. Neuroethologists have uncovered a number of systems that rely on such command neurons (Nolen and Hoy, 1984).

On the other hand, modules are not fault tolerant. A modular system requires massive redundancy to cope with the complete or even partial failure of any one module. Most significant to our discussion, though, is that modules increase the average hop distance: the bigger the module, the greater the hop distance. In other words, a smaller ratio of interface neurons to hidden neurons results in a greater average distance between any two neurons in different modules, subsequently decreasing the reusability of that module's neurons. Additionally, larger modules tend to implement more complex and therefore less abstracted units of computation, which leads to a decreased reusability of the module as a whole. Because of these wasteful features, the brain is not likely to pursue a strategy of using more than a few large modules.

3.05.3.3 Exploring the Effects of Uncoordinated Scaling in a Neural Net

Because neither nearest-neighbor nor modular wiring strategies seem likely to give a full account of the wiring of the mammalian brain, from the arguments above, and because the units of small-world networks have properties quite unlike neurons, we have chosen to employ a standard,

three-layered neural net architecture to explore the effects of scaling on network configuration. We address the effect of scaling up any part of a network architecture on the properties of the rest of the network; our interest is to gain some insight as to why the sizes of the various sensory surfaces, motor interfaces, and brain parts are so closely correlated as species scale in evolution when energetic efficiency would suggest part-by-part independence should be preferred. Prefacing with our conclusion, we find that scaling any kind of subnet within a neural net also causes a scaling response in all neighboring subnets, and arguably in all the subnets within a system.

3.05.3.3.1 **The model employed** We used a standard neural net architecture of either 64 or 128 nodes, limiting the number of connections to a unit in the net to 10. First, we trained a 64-node neural net on either one of two arbitrary tasks chosen not because of their similarity to any known neural process, but because they have been well studied. These two tasks are character recognition and graph fragment completion, two of the first tasks to be successfully implemented by artificial neural networks and still in use today (i.e., the US post office uses neural nets to read zip codes). No interesting differences between these two tasks emerged in net configuration and we will not address this aspect further. In order to understand the allocation of function within the neural net, we made lesions in one node at a time and measured the decrease in firing rate across all other nodes.

The absolute value in the change of firing rate across all nodes has very high contrast consequent to this lesion, with a small group of nodes significantly more affected by the lesion than all the others (Figure 7), revealing subnet contours (Figure 8).

When a subnet node is in a feedforward position, the nodes afferent to it are not affected by the lesion. Nodes affected by a lesion are fewer in number than would be expected by random, proportional interconnections.

Through single-node lesion testing, we found nine subnets in the 64-node network. We repeated the experiment with a 128-node network performing the same character recognition task (which had a proportionally scaled input layer) and found 13 subnets. Though not surprising, the first point to take note of is that average subnet size increased in response to an increase in the number of nodes in the hidden layer. With the task remaining constant, while the number of nodes doubled (128:64), the number of subnets did not double (13:9). Average subnet size went from 7.1 nodes to 9.8 nodes.

Suppose that the body becomes bigger, generating a larger S1 to maintain sensory acuity (Figure 8). Does this have consequences for the rest of the brain (or net) that the larger S1 is patched into? To explore the effects an enlarged S1 required by an enlarged body on the structure of the rest of the network, we artificially enlarged sections of the 64-node network to see whether or not the average subnet size increased. For each replication of the experiment, we grafted just one subnet from the 128-node network into an untrained pool of 64 nodes minus the number of nodes in the grafted subnet, thus entering a subnet of an average of 9.8 nodes, fixed in size, into a neural net that normally generates 7.1-node subnets. In each replication of a single subnet graft, we permuted over every possible placement of the grafted subnet into the 64-node net, principally to avoid any configuration that would artificially isolate any nodes and decrease the net size artifactually, examining only the best results from the permutations, in terms of task performance (i.e., percentage of characters correctly identified). After grafting, each 64-node net was trained, but with the provision that the grafted

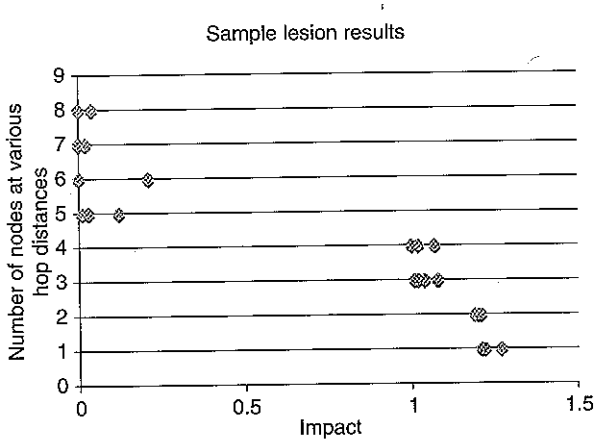

Figure 7 Example of the effects of deletion of single nodes to show connectivity structure within a neural net.

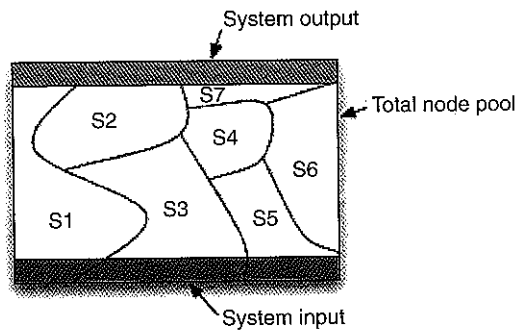

Figure 8 Diagram of input, output, and subnet structure in a trained neural net.

subnets were immutable. Grafted subnets contributed to the solution of the task, with the initially untrained net reconfiguring around the grafted component. For each neural net with a scaled-up graft, we found a smaller number of total subnets (and therefore larger average subnet size) than was the result in the original 64-node net experiment. There was an average decrease of 1.8 subnets to a mean of 7.2, with eight nodes per subnet including the grafted subnet and an average performance drop of 17% (failed character recognition or graph fragment completion). The decrease in number of subnets is significantly greater than the simple decrease expected by a subtraction of a subnet of 9.8 nodes; we trained nets on the same two tasks with 64 minus N nodes, where N is the size of each graft and is on average 9.8. Even when normalized to 64 minus N over 64, the control group produced a larger number (10) of smaller sized subnets that on average contained 5.4 nodes.

In summary, we found that when one subnet is scaled up in number of nodes, the other subnets in the network will scale up as well when trained, propagating the size of the new, larger unit throughout the net at the expense of subnet count as well as task performance. To maintain performance, we presume that other nodes would have to have been added in compensation, though we have not yet performed this experiment. Below, we will argue that the larger, grafted subnet is propagating its size through the net because of problems in input/output (IO) matching.

3.05.3.3.2 Possible causes of network scaling behavior Within any pool of nodes, such as the 64- and 128-node nets we used, there is always an interface to put input into the net, and register its output. However, subnets within that pool also have their own interfaces used to communicate across subnet boundaries. Subnet communication sends output from one subnet and feeds it to the input of another. As such, the number of subnet output neurons in the upstream subnet is tightly coupled with the number of subnet input neurons in the downstream subnet. Increasing the IO layer of any one subnet should cause subnets that are connected to the scaled subnet to widen their IO layers as well.

A quick look at the feedback nature of the nervous system suggests that IO matching will cascade (Figure 9). Potentially, in a strictly feedforward network, IO matching might only occur between the superscaled subnet and its immediate neighbors. However, any subnets that communicate not only with the scaled subnet, but also with each other would also have to scale their output layers to maintain compatibility with each other, causing a systemwide IO width increase cascade. Multiple types of compensation for systemwide scaling could possibly result, reconfiguring subnets in various ways. Even so, scaling any single component will inevitably tend toward a decrease in the number of resources dedicated to providing computational services as compared to the number of resources spent on wiring in the new, larger component.

At large scales, the increase in the number of subnets causes another kind of fragmentation of net structure when engaged in the same tasks. To maintain a relatively stable and low number of average hops, it is necessary for subnets to be directly connected to a number of other subnets that is proportional to the total number of nodes in the pool. This means as the pool grows larger, subnets will on average experience an increase in the number of other subnets to which they are directly connected. Ordinarily, downstream subnets converge on a single input layer, and upstream ones receive their input for multiple terminations of the same output layer. However, it may be the case that not all downstream subnets are sending information in the same encoding, or it may also be the case that some downstream subnets are transmitting information that could clobber other senders' output. In both cases, there is a need for a dedicated set of interfaces to the same subnet, such that input from different subnets may be coming in on different nodes.

As the number of nodes in a pool grows, so too does the likelihood that within any given set of

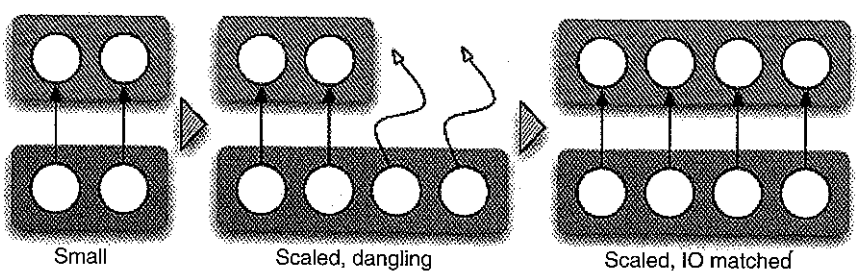

Figure 9 Diagram of the cascade of the magnitude of input/output (IO) units within a neural net.

subnets connected to the same subnet, some of the members of the set will require their own dedicated interface on the subnet to which they are connected. The result is a larger number of neurons in the interface layers and a smaller number of neurons in the hidden layers, increasing the number of nodes involved in IO, leaving fewer nodes available for actual computation. This phenomenon, like IO matching, makes subnets less efficient as the pool to which they belong scales up.

We suspect that simply inserting more neurons between the input and output neurons to improve processing power is not a viable way around the cascading scale effect, though we intend to investigate this kind of manipulation in future work. First, some subnets may already exist that cannot be significantly upgraded because they are at or close to the apex of the resource consumption to the computational power ratio. Secondly, certain important types of subnets, such as those that store memory, would require an increase in the IO layer to provide greater capacity, and more memory requires a larger interface to address that memory. Here too, the larger interface would trigger a cascading scale effect; any subnet interested in storing or retrieving information would need the ability to handle the memory's address space.

We believe that the decrease in both task performance and the number of modules as the result of grafting superscaled modules suggests why evolving brains scale in a coordinated manner. When a system is presented with a scaled-up component, it undergoes a systemwide reorganization that both changes underlying intermodule wiring strategies and the number and function of the modules themselves. The scaling of an area of the brain such as the somatosensory strip, S1, would have a profound impact on the rest of the brain. This cortical area participates not only in somatosensory analysis, but also is activated during word-finding tasks with a somatosensory component (Pulvermuller, 2001). The underlying mechanism that drives the correlation between behavioral complexity and relative but not absolute brain size may be the scaling properties of neural networks.

3.05.4 Overview: Distinguishing Developmental and Computational Structure from Constraint

3.05.4.1 What Are the Units of Brain Architecture in Development, Evolution, and Mature Function?

Developmental and evolutionary biologists and computational scientists attempting to understand the brain are all engaged in parallel quests to describe the fundamental units of brain structure and a syntax for their interaction. Each addresses a separate kind of change in the brain, over evolution, development, or while functioning. Much is to be gained by understanding how the requirements of each aspect of brain change reinforce and constrain each other. Much confusion is the result, however, if the structure in variation of these different types of brain change is arbitrarily assumed to be the same, or if terms are exported and imported between areas of study without particular attention to their referents. The unit of a cortical area is a good example of this. In the initial stages of studying mature cortex physiology, the cortical area was viewed as an extremely important level of analysis, in that each area was hypothesized to take its input and perform a particular transform of it before passing it up to the next, from area 17 to 18, to 19, and so on (Hubel and Wiesel, 1998). As is well known, neither the circuitry (DeYoe and Van Essen, 1988) nor the functional allocation proved to be very well described this way (for example, Schiller, 1993). Although debate persists for some particular cases (for example, Gauthier *et al.*, 1999; Haxby *et al.*, 2001; Grill-Spector *et al.*, 2004), most agree that perceptual or cognitive functions of even minor complexity are rarely associated uniquely with particular cortical areas and are usually highly distributed across the cortical surface (one example each from four independent functional domains: Andersen, 1995; Duncan and Owen, 2000; Haxby *et al.*, 2001; Pulvermuller, 2001). So, while the cortical area remains an extremely important aspect of cortical syntax, containing topographically ordered sensory and computed maps and highly constrained thalamocortical and downstream connectivity, comparison of present box diagrams of information flow for disparate functions would show much overlap and subdivision.

The researchers examining the developmental neurobiology of the cortex also began by concentrating on the cortical area, the holy grail of this enterprise to determine if particular cortical areas were genetically unique in the absence of patterning input. The essential hypothesis to be tested was whether the structure of mature cortical areas could be identified in the initial deployment of cells. Thus, researchers were very interested to determine if location in the cortical germinal zone was faithfully transmitted to the mature cortex, to determine if the nature of cortical areas could be fixed at the time of neurogenesis (Austin and Cepko, 1990; Rakic, 1990), or whether the properties of cortical areas could survive embryonic transplant

(O'Leary and Stanfield, 1989) or thalamic rewiring (Pallas, 2001). Though the picture is very far from complete, we now know that the primary sensory and motor areas may be identified by the graded pattern of gene expression that serves both to organize the polarity of cortical maps and direct early thalamocortical and intracortical connectivity (Ragsdale and Grove, 2001). Whether any comparable specification exists for areas other than the primary sensory and motor regions is not known, but it appears unlikely, and even within the most specified regions, much residual plasticity is retained (Kingsbury and Finlay, 2001). Developmental neurobiology has therefore produced evidence for unique specification of only a few of the multiple areas that are seen in the mature brain.

Comparative neurobiologists found evidence for a conserved set of primary regions across mammals (Krubitzer, 1995), but as this paper has argued, there is reasonable evidence that cortical areas proliferate predictably as a function of cortex size, and very little evidence to suggest that the relative size or number of cortical areas is related to niche-specific requirements or special adaptation (Nudo and Masterton, 1990; Kaskan *et al.*, 2005). We suggest that the comparative evidence is consistent with essentially three types of cortical areas, a set of primary sensory regions that are genetically specified and conserved across mammalian brains, a set arising from axonal sorting interactions in relatively small brains and similar to other features of local cortical topography such as ocular dominance columns and cytochrome oxidase blobs, and a final set arising from some unknown mechanism at very large scales.

The relationship of the concept of a cortical area in evolution, development, and mature function is subtle, important, and complex, but it is no longer the simple concept that the cortical area is the computational unit of the functioning cortex, the genetically specified fundamental unit of the developing cortex, and the selected-on unit of cortical evolution. Are there other candidates for fundamental units of brain development and evolution?

The developmental radial unit and its outcome in the mature brain (Rakic, 1990), the cortical column, is certainly a central aspect of the story, but our first explorations in the computational aspects of scaling suggest we might look for units at an intermediate level of analysis. Subnets emerged in neural nets trained on generic tasks, revealed by their relatively greater interconnectivity. These subnets have properties in computer simulations that demonstrate how cascading IO scaling is also likely to be a problem in real-world brain scaling. Subnets may be an essential unit of a neural wiring strategy, using modularity at a micro scale, where clusters of neurons implement logical units of computation, rather than the higher-level behaviors that have been the focus of most modular accounts of the brain. The demonstration of similar subnet structure in the brain and investigation of their scaling properties in brains of different sizes would be useful: possible candidates might be like the axonal architectures that link the cortical representations of extended visual contours or other spatial structures that can appear either early in development (Fitzpatrick, 2000) or later as a result of learning (Gilbert *et al.*, 1996).

A particular subnet need not be restricted to use within a single class of behavior; it can be used concurrently across different and functionally unrelated subnets. To grow larger, a subnet would recruit neurons from subnets with which it already shares neurons. Conversely, to contract, a module would relinquish neurons shared with neighboring micromodules increasing the number of neurons dedicated to a single module. Subnets might also be replicated across the brain, especially when it is cheaper to replicate than connect; the probability a subnet will replicate is inversely related to its size and directly related to the number of other subnets that depend on it.

3.05.4.2 Allometric Constraints May Show Us the Scaling Properties of Brain Networks

At this point, we may turn around to look at some of the most highly conserved aspects of mammalian brain scaling and raise them as potential points of interest for their network scaling properties. At the most basic level of proliferation, the change in number of motor neurons will be less than the proliferation of secondary sensory neurons as brains enlarge. Why?

Looking at the proliferation of the thalamus, it is quite notable that the primary sensory regions, the lateral geniculate, medial geniculate, and ventrobasal nuclei have their terminal birth dates earliest and proliferate at a greatly reduced rate compared to the rest of the thalamus and, of course, with respect to the cortex. For the pulvinar and frontally connecting regions of the thalamus, neurons are even conscripted in large brains from new regions to expand their size. Our initial work would suggest that the primary thalamic regions might be kept small, perhaps even to some detriment to acuity, to minimize the number of neurons wasted on IO functions and allow the proliferation of large numbers of

reusable subnets (Gilbert *et al.*, 1996). Finally, it is clear that the size of large regions of the brain, cortex, olfactory bulb, hippocampus, and cerebellum may enlarge differentially, usually at a rate characteristic of whole radiations. How does the wiring between large regions stay exempt from the scaling and wiring constraints we have described thus far?

3.05.4.3 Where Do We Look for Species-Specific Adaptations in the Brain?

We have described several kinds of sources for species differences in the brain already, including the residual variation left after allometrically predictable variation is taken into account, the obvious possibility of modifications of neuronal architecture and function, and our two-hit model for placing new functions in areas destined to become disproportionately large in large brains by virtue of their developmental placement. We end with a plea that further investigation of the nature of brain scaling and neuroanatomical evolution abandon models of brain function that have not been used in sophisticated analysis of brain function for the last 25 years, and consider the network structure of the brain, and how computations might be deployed over networks of different scales. Hypothesis development in this next phase of understanding brain evolution will require that the hypotheses be explicit about the nature of networks and the nature of the computations embodied there.

References

Aiello, L. C. and Wheeler, P. 1995. The expensive-tissue hypothesis: The brain and digestive system in human and primate evolution. *Curr. Anthropol.* 36, 199–221.

Airey, D. C., Robbins, A. I., Enzinger, K. M., Wu, F., and Collins, C. E. 2005. Cortical map variation predicts inbred mouse strain identity. *BMC Neurosci.* 6, 18.

Albritton, C. C. 1989. Catastrophic Episodes in Earth History. Chapman and Hall.

Alvarez, L. W., Alvarez, W., Asaro, F., and Michel, H. V. 1980. Extraterrestrial cause for the Cretaceous–Tertiary extinction. *Science* 208, 1095–1108.

Andersen, R. A. 1995. Encoding of intention and spatial location in the posterior parietal cortex. *Cereb. Cortex* 5, 457–469.

Austin, C. P. and Cepko, C. L. 1990. Cellular migration patterns in the developing mouse cerebral cortex. *Development* 110, 713–732.

Barkow, J. H., Cosmides, L., and Tooby, J. 1992. The Adapted Mind: Evolutionary Psychology and the Generation of Culture. Oxford University Press.

Barton, R. A. 1998. Visual specialization and brain evolution in primates. *Proc. R. Soc. Lond: Biol. Sci.* 265, 1933–1937.

Barton, R. A. and Harvey, P. H. 2000. Mosaic evolution of brain structure in mammals. *Nature* 405, 1055–1058.

Barton, R. A., Purvis, A., and Harvey, P. H. 1995. Evolutionary radiation of visual and olfactory brain systems in primates, bats and insectivores. *Philos. Trans. R. Soc. Lond. B* 348, 381–392.

Bayer, S. A. 1980. Development of the hippocampal region in the rat. I: Neurogenesis examined with 3H-thymidine autoradiography. *J. Comp. Neurol.* 190, 87–114.

Bayer, S. A. 1983. [3H] Thymidine-radiographic studies of neurogenesis in the rat olfactory bulb. *Exp. Brain Res.* 50, 329–340.

Brodmann, K. 1909. Vergleichende Lokalisationslehre de Grosshirnrinde in Ihren Prinzipien dargestellt auf Grund des Zellenbaues. Barth.

Bronchti, G., Heil, P., Sadka, R., Hess, A., Scheich, H., and Wollberg, Z. 2002. Auditory activation of 'visual' cortical areas in the blind mole rat (*Spalax ehrenbergi*). *Eur. J. Neurosci.* 16, 311–329.

Callaerts, P., Halder, G., and Gehring, W. J. 1997. Pax-6 in development and evolution. *Ann. Rev. Neurosci.* 20, 483–532.

Calvert, G. A., Bullmore, E. T., Brammer, M. J., *et al.* 1997. Activation of auditory cortex during silent lipreading. *Science* 276, 593–596.

Catania, K. C. and Kaas, J. H. 1997. Somatosensory fovea in the star-nosed mole: Behavioral use of the star in relation to innervation patterns and cortical representation. *J. Comp. Neurol.* 387, 215–233.

Chomsky, N. 1975. Reflections on Language. Pantheon Books.

Clancy, B., Darlington, R. B., and Finlay, B. L. 1999. The course of human events: Predicting the timing of primate neural development. *Dev. Sci.* 3, 57–66.

Clancy, B., Darlington, R. B., and Finlay, B. L. 2001. Translating developmental time across mammalian species. *Neuroscience* 105, 7–17.

Clark, D. A., Mitra, P. P., and Wang, S. S. H. 2001. Scalable architecture in mammalian brains. *Nature* 411, 189–193.

Cooper, H. M., Herbin, M., Nevo, E., and Negroni, J. 1995. Neuroanatomical consequences of microphthalmia in mammals. *Vis. Adapt.* 6, 127–139.

Darlington, R. B., Dunlop, S. A., and Finlay, B. L. 1999. Neural development in metatherian and eutherian mammals: Variation and constraint. *J. Comp. Neurol.* 411, 359–368.

Dawkins, R. 1976. The Selfish Gene. Oxford University Press.

Dawkins, R. 1986. The Blind Watchmaker. W. W. Norton.

de Winter, W. and Oxnard, C. E. 2001. Evolutionary radiations and convergences in the structural organization of mammalian brains. *Nature* 409, 710–714.

DeYoe, E. A. and Van Essen, D. C. 1988. Concurrent processing streams in monkey visual cortex. *Trends Neurosci.* 11, 219–226.

Dickson, B. J. 2002. Molecular mechanisms of axon guidance. *Science* 298, 1959–1964.

Dominy, N. J. and Lucas, P. W. 2001. Ecological importance of trichromatic vision to primates. *Nature* 410, 363–366.

Duboule, D. and Dollé, P. 1989. The structural and functional organisation of the mouse Hox gene family resembles that of *Drosophila* homeotic genes. *EMBO J.* 8, 1497–1505.

Duncan, J. and Owen, A. M. 2000. Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends Neurosci.* 23, 475–481.

Falchier, A., Renaud, L., Barone, P., and Kennedy, H. 2001. Extensive projections from the primary auditory cortex and polysensory area STP to peripheral area V1 in the macaque. In: 31st Annual Meeting for the Society for Neuroscience Society for Neuroscience.

Field, D. J. 1994. What is the goal of sensory coding? *Neural Comput.* 6, 559–601.

Finlay, B. L. and Darlington, R. B. 1995. Linked regularities in the development and evolution of mammalian brains. *Science* 268, 1578–1584.

Finlay, B. L. and Pallas, S. L. 1989. Control of cell number in the developing visual system. *Prog. Neurobiol.* 32, 207–234.

Finlay, B. L., Hersman, M. N., and Darlington, R. B. 1998. Patterns of vertebrate neurogenesis and the paths of vertebrate evolution. *Brain Behav. Evol.* 52, 232–242.

Finlay, B. L., Darlington, R. B., and Nicastro, N. 2001. Developmental structure in brain evolution. *Behav. Brain Sci.* 24, 263–307.

Finlay, B. L., Cheung, D., and Darlington, R. B. 2005a. Developmental constraints on or developmental structure in brain evolution? In: Processes of Change in Brain and Cognitive Development (*eds.* Y. Munakata and M. Johnson), vol. XXI. Oxford University Press.

Finlay, B. L., Silveira, L. C. L., and Reichenbach, A. 2005b. Comparative aspects of visual system development. In: The Structure, Function and Evolution of the Primate Visual System (*ed.* J. Kremers). Wiley.

Fitzpatrick, D. 2000. Seeing beyond the receptive field in primary visual cortex. *Curr. Opin. Neurobiol.* 10, 438–443.

Fodor, J. A. 1983. The Modularity of Mind. MIT Press.

Fodor, J. A. 1992. Precis of the modularity of mind. *Behav. Brain Sci.* 8, 1–42.

Fodor, J. A. 2000. The Mind Doesn't Work That Way. MIT Press.

Frahm, H. K., Stephan, H., and Baron, G. 1984. Comparisons of brain structure volumes in insectivora and primates. V: Area striata. *J. Hirnforsch.* 25, 537–557.

Gao, W. J. and Pallas, S. L. 1999. Cross-modal reorganization of horizontal connectivity in auditory cortex without altering thalamocortical projections. *J. Neurosci.* 19, 7940–7950.

Gauthier, I., Behrmann, M., and Tarr, M. J. 1999. Can face recognition really be dissociated from object recognition? *J. Cogn. Neurosci.* 11, 349–370.

Gerhart, J. and Kirschner, M. 1997. Cells, Embryos and Evolution. Blackwell Science.

Gibson, K. R. 2002. Evolution of human intelligence: The roles of brain size and mental construction. *Brain Behav. Evol.* 59, 10–20.

Gilbert, C. D., Das, A., Ito, M., Kapadia, M., and Westheimer, G. 1996. Spatial integration and cortical dynamics. *Proc. Natl. Acad. Sci. USA* 93, 615–622.

Gilissen, E. and Zilles, K. 1995. The relative volume of the primary visual cortex and its intersubject variability among humans: A new morphometric study. *C. R. Acad. Sci. Paris* 320, 897–902.

Glendenning, K. K. and Masterton, R. B. 1998. Comparative morphometry of mammalian central auditory systems: Variation in nuclei and form of the ascending system. *Brain Behav. Evol.* 51, 59–89.

Gould, E., Reeves, A. J., Graziano, M. S. A., and Gross, C. G. 1999. Neurogenesis in the neocortex of adult primates. *Science* 286, 548–552.

Gould, S. J. 1975. Allometry in primates, with emphasis on scaling and the evolution of the brain. In: Approaches to Primate Paleobiology, vol. 5, pp. 244–292. Karger.

Gould, S. J. 1977. Ontogeny and Phylogeny. Harvard University Press.

Gould, S. J. 1980. The Panda's Thumb. W. W. Norton.

Graham, A., Papalopulu, N., and Krumlauf, R. 1989. The murine and drosophil homeobox gene complex have common features of organisation and expression. *Cell* 57, 367–378.

Greenough, W. T. and Bailey, C. H. 1988. The anatomy of a memory: Convergence of results across a diversity of tests. *Trends Neurosci.* 11, 142–147.

Grill-Spector, K., Knouf, N., and Kanwisher, N. 2004. The fusiform face area subserves face perception, not generic within-category identification. *Nat. Neurosci.* 7, 555–562.

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., and Pietrini, P. 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293, 2425–2430.

Hof, P. R., Glezer, I. I., Nimchinsky, E. A., and Erwin, J. M. 2000. Neurochemical and cellular specializations in the mammalian neocortex reflect phylogenetic relationships: Evidence from primates, cetaceans, and artiodactyls. *Brain Behav. Evol.* 55, 300–310.

Hofman, M. A. 1989. On the evolution and geometry of the brain in mammals. *Prog. Neurobiol.* 32, 137–158.

Hubel, D. H. and Wiesel, T. N. 1962. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.* 160, 106–154.

Hubel, D. H. and Wiesel, T. N. 1968. Receptive fields and functional architecture of monkey striate cortex. *J. Physiol.* 195, 215–243.

Hubel, D. H. and Wiesel, T. N. 1998. Early exploration of the visual cortex. *Neuron* 20, 401–412.

Jerison, H. J. 1973. Evolution of the Brain and Intelligence. Academic Press.

Jerison, H. J. 1997. Evolution of prefrontal cortex. In: Development of the Prefrontal Cortex: Evolution, Neurobiolgy and Behavior (*eds.* N. A. Krasnegor, G. R. Lyon, and P. S. Goldman-Rakic), pp. 9–26. Pall H. Brooks.

Jolicoeur, P., Pirlot, P., Baron, G., and Stephan, H. 1984. Brain structure and correlation patterns in insectivora, chiroptera and primates. *Syst. Zool.* 33, 14–29.

Jones, E. G. 1998. Viewpoint: The core and matrix of thalamic organization. *Neuroscience* 85, 331–345.

Kaas, J. H. 1987. The organization of neocortex in mammals: Implications for theories of brain function. *Annu. Rev. Psychol.* 38, 129–151.

Kaas, J. H. 1997. Topographic maps are fundamental to sensory processing. *Brain Res. Bull.* 44, 107–112.

Kaas, J. H. 2000. Why is brain size so important: Design problems and solutions as neocortex get bigger or smaller. *Brain Mind* 1, 1–25.

Kaskan, P. and Finlay, B. L. 2001. Encephalization and its developmental structure: How many ways can a brain get big? In: Evolutionary Anatomy of the Primate Cerebral Cortex (*ed.* T. Sanderson), pp. 14–29. Cambridge University Press.

Kaskan, P., Franco, C., Yamada, E., Silveira, L. C. L., Darlington, R., and Finlay, B. L. 2005. Peripheral variability and central constancy in mammalian visual system evolution. *Proc. R. Soc. Biol. Sci.* 272, 91–100.

Kaushal, D., Contos, J. A. J., Treuner, K., *et al.* 2003. Alteration of gene expression by chromosome loss in the postnatal mouse brain. *J. Neurosci.* 23, 5599–5606.

Keimel, K. and Roth, W. 1992. Ordered Cones and Approximation. Springer.

Kiltie, R. A. 2000. Scaling of visual acuity with body size in mammals and birds. *Funct. Ecol.* 14, 226–234.

Kingsbury, M. A. and Finlay, B. L. 2001. The cortex in multidimensional space: Where do cortical areas come from? *Dev. Sci.* 4, 125–156.

Knudsen, A. G. 2001. Two genetic hits (more or less) to cancer. *Nat. Rev. Cancer* 1, 157–162.

Krubitzer, L. 1995. The organization of neocortex in mammals: Are species differences really so different? *Trends Neurosci.* 18, 408–417.

Leutenegger, W. 1982. Encephalization and obstetrics in primates. In: Primate Brain Evolution: Methods and Concepts (*eds.* E. Armstrong and D. Falk), pp. 43–96. Plenum.

Manev, R. and Manev, H. 2004. Just a few new neurons may help maintain brain's small-world network. *Med. Hypotheses* 64, 114–117.

Manger, P., Sum, M., Szymanski, M., Ridgway, S., and Krubitzer, L. 1998. Modular subdivisions of dolphin insular cortex: Does evolutionary history repeat itself? *J. Cogn. Neurosci.* 10, 153–166.

McClelland, J. L., McNaughton, B. L., and O'Reilly, R. C. 1995. Why there are complementary learning systems in the hippocampus and neocortex: Insights from successes and failures of connectionist models of learning and memory. *Psychol. Bull.* 102, 419–457.

Merzenich, M. M., Schreiner, C., Jenkins, W., and Wang, X. Q. 1993. Neural mechanisms underlying temporal integration, segmentation, and input sequence representation – some implications for the origin of learning disabilities. In: Temporal Information Processing in the Nervous System (*eds.* P. Tallal, A. M. Galaburda, R. R. Llinas, and C. Voneuler), vol. 682, pp. 1–22. Academic Sciences.

Milo, R., Itzkovitz, S., Kashtan, N., *et al.* 2004. Superfamilies of evolved and designed networks. *Science* 303, 1538–1542.

Moller, P. and Hurlbert, A. C. 1996. Psychophysical evidence for fast region-based segmentation processes in motion and color. *Proc. Natl. Acad. Sci. USA* 93, 7421–7426.

Newman, M. E. J. 2003. The structure and function of complex networks. *SIAM Rev.* 45, 167–256.

Nickle, D. A. and Heymann, E. W. 1996. Predation on orthoptera and other orders of insects by tamarin monkeys, *Saguinus mystac mystax* and *Saguinus fuscicollis nigrifrons* (Primates: Callitrichidae) in north-eastern Peru. *J. Zool. Lond.* 239, 799–819.

Nolen, T. and Hoy, R 1984. Initiation of behavior by single neurons: The role of behavioral context. *Science* 226, 992–994.

Nolfi, S. and Floreano, D. 2002. Synthesis of autonomous robots through evolution. *Trends Cogn. Sci.* 6, 31–37.

Nudo, R. J. and Masterton, R. B. 1990. Descending pathways to the spinal cord. IV: Some factors related to the amount of cortex devoted to the corticospinal tract. *J. Comp. Neurol.* 296, 584–597.

Olavarria, J. and Montero, V. 1990. Elaborate organization of visual cortex in hamster. *Neurosci. Res.* 8, 40–47.

O'Leary, D. D. M. and Stanfield, B. B. 1989. Selective elimination of axons extended by developing cortical neurons is dependent on regional locale: Experiment utilizing fetal cortical transplants. *J. Neurosci.* 9, 2230–2246.

Pallas, S. L. 2001. Intrinsic and extrinsic factors that shape neocortical specification. *Trends Neurosci.* 24, 417–423.

Parnas, D. L. 1972. A technique for software module specification with examples. *Commun. ACM* 15, 330–336.

Persons, M. H., Fleishman, L. J., Frye, M. A., and Stimphil, M. E. 1999. Sensory response patterns and the evolution of visual signal design in anoline lizards. *J. Comp. Physiol. A* 184, 585–607.

Petersen, S. E. and Fiez, J. A. 1992. Processing of single words studied with positron emission tomography. *Annu. Rev. Neurosci.* 16, 509.

Pulvermuller, F. 2001. Brain reflections of words and their meaning. *Trends Cognitive Sci.* 5, 517–524.

Purvis, A. and Rambaut, A. 1995. Comparative analysis by independent contrasts (CAIC): An Apple Macintosh application for analysing comparative data. *Comput. Appl. Biosci.* 11(3), 247–251.

Qi, H.-X. and Kaas, J. H. 2004. Myelin stains reveal an anatomical framework for the representation of digits in somatosensory area 3b of macaque monkeys. *J. Comp. Neurol.* 477, 172–187.

Radman, M., Matic, I., and Taddei, F. 1999. Evolution of evolvability. *Ann. NY Acad. Sci.* 870, 146–155.

Ragsdale, C. W. and Grove, E. A. 2001. Patterning the mammalian cerebral cortex. *Curr. Opin. Neurobiol.* 11, 50–58.

Rakic, P. 1988. Specification of cerebral cortical areas. *Science* 241, 170–176.

Rakic, P. 1990. Critical cellular events in cortical evolution: Radial unit hypothesis. In: The Neocortex: Ontogeny and Phylogeny (*eds.* B. L. Finlay, G. Innocenti, and H. Scheich), pp. 21–32. Plenum.

Rakic, P. 2002. Neurogenesis in adult primate neocortex: An evaluation of the evidence. *Nat. Rev. Neurosci.* 3, 65–71.

Rakic, P., Suner, I., and Williams, R. 1991. A novel cytoarchitectonic area induced experimentally within the primate visual cortex. *Proc. Natl. Acad. Sci. USA* 88, 2083–2087.

Reep, R., Darlington, R. B., and Finlay, B. L. 2007. The limbic system in mammalian brain evolution. *Brain Behav. Evol.* (in press).

Regan, B. C., Julliot, C., Simmen, B., Vienot, F., Charles-Dominique, P., and Mollon, J. D. 2001. Fruits, foliage and the evolution of primate colour vision. *Philos. Trans. R. Soc. Lond. Biol. Sci.* 356, 229–283.

Rockland, K. S. 2001. Calcarine area V1 as a multimodal convergence area. In: 31st Annual Meeting for the Society for Neuroscience. Society for Neuroscience.

Rubenstein, J. L. R., Martinez, S., Shimamura, K., and Puelles, L. 1994. The embryonic vertebrate forebrain: The prosomeric model. *Science* 266, 578–579.

Ryan, M. J. and Rand, A. S. 1995. Female responses to ancestral advertisement calls in tungara frogs. *Science* 269, 390–392.

Ryan, R. J. 1998. Sexual selection, receiver biases and the evolution of sex differences. *Science* 281, 1999–2003.

Sadato, N., Pascual-Leone, A., Grafman, J., *et al.* 1996. Activation of the primary visual cortex by Braille reading in blind subjects. *Nature* 380, 526–528.

Sauerwein, H. and Lassonde, M. 1994. Cognitive and sensorimotor functioning in the absence of the corpus callosum: Neuropsychological studies in callosal agenesis and callosotomized patients. *Behav. Brain Res.* 64, 229–240.

Scharff, C. 2000. Chasing fate and function of new neurons in adult brains. *Curr. Opin. Neurobiol.* 10, 774–783.

Schiller, P. H. 1993. The effects of V4 and middle temporal (MT) area lesions on visual performance in the rhesus monkey. *Vis. Neurosci.* 10, 717–746.

Schmidt-Nielsen, K. 1984. Scaling: Why Is Animal Size So Important? Cambridge University Press.

Schüz, A. and Miller, R. 2002. Cortical Areas: Unity and Diversity. Taylor and Francis.

Semendeferi, K., Lu, A., Schenker, N., and Damasio, H. 2002. Humans and great apes share a large frontal cortex. *Nat. Neurosci.* 5, 272–276.

Silveira, L. C. L., Pincanco-Diniz, C. W., Sampaio, L. F. S., and Oswaldo-Cruz, E. 1989. Retinal ganglion cell distribution in the cebus monkey: A comparison with the cortical magnification factors. *Vis. Res.* 29, 1471–1483.

Stein, E. and Tessier-Lavigne, M. 2001. Hierarchical organization of guidance receptors: Silencing of netrin attraction by slit through a Robo/DCC receptor complex. *Science* 291, 1928–1938.

Stephan, H., Baron, G., and Frahm, H. D. 1981. New and revised data on volumes of brain structures in insectivores and primates. *Folia Primatol.* 53, 1–29.

Stevens, C. F. 2001. An evolutionary scaling law for the primate visual system and its basis in cortical function. *Nature* 411, 193–195.

Suga, N., Kuzirai, K., and O'Neill, W. E. 1981. How biosonar information is represented in the bat cerebral cortex. In: Neuronal Mechanisms of Hearing (*eds.* J. Syka and L. Aitken). Plenum.

Swindale, N. V. 1980. A model for the formation of ocular dominance stripes. *Proc. R. Soc. Biol. Sci.* 208, 243–264.

Swindale, N. V. 2001. Keeping the wires short: A singularly difficult problem. *Neuron* 29, 316–317.

Takahashi, T., Nowakowski, R. S., and Caviness, V. S. 1997. The mathematics of neocortical neuronogenesis. Dev. *Neurosci.* 19, 17–22.

Valverde, F. 1990. Aspects of phylogenetic variability in neocortical intrinsic organization. In: The Neocortex: Ontogeny and Phylogeny (*eds.* B. L. Finlay, C. Innocenti, and H. Scheich), pp. 87–102. Plenum.

Van Essen, D. C., Newsome, W. T., and Maunsell, J. H. 1984. The visual field representation in striate cortex of the macaque monkey: Asymmetries, anisotropies, and individual variability. *Vis. Res.* 24, 429–448.

Watts, D. J. and Strogatz, S. H. 1998. Collective dynamics of 'small world' networks. *Nature* 393, 440–442.

Watts, D. S. 1999. Small worlds: The dynamics of networks between order and randomness. Princeton University Press.

Wilkins, A. S. 2001. The Evolution of Developmental Pathways. Sinauer Associates.

Xiong, M. and Finlay, B. L. 1996. What do developmental mapping rules optimize? Prog. *Brain Res.* 112, 350–361.

Zhang, K. and Sejnowski, T. J. 2000. A universal scaling law between gray matter and white matter of cerebral cortex. *Proc. Natl. Acad. Sci. USA* 97, 5621–5626.

Zilles, K. 1985. The Cortex of the Rat. Springer.

Further Reading

DeYoe, E. A. and Van Essen, D. C. 1988. Concurrent processing streams in monkey visual cortex. *Trends Neurosci.* 11, 219–226.

Finlay, B. L., Darlington, R. B., and Nicastro, N. 2001. Developmental structure in brain evolution. *Behav. Brain Sci.* 24, 263–307.

Gerhart, J. and Kirschner, M. 1997. Cells, Embryos and Evolution. Blackwell.

Jerison, H. J. 1973. Evolution of the Brain and Intelligence. Academic Press.

Kaas, J. H. 1997. Topographic maps are fundamental to sensory processing. *Brain Res. Bull.* 44, 107–112.

Kaskan, P., Franco, C., Yamada, E., Silveira, L. C. L., Darlington, R., and Finlay, B. L. 2005. Peripheral variability and central constancy in mammalian visual system evolution. *Proc. R. Soc. Biol. Sci.* 272, 99–100.

Krubitzer, L. 1995. The organization of neocortex in mammals: Are species differences really so different? *Trends Neurosci.* 18, 408–417.

McClelland, J. L., McNaughton, B. L., and O'Reilly, R. C. 1995. Why there are complementary learning systems in the hippocampus and neocortex: Insights from successes and failures of connectionist models of learning and memory. *Psychol. Bull.* 102, 419–457.

Milo, R., Itzkovitz, S., Kashtan, N., *et al.* 2004. Superfamilies of evolved and designed networks. *Science* 303, 1538–1542.

Pallas, S. L. 2001. Intrinsic and extrinsic factors that shape neocortical specification. *Trends Neurosci.* 24, 417–423.

Ragsdale, C. W. and Grove, E. A. 2001. Patterning the mammalian cerebral cortex. *Curr. Opinion Neurobiol.* 11, 50–58.

Rakic, P. 1990. Critical cellular events in cortical evolution: Radial unit hypothesis. In: The Neocortex: Ontogeny and Phylogeny (*eds.* B. L. Finlay, G. Innocenti, and H. Scheich), pp. 21–32. Plenum.

Rubenstein, J. L. R., Martinez, S., Shimamura, K., and Puelles, L. 1994. The embryonic vertebrate forebrain: The prosomeric model. *Science* 266, 578–579.

Ryan, R. J. 1998. Sexual selection, receiver biases and the evolution of sex differences. *Science* 281, 1999–2003.

Sadato, N., Pascual-Leone, A., Grafman, J., *et al.* 1996. Activation of the primary visual cortex by Braille reading in blind subjects. *Nature* 380, 526–528.

Schüz, A. and Miller, R. 2002. Cortical Areas: Unity and Diversity. Taylor and Francis.

Stephan, H., Baron, G., and Frahm, H. D. 1981. New and revised data on volumes of brain structures in insectivores and primates. *Folia Primatol.* 53, 1–29.