The evolution of Fox genes and their role in development and disease

Sridhar Hannenhalli[*‡] *and Klaus H. Kaestner*[*§]

Abstract | The forkhead box (Fox) family of transcription factors, which originated in unicellular eukaryotes, has expanded over time through multiple duplication events, and sometimes through gene loss, to over 40 members in mammals. Fox genes have evolved to acquire a specialized function in many key biological processes. Mutations in Fox genes have a profound effect on human disease, causing phenotypes as varied as cancer, glaucoma and language disorders. We summarize the salient features of the evolution of the Fox gene family and highlight the diverse contribution of various Fox subfamilies to developmental processes, from organogenesis to speech acquisition.

The forkhead box, or Fox, gene family of transcriptional regulators is an evolutionarily ancient gene family that is named after the *Drosophila melanogaster* gene fork head (*fkh*). Mutations in *fkh* cause defects in head fold involution during embryogenesis, resulting in a characteristic spiked head appearance in adult flies[1]. Hundreds of Fox genes have been identified in species ranging from yeasts to humans, and have been classified into subfamilies, such as FoxA and FoxP. Genetic analyses have shown that many of these genes have important biological functions in multiple species, from control of the cell cycle to differentiation of epithelia, and from placental development to formation of the inner ear. The evolutionary conservation of the crucial DNA-binding domain between orthologous members of the Fox gene family is remarkable; for example, there is 90% amino acid similarity between the *D. melanogaster* Fork head and the human FOXA1 protein. Several Fox genes are mutated in human disease, with phenotypes ranging from defective T cell differentiation to speech impediments[2,3]. Recent findings on the contribution of FOXA1-mediated gene regulation in breast and prostate cancer (see below) further show the large contribution of this gene family to human health[4,5].

In this Review, we first describe the evolution and phylogeny of this fascinating gene family. It is impossible to review in detail the contribution of all of the Fox genes to development and function, and for a short summary sketch on the entire gene family we refer the reader to two recent 'snapshots' (REFS 2,3). However, to indicate the breadth of function of the Fox gene family in more detail, we focus on three Fox classes — FoxO, FoxA and FoxP — because each of these classes shows a unique and important aspect of the diverse biology of this gene family. Over the past 20 years, genetic studies in organisms ranging from flies to humans have informed us of the essential biological functions of these representative Fox genes. Although in many cases the molecular details of how these factors select and control their targets, as well as the upstream regulation of these factors, remain unknown, their mutational analysis is nearing completion and we have a reasonable understanding of the biological pathways and disease processes that are controlled by these gene families.

Discovery

The unifying feature of Fox proteins is the ~100-residue forkhead (FKH) DNA-binding domain, which is highly conserved across all members of the Fox family (FIGS 1,2). A remarkable convergence of scientific discoveries on two continents and in two species led to the identification of the first members of this family of DNA-binding proteins 20 years ago. In 1989, the gene responsible for the *fork head* phenotype mentioned above was cloned in *D. melanogaster*, but the sequence showed no resemblance to any protein motif that was known at the time[1]. Shortly afterwards, the cDNA-encoding hepatocyte nuclear factor 3α (HNF3α), now known as FOXA1, was cloned in the rat[6]. Again, the sequence of FOXA1 showed no similarity to the DNA-binding motifs that were known at the time, such as those of nuclear receptors or zinc-finger transcription factors. However, Lai and colleagues[6] delimited the DNA-binding region of FOXA1 to amino acids 124–288 of the protein. In 1990, Weigle and Jäckle[7] noticed the astonishing similarity of the central 110 amino acids of the *D. melanogaster* and mammalian Fox proteins, which are completely contained in the

*Department of Genetics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104, USA.
‡Penn Center for Bioinformatics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104, USA.
§Institute for Diabetes, Obesity and Metabolism, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104, USA.
e-mails: sridharh@pcbi.upenn.edu; Kaestner@mail.med.upenn.edu
doi:10.1038/nrg2523
Published online
10 March 2009

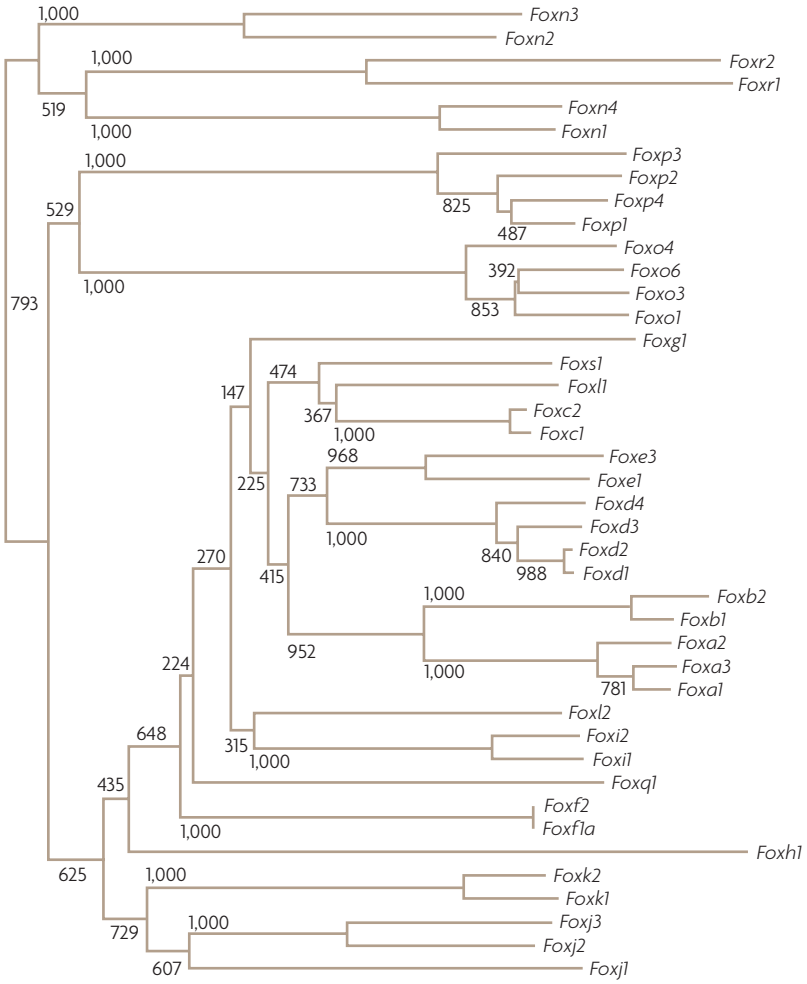

Figure 1 | **Evolutionary tree of mouse forkhead box (Fox) genes.** A neighbour-joining tree is shown that is based on the protein sequence of the forkhead domain. The relationships shown in the tree are based on multiple alignment (FIG. 2), using ClustalX as the alignment tool. Each branch is annotated with a bootstrap value that was based on 1,000 samples. The branch lengths are proportional to the mutation rate. For a recent phylogenetic analysis of Fox genes, see REFS 13,14,22.

DNA-binding domain identified by Lai and colleagues[6], and suggested the name forkhead domain for this characteristic DNA-binding motif.

The canonical FKH domain consists of three α-helices, three β-sheets and two 'wing' regions that flank the third β-sheet. The structure of the DNA-bound FKH domain has been solved for several Fox proteins: FOXA1 (REF. 8), FOXA3 (REF. 9), FOXD3 (REF. 10), FOXK1 (REF. 11) and FOXP2 (REF. 12). Because of the butterfly-like winged structure adopted by the DNA-bound Fox proteins, the FKH domain has also been termed the winged-helix domain; however, the winged-helix structure is not unique to Fox proteins. Most Fox proteins bind to DNA as monomers, contacting their target sequences by the third α-helix, and by flanking residues and the two wings.

Classification and phylogeny

Since the cloning of *fkh* in *D. melanogaster*, hundreds of Fox genes have been discovered in numerous species, and were assigned a confusing range of names by their discoverers. To provide a system for the study of Fox genes, in 2000, the Fox nomenclature committee proposed a classification system based on a phylogenetic analysis of 172 Fox proteins in 14 species (including multiple isoforms for some proteins)[13]. Based on similarities in the FKH domain, Fox proteins were divided into 15 classes from FoxA to FoxO. The phylogenetic tree has since been expanded and four more classes — from FoxP to FoxS — have been catalogued (see Further information for a link to the Index of winged-helix proteins). Thus, 19 clades of Fox proteins are now known. In a recent analysis that focused on the early expansion of the Fox gene family, *Foxj1* was inferred to be present in the most recent common ancestor of fungi and metazoans[14].

The overall classification scheme has remained generally robust against repeated phylogenetic analysis by different authors working on different sets of proteins and species[13,15–21]. However, there are some exceptions. Some researchers have proposed to rename FOXR1 and FOXR2 as FOXN5 and FOXN6, respectively[15], and FOXL1 and FOXL2 are not grouped together in several phylogenetic analyses[15,21]. Similarly, some analyses separate FOXJ1 from FOXJ2 and FOXJ3, and FOXN1 and FOXN4 from FOXN2 and FOXN3 (REF. 14). Moreover, certain species-specific Fox genes that could not be unambiguously assigned to a pre-existing class have been assigned to their own class, such as FoxX and FoxY in sea urchins[21]. The FKH domain is the only part of the Fox peptide sequence that can be confidently aligned across all classes; therefore, phylogenetic inferences are made solely on the basis of this short domain. The precise phylogenetic relationship can sometimes be resolved by carrying out a more detailed analysis of a subset of proteins, such as for FoxA, FoxB and other unresolved FoxA- or FoxB-like proteins in sea urchins[21], or by using peptide regions outside the FKH domain. For more comprehensive evolutionary analyses of Fox genes from multiple species, we refer the reader to recent articles[13,14,22].

Evolution

Functional diversity. Despite the similarity in their DNA-binding domains, and consequently their DNA-recognition motifs, various Fox proteins have evolved distinct roles. Fox protein sequences share little more sequence similarity than the FKH domain. When gene families expand through duplication events, newly created members often evolve and acquire distinct functions[23,24]. However, overlapping and redundant functions can be maintained over long evolutionary periods; for example, FOXA1 and FOXA2 cooperate during liver and lung morphogenesis[25,26] (discussed in detail below). The functional diversity among Fox proteins is achieved partially through differences in interaction partners, such as modifying enzymes and cofactors, and partially through differences in the spatio-temporal expression patterns of the Fox genes. An example of spatio-temporal regulation of a Fox gene is the expression and function of *Foxa3* in testes, as the deletion of *Foxa3* results in subfertility in males, which is a clear evolutionary disadvantage[27]. In general, paralogous transcription factors with a similar DNA recognition motif tend to have divergent

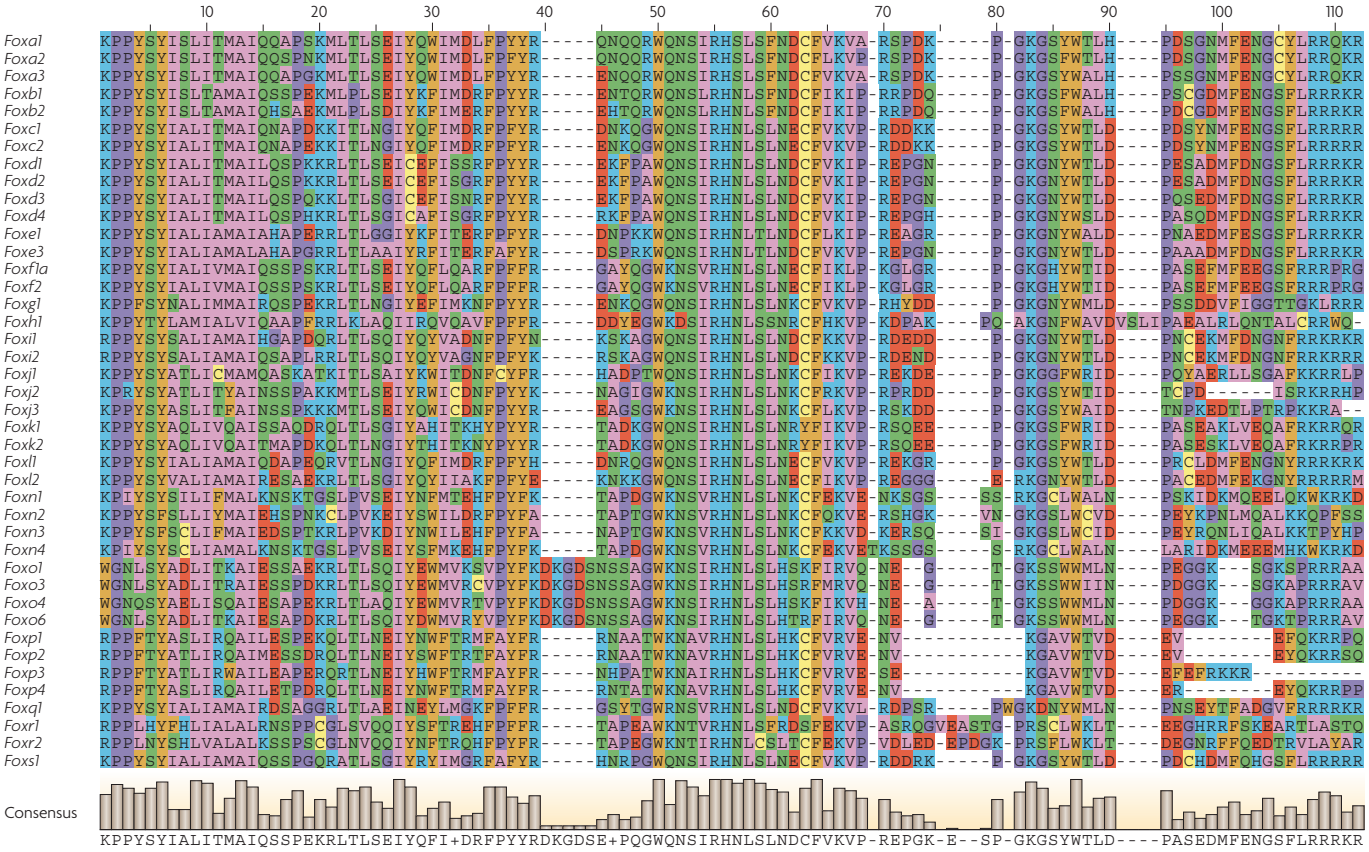

Figure 2 | **Alignment of forkhead box (Fox) genes in mice.** All protein sequences that contain the forkhead (FKH) domain (Pfam identification number PF00250) were extracted from ENSEMBL v50 (see Further information), and the longest isoform for each gene was used. The *Foxl1* sequence (RefSeq identification number NP_032050) was obtained from the National Center for Biotechnology Information (NCBI). The protein sequences were aligned using the multiple alignment tool T-Coffee. The alignment was viewed using the alignment editor Jalview. The bottom panel shows the FKH domain as recorded in the Pfam database. The FKH domain is the only consistently conserved portion of the protein across all members of the family, whereas there are limited similarities in other regions among Fox subfamilies. The colour coding of the amino acids is based on the physicochemical properties provided in Jalview.

expression domains, such that in a particular cellular context a maximum of one of the paralogues is expressed at a high level[28]. Although many experimental studies have shown the precise functions of a subset of Fox proteins in several species, a broader understanding of the evolution and functional diversification of the Fox family since the rise of eukaryotes approximately 1 billion years ago remains an important long-term goal.

Selective constraints. Because of their crucial role in development and other key biological processes, Fox genes are likely to be under strict evolutionary constraints, either in their coding sequence or in their genomic location and *cis*-regulatory elements. A standard method to assess whether a protein-coding gene has evolved under positive selection or purifying selection is to compare the non-synonymous (K_A) and synonymous (K_S) mutation rates using the K_A/K_S ratio test. A K_A/K_S of ~1 is indicative of a neutrally evolving sequence, a K_A/K_S of <1 suggests that evolution is under purifying selection and a K_A/K_S of >1 suggests that positive selection is occurring. For example, *Foxl2*, which is involved in ovarian development, was shown to be evolving under purifying selection[29]. *FOXP2*, which is discussed in detail below, was the first gene that was identified to be relevant to human speech. Human-specific changes in the amino-acid sequence of FOXP2, as well as the polymorphism pattern, and cosegregation of a point mutation in *FOXP2* in a family with speech impairment, are consistent with adaptive evolution of *FOXP2* during human evolution[30]. However, *Foxg1*, a potential determinant of forebrain size in vertebrates, shows no evidence of evolution under selection[31]. A global analysis of the evolution of the entire Fox gene family across several species using a different technique that uses site-specific shifts in substitution rates across clades found specific sites in the FKH domain that seemed to underlie the selective divergence between Fox subclasses[22].

Genomic organization. Functionally related genes with coordinated regulation are sometimes organized into evolutionarily conserved clusters in the genome: a prime example is the homeobox (Hox) gene cluster[32]. Although no such global clustering exists for the Fox genes, one cluster that includes *Foxl1*, *Foxc1*, *Foxf2* and *Foxq1* has

Positive selection
An evolutionary process by which beneficial alleles (alleles that result in increased fitness of the organism) become more frequent in a population.

Purifying selection
An evolutionary process by which deleterious alleles (alleles that result in reduced fitness of the organism) become less frequent in the population, thereby making sequences in which this process occurs more similar compared with those from different species or from individuals of the same species.

Box 1 | **Fox genes in human disease**

Although the Fox gene family was first discovered in model organisms, work over the past 20 years has shown the importance of this gene family in human diseases. The Online Inheritance in Man (OMIM) database lists four Fox genes that when mutated cause human disease. These are *FOXC1*, *FOXC2*, *FOXP2* and *FOXP3*.

A spectrum of dominantly inherited glaucoma phenotypes, including iris hypoplasia and Rieger syndrome, is caused by mutations in *FOXC1* (REF. 90). Mutations in the related *FOXC2* gene cause lymphedema–distichiasis syndrome, an autosomal dominant disorder that presents with lymphedema of the limbs, with variable age at onset, and double rows of eyelashes[91]. Mutations in *FOXP2* lead to language acquisition defects[79] (discussed in detail in the main text). Interestingly, human FOXP2 differs from chimpanzee and gorilla FOXP2 by only two amino acids; this creates a potential protein kinase C phosphorylation site, which might be important for the adaptation needed for speech development[30]. *FOXP3* has become prominent because it is a marker of $CD25^{+}CD4^{+}$ regulatory T cells. Mutations in *FOXP3* cause X-linked immunodysregulation, polyendocrinopathy and enteropathy (IPEX) in humans and the equivalent 'scurfy' phenotype in mice[92,93].

been maintained since the period of the early bilaterians to the present day in flies, mammals and amphioxus[19]. This cluster is likely to have been duplicated in vertebrates, with the subsequent loss of *Foxq1* from one cluster and *Foxl1* from the other. *Foxl1*, *Foxc1*, *Foxf2* and *Foxq1* have well-defined sequential expression patterns during embryonic development in amphioxus endomesoderm, and it has been suggested that their genomic clustering has been selectively maintained during 500 million years of evolution. Although the molecular mechanisms of the coordinate regulation of these genes have not been worked out, it is clear that at least FOXL1 and FOXF2 are targets of hedgehog signalling, as there are multiple well-conserved glioma-associated (Gli) binding sites in their *cis*-regulatory domains[33]. Interestingly, the regulation of Fox genes by hedgehog signalling was first established for *Foxf1*, a close relative of *Foxf2* that is located on a different chromosome[34].

Because of the size and breadth of function of the Fox gene family, it is impossible to cover all of the classes in detail. In the following sections, we describe important aspects of three Fox classes — FoxO, FoxA and FoxP — that provide examples of the unique aspects of the diverse biology of the Fox gene family. FoxO has become prominent because FoxO proteins were discovered to be central mediators in the insulin signalling system, which is important in human health (BOX 1) and the determination of life span, at least in invertebrates. The FoxA class was selected because it contains the founding members of the whole gene family, fork head and *HNF3*, and because it regulates processes from earliest organ development to nuclear hormone receptor action in breast cancer. Finally, we selected the FoxP class because of its unique and fascinating biological role in language acquisition.

FoxO: insulin signalling, diabetes and ageing

Glucose homeostasis. The discovery that the *Caenorhabditis elegans* orthologue of mammalian FoxO proteins, DAF-16, is an essential mediator of the effects of insulin on longevity stimulated research on the mammalian FoxO genes as the final executors of the transcriptional repression mediated by insulin or insulin-like growth factor 1 (IGF1)[35–41]. Although the physiological role of insulin signalling in nematodes and mammals is different, important components of this signalling pathway, including the FoxO proteins, have been conserved. In mammals, insulin action has traditionally been thought to primarily promote glucose uptake into adipose tissue and muscle. More recently, the ability of insulin to suppress hepatic glucose output has been recognized as an additional essential function of this hormone[42]. Insulin regulates the transcriptional activity of hundreds of genes involved in glucose and lipid metabolism in the liver. The repression of genes encoding gluconeogenic enzymes was the first paradigm of insulin-mediated transcriptional activity and led to the identification of the negatively acting insulin response sequence (IRS) with the consensus T(G/A)TTT(T/G)(G/T), which is strikingly similar to the FoxA binding site[43]. *In vitro* DNA-binding assays and transfection experiments showed that both mammalian FoxO and FoxA proteins can bind to this IRS and mediate transcriptional activation[44–47]. In mammalian systems, insulin action and the FoxO proteins are linked by the mechanism shown in FIG. 3.

Although they are normally nuclear, all three FoxO proteins (formerly known as FKHR, FKHRL1 and AFX) are targets for phosphorylation by Akt (also known as protein kinase B) kinases, leading to nuclear exclusion and loss of DNA binding. Therefore, an elegant mechanism for insulin-dependent gene inhibition exists (FIG. 3): binding of insulin to its receptor on the cell surface initiates phosphatidylinositol 3-kinase (PI3K) and Akt activation, followed by FoxO phosphorylation and nuclear exclusion. This nuclear exclusion then results in the deactivation of FoxO targets, which in the liver includes the genes that encode gluconeogenic enzymes. Haploinsufficiency for *Foxo1* restores insulin sensitivity in insulin-resistant mice by reducing hepatic expression of gluconeogenic enzymes, and therefore this model is supported by genetic evidence[48]. Furthermore, when *Foxo1* was conditionally ablated in mouse hepatocytes, stimulation of hepatic glucose production by both glycogenolysis and gluconeogenesis was impaired[49]. Ablation of *Foxo1* in hepatocytes could also reverse most of the metabolic effects of simultaneous ablation of *Irs1* and *Irs2*, which are major mediators of insulin and IGF1-receptor signalling[50].

Cancer. In addition to their role in the regulation of glucose homeostasis, FoxO genes also have a crucial role in human cancer. Fusion proteins of FOXO1 and PAX3 or PAX7 are found in rhabdomyosarcoma[51], and fusion proteins of FOXO3 or FOXO4 with the *MLL* (mixed lineage leukaemia) gene are found in acute lymphocytic leukaemia[52] (BOX 2). In these fusion proteins, juxtaposition of the strong transcriptional activation domains of the FoxO proteins with the DNA-binding domains of the Pax proteins results in the inappropriate activation of a whole battery of proliferative genes and, ultimately, cancer[53]. Interestingly, however, the level of fusion protein expression also matters; if expression is too high, growth suppression ensues[54]. In addition, in genetic studies in mice, ablation of *Foxo1*, *Foxo3* and *Foxo4* led to a tumour-prone condition and the development of thymic lymphomas and haemangiomas, indicating that the FoxO proteins

Gluconeogenic
This term describes processes that relate to gluconeogenesis, the process of synthesizing glucose from amino acids and glucose in the liver and kidney in response to fasting.

Haploinsufficiency
A condition in a diploid organism in which a single functional copy of a gene results in a phenotype such as a disease. In this case, having only 50% of gene function (50% of the protein levels present in the wild-type state) is not sufficient to fulfil the needs of the cells or the organism.

Glycogenolysis
The breakdown of glycogen that occurs during fasting. Glycogen is the storage form of carbohydrates used in animals and is broken down to liberate glucose during times of fasting.

are tumour suppressors[45]. Therefore, the FoxO proteins are also major mediators of the activation of the PI3K and Akt signalling pathways in cancer.

FoxA proteins as pioneer factors

The *D. melanogaster* gene *fkh* was found by positional cloning of a mutation that caused homeotic transformations of the ectodermal portions of the foregut and hindgut[1]. On the basis of their expression pattern and the similarity of the null phenotypes to the *fork head* phenotype, it was suggested that the FoxA genes, the orthologues of *fkh* in mammals, have a conserved role in the development of the derivatives of the primitive gut[55]. However, studies of the expression patterns of these genes, as well as gene targeting experiments, in mice have shown that the FoxA clade functions in many developmental and physiological processes in addition to gut development. The phenotype of *Foxa2*-null mice is particularly striking: they die shortly after gastrulation and lack the embryonic structure of the node, which is essential for the proper development of the notochord and the patterning of the neural tube[56,57].

Chromatin remodelling. The winged-helix structure of the FKH domain is not unique to the Fox genes. In fact, when the three-dimensional structure of the FKH domain was first solved, it was noted that it resembled the structure of the linker histone H1 (REF. 9). The importance of this relatedness is shown by the binding of FoxA proteins, at least *in vitro*, to highly compacted chromatin, even in the absence of the SWI–SNF chromatin remodelling complex[58]. The mechanism of this activity is almost certainly due to the ability of the carboxy-terminal domain of FoxA to interact with the core histones H3 and H4 (REF. 58). Consequently, the FoxA proteins have been proposed to act as 'pioneer' transcription factors, displacing linker histones from compacted chromatin and facilitating the binding of other transcription factors. Genetic evidence for such a role is provided by the observation that mouse foregut endoderm that lacks *Foxa1* and *Foxa2* loses its ability to respond to inductive signals to activate the hepatic transcriptional programme[25].

Binding partners of nuclear receptors. Several genetic and functional genomics studies have expanded our understanding of the role of the FoxA proteins by showing that they are also pioneer factors for nuclear hormone receptors in various settings. For example, target occupancy of the glucocorticoid receptor, which is activated during fasting, depends partially on the presence of FOXA2 (REF. 59). Similarly, androgen response elements and binding sites for FOXA1 and/or FOXA2 are frequently co-localized in the *cis*-regulatory elements of prostate-specific genes, and the androgen receptor and the FoxA protein physically interact[60–63]. Recent studies on the large-scale identification of oestrogen receptor binding sites showed that there was a striking proximity between the locations of oestrogen response elements and binding sites for FOXA1 (REFS 4,5). Additional *in vitro* studies confirmed that at least part of the oestrogen response in cultured breast cancer cell lines is FOXA1-dependent[4].

Therefore, a new paradigm is emerging in which the binding of FoxA-type transcription factors in the vicinity of a response element for a nuclear hormone receptor is important for hormone receptor binding. It remains to be explored whether FOXA1–3, and possibly other Fox proteins, enable or facilitate the binding of additional nuclear hormone receptors or other transcription factors in other tissues or physiological states. It is also not known whether nuclear hormone receptor binding is mediated by alterations in chromatin structure or by another mechanism.

Mammalian FoxA genes: overlapping and unique functions in organ development. The functions of the FoxA subclass in mammalian development and physiology have been studied extensively using genetic approaches (for a detailed review, see REF. 64). The general principle that has emerged from these studies is that individual FoxA genes have taken on distinct roles by diversifying their expression patterns (for example, the unique functions of *Foxa2* in the embryonic node[56,57] or in bile acid homeostasis in the liver[65] and of *Foxa3* in male fertility[27]). However, it is also clear that for their principal developmental functions, *Foxa1* and *Foxa2* have retained overlapping functions. This has been shown most clearly for liver and lung development.

Foxa1 and *Foxa2*, but not *Foxa3*, are expressed in the endoderm domain that gives rise to the lung buds, where they continue to be expressed in the pulmonary epithelium until adulthood[55,66–69]. Although the lungs of *Foxa1*-null mice seem normal and are functional at birth, lung development is delayed during fetal life in these mice[70–72]. Similarly,

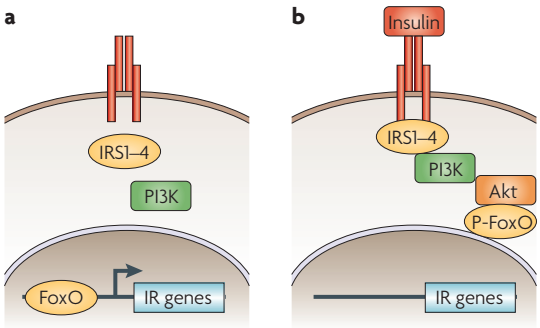

Figure 3 | **The insulin, Akt and FoxO pathway. a** | In the absence of insulin binding, FoxO is a transcriptional activator of multiple insulin-responsive (IR) genes. **b** | Engagement of the phosphatidylinositol 3-kinase (PI3K) pathway in response to insulin or other factors leads to the activation of Akt and to the phosphorylation (P) of FoxO proteins on crucial serine and threonine residues (Thr24, Ser256 and Ser319 in FOXO1)[45–47]. Phosphorylation seems to be sequential and Ser256 is phosphorylated first. Ser256 phosphorylation is sufficient for reduced DNA- binding affinity, presumably because this residue is located in the basic region of the DNA-binding domain, but nuclear exclusion also requires phosphorylation of the other two residues, Thr24 and Ser319 (REF. 54). Therefore, an elegant mechanism for insulin-dependent gene inhibition exists: insulin binding to its receptor on the cell surface initiates PI3K and Akt activation, followed by FoxO phosphorylation, nuclear exclusion and loss of transcriptional activation.

Linker histone
The major protein component of chromatin. Linker histones allow compaction of DNA, but also play important parts in gene regulation. The linker histone H1 binds to the DNA strands as they emerge from the nucleosome particle, which is assembled by the core histones.

Box 2 | **Forkhead box (Fox) genes in cancer**

Several Fox gene family members play important parts in the aetiology of cancer. The first evidence for this causative relationship was provided by the discovery by Vogt and colleagues that the oncogene of avian sarcoma virus 31, originally termed '*qin*', contains a forkhead-type winged-helix domain and is closely related to the mammalian *Foxg1* gene[94]. Around the same time, Barr and colleagues found that fusions of FOXO1 with PAX3 led to alveolar rhabdomyosarcoma[51]. In these fusion proteins, an artificial transcription factor is created that combines the strong transcriptional activation domain of FOXO1 with the DNA-binding domains of PAX3, leading to inappropriate activation of growth-promoting genes. In contrast to this 'neomorphic' function of the FoxO and Pax fusion proteins, FOXM1 has been shown to have pro-proliferative function on its own. *FOXM1* is expressed in all cycling cells, and is therefore normally found at high levels in the intestine, thymus and testes[95]. However, FOXM1 is activated when quiescent cells re-enter the cell cycle; for example, after partial hepatectomy. FOXM1 directly activates multiple cyclin, cyclin-dependent kinase and anti-apoptotic genes. As might be expected, *FOXM1* expression is upregulated in dozens of human cancers, and high FOXM1 levels correlate with poor prognosis[96]. Because FOXM1 is inhibited by cyclin-dependent kinase inhibitor 2A (CDKN2A; also known as ARF), a cell-penetrating CDKN2A peptide has been proposed as an anti-tumour therapeutic[97].

mice in which the *Foxa2* genes are conditionally ablated in the respiratory epithelium are normal[73]. However, when mice that were deficient for both *Foxa1* and *Foxa2* in the lung epithelium were derived, these embryos had severe defects in branching morphogenesis and the epithelium never matured, owing in part to the loss of *Shh* (sonic hedgehog) expression, which is a FoxA target[26]. Similarly, although either *Foxa1*- or *Foxa2*-deficient endoderm can specify a liver, embryos that lack both *Foxa1* and *Foxa2* in the endoderm are completely deficient in hepatic specification, as neither liver bud development nor expression of the earliest liver marker gene, alpha fetoprotein (*Afp*), occurs[25]. Although exposure of ventral foregut endoderm to fibroblast growth factor 2 in explant culture normally results in the induction of liver marker genes, such as albumin 1 (*Alb1*) and transthyretin (*Ttr*), the *Foxa1*- and *Foxa2*-null endoderm lacks this capacity, showing that *Foxa1* and *Foxa2* are required for the induction of liver specification by inductive signals[25]. Currently, the *Foxa1*- and *Foxa2*-deficient mouse is the only known model of a vertebrate that completely lacks a liver.

FoxA in the worm: from pharyngeal development to longevity. Exciting insights into the function of the FoxA clade have also been provided by the study of the FoxA orthologue in *C. elegans*, termed *pha-4* (*pharynx development 4*). Over 15 years ago, Mango and colleagues showed that worms lacking *pha-4* activity had a complete absence of pharyngeal structures[74]. Subsequently, they showed that without *pha-4* the prospective pharyngeal endoderm adopts ectodermal fates, and that conversely, forced expression of *pha-4* could instruct pharyngeal differentiation[75]. Using expression profiling, it was later shown that almost all pharyngeal gene expression is *pha-4*-dependent. By analysing selected PHA-4-target promoters, the authors showed that the pharyngeal levels of PHA-4 increase over time, and therefore proposed that the temporal sequence of PHA-4-target activation depends on the affinity of PHA-4 for the specific binding site in a particular target promoter[76].

More recently, a new and exciting function has been reported for *pha-4* in regulating the increase in longevity that is induced by diet restriction[77]. Severe reductions in food intake increase lifespan in some species, including nematodes[78]. Diet-restricted *pha-4* mutants remained short-lived, showing that the PHA-4 pathway is essential for longevity under these conditions[77]. It will be exciting to explore whether this striking function of *pha-4* extends to its orthologues in mammals.

FoxP2: vocal learning

Speech acquisition in humans. Possibly the most fascinating Fox gene is *FOXP2*, which has been shown to be required for speech acquisition in humans. Arguably, the most important transition in the evolution of modern humans was the acquisition of language and speech, which, owing to its uniqueness to humans, is difficult to study experimentally. The first molecular insight into this problem was provided by the analysis of a multigenerational family with a monogenic disorder that affected speech. Formulating words requires the production of complex sequences of muscle activation in the mouth; impairment of this process results in verbal dyspraxia, which is characterized by both expressive and receptive language deficits. Positional cloning of the causative mutation identified a point mutation in the winged-helix domain of FOXP2 in this family[79]. Expression of FOXP2 in humans and mice was mapped to brain regions that are important in motor control, and the sites of pathology that were identified by imaging patients with verbal dyspraxia correlate well with the early expression domains of FOXP2 (REF. 80). Further evidence for *FOXP2* as the causative gene in verbal dyspraxia was provided by the identification of a truncating mutation in *FOXP2* that segregates with the disorder[81]. But how can verbal dyspraxia be modelled in animals, given that they do not speak? The mouse null mutation for *Foxp2* has been derived, but provides only limited insights. Pups deficient for *Foxp2* emit fewer isolation calls when they are removed from their mother; however, this is an innate behaviour and not a learned vocalization, and the structure of the calls was not changed[82].

Vocal learning in birds. Fortunately for experimental biologists, there are vocal learners other than humans in the animal kingdom[83]. Vocal learners, such as songbirds, alter their innate vocals by imitating mature birds: they have to learn how to sing. Scharff and colleagues exploited the zebra finch model to investigate the function of *FOXP2* further[84]. They used lentiviral RNAi to suppress FOXP2 expression in 'area X', the song nucleus in the basal ganglia. When RNAi-treated juvenile finches were tutored by an adult bird, they had a reduced ability to mimic the song pattern of the tutor compared with control birds. *FOXP2*-deficient finches copied fewer syllables, and those syllables that were copied were imitated with less precision. In studies with zebra finches, *FOXP2*-deficient birds could still produce a normal repertoire of syllables, which suggested that the central defect in human verbal dyspraxia might not be in defective motor performance owing to altered brain structures, but might

be caused by defective motor learning. A caveat to this interpretation is that in patients with verbal dyspraxia, the *FOXP2* mutation is present from conception, whereas in the zebra finch model FOXP2 levels were manipulated in juveniles. Nevertheless, these data provide further evidence for the hypothesis that, during evolution, ancestral neural systems and genetic networks were adapted in the development of human speech.

Imprinting status. As mentioned above, *FOXP2* has undergone adaptive evolution during human evolution. An interesting facet to this story is the fact that *FOXP2* seems to be imprinted and there is higher *FOXP2* expression from the paternal chromosome[85]. According to the conflict theory of evolution of imprinting, this pattern of gene expression is thought to favour selfish behaviours in offspring[86]. It is possible that a gene that favours early language acquisition strengthens the relationship between offspring and mother, and the mother will then provide more resources to a faster-learning child. This concept links the imprinted status of *FOXP2* to the evolutionary theory of acquisition of language[86].

The non-forkhead portions of even the closely related members of the FoxP clade have diverged significantly and, consistent with this, various FoxP genes have dramatically different functions. For example, *FOXP1* and *FOXP4* play essential parts in cardiac morphogenesis[87,88], and *FOXP3* is essential for the programming of regulatory T cells[89]. Important questions to resolve in the future are how these related transcription factors recognize their unique targets and how the evolution of the non-forkhead domains of the FoxP proteins is directed by the requirements for these proteins to interact with specific protein partners.

Conclusion

In the 20 years since its discovery, the Fox gene family has expanded both in number and prominence. Thousands of papers have focused on the cloning, characterization and genetic analysis of hundreds of Fox family members in dozens of species. There can be no question that the Fox gene family is an ancient class of DNA-binding transcription factors that has an astonishing array of functions in development, physiology, cancer and cognition. Future research will need to focus on improving our understanding of how Fox genes select from their thousands of potential binding sites in the genome to exert their specific effects and on how these important genes are regulated, both at the transcription level and by post-translational modifications.

1. Weigel, D., Jurgens, G., Kuttner, F., Seifert, E. & Jackle, H. The homeotic gene *fork head* encodes a nuclear protein and is expressed in the terminal regions of the *Drosophila* embryo. *Cell* **57**, 645–658 (1989).
This paper describes the cloning of the gene responsible for the *fork head* mutation in *D. melanogaster*, after which the Fox gene family is named.
2. Tuteja, G. & Kaestner, K. H. SnapShot: forkhead transcription factors I. *Cell* **130**, 1160 (2007).
3. Tuteja, G. & Kaestner, K. H. Forkhead transcription factors II. *Cell* **131**, 192 (2007).
4. Carroll, J. S. *et al.* Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. *Cell* **122**, 33–43 (2005).
5. Laganiere, J. *et al.* Location analysis of estrogen receptor α target promoters reveals that FOXA1 defines a domain of the estrogen response. *Proc. Natl Acad. Sci. USA* **102**, 11651–11656 (2005).
6. Lai, E. *et al.* HNF-3A, a hepatocyte-enriched transcription factor of novel structure is regulated transcriptionally. *Genes Dev.* **4**, 1427–1436 (1990).
This paper is the first characterization of a Fox gene as a DNA-binding protein and transcriptional regulator.
7. Weigel, D. & Jackle, H. The fork head domain: a novel DNA binding motif of eukaryotic transcription factors? *Cell* **63**, 455–456 (1990).
8. Cirillo, L. A. & Zaret, K. S. Specific interactions of the wing domains of FOXA1 transcription factor with DNA. *J. Mol. Biol.* **366**, 720–724 (2007).
9. Clark, K. L., Halay, E. D., Lai, E. & Burley, S. K. Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. *Nature* **364**, 412–420 (1993).
This paper describes the first molecular structure of the DNA-binding domain of a Fox protein and shows that it resembles those of linker histones.
10. Jin, C., Marsden, I., Chen, X. & Liao, X. Dynamic DNA contacts observed in the NMR structure of winged helix protein–DNA complex. *J. Mol. Biol.* **289**, 683–690 (1999).
11. Tsai, K. L. *et al.* Crystal structure of the human FOXK1a–DNA complex and its implications on the diverse binding specificity of winged helix/forkhead proteins. *J. Biol. Chem.* **281**, 17400–17409 (2006).
12. Stroud, J. C. *et al.* Structure of the forkhead domain of FOXP2 bound to DNA. *Structure* **14**, 159–166 (2006).
13. Kaestner, K. H., Knochel, W. & Martinez, D. E. Unified nomenclature for the winged helix/forkhead transcription factors. *Genes Dev.* **14**, 142–146 (2000).
This paper provides the first systematic naming scheme of Fox genes on the basis of protein similarities in the forkhead domain.
14. Larroux, C. *et al.* Genesis and expansion of metazoan transcription factor gene classes. *Mol. Biol. Evol.* **25**, 980–996 (2008).
This paper provides a phylogenetic analysis of Fox genes, focusing on the early expansion of this family.
15. Katoh, M. Human FOX gene family. *Int. J. Oncol.* **25**, 1495–1500 (2004).
16. Lee, H. H. & Frasch, M. Survey of forkhead domain encoding genes in the *Drosophila* genome: classification and embryonic expression patterns. *Dev. Dyn.* **229**, 357–366 (2004).
17. Leveugle, M., Prat, K., Popovici, C., Birnbaum, D. & Coulier, F. Phylogenetic analysis of *Ciona intestinalis* gene superfamilies supports the hypothesis of successive gene expansions. *J. Mol. Evol.* **58**, 168–181 (2004).
18. Magie, C. R., Pang, K. & Martindale, M. Q. Genomic inventory and expression of Sox and Fox genes in the cnidarian *Nematostella vectensis*. *Dev. Genes Evol.* **215**, 618–630 (2005).
19. Mazet, F., Amemiya, C. T. & Shimeld, S. M. An ancient Fox gene cluster in bilaterian animals. *Curr. Biol.* **16**, R314–R316 (2006).
20. Solomon, K. S., Logsdon, J. M. Jr & Fritz, A. Expression and phylogenetic analyses of three zebrafish FoxI class genes. *Dev. Dyn.* **228**, 301–307 (2003).
21. Tu, Q., Brown, C. T., Davidson, E. H. & Oliveri, P. Sea urchin forkhead gene family: phylogeny and embryonic expression. *Dev. Biol.* **300**, 49–62 (2006).
22. Mazet, F., Yu, J. K., Liberles, D. A., Holland, L. Z. & Shimeld, S. M. Phylogenetic relationships of the Fox (forkhead) gene family in the Bilateria. *Gene* **316**, 79–89 (2003).
This paper provides a detailed phylogenetic analysis of Fox genes in the Bilateria, including an analysis of natural selection based on site-specific shifts in mutation rates between clades.
23. Lynch, M. & Conery, J. S. The evolutionary fate and consequences of duplicate genes. *Science* **290**, 1151–1155 (2000).
24. Ohno, S. *Evolution by Gene Duplication* (Allen and Unwin, London, 1970).
25. Lee, C. S., Friedman, J. R., Fulmer, J. T. & Kaestner, K. H. The initiation of liver development is dependent on Foxa transcription factors. *Nature* **435**, 944–947 (2005).
26. Wan, H. *et al.* Compensatory roles of Foxa1 and Foxa2 during lung morphogenesis. *J. Biol. Chem.* **280**, 13809–13816 (2005).
27. Behr, R., Sackett, S. D., Bochkis, I. M., Le, P. P. & Kaestner, K. H. Impaired male fertility and atrophy of seminiferous tubules caused by haploinsufficiency for Foxa3. *Dev. Biol.* **306**, 636–645 (2007).
28. Singh, L. N. & Hannenhalli, S. Functional diversification of paralogous transcription factors via divergence in DNA binding site motif and in expression. *PLoS ONE* **3**, e2345 (2008).
29. Cocquet, J. *et al.* Structure, evolution and expression of the FOXL2 transcription unit. *Cytogenet. Genome Res.* **101**, 206–211 (2003).
30. Enard, W. *et al.* Molecular evolution of *FOXP2*, a gene involved in speech and language. *Nature* **418**, 869–872 (2002).
This paper provides some evidence for adaptive evolution of *FOXP2* in humans, although an alternative explanation involving relaxed constraint could not be ruled out.
31. Bredenkamp, N., Seoighe, C. & Illing, N. Comparative evolutionary analysis of the FoxG1 transcription factor from diverse vertebrates identifies conserved recognition sites for microRNA regulation. *Dev. Genes Evol.* **217**, 227–233 (2007).
32. Kappen, C., Schughart, K. & Ruddle, F. H. Organization and expression of homeobox genes in mouse and man. *Ann. NY Acad. Sci.* **567**, 243–252 (1989).
33. Madison, B. B., McKenna, L. B., Dolson, D., Epstein, D. J. & Kaestner, K. H. FOXF1 and FOXL1 link hedgehog signaling and the control of epithelial proliferation in the developing stomach and intestine. *J. Biol. Chem.* 2 Dec 2008 (doi: 10.1074/jbc.M808103200).
34. Mahlapuu, M., Enerback, S. & Carlsson, P. Haploinsufficiency of the forkhead gene *Foxf1*, a target for sonic hedgehog signaling, causes lung and foregut malformations. *Development* **128**, 2397–2406 (2001).
35. Brunet, A. *et al.* Akt promotes cell survival by phosphorylating and inhibiting a Forkhead transcription factor. *Cell* **96**, 857–868 (1999).
36. Kops, G. J. *et al.* Direct control of the Forkhead transcription factor AFX by protein kinase B. *Nature* **398**, 630–634 (1999).
37. Cahill, C. M. *et al.* Phosphatidylinositol 3-kinase signaling inhibits DAF-16 DNA binding and function via 14-3-3-dependent and 14-3-3-independent pathways. *J. Biol. Chem.* **276**, 13402–13410 (2001).
38. Ogg, S. *et al.* The Fork head transcription factor DAF-16 transduces insulin-like metabolic and longevity signals in *C. elegans*. *Nature* **389**, 994–999 (1997).

39. Paradis, S. & Ruvkun, G. *Caenorhabditis elegans* Akt/PKB transduces insulin receptor-like signals from AGE-1 PI3 kinase to the DAF-16 transcription factor. *Genes Dev.* **12**, 2488–2498 (1998).
40. Tissenbaum, H. A. & Ruvkun, G. An insulin-like signaling pathway affects both longevity and reproduction in *Caenorhabditis elegans*. *Genetics* **148**, 703–717 (1998).
41. Lin, K., Dorman, J. B., Rodan, A. & Kenyon, C. *daf-16*: an HNF-3/forkhead family member that can function to double the life-span of *Caenorhabditis elegans*. *Science* **278**, 1319–1322 (1997).
References 38–41 establish the link between lifespan, insulin signalling and the FoxO (DAF-16 in nematodes) class of transcription factors. These findings formed the basis of the subsequent demonstration that FoxO proteins are important mediators of insulin signalling in mammals.
42. Michael, M. D. *et al.* Loss of insulin signaling in hepatocytes leads to severe insulin resistance and progressive hepatic dysfunction. *Mol. Cell* **6**, 87–97 (2000).
43. O'Brien, R. M. *et al.* Hepatic nuclear factor 3- and hormone-regulated expression of the phosphoenolpyruvate carboxykinase and insulin-like growth factor-binding protein 1 genes. *Mol. Cell. Biol.* **15**, 1747–1758 (1995).
44. Ayala, J. E. *et al.* Conservation of an insulin response unit between mouse and human glucose-6-phosphatase catalytic subunit gene promoters: transcription factor FKHR binds the insulin response sequence. *Diabetes* **48**, 1885–1889 (1999).
45. Guo, S. *et al.* Phosphorylation of serine 256 by protein kinase B disrupts transactivation by FKHR and mediates effects of insulin on insulin-like growth factor-binding protein-1 promoter activity through a conserved insulin response sequence. *J. Biol. Chem.* **274**, 17184–17192 (1999).
46. Hall, R. K. *et al.* Regulation of phosphoenolpyruvate carboxykinase and insulin-like growth factor-binding protein-1 gene expression by insulin. The role of winged helix/forkhead proteins. *J. Biol. Chem.* **275**, 30169–30175 (2000).
47. Rena, G., Guo, S., Cichy, S. C., Unterman, T. G. & Cohen, P. Phosphorylation of the transcription factor forkhead family member FKHR by protein kinase B. *J. Biol. Chem.* **274**, 17179–17183 (1999).
48. Nakae, J. *et al.* Regulation of insulin action and pancreatic β-cell function by mutated alleles of the gene encoding forkhead transcription factor Foxo1. *Nature Genet.* **32**, 245–253 (2002).
49. Matsumoto, M., Pocai, A., Rossetti, L., Depinho, R. A. & Accili, D. Impaired regulation of hepatic glucose production in mice lacking the forkhead transcription factor Foxo1 in liver. *Cell Metab.* **6**, 208–216 (2007).
50. Dong, X. C. *et al.* Inactivation of hepatic Foxo1 by insulin signaling is required for adaptive nutrient homeostasis and endocrine growth regulation. *Cell Metab.* **8**, 65–76 (2008).
51. Galili, N. *et al.* Fusion of a fork head domain gene to *PAX3* in the solid tumour alveolar rhabdomyosarcoma. *Nature Genet.* **5**, 230–235 (1993).
52. Parry, P., Wei, Y. & Evans, G. Cloning and characterization of the t(X;11) breakpoint from a leukemic cell line identify a new member of the forkhead gene family. *Genes Chromosomes Cancer* **11**, 79–84 (1994).
53. Barr, F. G. The role of chimeric paired box transcription factors in the pathogenesis of pediatric rhabdomysarcoma. *Cancer Res.* **59** (Suppl.), 1711–1715 (1999).
54. Xia, S. J. & Barr, F. G. Analysis of the transforming and growth suppressive activities of the PAX3–FKHR oncoprotein. *Oncogene* **23**, 6864–6871 (2004).
55. Lai, E., Prezioso, V. R., Tao, W. F., Chen, W. S. & Darnell, J. E. Jr. Hepatocyte nuclear factor 3 alpha belongs to a gene family in mammals that is homologous to the *Drosophila* homeotic gene fork head. *Genes Dev.* **5**, 416–427 (1991).
56. Ang, S. L. & Rossant, J. HNF-3β is essential for node and notochord formation in mouse development. *Cell* **78**, 561–574 (1994).
57. Weinstein, D. C. *et al.* The winged-helix transcription factor HNF-3β is required for notochord development in the mouse embryo. *Cell* **78**, 575–588 (1994).
58. Cirillo, L. A. *et al.* Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Mol. Cell* **9**, 279–289 (2002).
59. Zhang, L., Rubins, N. E., Ahima, R. S., Greenbaum, L. E. & Kaestner, K. H. Foxa2 integrates the transcriptional response of the hepatocyte to fasting. *Cell Metab.* **2**, 141–148 (2005).
60. Gao, N. *et al.* Forkhead box A1 regulates prostate ductal morphogenesis and promotes epithelial cell maturation. *Development* **132**, 3431–3443 (2005).
61. Gao, N. *et al.* The role of hepatocyte nuclear factor-3α (forkhead box A1) and androgen receptor in transcriptional regulation of prostatic genes. *Mol. Endocrinol.* **17**, 1484–1507 (2003).
62. Yu, X. *et al.* Foxa1 and Foxa2 interact with the androgen receptor to regulate prostate and epididymal genes differentially. *Ann. NY Acad. Sci.* **1061**, 77–93 (2005).
63. Yu, X. *et al.* The role of forkhead box A2 to restrict androgen-regulated gene expression of lipocalin 5 in the mouse epididymis. *Mol. Endocrinol.* **20**, 2418–2431 (2006).
64. Friedman, J. R. & Kaestner, K. H. The Foxa family of transcription factors in development and metabolism. *Cell. Mol. Life Sci.* **63**, 2317–2328 (2006).
65. Bochkis, I. M. *et al.* Hepatocyte-specific ablation of *Foxa2* alters bile acid homeostasis and results in endoplasmic reticulum stress. *Nature Med.* **14**, 828–836 (2008).
66. Ang, S. L. *et al.* The formation and maintenance of the definitive endoderm lineage in the mouse: involvement of HNF3/forkhead proteins. *Development* **119**, 1301–1315 (1993).
67. Besnard, V., Wert, S. E., Hull, W. M. & Whitsett, J. A. Immunohistochemical localization of Foxa1 and Foxa2 in mouse embryos and adult tissues. *Gene Expr. Patterns* **5**, 193–208 (2004).
68. Kaestner, K. H., Hiemisch, H., Luckow, B. & Schutz, G. The HNF-3 gene family of transcription factors in mice: gene structure, cDNA sequence, and mRNA distribution. *Genomics* **20**, 377–385 (1994).
69. Monaghan, A. P., Kaestner, K. H., Grau, E. & Schutz, G. Postimplantation expression patterns indicate a role for the mouse forkhead/HNF-3α, β and γ genes in determination of the definitive endoderm, chordamesoderm and neuroectoderm. *Development* **119**, 567–578 (1993).
70. Besnard, V., Wert, S. E., Kaestner, K. H. & Whitsett, J. A. Stage-specific regulation of respiratory epithelial cell differentiation by Foxa1. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **289**, L750–L759 (2005).
71. Kaestner, K. H., Katz, J., Liu, Y., Drucker, D. J. & Schutz, G. Inactivation of the winged helix transcription factor HNF3α affects glucose homeostasis and islet glucagon gene expression *in vivo*. *Genes Dev.* **13**, 495–504 (1999).
72. Shih, D. Q., Navas, M. A., Kuwajima, S., Duncan, S. A. & Stoffel, M. Impaired glucose homeostasis and neonatal mortality in hepatocyte nuclear factor 3α-deficient mice. *Proc. Natl Acad. Sci. USA* **96**, 10152–10157 (1999).
73. Wan, H. *et al.* Foxa2 regulates alveolarization and goblet cell hyperplasia. *Development* **131**, 953–964 (2004).
74. Mango, S. E., Lambie, E. J. & Kimble, J. The *pha-4* gene is required to generate the pharyngeal primordium of *Caenorhabditis elegans*. *Development* **120**, 3019–3301 (1994).
75. Horner, M. A. *et al. pha-4*, an *HNF-3* homolog, specifies pharyngeal organ identity in *Caenorhabditis elegans*. *Genes Dev.* **12**, 1947–1952 (1998).
76. Gaudet, J. & Mango, S. E. Regulation of organogenesis by the *Caenorhabditis elegans* FoxA protein PHA-4. *Science* **295**, 821–825 (2002).
77. Panowski, S. H., Wolff, S., Aguilaniu, H., Durieux, J. & Dillin, A. PHA-4/Foxa mediates diet-restriction-induced longevity of *C. elegans*. *Nature* **447**, 550–555 (2007).
78. Mair, W. & Dillin, A. Aging and survival: the genetics of life span extension by dietary restriction. *Annu. Rev. Biochem.* **77**, 727–754 (2008).
79. 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).
80. 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).
81. 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).
82. 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).
83. Scharff, C. & Haesler, S. An evolutionary perspective on FoxP2: strictly for the birds? *Curr. Opin. Neurobiol.* **15**, 694–703 (2005).
84. Haesler, S. *et al.* Incomplete and inaccurate vocal imitation after knockdown of *FoxP2* in songbird basal ganglia nucleus area X. *PLoS Biol.* **5**, e321 (2007).
85. Feuk, L. *et al.* Absence of a paternally inherited *FOXP2* gene in developmental verbal dyspraxia. *Am. J. Hum. Genet.* **79**, 965–972 (2006).
86. Crespi, B. J. Sly *FOXP2*: genomic conflict in the evolution of language. *Trends Ecol. Evol.* **22**, 174–175 (2007).
87. Li, S., Zhou, D., Lu, M. M. & Morrisey, E. E. Advanced cardiac morphogenesis does not require heart tube fusion. *Science* **305**, 1619–1622 (2004).
88. Wang, B. *et al. Foxp1* regulates cardiac outflow tract, endocardial cushion morphogenesis and myocyte proliferation and maturation. *Development* **131**, 4477–4487 (2004).
89. Fontenot, J. D., Gavin, M. A. & Rudensky, A. Y. *Foxp3* programs the development and function of $CD4^+CD25^+$ regulatory T cells. *Nature Immunol.* **4**, 330–336 (2003).
90. Nishimura, D. Y. *et al.* The forkhead transcription factor gene *FKHL7* is responsible for glaucoma phenotypes which map to 6p25. *Nature Genet.* **19**, 140–147 (1998).
91. Fang, J. *et al.* Mutations in *FOXC2* (*MFH-1*), a forkhead family transcription factor, are responsible for the hereditary lymphedema–distichiasis syndrome. *Am. J. Hum. Genet.* **67**, 1382–1388 (2000).
92. Chatila, T. A. *et al. JM2*, encoding a fork head-related protein, is mutated in X-linked autoimmunity-allergic disregulation syndrome. *J. Clin. Invest.* **106**, R75–R81 (2000).
93. Wildin, R. S. *et al.* X-linked neonatal diabetes mellitus, enteropathy and endocrinopathy syndrome is the human equivalent of mouse scurfy. *Nature Genet.* **27**, 18–20 (2001).
94. Li, J. & Vogt, P. K. The retroviral oncogene qin belongs to the transcription factor family that includes the homeotic gene fork head. *Proc. Natl Acad. Sci. USA* **90**, 4490–4494 (1993).
95. Costa, R. H., Kalinichenko, V. V., Holterman, A. X. & Wang, X. Transcription factors in liver development, differentiation, and regeneration. *Hepatology* **38**, 1331–1347 (2003).
96. Wierstra, I. & Alves, J. FOXM1, a typical proliferation-associated transcription factor. *Biol. Chem.* **388**, 1257–1274 (2007).
97. Kalinichenko, V. V. *et al.* Foxm1b transcription factor is essential for development of hepatocellular carcinomas and is negatively regulated by the p19ARF tumor suppressor. *Genes Dev.* **18**, 830–850 (2004).

Acknowledgements

The related work in the Kaestner laboratory is supported by the National Institutes of Health (NIH) grants DK-42910, DK-53839 and DK-55342, and in the Hannenhalli laboratory by NIH grant R01GM085226. The authors thank L. N. Singh for his help with preparing some of the figures. We apologize to all our colleagues whose important contributions could not be cited owing to space constraints.

DATABASES

Entrez Gene: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene
fkh | *Foxa1* | *Foxa2* | *Foxa3* | *FOXC1* | *FOXC2* | *Foxo1* | *FOXP1* | *FOXP2* | *FOXP3* | *FOXP4* | *Irs1* | *Irs2* | *pha-4*
OMIM: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM
IPEX | lymphedema–distichiasis | Rieger syndrome
UniProtKB: http://www.uniprot.org
CDKN2A | DAF-16 | FOXA1 | FOXA2 | FOXM1 | FOXO1 | FOXO3 | FOXO4 | IGF1 | PAX3 | PAX7

FURTHER INFORMATION

Sridhar Hannenhalli's laboratory homepage: http://www.med.upenn.edu/apps/faculty/index.php/g306/p1999560
Klaus H. Kaestner's laboratory homepage: http://www.med.upenn.edu/kaestnerlab
ClustalX multiple sequence alignment: http://www.clustal.org
ENSEMBL v50: http://www.ensembl.org
Index of winged helix proteins: http://www.biology.pomona.edu/fox
Jalview multiple alignment editor: http://www.jalview.org
NCBI: http://www.ncbi.nlm.nih.gov
Pfam database: http://www.pfam.org
T-Coffee multiple sequence alignment package: http://www.tcoffee.org/Projects_home_page/t_coffee_home_page.html

ALL LINKS ARE ACTIVE IN THE ONLINE PDF