The eloquent ape: genes, brains and the evolution of language

Simon E. Fisher and Gary F. Marcus‡*

Abstract | The human capacity to acquire complex language seems to be without parallel in the natural world. The origins of this remarkable trait have long resisted adequate explanation, but advances in fields that range from molecular genetics to cognitive neuroscience offer new promise. Here we synthesize recent developments in linguistics, psychology and neuroimaging with progress in comparative genomics, gene-expression profiling and studies of developmental disorders. We argue that language should be viewed not as a wholesale innovation, but as a complex reconfiguration of ancestral systems that have been adapted in evolutionarily novel ways.

Linguistics
The scientific study of language.

Phonetics
The study of the production, perception and physical properties of speech sounds.

Phonology
The study of the sound systems of languages and the ways in which speech sounds can be combined.

Cerebral cortex
The outer layer of the mammalian cerebrum, which is involved in processes such as sensation, perception, cognition and (in humans) language.

**Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK.*
‡Department of Psychology, New York University, 6 Washington Place, New York 10003, USA.
Correspondence to S.E.F.
e-mail: simon.fisher@well.ox.ac.uk
doi:10.1038/nrg1747

Human language seems to be unique in the natural world. Non-human communication is predominantly restricted to simple messages such as alarm calls and identification signals, with little in the way of complex structure[1,2]. By contrast, the average human has access to a vocabulary of tens of thousands of words and can, guided by an intricate set of structural rules, assemble them into a potentially infinite number of meaningful sentences, referring not only to the here and now, but also to the past, the future and the abstract[3,4]. Remarkably, this rich linguistic system is usually acquired without conscious effort or formal instruction. The drive to acquire language is so robust that a lack of aural input does not necessarily abate it; deaf babies who are exposed to sign language babble using their hands[5], and deaf children who have had little access to sign-language input can develop language-like gesture systems[6]. In comparison, no other living primate naturally acquires more than a few signals, and these are combined in rudimentary ways[2].

Given such sharp distinctions between communication in humans and that found in other species, language has often been investigated as an isolated phenomenon. Experts in linguistics have studied aspects of language that include its sound systems (phonetics and phonology), the ways in which words can be put together from smaller meaningful units (morphology), and the principles that govern sentence construction (syntax) and meaning (semantics), with little or no reference to the biology or psychology of non-human species. Similarly, neuroscientists who seek to understand the neural basis of human communication have tended to focus their attention on two regions of the cerebral cortex that were thought to provide specialized human-specific substrates for processing language — Broca's area (commonly described as the seat of grammar) and Wernicke's area (described as the seat of meaning and sound structure)[7,8].

These efforts have proved effective for many purposes, such as clarifying the nature of language (BOX 1) and probing electrophysiological activity in the brain during the production or comprehension of a sentence. Even so, although researchers have made progress by studying language purely on its own terms, it does not follow that language should be studied in this way. Few if any phenotypic traits are entirely without precedent. The avian wing, for example, can be thought of as a specially modified version of the basic vertebrate forelimb — an idea that is supported by a well-described molecular and developmental basis[9]. As suggested by Darwin over a century ago[10], the behavioural and cognitive peculiarities of *Homo sapiens* — including our extraordinary capacity for language — should be similarly explicable as the product of descent with modification[17]. Here we argue that with recent progress across many disciplines — including genetics and genomics, which are the focus of this article — the scientific community is finally approaching a position in which it can fulfil Darwin's promise.

Approaching language evolution

The search for the origins of language is far from new. A whole host of different (often conflicting) hypotheses have been proposed[11], which have been framed with respect to a wide range of questions. Can language be explained by the same kind of adaptive evolution that has shaped other traits[12]? Was language itself subject

Box 1 | What is language?

Although linguistic functions seem to be seamless, operating largely below the level of conscious awareness, the act of communication requires the extensive coordination of a broad range of mechanisms. When a person speaks, abstract thoughts are automatically transformed into rapid and intricately synchronized articulatory gestures that involve, among others, muscles of the tongue, lips and jaw. The resulting modulations in the shape of the vocal tract, coupled with precisely timed onsets and offsets for the vibration of the larynx, translate into ordered sequences of temporally and acoustically complex sounds[26]. The listener extracts meaning from this stream of raw acoustic signals, despite variations in pitch, rate and accent, without any obvious acoustic markers to signal boundaries between words[129]. Speaker and listener might switch roles many times, and can have a meaningful conversation without any previous explicit agreement about what particular sounds should signify[20]. In short, language is a rich computational system that simultaneously coordinates syntactic, semantic, phonological and pragmatic representations with each other, motor and sensory systems, and both the speaker's and listener's knowledge of the world. As such, tracing the genetic origins of language will require an understanding of a great number of sensory, motor and cognitive systems, of how they have changed individually, and of how the interactions between them have evolved.

to selective forces, or did it emerge as a secondary by-product of other properties, such as a larger and more complex brain, with greater computational resources[13]? Is language the consequence of a single radical macromutation[14,15], or was it honed in successive steps[12]? What selective advantages might be associated with this trait? Suggestions have ranged from enhanced communication of information[12] to improved organization of internal thought[16], sexual selection[18] and increased social cohesion[19]. What came first — a means for the fine articulation of the vocal tract (speech) or a means for combining individual communicative elements and coordinating them with meaning (language)? Or did the two co-evolve[20]?

Until recently, relevant empirical investigations were mainly restricted to three domains — archaeological studies, linguistic reconstructions of intermediate forms of language and computational modelling of constraints on language evolution. These approaches have yielded interesting findings, but each has been hampered by uncertainty. Archaeological approaches are limited because cognitive systems do not leave any direct physical fossil record. Although studies of fossilized hominin skeletons have provided evidence about the position of the larynx[21], degree of tongue innervation[22] and sophistication of breathing control[23] during human evolution, the significance of these changes for the emergence of language remains highly controversial[24–26]. Putative precursors of language systems — which are based on studying aspects of modern usage[27,28] and the ways that newly formed languages develop[27,29] — are not open to independent verification. Mathematical and computational approaches[30,31] face similar problems. For example, studies have identified circumstances under which a language that has a lexicon but no rules for combining words into sentences could evolve into a system that contains rules for constructing new sentences to describe novel situations[30]. However, at present there is no way to validate the core assumption that lexicons evolved before grammar.

Against this daunting backdrop we see several reasons to be optimistic. First, contemporary data have highlighted the flaws in traditional views of the neurological bases of language[8,32,33] (BOX 2). Because the classical model that revolves around Broca's and Wernicke's areas invokes neural substrates that are unique to language and to humans, it unduly minimizes the possibility of understanding language origins through studies of animals or other cognitive systems. However, neither Broca's nor Wernicke's area is devoted entirely to language processing[34,35] and, in fact, these substrates might not be human-specific[36–38]. It is also now generally accepted that language capacity involves a complex network of cortical and subcortical circuits that is broadly distributed across the brain[32,39] (BOX 2).

Second, although non-human primate communication shows qualitative differences from human language, studies have established that most components of language show some degree of continuity with other species. For example, the human vocal tract supports a wider repertoire of speech sounds than could be produced by other primates[26], but the capacity to create richly modulated formants is not unique to humans[40]. Likewise, many animals and birds can distinguish different human speech sounds, and adult tamarin monkeys can discriminate between the distinctive rhythmic properties of different languages[41]. Debate continues about exactly how much of the machinery of language is species- or language-specific; for example, opinion is divided over whether recursion represents the only component that is genuinely new to the human species[42,43]. Nevertheless, views that consider language to be fully independent of ancestral systems are no longer tenable, and there is a growing recognition that cognitive, physiological, neuroanatomical and genetic data from non-speaking species can greatly inform our understanding of the nature and evolution of language[32,42–44].

A third principal reason for optimism comes from developments in molecular genetics, including large-scale comparative genomics[45], investigations of gene expression[46] and explorations of specific genes that have been suggested by studies of developmental disorders[47]. As described below, these advances collectively offer new types of empirical data for addressing hypotheses about how humans diverged from other primates.

Comparing primate genomes

One new avenue seeks to investigate the origins of language by comparing the genomic sequences of humans and other closely related species. Although we currently lack adequate genetic material from extinct hominin species[48] (but see REF. 49), the complete draft genome sequence of *Pan troglodytes*, the closest extant primate cousin of *H. sapiens*, yields a catalogue of almost every sequence difference that distinguishes a human from a chimpanzee[50]. Furthermore, genomic sequencing of the rhesus macaque and orangutan is well underway.

Surprisingly, it seems that most human–chimpanzee genomic differences have accumulated through genetic drift during the several million years since the two

Macromutation
A single evolutionary event that has large-scale effects on a phenotype, which involves concurrent alteration of numerous characteristics.

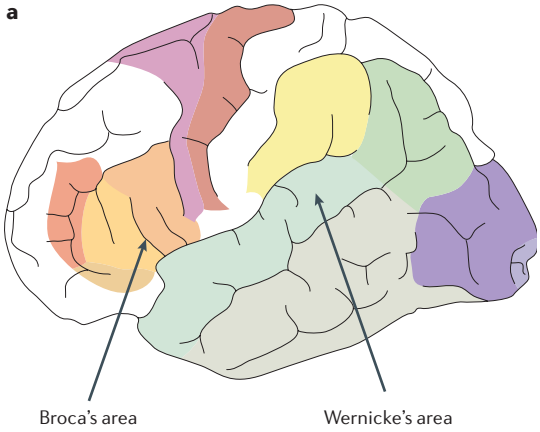
Sexual selection
The evolution of a trait as a consequence of competition among members of one sex (usually males) for fertilization opportunities with the other sex.

Hominin
Modern humans and all extinct human-like ancestors and their relatives that existed following divergence from other ape lineages. Includes all species of the genera *Homo* and *Australopithecus*, and other ancient forms such as *Paranthropus* and *Ardipithecus*.

Larynx
The specialized upper portion of the respiratory tract that houses the vocal cords — folds of mucous membrane that provide the source for vocal sounds. The low position of the larynx in adult humans allows a rich phonetic repertoire, but its significance for language evolution remains a matter of debate.

Lexicon
The vocabulary and word forms of a language.

Box 2 | **Evolving views of the neurological basis of language**



Popular accounts of the neurological basis of language often begin with two discrete and lateralized regions of the cortex: Broca's area in the inferior frontal gyrus and Wernicke's area in the superior temporal gyrus, and the primary connection between them — a cortical fibre tract that is known as the arcuate fasciculus. (Precise boundaries for Broca's and Wernicke's areas remain a matter of debate, but their approximate location with respect to a surface side view of the left side of a human brain is shown in panel **a**.) In a once prominent model, these areas represented the key language-specific substrates, serving distinctive functions such as speech production and/or grammar (Broca's area) and meaning or comprehension (Wernicke's area). This classical view had its roots in early studies of brain lesions, and still resonates today in the names of two common forms of aphasia[7]. Broca's aphasia involves poorly articulated effortful speech with few words, whereas Wernicke's aphasia involves fluent speech but disrupted content, which is accompanied by deficits in language comprehension.

Most contemporary researchers see the traditionally defined language areas as part of a larger network that remains poorly understood, involving diverse regions of the cortex (coloured areas in panel **a**, as discussed in REF. 39) and subcortical structures such as the striatum, thalamus and cerebellum (see panel **b** for locations of these structures with respect to a sagittal cross-section through a human brain)[8,32,33,39]. People with lesions in Broca's area do not necessarily develop symptoms of Broca's aphasia, and those with damage to Wernicke's area do not always suffer from Wernicke's aphasia[8,33]. These and other types of aphasia might result from damage to various cortical regions, or to subcortical structures, particularly the striatum[32,74,130]. Functional neuroimaging studies strengthen the view that different aspects of language processing involve a wide range of structures (panel **a**)[39]. Furthermore, comparative studies suggest that well-documented left–right hemispheric asymmetries in Broca's and Wernicke's areas might be present in the corresponding regions of other great apes[36,37], and a recent report proposed the existence of a homologue of Broca's area in the macaque monkey[38]. Finally, Broca's and Wernicke's areas are known to be active in a range of cognitive domains. For example, Broca's area seems to be involved in imitation, motor control and music cognition[34,35]. At the neural level, as at the cognitive level, language might be seen as the product of a coordinated mixture of mechanisms, some specialized for language, others not. Panel **b** is adapted from *Nature Reviews Neuroscience* REF. 131 © (2005) Macmillan Magazines Ltd.

Aphasia
An inability to produce and/or comprehend language that is due to brain injury or disease.

Subcortical
Describes the brain structures that are below the cerebral cortex.

Striatum
Part of the group of interconnected subcortical nuclei that are known as the basal ganglia. The striatum comprises two nuclei — the caudate and putamen — and is involved in the planning and modulation of movement pathways, as well as a range of other cognitive processes.

Thalamus
A forebrain structure that is located beneath the cerebral hemispheres and that modulates and relays sensory signals to and from the cerebral cortex.

Cerebellum
A multilayered structure in the vertebrate hindbrain that comprises a complex mixture of different cell types. The cerebellum modulates the force and range of movements, maintains balance and is involved in motor learning.

Hemispheric asymmetries
These are differences in the structure or function of the left and right hemisphere counterparts of a particular brain region.

Motor control
The ability to direct and coordinate muscle movements.

Formants
Peaks in the acoustic energy spectrum that result from the resonant frequencies of vocal tracts.

Recursion
A process by which ever more complex elements are generated through the repeated recombination of simpler elements.

lineages diverged[50]. To determine which specific changes have shaped the distinctive features of human biology, researchers have begun by searching coding regions for evidence of accelerated amino-acid change that has exceeded that expected from local mutation rates (BOX 3). These studies have found that genes which are specifically or maximally expressed in the brain tend to show a much lower degree of amino-acid divergence than other genes[50–52]. A probable explanation is that the proteins involved in nervous-system biology are usually subject to strong functional constraint[52]. Despite the overall pattern of reduced divergence, a subset of genes that are implicated in brain development seem to have evolved more rapidly in primates than would be expected on the basis of studies of rodent species[50,52,53]. Further work is needed to determine whether this observation reflects the action of positive selection, but detailed investigation of such outliers might provide candidates for involvement in human brain evolution.

The combination of primate sequences and within-species diversity data from human populations offers a potential route for identifying signatures of recent selection anywhere in the genome (BOX 3). Using chimpanzee data it is often possible to determine which allele of a human SNP (single nucleotide polymorphism) represents the state that was present in the common human–chimpanzee ancestor. Preliminary analyses of 120,000 validated human SNPs[50] highlighted 7 large genomic regions that show a reduced rate of diversity[54] and a large proportion of high-frequency, non-ancestral alleles[55]. These features are suggestive of a selective sweep during human evolution, and so could be another source of candidate genes for exploring the emergence of traits such as language.

At the same time, it is important to realize that access to a complete chimpanzee genome sequence is no panacea. Although early reports of a high degree of genetic overlap still stand[56], the genomes of humans

Box 3 | **Signatures of selection**

The effects of natural selection can sometimes be inferred from patterns of sequence divergence between and/or within species. For example, consider the ratio between the number of non-synonymous substitutions (K_A) and synonymous substitutions (K_S) in a gene during a specific evolutionary period. Assuming that K_S provides an index of the random mutation rate, the K_A/K_S ratio measures whether the rate of protein evolution differs from the rate expected under neutral drift. If $K_A > K_S$, this is taken to indicate accelerated amino-acid change, which might be due to positive selection. Conversely, if $K_A < K_S$, this suggests purifying selection. Substitutions can occur on any lineage of a phylogenetic tree, and selective pressures can differ for different branches, so the comparison of only two species is often inadequate. The use of outgroups allows the inference of ancestral states and lineage-specific analyses. In human–chimpanzee comparisons, outgroups allow the derivation of lineage-specific K_A/K_S ratios and the identification of positive selection that occurred on the human lineage following its split from the chimpanzee.

Genuine positive selection can remain undetected by standard K_A/K_S tests. For example, when a protein is under tight functional constraint some regions might experience intensive purifying selection while others concurrently undergo positive selection, yielding K_A/K_S ratios that are well below 1. As such, researchers have sometimes used 'sliding windows' that cover different regions of a single protein[79–81] or have exploited maximum likelihood methods that allow for K_A/K_S variation at different sites[132]. Where possible, a more robust solution is to compare between-species divergence with the degree of within-species polymorphism for the gene in question[133]. A rate of protein change during evolution of the human lineage that exceeds the rate expected from current levels of polymorphism for that protein in modern human populations represents a stringent indicator of positive selection.

Within-species variation data have the advantage that they allow positive selection to be detected anywhere in the genome, not just in coding regions. When an advantageous mutation spreads through a population through a selective sweep, linked genomic regions that surround the selected site are also affected, yielding a skewed pattern of polymorphism around that site[54,55]. This pattern typically involves an excess of both high- and low-frequency changes: high-frequency changes result from nearby neutral variants 'hitchhiking' along with the selected mutation[55]; low-frequency changes result from a new build up of polymorphisms that occurs after the mutation has become prevalent[54].

No single technique is ideal: between-species K_A/K_S analyses are simple to apply but are limited in scope and can miss true cases of positive selection, whereas methods that exploit within-species comparisons are applicable to any genomic region but depend on the availability of extensive polymorphism data. As described in the main text, each method has proved valuable in studying the evolution of the human brain.

and chimpanzees are replete with differences. By itself, the between-species 1.23% substitution rate in single-copy genomic regions corresponds to over 35 million changes[50], and it is accompanied by at least 5 million indel events, which represent another 3% or so of genomic divergence[50]. Further sources of change that might prove to be important include differences in non-coding RNAs[57], DNA methylation[58], chromosomal rearrangements[50] and altered gene-copy numbers that result from lineage-specific duplications or losses[59,60]. Our ability to distinguish functional adaptive change from the considerable amount of background noise remains limited, particularly for non-coding regions[61], and defining the role of any given gene represents a major undertaking.

In terms of understanding the origins of language, it is also worth noting that many of the adaptive changes in the human genome might be unrelated to this trait; human biology is distinctive for a range of anatomical, metabolic, biomedical and behavioural characteristics[45]. Some — such as bipedalism, increased relative brain size and advanced tool use — might be defining features of the human condition. However, other traits — such as the human-specific susceptibility to malaria — simply involve adaptive differences of a kind that are commonly found between closely related species[45]. Identifying the genomic contributions to language evolution will ultimately depend not only on evidence of positive selection, but also on a demonstration that variation in a given candidate gene affects linguistic capacities.

Expressive apes and expression arrays

Humans and chimpanzees differ not just in the genes that they carry, but also in how these genes are expressed. With the advent of high-throughput techniques for characterizing gene-expression profiles, such as hybridization of RNA to oligonucleotide microarrays[46], species-wide (and organism-wide) differences in spatial and temporal regulation of gene expression can now be directly investigated. This approach is in its infancy, but recent studies of neural tissue from equivalent regions of different primate species, or distinct regions of the same species, support several preliminary conclusions.

First, convergent data from many studies indicate an acceleration of neural gene-expression changes during human evolution[62–66]. It should be emphasized that human and chimpanzee brains show considerably less absolute divergence than other tissues, such as the liver and the heart, both in terms of the number of genes that are differentially expressed and the magnitudes of the differences[63–65], which is probably due to higher functional constraint in neural tissue[52]. However, alteration of neural gene expression on the branch leading to humans seems to have been more dramatic than that found for the chimpanzee branch during the same period. A parallel (although less significant) tendency for human-lineage acceleration has been proposed for rates of amino-acid change in brain-expressed genes[52,53]. So, human cognitive evolution might have involved a complex combination of changes in the regulation of gene expression and in protein structure.

Second, most studies report that the above acceleration of expression differences is associated with a general upregulation of expression in the human cortex compared with that found in the chimpanzee[63–65] (although see REFS 66,67 for alternative interpretations). Upregulated genes tend to be enriched in genomic regions that have been recently duplicated in human evolution[68], but the functional importance of this finding is not known.

Third, coincident with cortical asymmetries in language function, a recent investigation identified marked expression differences between the left and right hemispheres in the early development of human embryonic brains, preceding the emergence of morphological distinctions[69]. In particular, LIM domain only 4 (*LMO4*), a gene that is linked to cortical patterning, is more highly expressed in part of the right cortex than the equivalent region of the left cortex at 12–14 weeks of gestation. Intriguingly, although *Lmo4* expression in the mouse cortex shows moderate asymmetry in individual brains, there is no consistent lateralization to one or other hemisphere; therefore, altered regulation of this gene during evolution might be relevant to the emergence of human asymmetries. Given that asymmetries of brain structure

Positive selection
When a novel allele that increases the fitness of an organism becomes more prevalent in the population.

Purifying selection
Selection against alleles that have harmful phenotypic effects, which leads to their loss from the population.

Outgroups
When two closely related species are compared, the status of the common ancestor (for example, at a site of substitution) could be deduced by including a more distant third species that branched from the parent group before the other two groups diverged.

Box 4 | **Big brains and language evolution**

Brain size and complexity vary considerably among different mammalian species. Differences in encephalization quotient — a measure of the relative brain weight, which is adjusted for overall body weight — are particularly marked among primates, with a significant trend for progressive brain expansion throughout the ~60 million years of primate evolution[76]. The most dramatic changes seem to have occurred in the human lineage in the past 2–3 million years, such that the encephalization quotient of a modern human is three to fourfold higher than that of any other great ape[77].

Many genes might participate in modulating brain size (for example, REFS 134,135), including genes that mediate processes such as cell proliferation and cell death. At this stage, two genes — *MCPH1* (the gene that encodes microcephalin) and *ASPM* (abnormal-spindle-like, microcephaly associated) have been the subject of the most intensive evolutionary study[78–82], following their association with primary microcephaly, which is a rare human condition of reduced brain size[77] (TABLE 1). Patterns of between-species and within-species divergence indicate that these two genes underwent positive selection during primate evolution, albeit with acceleration of coding change peaking at different times. In each case 'mosaic' patterns of selection have been observed; certain domains evolved under intensive positive selection, whereas others have remained tightly conserved[79–82]. Both *MCPH1* and *ASPM* are evolutionarily ancient, with orthologues that are likely to be present in all chordates, and their roles in primate evolution seem to have extended over many millions of years, involving a succession of adaptive events[78–82]. As such, MCPH genes provide apt examples of Darwinian selection in human brain evolution, whereby novel configurations and variants of ancient elements can fuel substantial adaptive change.

However, although modifications of these or other similar genes that impact on gross features of brain development might have provided significant preconditions for more sophisticated behaviours, it is unlikely that they were sufficient alone for the emergence of language. For example, despite a stark reduction in cortical volume, which can be as little as a third of that found in normal individuals, children with primary microcephaly have an overtly normal neuroanatomical architecture, show only mild-to-moderate mental retardation and reach many developmental milestones[76,77]. In our view the honing of traits such as language probably depended not just on increased 'raw materials' in the form of a more ample cortex, but also on more specific modifications of particular neural pathways.

Functional constraint
The degree to which changes in gene sequence are tolerated. For genes that have higher functional constraints a larger proportion of potential mutations are deleterious, reducing the substitution rate.

Maximum likelihood
A statistical method that is commonly used to make inferences about the most likely value of one of more parameters that underlie a given data set.

Selective sweep
Occurs when an allele increases in frequency as a consequence of positive selection and concurrently eliminates neutral variation at linked chromosomal sites.

Indel
A difference between sequences of related genomes that results from an insertion or deletion event; a term that is especially used when the evolutionary direction of the change is unknown.

have been suggested for other primates[36,37], and that functional asymmetry is associated with several aspects of cognition (such as spatial and facial recognition[70]), it is unclear how relevant these results are to language, but they clearly represent another avenue that is worth pursuing.

Fourth, studies have yet to identify any specific human gene (or set of genes) that is uniquely expressed in language-related regions of adult brains. An investigation of three human brains found that language-related substrates in the cerebral cortex showed similar expression profiles to those of other cortical areas[68]; differences between individuals tended to outweigh differences between regions within any single individual. Expression profiles did not differ significantly between Broca's area and the corresponding area of the right hemisphere[68], despite well-documented asymmetries of cytoarchitecture[71] and function[72]. Moreover, comparison to homologous brain regions in chimpanzees indicated that the vast majority of expression differences between these species are common to all cortical regions, rather than being localized to areas that are related to linguistic function[68].

So far, neural expression profiling in humans and primates has lacked the power to detect effects that involve subsets of cells, which is problematic as each region of the cortex comprises a highly complicated mix of distinctive cell types. In addition, it is difficult to identify clearly which expression changes were adaptive and which were selectively neutral[67]. Finally, for ethical reasons, these investigations have exploited only tissues from autopsies, as it is currently unfeasible to characterize dynamic expression profiles in the functioning human brain. As we discuss elsewhere in this review, comparative analyses of neural expression patterns in other species, particularly song-birds[73–75], are not subject to the same limitations as human–primate studies, and might provide further entry points into language-related mechanisms.

Gene-driven strategies

An alternative approach to large-scale comparative studies begins by pinpointing genes that are judged likely to influence human language, and involves targeting these genes for detailed evolutionary investigation. The exact nature of the neuromolecular pathways that underlie language remains mysterious, so discovering these genes is far from trivial, but some progress has been made through positional cloning studies of human neurodevelopmental disorders[47]. If mutation of a particular gene is implicated in neural abnormalities, then sequence variation in the gene can, in principle, influence relevant developmental outcomes[76]. Therefore, a close examination of between- and within-species diversity is warranted, as it can reveal whether alteration of the gene in question was involved in human evolution. For example, studies of primary microcephaly[77] — a disorder of reduced brain size — have suggested mechanisms that could have contributed to cortical expansion during primate evolution[78–82] (BOX 4), whereas other rare syndromes (such as Joubert syndrome 3 (JBTS3)[83] and bilateral frontoparietal polymicrogyria (BFPP)[84]) might provide clues to adaptive changes in brain organization (TABLE 1). Genetically mediated increases in brain size and overall organization have clearly been important in human evolution, but they do not themselves adequately explain language origins. For more direct insights it is worth focusing on neurodevelopmental disorders that primarily affect speech and/or language skills[47].

Insights from complex disorders. Although language acquisition seems to be universal across the human species[4], a significant proportion of children have language-related deficits that cannot be explained by a known cause, such as mental retardation, cerebral palsy, autism or hearing impairments[85]. There is considerable phenotypic heterogeneity among these children and the diagnosis of subtypes remains controversial[85], but genetic factors make a strong contribution[86]. Identification of the relevant genes is proving to be challenging, especially given that most speech and language disorders have a multifactorial basis[47]. Nevertheless, advances in genotyping technology and statistical methods, coupled with sophisticated phenotypic characterization, are delivering encouraging results. For common forms of disorder, genetic studies point to several chromosomal intervals that might harbour risk variants that are involved in specific language impairment (SLI; intervals

Table 1 | **Insights into primate brain evolution from genetic studies of human neurodevelopmental disorders**

Syndrome	Clinical features	Positional cloning	Functional insights	Insights for primate evolution	Refs
Brain size/organization					
MCPH	Head circumference at birth is >3 SD below the mean, but neuroanatomy is otherwise normal; mild–moderate mental retardation	6 MCPH loci have been mapped; mutations in 4 genes have been identified — microcephalin (*MCPH1*), *ASPM* (*MCPH5*), *CDK5RAP2* (*MCPH3*) and *CENPJ* (*MCPH6*)	Microcephalin contains three BRCA1 C-terminal domains; that are involved in cell-cycle regulation and/or DNA repair; ASPM is an orthologue of the abnormal spindle protein in fruitflies; it is essential for organizing spindle structures during neuroblast division; CDK5RAP2 and CENPJ influence centrosomal microtubule production during neurogenesis	K_A/K_S ratios for *MCPH1* and *ASPM* show accelerated change along the lineages leading from the common primate ancestor to humans; *MCPH1* ratios peaked before the greater–lesser ape divergence; *ASPM* ratios peaked on the human lineage after the split from chimpanzees; for both genes, between-species primate divergence exceeds within-species human divergence, which provides further support for positive selection	77–82
JBTS	Complex abnormalities of cerebellar vermis/brainstem; clumsiness, abnormal motor function, developmental delay, behavioural anomalies	One brain-specific form maps to 6q23 (*JBTS3*) and is caused by mutations of *AHI1*	AHI1 is a cytoplasmic adaptor protein that might modulate crossing of axonal fibres	*AHI1* shows $K_A/K_S>1$ in the human lineage; changes to 12 amino acids (of 1,196 residues) occurred in this lineage after the split from chimpanzees, and are fixed in modern humans; relevant to emergence of human-specific motor behaviours?	83
BFPP	Abnormal layering or excessive folding of the cortex, mainly in the frontal lobes; abnormal gait, mental retardation, seizures and language impairment	Mutations have been identified in *GPR56* on 16q13–21	GPR56 is an orphan G-protein-coupled receptor that is involved in regulating cortical patterning, particularly in the frontal lobes	No specific evaluation of primate evolution has been reported, but animals that have a cerebral cortex carry a unique variable GPR56 N-terminal region of unknown function	84
Language development					
SLI	Significant discrepancy between verbal and non-verbal skills despite adequate educational opportunity and in the absence of another neurological condition (for example, cerebral palsy or autism)	Genome-wide scans have highlighted regions on 13q21 (*SLI3*), 16q24 (*SLI1*) and 19q13 (*SLI2*)	Mapped intervals span large numbers of candidate genes with diverse functions; many might relate to neural development, for example, the *SLI1* region contains *USP10*, which encodes a protein that is involved in synaptic growth	Evolutionary comparisons might help to prioritize genes for mutation testing in linked chromosomal regions; *SLI1* and *SLI2* intervals contain genes (including *USP10*) that show an increased copy number in the human lineage	85–88
DD	Difficulties in reading or spelling despite overtly normal verbal skills and adequate educational opportunity; often associated with subtle impairments in language processing	Multiple replicated loci for example, 2p12–16 (*DYX3*), 3p12–q13 (*DYX5*), 6p22.2 (*DYX2*), 15q21 (*DYX1*) and 18p11 (*DYX6*)	Specific genes have recently been proposed as strong candidates; for example, *DYX1C1* (15q21), which encodes a protein with TPR protein-interaction motifs, and *KIAA0319* (6p22.2), which encodes a protein with PKD cell-adhesion domains	DD is associated with deficits in linguistic pathways; analyses of risk genes might shed light on evolution; *DYX1C1* shows coding changes in different primates, but no formal evidence of selection; *KIAA0319* evolution has not yet been studied	90,92–97
AD	Syndrome characterized by deficits in reciprocal social interaction and communication, which is accompanied by repetitive and stereotyped behaviours	Many chromosomal regions have been implicated; a subset (for example, 3q, 7q or 13q21) might relate to language	Numerous diverse brain-related genes have been tested for involvement in this disorder; there is no clear link as yet between any particular gene and AD-associated language deficits	No direct insights into language origins have emerged as yet from studies of AD; this area might be the most informative for understanding social aspects of communication	91,98–100
DVD	Impaired learning and production of sequences of mouth movements; often accompanied by expressive and receptive language or grammar deficits	One Mendelian form of the disorder (SPCH1) is caused by mutations in *FOXP2* on 7q31	FOXP2 is a transcription factor that might regulate development or function of distributed circuits in diverse brain regions, including the cortex, striatum, thalamus and cerebellum	*FOXP2* is highly conserved in vertebrates, but underwent accelerated coding change on the human lineage after the split from chimpanzees; patterns of intronic polymorphism indicate recent selection	101–123

Examples are given for two broadly defined classes of human disorder — the first is related to gross features of brain development (organization/size), and the second is more directly relevant to aspects of language development. This is intended as an illustration of strategy; it is not an exhaustive list of all findings in this area. See the main text for a discussion. AD, autistic disorder; *AHI1*, Abelson helper integration site; *ASPM*, abnormal-spindle-like, microcephaly associated; BFPP, bilateral frontoparietal polymicrogyria; BRCA1, breast cancer 1, early onset; *CDK5RAP2*, CDK5 regulatory-subunit-associated protein 2; *CENPJ*, centromere protein J; DD, developmental dyslexia; *DYX1C1*, dyslexia susceptibility 1 candidate 1; DVD, developmental verbal dyspraxia; *FOXP2*, forkhead box P2; JBTS, Joubert syndrome; MCPH, primary microcephaly; PKD, polycystic kidney disease; SD, standard deviation; SLI, specific language impairment; SPCH1, speech-language disorder 1; TPR, tetratricopeptide repeat; *USP10*, ubiquitin-specific peptidase 10.

SLI1–SLI3) (REFS 47,87,88) (TABLE 1) and in speech-sound disorders (SSD)[89]. Quantitative trait analyses of *SLI1* show that this region strongly influences a child's ability to repeat pronounceable nonsense words[88] — a skill that represents a highly heritable behavioural marker of SLI[86]. Intriguingly, the *SLI1* and *SLI2* regions contain genes that show increased copy number in the human lineage[59], and that can therefore be prioritized for study (TABLE 1). It is too early to tell whether this is merely coincidence, but this does highlight how the integration of data from positional cloning efforts and comparative genomics can generate new testable hypotheses, even before identifying actual risk genes.

Other heritable neurodevelopmental syndromes, such as developmental dyslexia (DD) and autistic disorder (AD), are relevant to language, although they too are characterized by genetic complexity, with multiple chromosomal regions highlighted by mapping studies[90,91] (TABLE 1). Dyslexia is not a linguistic disorder *per se* — it is diagnosed when a child with overtly normal language has unexplained difficulties with learning to read and/or spell[92]. (Reading and spelling, unlike spoken language, do not develop naturally or universally, and instead require extensive tuition.) However, most dyslexic people have subtle underlying deficits in language processing, particularly with handling phonemes[93]. As such, emerging genetic discoveries about the aetiology of dyslexia (for example, see REFS 94–97) might inform our understanding of the basis of language. Similarly, although autism is not primarily a language disorder, deficits in the area of communication represent an important diagnostic feature, along with problems in social interaction and repetitive or stereotyped behaviours[98]. Many autistic children are completely non-verbal, and those that do acquire language competence almost always show pervasive deficits in their use of pragmatics[98]. So, the relevant susceptibility genes, once identified, could be informative for understanding the evolution of social cognition and how this relates to language origins. Some of the putative loci that are linked to autism have been proposed as being relevant to language[99,100]; these studies have been based on analysing subsets of children with language delay, using language-related measures as endophenotypes and/or observing concordant mapping in other disorders.

Insights from a Mendelian disorder: the* FOXP2 *gene. The first direct evidence of a specific gene that influences speech and language acquisition has come not from complex traits, but from an unusual autosomal dominant form of communication disorder[101] that is caused by mutation of the forkhead box P2 (*FOXP2*) gene, which encodes a forkhead box transcription factor[102]. The consequences of *FOXP2* disruption differ from typical SLI in that they include prominent difficulties in learning and producing sequences of movements that involve the mouth and lower part of the face[103]. Affected individuals have problems with speech articulation (developmental verbal dyspraxia or DVD), which are accompanied by wide-ranging deficits in many aspects of language and grammar[104,105]. Crucially, although general intelligence varies among individuals who carry the same *FOXP2* mutation, speech and language deficits are always evident, even for children with normal non-verbal intelligence[105]. Moreover, the associated problems with processing language and grammar are not exclusively tied to speech — they are evident in the written domain and occur for comprehension as well as expression[105]. (For more detailed discussion see REFS 47,106,107.)

The link between *FOXP2* and disordered language was initially identified through genetic studies of a large three-generational family (known as KE)[108,109], in which affected members carry a heterozygous missense mutation that alters the DNA-binding domain of the FOXP2 protein[102] (FIG. 1a). The KE substitution markedly affects the function of the encoded protein (J. Nicôd, S.C. Vernes, F.M. Elahi, A.M. Coupe, L.E. Bird and S.E.F., unpublished observations). *FOXP2* mutations are not a predominant cause of language impairment[110]; however, a second heterozygous point mutation in *FOXP2* was recently identified that co-segregates with speech and language deficits in another family[111]. This nonsense mutation severely truncates the protein, deleting essential functional motifs, including protein–protein interaction domains, the DNA-binding domain and suspected nuclear localization signals[111]. Independent chromosomal aberrations (including translocations and deletions) that disrupt *FOXP2* are associated with speech and language deficits[102,109,112].

A naive view of language evolution might predict that genes such as *FOXP2* are unique to humans, or at least are substantially different in non-speaking species. Instead, the *FOXP2* sequence is highly conserved, even in distantly related vertebrate species[113–115]. Despite this conservation, there has been a profound (>60-fold) increase in amino-acid substitution rate in the human lineage[114] — of 3 substitutions that distinguish the human FOXP2 protein from its mouse counterpart, 2 occurred on the human branch after splitting from the chimpanzee[113] (FIG. 1b). (The acceleration is unlikely to be due to relaxation of functional constraint, because the FOXP2 protein shows little polymorphism in current human populations[110].) Examination of the human within-species variability in intronic sequences that flank the substitutions indicates a selective sweep (BOX 3) that probably occurred within the past 200,000 years[113,114]. In short, one (or both) of the amino-acid substitutions or unidentified regulatory sequences in the flanking introns seem to have been subject to positive selection in recent human history. These studies also dispel the idea that genetically mediated language impairment must be atavistic in nature. Pathological mutations of *FOXP2* (REFS 102,111) are distinct from the putative adaptive changes that occurred during human evolution[113,114] (FIG. 1a) and do not reflect reversions to the ancestral state.

Molecular windows into language origins

The discovery of language-related genes offers new routes for addressing old questions about human evolution. In the case of *FOXP2*, the human version of the gene seems to influence the development and function

Encephalization quotient
A measure of the relative brain size, in which the brain weight is compared with that of the average living mammal of equal body weight.

Cytoarchitecture
The cellular composition of a bodily structure. In neuroscience the term is used to refer to local differences in the arrangement of nerve cells in particular regions of the brain.

Speech-sound disorder
An inability to produce speech sounds that would be expected on the basis of age and dialect, but in the absence of an obvious cause (such as cerebral palsy or hearing impairment). This disorder might occur in isolation or together with other linguistic deficits.

Pragmatics
Social aspects of communication — in particular the influence of context on the interpretation of meaning.

Endophenotype
A measurable intermediate trait that is assumed to provide a closer link to the neurobiological substrate of a disorder.

Forkhead box
An 80–100 amino-acid motif that is found in a similar form in every member of the forkhead box family of transcription factors. It forms a winged-helix structure that allows the protein to bind to DNA.

Nuclear localization signals
Short stretches of amino acids that help to mediate the transport of proteins to the nucleus of the cell.

Atavism
The reappearance in an organism of characteristics that were typical of the organism's remote ancestors.

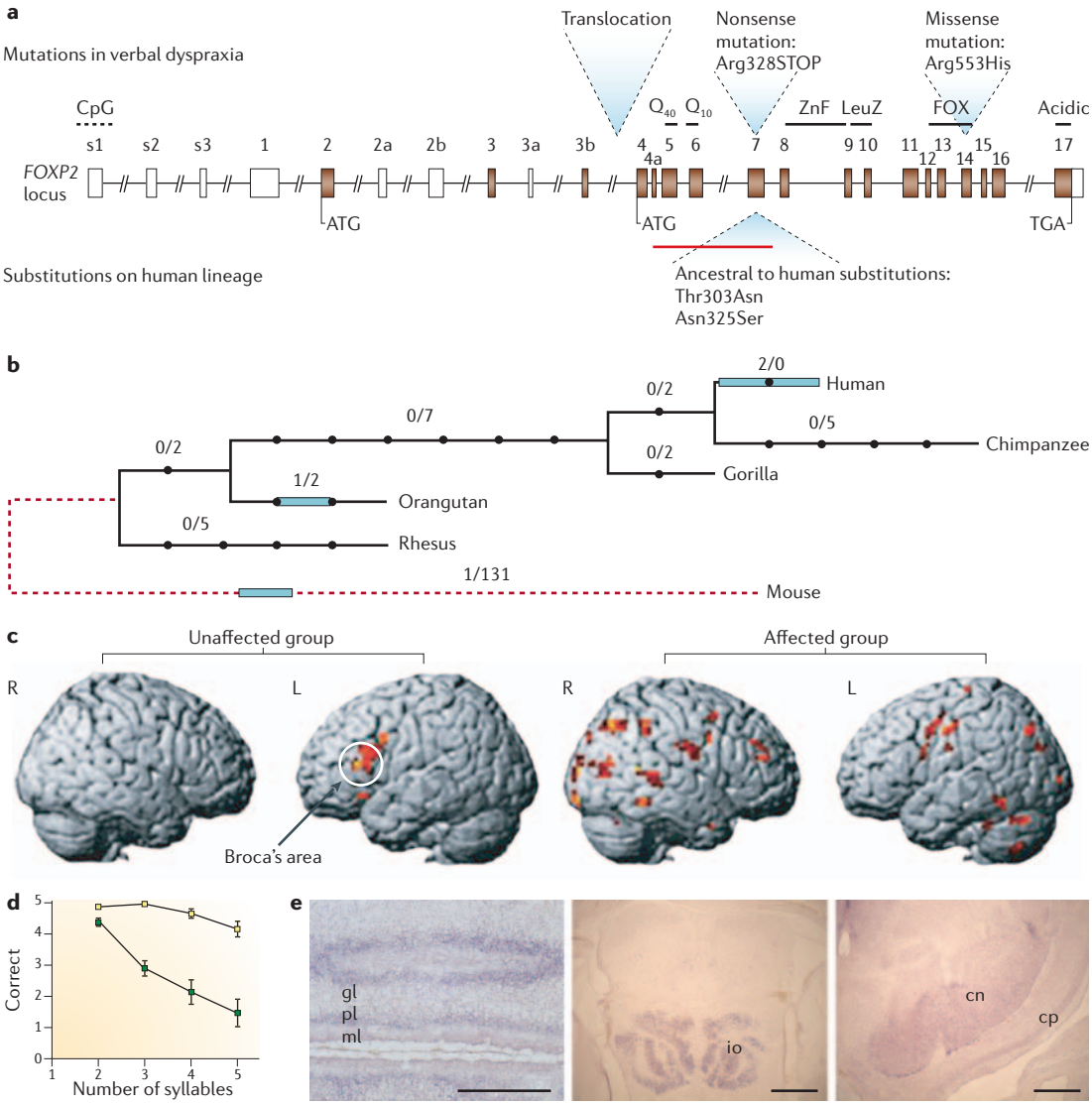


Figure 1 | **A multidisciplinary perspective on language evolution. a** | Genetics — the genomic structure of human forkhead box P2 (*FOXP2*), showing the location of mutations that cause verbal dyspraxia, which are distinct from sites of evolutionary substitution in the human lineage (filled rectangles, coding exons; white rectangles, non-coding exons). The red bar indicates genomic regions that show evidence of a selective sweep[113,114]. Exons encode polyglutamine tracts (Q40 and Q10), a zinc-finger motif (ZnF), a leucine zipper (LeuZ), the forkhead domain (FOX) and an acidic C-terminus (Acidic). s1–s3 are alternatively spliced untranslated 5′ exons. Adapted, with permission, from REF. 111 © (2005) University of Chicago Press. **b** | Evolution — nucleotide substitutions in the *FoxP2* coding region for different lineages during primate evolution, shown as non-synonymous over synonymous substitutions (horizontal bars, nucleotide changes over time; shaded bars, amino-acid changes). Reproduced, with permission, from *Nature* REF. 113 © (2002) Macmillan Magazines Ltd. **c** | Neuroimaging — humans carrying disrupted FOXP2 show functional abnormalities when carrying out a language task, even when producing verb forms mentally rather than aloud. The anomalies involve underactivation of Broca's area and bilateral activation in multiple cortical regions. The diagram shows the group average activation in the unaffected and affected members of the KE family, which is displayed at a threshold of $P<0.05$, corrected for multiple comparisons (L, left hemisphere; R, right hemisphere). Reproduced, with permission, from *Nature Neuroscience* REF. 117 © (2003) Macmillan Magazines Ltd. **d** | Neuropsychology — *FOXP2* disruption leads to difficulties with coordinating speech. Affected KE family members (green squares) perform worse than unaffected members (yellow squares) on word-repetition tests that involve simple articulation patterns (error bars, standard error of the mean). Impairment increases with syllable length. Similar results are seen when repeating nonsense words, with greatest deficits on multisyllabic words that have complex articulation patterns. Adapted, with permission, from REF. 105 © (2002) Guarantors of Brain. **e** | Molecular neuroscience — example sites of high *Foxp2* mRNA expression in transverse sections from a newborn mouse brain (scale bars represent 0.5 mm). In the cerebellum (left panel) *Foxp2* expression is limited to Purkinje cells (pl), and absent from molecular (ml) and granular (gl) layers. In the medulla (middle panel) *Foxp2* is expressed in the inferior olivary nucleus (io). In the forebrain (right panel) there is strong expression in the caudate nucleus (cn) and the deepest layers of the cortical plate (cp). Neural expression patterns for this gene are highly conserved in all vertebrate species that have been studied, which range from humans to zebrafish. Reproduced, with permission, from REF. 118 © (2003) Guarantors of Brain.

not only of classical language-related areas of the cortex, but also of other cortical regions, as well as subcortical structures. Although the brains of individuals who carry *FOXP2* mutations are overtly normal, they have subtle but significant structural anomalies[103,116] — most notably, reduced grey matter in the inferior frontal gyrus (including Broca's area), caudate nucleus and cerebellum[103,116]. Moreover, *FOXP2* disruption leads to functional abnormalities in Broca's area and the striatum during language tasks[103,117]; these tasks elicit diffuse bilateral activation of cortical regions that remain inactive in normal controls[117] (FIG. 1c). *FOXP2* expression during human fetal development shows intriguing overlaps with sites of adult pathology, particularly in the developing cortical plate, caudate nucleus and cerebellum[118]. Therefore, molecular and neuroimaging data independently implicate human *FOXP2* in the development of distributed circuits that involve the frontal cortex, striatum and cerebellum. The networks in question are important for the learning and production of speech sequences, and might account for the articulatory deficits that are associated with damage to this gene (FIG. 1d). Furthermore, the expression data hint that problems with sequencing mouth movements and impaired linguistic development in *FOXP2*-associated disorders might reflect partly independent, pleiotropic effects of a gene that participates in the patterning of both motor and language systems. These findings are consistent with contemporary views of the neurological basis of language (BOX 2), and fit well within a perspective in which linguistic circuitry evolved through descent with modification from ancestral networks that support other cognitive tasks.

FoxP2 *in non-linguistic species*. Investigations of *FoxP2* orthologues further support this model. Vertebrates as diverse as humans[118], rodents[118–120], birds[121,122], reptiles[121] and fish[123] demonstrate concordant regulation of *FoxP2* expression in corresponding brain regions, particularly the cortex (the pallium in non-mammalian species), striatum, thalamus and cerebellum, with conserved patterns of sublocalization (FIG. 1e). Given the similarities in expression and protein structure between distant modern orthologues, it seems highly likely that an ancestral form of *FoxP2* had essential neurodevelopmental functions in the common ancestor of mammals, birds, reptiles and fish. Based on neural expression data, this earlier orthologue might have primarily influenced the circuitry involved in processing sensory information, sensory-motor integration and control of skilled coordinated movements[118–123].

It has been suggested that the ancient neural functions of *FoxP2* have been co-opted to subserve aspects of vocal communication in several species, not just our own[75,121,122]. Many species use innately specified calls, but a few acquire new vocalizations through the imitation of peers, including three groups of birds (parrots, hummingbirds and song-birds) and at least three groups of mammals (humans, cetaceans and bats). This rare trait of vocal learning, which is defined as the ability to imitate and modify sounds, including learning to sequence individual sound units into new combinations, could be viewed as the behavioural substrate for spoken language[74]. In birds and mammals, different vocal-learning species are phylogenetically separated by non-learners, which indicates that this trait might have evolved independently on several occasions[74]. Bird species that are capable of vocal learning have a similar system of discrete brain structures that participate in this trait and show robust alterations in gene-expression patterns during the learning and production of song[74,75]. These structures have not been detected in non-learners, although homologues might exist in some basic form. All birds, regardless of vocal-learning ability, express *FoxP2* in sensory-motor circuitry, and in most regions mRNA levels are consistent across avian species[121,122]. However, there seem to be species-specific differences in expression in the song system of vocal learners, which might relate to variability in vocal plasticity[121]. In the zebra finch, a striatal nucleus, known as Area X, shows higher *FoxP2* levels than the surrounding tissue, but only at the developmental stage when birds learn to imitate song. In adult canaries, expression of this gene in Area X varies with the season: *FoxP2* levels peak in the months when song shows most plasticity. In the zebra finch, it is only the male that learns to sing, and although *FoxP2* itself has similar expression patterns in males and females, the encoded protein might interact directly with the product of *FoxP1* (REF. 124), a closely related gene that has sexually dimorphic expression in the song system[122]. Intriguingly, although there is no evidence of positive selection of FoxP2 protein-coding changes in different avian lineages[115], the above data indicate significant alteration in regulating *FoxP2* in the evolution of song systems.

Combining what is known about avian and mammalian orthologues of *FOXP2*, a compelling hypothesis is that earlier forms of the gene were important for shaping cortical and subcortical sensory-motor networks; circuits which were subsequently recruited, on more than one occasion, to subserve learning and production of complex combinatorial sequences of movements[75]. Just as avian wings are unlike other vertebrate limbs in their ability to support flight, but remain true to their heritage in terms of basic structural properties[9], *FOXP2* seems to have developed its roles in supporting language in humans (and possibly vocal learning in song-birds) while strongly reflecting its ancestral functions.

Future prospects

Further advances in comparative genomics, expression profiling and, in particular, identification of susceptibility factors in developmental language disorders should allow the discovery of language-related genes other than *FOXP2*. It is difficult to predict the nature of these genes. Variation in diverse regulatory factors, receptors, signalling molecules, structural proteins and other functional categories of protein can influence brain organization and function[44]. However, we predict that, in most cases, language-related functions will involve modifications of ancestral roles, which can be defined through multidisciplinary studies. We should not expect such genes to be limited to influencing the brain. For example, as with

Purkinje cells
The output neurons of the cerebellum, which integrate complex inputs and project to the deep motor nuclei of the brain.

Inferior olivary nucleus
A precerebellar nucleus that provides direct input to the Purkinje cells through a network of climbing fibres. Olivocerebellar circuits have a crucial role in controlling movement.

Area X
A striatal nucleus that is present in the song system in the brains of vocal-learning birds.

many transcription factors, *FOXP2* expression is not confined to one particular tissue[102], but is re-used in different contexts (including regions of the developing heart, lung and gut[125]). Functional genetic approaches should allow the teasing out of the properties of language-related genes that are important for language. One possibility would be to compare the behaviour of human and chimpanzee orthologues in model systems.

To the extent that neural functions of language-related genes reflect the modifications of pathways that are present in a common mammalian ancestor, studies of gene function in rodents provide a new means for investigating the origins of language at many levels: molecular, anatomical, developmental and behavioural. Mouse pups make innate calls (both audible and ultrasonic), which are important for mother–offspring interactions[126], and adult males emit ultrasonic vocalizations on encountering female mice or their pheromones[127]. Furthermore, a recent report has demonstrated that the vocalizations of adult male mice are surprisingly rich and have the characteristics of song, with distinctive syllable types that are produced in regular patterns, although, crucially, it has not yet been determined whether these songs are innate or need to be acquired through imitation[127]. Even so, it is important to avoid the simplistic view that rodents can offer a direct model of human communication; behavioural data from knockout studies cannot always be taken at face value. For example, based on an observation that *Foxp2*-deficient pups produced fewer ultrasonic vocalizations than wild-type littermates when artificially separated from their mother, one recent study proposed that these mice demonstrate vocal communication deficits, supposedly paralleling those found in human speech disorders[128]. However, mother–offspring communication in these mice seemed to be generally intact; mothers provided care for all pups regardless of *Foxp2* status and any calls that were produced by mutant pups had normal characteristics[128].

Continuing investigations of song-birds and other vocal learners are likely to continue to yield important insights[74,75]. Although vocal learning in other species is best viewed as analogous, rather than homologous, to aspects of human communication, it is becoming apparent that evolution of the relevant neural systems might be subject to common developmental constraints[74,75]. As such, the powerful strategies that are available for studying avian song systems might offer vital clues to genetic pathways that are involved in human language[73–75].

Understanding the structure and evolution of a trait that is as complicated and unusual as language clearly requires an intensive multidisciplinary effort — synthesizing work from fields as diverse as genetics, linguistics, psychology, neuroscience, anthropology and developmental biology. This will be true both at the level of understanding individual genes, and in describing the system as a whole. In this review, *FOXP2* is necessarily the gene to which we have devoted the most attention. Although this gene is likely to be representative only in some ways but not in others, and is plainly just one piece of the evolutionary puzzle, it provides an exciting taste of what might surface next in investigations of language origins.

1. Terrace, H. S., Petitto, L. A., Sanders, R. J. & Bever, T. G. in *Children's Language* Vol. 2 (ed. Nelson, K. E.) 371–496 (Gardner, New York, 1980).
2. Hauser, M. D. *The Evolution of Communication* (MIT Press, Cambridge, Massachusetts, 1996).
3. Chomsky, N. *Aspects of the Theory of Syntax* (MIT Press, Cambridge, Massachusetts, 1965).
4. Pinker, S. *The Language Instinct* (Allen Lane, London, 1994).
5. Petitto, L. A. & Marentette, P. F. Babbling in the manual mode: evidence for the ontogeny of language. *Science* **251**, 1493–1496 (1991).
6. Goldin-Meadow, S. & Mylander, C. Spontaneous sign systems created by deaf children in two cultures. *Nature* **391**, 279–281 (1998).
7. Damasio, A. R. Aphasia. *N. Engl. J. Med.* **326**, 531–539 (1992).
8. Poeppel, D. & Hickok, G. Towards a new functional anatomy of language. *Cognition* **92**, 1–12 (2004). **The introduction to a collection of articles that describe contemporary views of the neurological basis of language, which highlights the limitations of classical models and lays out an agenda for future research in this area.**
9. Pennisi, E. Bird wings really are like dinosaurs' hands. *Science* **307**, 194 (2005).
10. Darwin, C. *The Descent of Man and Selection in Relation to Sex* (John Murray, London, 1871).
11. Christiansen, M. H. & Kirby, S. *Language Evolution* (Oxford Univ. Press, Oxford; New York, 2003).
12. Pinker, S. & Bloom, P. Natural language and natural selection. *Behav. Brain Sci.* **13**, 707–726 (1990).
13. Gould, S. J. Evolution: the pleasures of pluralism. *NY Rev. Books* **44**, 47–52 (1997).
14. Crow, T. J. Schizophrenia as the price that *Homo sapiens* pays for language: a resolution of the central paradox in the origin of the species. *Brain Res. Brain Res. Rev.* **31**, 118–129 (2000).
15. Klein, R. G. & Edgar, B. *The Dawn of Human Culture* (Wiley, New York, 2002).
16. Dennett, D. C. *Darwin's Dangerous Idea: Evolution and the Meanings of Life* (Simon & Schuster, New York, 1995).
17. Marcus, G. F. Before the word. *Nature* **431**, 745 (2004).
18. Miller, G. F. *The Mating Mind: How Sexual Choice Shaped the Evolution of Human Nature* (Doubleday, New York, 2000).
19. Dunbar, R. Evolution of the social brain. *Science* **302**, 1160–1161 (2003).
20. Liberman, A. M. & Whalen, D. H. On the relation of speech to language. *Trends Cogn. Sci.* **4**, 187–196 (2000).
21. Arensburg, B. *et al.* A Middle Palaeolithic human hyoid bone. *Nature* **338**, 758–760 (1989).
22. Kay, R. F., Cartmill, M. & Balow, M. The hypoglossal canal and the origin of human vocal behavior. *Proc. Natl Acad. Sci. USA* **95**, 5417–5419 (1998).
23. MacLarnon, A. M. & Hewitt, G. P. The evolution of human speech: the role of enhanced breathing control. *Am. J. Phys. Anthropol.* **109**, 341–363 (1999).
24. Lieberman, P., Laitman, J. T., Reidenberg, J. S., Landahl, K. & Gannon, P. J. Folk psychology and talking hyoids. *Nature* **342**, 486–487 (1989).
25. DeGusta, D., Gilbert, W. H. & Turner, S. P. Hypoglossal canal size and hominid speech. *Proc. Natl Acad. Sci. USA* **96**, 1800–1804 (1999).
26. Fitch, W. T. The evolution of speech: a comparative review. *Trends Cogn. Sci.* **4**, 258–267 (2000). **An accessible overview of the ways in which comparisons between diverse species can help to shed light on the origins of speech.**
27. Bickerton, D. *Language and Species* (Univ. Chicago Press, Chicago, 1990).
28. Jackendoff, R. Possible stages in the evolution of the language capacity. *Trends Cogn. Sci.* **3**, 272–279 (1999).
29. Senghas, A., Kita, S. & Ozyurek, A. Children creating core properties of language: evidence from an emerging sign language in Nicaragua. *Science* **305**, 1779–1782 (2004).
30. Nowak, M. A., Komarova, N. L. & Niyogi, P. Evolution of universal grammar. *Science* **291**, 114–118 (2001).
31. Kirby, S. Natural language from artificial life. *Artif. Life* **8**, 185–215 (2002).
32. Lieberman, P. On the nature and evolution of the neural bases of human language. *Am. J. Phys. Anthropol.* Suppl. **35**, 36–62 (2002).
33. Dronkers, N. F., Wilkins, D. P., Van Valin, R. D. Jr, Redfern, B. B. & Jaeger, J. J. Lesion analysis of the brain areas involved in language comprehension. *Cognition* **92**, 145–177 (2004).
34. Nishitani, N. & Hari, R. Temporal dynamics of cortical representation for action. *Proc. Natl Acad. Sci. USA* **97**, 913–918 (2000).
35. Maess, B., Koelsch, S., Gunter, T. C. & Friederici, A. D. Musical syntax is processed in Broca's area: an MEG study. *Nature Neurosci.* **4**, 540–545 (2001).
36. Gannon, P. J., Holloway, R. L., Broadfield, D. C. & Braun, A. R. Asymmetry of chimpanzee planum temporale: humanlike pattern of Wernicke's brain language area homolog. *Science* **279**, 220–222 (1998).
37. Cantalupo, C. & Hopkins, W. D. Asymmetric Broca's area in great apes. *Nature* **414**, 505 (2001).
38. Petrides, M., Cadoret, G. & Mackey, S. Orofacial somatomotor responses in the macaque monkey homologue of Broca's area. *Nature* **435**, 1235–1238 (2005).
39. Demonet, J. F., Thierry, G. & Cardebat, D. Renewal of the neurophysiology of language: functional neuroimaging. *Physiol. Rev.* **85**, 49–95 (2005).
40. Riede, T., Bronson, E., Hatzikirou, H. & Zuberbuhler, K. Vocal production mechanisms in a non-human primate: morphological data and a model. *J. Hum. Evol.* **48**, 85–96 (2005).
41. Ramus, F., Hauser, M. D., Miller, C., Morris, D. & Mehler, J. Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science* **288**, 349–351 (2000).

42. Hauser, M. D., Chomsky, N. & Fitch, W. T. The faculty of language: what is it, who has it, and how did it evolve? *Science* **298**, 1569–1579 (2002).
43. Pinker, S. & Jackendoff, R. The faculty of language: what's special about it? *Cognition* **95**, 201–236 (2005).
References 42 and 43 provide in-depth discussions of the relationship between language and other aspects of cognition, and the extent to which the evolution of human communication depended on language- and human-specific adaptations.
44. Marcus, G. F. *The Birth of the Mind: How a Tiny Number of Genes Creates the Complexities of Human Thought* (Basic Books, New York, 2004).
45. Olson, M. V. & Varki, A. Sequencing the chimpanzee genome: insights into human evolution and disease. *Nature Rev. Genet.* **4**, 20–28 (2003).
46. Preuss, T. M., Caceres, M., Oldham, M. C. & Geschwind, D. H. Human brain evolution: insights from microarrays. *Nature Rev. Genet.* **5**, 850–860 (2004).
47. Fisher, S. E., Lai, C. S. & Monaco, A. P. Deciphering the genetic basis of speech and language disorders. *Annu. Rev. Neurosci.* **26**, 57–80 (2003).
48. Pääbo, S. *et al.* Genetic analyses from ancient DNA. *Annu. Rev. Genet.* **38**, 645–679 (2004).
49. Noonan, J. P. *et al.* Genomic sequencing of Pleistocene cave bears. *Science* **309**, 597–599 (2005).
50. The Chimpanzee Sequencing and Analysis Consortium. Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* **437**, 69–87 (2005).
In this landmark paper, the availability of a complete draft sequence of the *Pan troglodytes* genome facilitated the first comprehensive comparison with that of *Homo sapiens*, as well as a large-scale search for signatures of selection. It also showed that most sequence differences between these species are probably due to neutral drift.
51. Nielsen, R. *et al.* A scan for positively selected genes in the genomes of humans and chimpanzees. *PLoS Biol.* **3**, e170 (2005).
52. Khaitovich, P. *et al.* Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees. *Science* **309**, 1850–1854 (2005).
This investigation makes use of the draft chimpanzee genome sequence in a thorough exploration of how human–chimpanzee divergence in expression levels and protein structure might differ for genes that are active in different tissues (brain, heart, liver, kidney and testis). Although neural genes diverge the least overall, they have tended to accumulate more changes on the human lineage than on the chimpanzee lineage.
53. Dorus, S. *et al.* Accelerated evolution of nervous system genes in the origin of *Homo sapiens*. *Cell* **119**, 1027–1040 (2004).
54. Tajima, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595 (1989).
55. Fay, J. C. & Wu, C. I. Hitchhiking under positive Darwinian selection. *Genetics* **155**, 1405–1413 (2000).
56. King, M. C. & Wilson, A. C. Evolution at two levels in humans and chimpanzees. *Science* **188**, 107–116 (1975).
57. Bentwich, I. *et al.* Identification of hundreds of conserved and nonconserved human microRNAs. *Nature Genet.* **37**, 766–770 (2005).
58. Enard, W. *et al.* Differences in DNA methylation patterns between humans and chimpanzees. *Curr. Biol.* **14**, R148–149 (2004).
59. Fortna, A. *et al.* Lineage-specific gene duplication and loss in human and great ape evolution. *PLoS Biol.* **2**, e207 (2004).
60. Cheng, Z. *et al.* A genome-wide comparison of recent chimpanzee and human segmental duplications. *Nature* **437**, 88–93 (2005).
61. Heissig, F. *et al.* Functional analysis of human and chimpanzee promoters. *Genome Biol.* **6**, R57 (2005).
62. Enard, W. *et al.* Intra- and interspecific variation in primate gene expression patterns. *Science* **296**, 340–343 (2002).
63. Cáceres, M. *et al.* Elevated gene expression levels distinguish human from non-human primate brains. *Proc. Natl Acad. Sci. USA* **100**, 13030–13035 (2003).
64. Gu, J. & Gu, X. Induced gene expression in human brain after the split from chimpanzee. *Trends Genet.* **19**, 63–65 (2003).
65. Hsieh, W. P., Chu, T. M., Wolfinger, R. D. & Gibson, G. Mixed-model reanalysis of primate data suggests tissue and species biases in oligonucleotide-based gene expression profiles. *Genetics* **165**, 747–757 (2003).
66. Uddin, M. *et al.* Sister grouping of chimpanzees and humans as revealed by genome-wide phylogenetic analysis of brain gene expression profiles. *Proc. Natl Acad. Sci. USA* **101**, 2957–2962 (2004).
67. Khaitovich, P., Pääbo, S. & Weiss, G. Toward a neutral evolutionary model of gene expression. *Genetics* **170**, 929–939 (2005).
68. Khaitovich, P. *et al.* Regional patterns of gene expression in human and chimpanzee brains. *Genome Res.* **14**, 1462–1473 (2004).
This study included a systematic comparative analysis of gene expression in Broca's area and other brain areas in human adults, including the corresponding region of the right hemisphere, but did not uncover any clear changes that might relate to language.
69. Sun, T. *et al.* Early asymmetry of gene transcription in embryonic human left and right cerebral cortex. *Science* **308**, 1794–1798 (2005).
70. Rossion, B. *et al.* Hemispheric asymmetries for whole-based and part-based face processing in the human fusiform gyrus. *J. Cogn. Neurosci.* **12**, 793–802 (2000).
71. Amunts, K., Schleicher, A., Ditterich, A. & Zilles, K. Broca's region: cytoarchitectonic asymmetry and developmental changes. *J. Comp. Neurol.* **465**, 72–89 (2003).
72. Musso, M. *et al.* Broca's area and the language instinct. *Nature Neurosci.* **6**, 774–781 (2003).
73. Clayton, D. F. Songbird genomics: methods, mechanisms, opportunities, and pitfalls. *Ann. NY Acad. Sci.* **1016**, 45–60 (2004).
74. Jarvis, E. D. Learned birdsong and the neurobiology of human language. *Ann. NY Acad. Sci.* **1016**, 749–777 (2004).
75. Scharff, C. & White, S. A. Genetic components of vocal learning. *Ann. NY Acad. Sci.* **1016**, 325–347 (2004).
References 73–75 demonstrate the promise of using contemporary molecular methods to shed light on vocal-learning mechanisms in song-birds, and discuss the relevance of this strategy for understanding human language.
76. Gilbert, S. L., Dobyns, W. B. & Lahn, B. T. Genetic links between brain development and brain evolution. *Nature Rev. Genet.* **6**, 581–590 (2005).
77. Woods, C. G., Bond, J. & Enard, W. Autosomal recessive primary microcephaly (MCPH): a review of clinical, molecular, and evolutionary findings. *Am. J. Hum. Genet.* **76**, 717–728 (2005).
78. Zhang, J. Evolution of the human *ASPM* gene, a major determinant of brain size. *Genetics* **165**, 2063–2070 (2003).
79. Evans, P. D. *et al.* Adaptive evolution of *ASPM*, a major determinant of cerebral cortical size in humans. *Hum. Mol. Genet.* **13**, 489–494 (2004).
80. Kouprina, N. *et al.* Accelerated evolution of the *ASPM* gene controlling brain size begins prior to human brain expansion. *PLoS Biol.* **2**, e126 (2004).
81. Evans, P. D., Anderson, J. R., Vallender, E. J., Choi, S. S. & Lahn, B. T. Reconstructing the evolutionary history of microcephalin, a gene controlling human brain size. *Hum. Mol. Genet.* **13**, 1139–1145 (2004).
82. Wang, Y. Q. & Su, B. Molecular evolution of microcephalin, a gene determining human brain size. *Hum. Mol. Genet.* **13**, 1131–1137 (2004).
83. Ferland, R. J. *et al.* Abnormal cerebellar development and axonal decussation due to mutations in *AHI1* in Joubert syndrome. *Nature Genet.* **36**, 1008–1013 (2004).
84. Piao, X. *et al.* G protein-coupled receptor-dependent development of human frontal cortex. *Science* **303**, 2033–2036 (2004).
85. Bishop, D. V. Genetic and environmental risks for specific language impairment in children. *Philos. Trans. R. Soc. Lond. B* **356**, 369–380 (2001).
86. Bishop, D. V., North, T. & Donlan, C. Nonword repetition as a behavioural marker for inherited language impairment: evidence from a twin study. *J. Child Psychol. Psychiatry* **37**, 391–403 (1996).
87. Bartlett, C. W. *et al.* Examination of potential overlap in autism and language loci on chromosomes 2, 7, and 13 in two independent samples ascertained for specific language impairment. *Hum. Hered.* **57**, 10–20 (2004).
88. The SLI Consortium. Highly significant linkage to the *SLI1* locus in an expanded sample of individuals affected by specific language impairment. *Am. J. Hum. Genet.* **74**, 1225–1238 (2004).
89. Stein, C. M. *et al.* Pleiotropic effects of a chromosome 3 locus on speech-sound disorder and reading. *Am. J. Hum. Genet.* **74**, 283–297 (2004).
90. Fisher, S. E. & DeFries, J. C. Developmental dyslexia: genetic dissection of a complex cognitive trait. *Nature Rev. Neurosci.* **3**, 767–780 (2002).
91. Wassink, T. H., Brzustowicz, L. M., Bartlett, C. W. & Szatmari, P. The search for autism disease genes. *Ment. Retard. Dev. Disabil. Res. Rev.* **10**, 272–283 (2004).
92. Demonet, J. F., Taylor, M. J. & Chaix, Y. Developmental dyslexia. *Lancet* **363**, 1451–1460 (2004).
93. Snowling, M. J. From language to reading and dyslexia. *Dyslexia* **7**, 37–46 (2001).
94. Taipale, M. *et al.* A candidate gene for developmental dyslexia encodes a nuclear tetratricopeptide repeat domain protein dynamically regulated in brain. *Proc. Natl Acad. Sci. USA* **100**, 11553–11558 (2003).
95. Scerri, T. S. *et al.* Putative functional alleles of *DYX1C1* are not associated with dyslexia susceptibility in a large sample of sibling pairs from the UK. *J. Med. Genet.* **41**, 853–857 (2004).
96. Francks, C. *et al.* A 77-kilobase region of chromosome 6p22.2 is associated with dyslexia in families from the United Kingdom and from the United States. *Am. J. Hum. Genet.* **75**, 1046–1058 (2004).
97. Cope, N. *et al.* Strong evidence that *KIAA0319* on chromosome 6p is a susceptibility gene for developmental dyslexia. *Am. J. Hum. Genet.* **76**, 581–591 (2005).
98. Tager-Flusberg, H., Joseph, R. & Folstein, S. Current directions in research on autism. *Ment. Retard. Dev. Disabil. Res. Rev.* **7**, 21–29 (2001).
99. Bradford, Y. *et al.* Incorporating language phenotypes strengthens evidence of linkage to autism. *Am. J. Med. Genet.* **105**, 539–547 (2001).
100. Alarcón, M., Yonan, A. L., Gilliam, T. C., Cantor, R. M. & Geschwind, D. H. Quantitative genome scan and ordered-subsets analysis of autism endophenotypes support language QTLs. *Mol. Psychiatry* **10**, 747–757 (2005).
101. Hurst, J. A., Baraitser, M., Auger, E., Graham, F. & Norell, S. An extended family with a dominantly inherited speech disorder. *Dev. Med. Child Neurol.* **32**, 352–355 (1990).
102. Lai, C. S., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F. & Monaco, A. P. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* **413**, 519–523 (2001).
The first demonstration that specific genetic mutations (a missense mutation and a translocation, both involving *FOXP2*) can lead to a developmental disruption of speech and language acquisition.
103. Vargha-Khadem, F. *et al.* Neural basis of an inherited speech and language disorder. *Proc. Natl Acad. Sci. USA* **95**, 12695–12700 (1998).
104. Gopnik, M. & Crago, M. B. Familial aggregation of a developmental language disorder. *Cognition* **39**, 1–50 (1991).
105. Watkins, K. E., Dronkers, N. F. & Vargha-Khadem, F. Behavioural analysis of an inherited speech and language disorder: comparison with acquired aphasia. *Brain* **125**, 452–464 (2002).
106. Marcus, G. F. & Fisher, S. E. *FOXP2* in focus: what can genes tell us about speech and language? *Trends Cogn. Sci.* **7**, 257–262 (2003).
107. Vargha-Khadem, F., Gadian, D. G., Copp, A. & Mishkin, M. *FOXP2* and the neuroanatomy of speech and language. *Nature Rev. Neurosci.* **6**, 131–138 (2005).
108. Fisher, S. E., Vargha-Khadem, F., Watkins, K. E., Monaco, A. P. & Pembrey, M. E. Localisation of a gene implicated in a severe speech and language disorder. *Nature Genet.* **18**, 168–170 (1998).
109. Lai, C. S. *et al.* The *SPCH1* region on human 7q31: genomic characterization of the critical interval and localization of translocations associated with speech and language disorder. *Am. J. Hum. Genet.* **67**, 357–368 (2000).
110. Newbury, D. F. *et al. FOXP2* is not a major susceptibility gene for autism or specific language impairment. *Am. J. Hum. Genet.* **70**, 1318–1327 (2002).
111. MacDermot, K. D. *et al.* Identification of FOXP2 truncation as a novel cause of developmental speech and language deficits. *Am. J. Hum. Genet.* **76**, 1074–1080 (2005).
112. Sarda, P. *et al.* [Interstitial deletion in the long arm of chromosome 7]. *Ann. Genet.* **31**, 258–261 (1988) (in French).

113. Enard, W. *et al.* Molecular evolution of *FOXP2*, a gene involved in speech and language. *Nature* **418**, 869–872 (2002).
114. Zhang, J., Webb, D. M. & Podlaha, O. Accelerated protein evolution and origins of human-specific features: FOXP2 as an example. *Genetics* **162**, 1825–1835 (2002).
Driven by the finding that *FOXP2* disruption is implicated in a human speech and language disorder, references 113 and 114 describe two independent targeted examinations of the primate evolution of this gene, each reaching the conclusion that *FOXP2* was subject to positive selection in recent human history.
115. Webb, D. M. & Zhang, J. *FoxP2* in song-learning birds and vocal-learning mammals. *J. Hered.* **96**, 212–216 (2005).
116. Belton, E., Salmond, C. H., Watkins, K. E., Vargha-Khadem, F. & Gadian, D. G. Bilateral brain abnormalities associated with dominantly inherited verbal and orofacial dyspraxia. *Hum. Brain Mapp.* **18**, 194–200 (2003).
117. Liégeois, F. *et al.* Language fMRI abnormalities associated with *FOXP2* gene mutation. *Nature Neurosci.* **6**, 1230–1237 (2003).
References 116 and 117 provide essential insights into the consequences of *FOXP2* disruption for the structure and function of the adult human brain, uncovering a distributed pathology that involves cortical and subcortical structures.
118. Lai, C. S., Gerrelli, D., Monaco, A. P., Fisher, S. E. & Copp, A. J. *FOXP2* expression during brain development coincides with adult sites of pathology in a severe speech and language disorder. *Brain* **126**, 2455–2462 (2003).
This study revealed striking overlaps between the sites of *FOXP2* expression in the human fetal brain and known sites of anomaly in a *FOXP2*-related disorder, previously highlighted by structural or functional neuroimaging. There was also remarkable concordance between the human neural expression pattern and that observed for rodent *Foxp2*, described here and in references 119 and 120.
119. Ferland, R. J., Cherry, T. J., Preware, P. O., Morrisey, E. E. & Walsh, C. A. Characterization of Foxp2 and Foxp1 mRNA and protein in the developing and mature brain. *J. Comp. Neurol.* **460**, 266–279 (2003).
120. Takahashi, K., Liu, F. C., Hirokawa, K. & Takahashi, H. Expression of *Foxp2*, a gene involved in speech and language, in the developing and adult striatum. *J. Neurosci. Res.* **73**, 61–72 (2003).
121. Haesler, S. *et al. FoxP2* expression in avian vocal learners and non-learners. *J. Neurosci.* **24**, 3164–3175 (2004).
122. Teramitsu, I., Kudo, L. C., London, S. E., Geschwind, D. H. & White, S. A. Parallel *FoxP1* and *FoxP2* expression in songbird and human brain predicts functional interaction. *J. Neurosci.* **24**, 3152–3163 (2004).
The molecular studies of song-bird brains in references 121 and 122 support the intriguing hypothesis that regulation of *FoxP1* and *FoxP2* in the song nuclei of vocal-learning species might be related to vocal plasticity.
123. Bonkowsky, J. L. & Chien, C. B. Molecular cloning and developmental expression of *foxP2* in zebrafish. *Dev. Dyn.* **234**, 740–746 (2005).
124. Wang, B., Lin, D., Li, C. & Tucker, P. Multiple domains define the expression and regulatory properties of Foxp1 forkhead transcriptional repressors. *J. Biol. Chem.* **278**, 24259–24268 (2003).
125. Shu, W., Yang, H., Zhang, L., Lu, M. M. & Morrisey, E. E. Characterization of a new subfamily of winged-helix/forkhead (Fox) genes that are expressed in the lung and act as transcriptional repressors. *J. Biol. Chem.* **276**, 27488–27497 (2001).
126. Ehret, G. Infant rodent ultrasounds — a gate to the understanding of sound communication. *Behav. Genet.* **35**, 19–29 (2005).
127. Holy, T. E & Guo Z. Ultrasonic songs of male mice. *PLoS Biol.* **3**, e386 (2005).
This is the first description of the properties of mouse song, which demonstrates unexpected similarities with bird song. These findings indicate that the neural underpinnings of song production and perception might be studied using the well-established techniques of genetic manipulation that are available for the mouse.
128. Shu, W. *et al.* Altered ultrasonic vocalization in mice with a disruption in the *Foxp2* gene. *Proc. Natl Acad. Sci. USA* **102**, 9643–9648 (2005).
129. Kuhl, P. K. Early language acquisition: cracking the speech code. *Nature Rev. Neurosci.* **5**, 831–843 (2004).
130. Crosson, B. *Subcortical Functions in Language and Memory* (Guilford, New York, 1992).
131. Jarvis, E. D. *et al.* Avian brains and a new understanding of vertebrate brain evolution. *Nature Rev. Neurosci.* **6**, 151–159 (2005).
132. Clark, A. G. *et al.* Inferring nonneutral evolution from human–chimp–mouse orthologous gene trios. *Science* **302**, 1960–1963 (2003).
133. McDonald, J. H. & Kreitman, M. Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature* **351**, 652–654 (1991).
134. Groszer, M. *et al.* Negative regulation of neural stem/progenitor cell proliferation by the *Pten* tumor suppressor gene *in vivo*. *Science* **294**, 2186–2189 (2001).
135. Chenn, A. & Walsh, C. A. Regulation of cerebral cortical size by control of cell cycle exit in neural precursors. *Science* **297**, 365–369 (2002).

Acknowledgements
We thank W. Enard for his helpful comments on this manuscript, and C. S. Lai and F. J. Liégeois for assistance with the figures. S.E.F. is a Royal Society Research Fellow, and his research is also supported by the Wellcome Trust, the 6th European Community Framework Programme and the Brain Sciences Initiative of the UK Medical Research Council. G.F.M. thanks the US National Institutes of Health and the Human Frontier Science Program for financial support.

Competing interests statement
The authors declare that they have no competing financial interests.

DATABASES
The following terms in this article are linked online to:
Entrez Gene: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene
ASPM | *FOXP2* | *LMO4* | *MCPH1*
OMIM: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM
BFPP | developmental verbal dyspraxia | JBTS3 | primary microcephaly | *SLI1* | *SLI2* | *SLI3* | SSD

FURTHER INFORMATION
Gary Marcus's web site: http://psych.nyu.edu/gary
Simon Fisher's web site: http://www.well.ox.ac.uk/~simon
Access to this interactive links box is free online.