In 2001, scientists characterized the first gene to be implicated in the cause of a speech and language disorder (FOXP2). Although FOXP2 was discovered using a unique family in which a severe speech and language disorder segregates in a monogenic fashion, at the time this discovery was heralded as “a milestone in understanding this uniquely human characteristic.” Approximately 1 year later, we discuss the impact of this gene discovery on the study of language and review the relevance of this gene to both specific language impairment and language aspects of the autistic phenotype. We also discuss recent molecular genetic advances made in the study of generalized specific language impairment.

Speech and language disorder in the KE family

The KE family is a rare pedigree in which approximately half the members (both male and female) are affected by a severe speech and language disorder inherited in a monogenic, autosomal dominant fashion (Fig. 1). Affected members of the KE family have an orofacial dyspraxia that limits control of fine motor movements of their lower face. This dyspraxia results in a speech and language disorder that is so severe that many younger members of the family have been taught a signing system to augment their spoken language [1]. Their speech difficulties are accompanied by widespread language impairments that are characterized by atypical grammar usage and deficits in both the expressive and receptive domains. These difficulties are evident in spoken and written language [1–3•].

Although speech and language impairments form a central aspect of the phenotype in affected members of the KE family, the exact nature of their deficit remains a matter of debate. Some researchers argue that their language difficulties are a secondary characteristic that arises as a direct result of restricted orofacial movement. However, although these limitations may obviously underlie the articulation and some expressive features of the deficit, it is not so clear how they could account for the receptive language impairments. Studies of children with cerebral palsy indicate that even a complete lack of expressive language abilities does not necessarily result in deficits in the receptive domain [4].

In addition to their language difficulties, many affected members of the KE family are reported to have significant nonverbal deficits. These include problems with nonverbal coding tasks [3•], the production of complex, sequential oral movements (eg, closing the lips, then opening the mouth, then protruding the tongue) [1,5], and the perception and production of rhythm sequences [6•]. Furthermore, many of the affected members have nonverbal IQ scores that are extreme enough to place them outside of the normal range. One study estimated the average performance IQ of affected individuals as 86 (range 71–111), with 6 obtaining scores below 85 [1]. Because such deficits preclude the existence of a discrepancy between verbal and nonverbal abilities, some researchers argue that the language problems segregating in the KE family may form part of more general learning difficulties. However, it should be noted that these non-
verbal deficits do not form an integral part of the phenotype and some affected members of the KE family have performance IQ within the normal range, whereas some unaffected individuals perform significantly below the expected mean [1]. Furthermore, an increasing number of studies demonstrate that it is not uncommon for the performance IQ of language-impaired individuals to decline over time, indicating that disorders of speech and language may limit the development of nonverbal intelligence as measured by these tests [3, 7].

In conclusion, members of the KE family have a disorder that is characterized by an orofacial dyspraxia and a severe speech and language impairment. Although affected individuals may show some nonverbal deficits, the relationship between the onset of the disorder and these additional deficits remains unclear.

**SPCH1 and chromosome 7q**

In 1998, a genome screen was performed on 27 members of the KE family [8]. This scan revealed a region of linkage on the long arm of chromosome 7, with marker D7S486 giving a maximal logarithm of the odds (LOD) score of 6.22. Haplotype analyses revealed critical recombinants that placed the SPCH1 locus (MIM605317) within a 5.6 cM critical region between markers D7S2459 (proximal) and D7S643 (distal) [8] (Fig. 2).

The publication of the SPCH1 findings led to the identification of a second patient, known as CS, with a de novo translocation involving a breakpoint within the SPCH1 region (t(5;7)(q22;q31.2)) [9]. CS was described as having an oral dyspraxia with language impairments in both the receptive and expressive domains. Linguistically, he was considered similar to the affected members of the KE family, and his nonverbal skills were reported to be within the normal range. Mapping of the CS translocation revealed the chromosome 7 breakpoint to lie within a brain-expressed transcript known as CAGH44 [9]. Subsequent investigations allowed the compilation of the full coding sequence of CAGH44 and revealed the presence of a forkhead box or winged-helix (fox) domain within the gene. Accordingly, the gene was renamed FOXP2 [10]. Mutation screening of the FOXP2 coding sequence identified a point mutation within the KE family that cosegregated perfectly with the speech and language disorder [10]. This mutation involves a single base transition (G→A) that results in an arginine to histidine substitution within exon 14 of the FOXP2 gene (R553H) [10].

Lai et al. postulate that the mutation identified in affected members of the KE family may result in an insufficient dosage of functional FOXP2 during embryogenesis that, in turn, may lead to the underdevelopment of brain areas critical for speech and language development [10].

**FOXP2 gene**

The FOX genes encode a large family of transcription factors, all of which possess a winged-helix, or forkhead box (fox), DNA-binding domain. The family consists of more than 100 members, which are found to be con-
served across species ranging from humans to mice, fruit flies, and yeast [11].

The forkhead domain typically consists of a stretch of 100 conserved amino acids that adopt a characteristic structure containing three α-helices and two large loops (or wings). In this structure, the third α-helix makes contact with the major groove of the target DNA [11]. The KE mutation involves a change within this third helix, at a residue that is highly conserved and lies adjacent to a residue that makes direct contact with the DNA [10••].

Lai et al. identified 19 exons spanning the FOXP2 coding region, two of which are alternatively spliced [10••]. The major splice form encodes a 715-amino acid protein that contains a characteristic fox domain (exons 12 to 14) and a 40-residue polyglutamine tract that is flanked by a smaller 10-residue polyglutamine stretch (exons 5 and 6) (Fig. 3).

**FOXP2 in autism**

The identification of the FOXP2 gene marked an exciting discovery, not only for researchers of language impairment but also for investigators in the field of autism.

Autism is a neurodevelopmental disorder characterized by deficits in reciprocal social interaction and communication, accompanied by repetitive and stereotyped behaviors and interests [12,13]. Language deficits form a major component of the autism diagnostic schedule and parallels are often drawn between severe forms of pragmatic language disorders and the autistic phenotype.

**Chromosome 7q in autism**

Over the last decade, several genome scans have been performed for autistic disorder, many of which have provided evidence for the existence of an autism susceptibility gene on chromosome 7q31, the location of which (AUTS1—MIM209850) coincides with the SPCH1 region [14–19]. Support for the role of 7q31 in both autism and language impairments is provided by reports of chromosome 7 translocations involving autistic or language disorder phenotypes [15,20,21]. Furthermore, it is reported that the linkage between chromosome 7q and autism can be strengthened by the stratification of samples according to the presence of severe language difficulties in the proband [22], or by the study of quantitative, language-related endophenotypes in autistic patients [23].

**FOXP2 in autism**

The apparent involvement of chromosome 7q31 in both autism and language impairment led to the proposal that the phenotypic similarities between the two disorders may reflect a shared genetic cause [24]. It, therefore, followed that FOXP2 may somehow be involved in the onset of autism. To date, two investigations have performed a mutation screen of FOXP2 within autistic patients. Newbury et al. screened the entire coding sequence of FOXP2 within 48 autistic probands who formed a subset of those cases used in the original identification of the AUTS1 locus, but found no mutations within the coding sequence [25]. Wassink et al. performed a screen of all exonic regions within 75 autistic individuals and identified changes in three autistic probands: two involved a reduction in the number of residues in the 40-residue polyglutamine stretch of FOXP2 (one heterozygous deletion of five glutamine residues and another of six) [26]. Until this point, the FOXP2 polyglutamine region was considered to be stable and not polymorphic because it contained a mixture of CAA and CAG codons and had been consistent in size in all control groups previously studied [10••,27]. Involvement of polyglutamine repeat regions in human neurodegenerative disease is well documented [28]. However, such cases invariably involve expansion of polyglutamine stretches rather than contraction. It is, therefore, likely that these changes represent rare polymorphisms and are not associated with the autism phenotype. The third change identified by Wassink et al. involved a C→T transition within exon 5 of FOXP2. This change was seen in a single proband and resulted in a synonymous change (G247G) and, therefore, would not be expected to contribute to the onset of autistic disorder.

Both of the above studies used markers (single nucleotide polymorphisms, microsatellites, or both) spread across the FOXP2 region to test for association and linkage within their autistic families (Fig. 3). However, neither investigation found any strong evidence for an association between autism and the FOXP2 gene.

---

**Figure 3. The FOXP2 coding sequence**

The FOXP2 gene consists of 19 exons, 2 of which (3a and 3b) are alternatively spliced. The forkhead domain is encoded by exons 12 to 14 and a polyglutamine tract is found to span exons 5 and 6. All exons are given to scale. Larger introns are shown with slashes with the sizes given above. Gray exons represent untranslated regions. Triangles and arrows give positions of markers used for association in Newbury et al. [25] and Wassink et al. [28] studies (see text). Gray shapes represent markers used by Newbury et al. White shapes represent markers analyzed by Wassink et al. Triangles represent microsatellites and arrows represent single nucleotide polymorphisms.
In summary, little evidence indicates a strong role for FOXP2 in autism and any mutations that have been found are not frequent enough to account for the strong level of linkage demonstrated between chromosome 7q and autistic disorder. It, therefore, is likely that the AUTS1 linkage is caused by a second gene (or genes) that, coincidentally, lies in a similar position to FOXP2.

Specific language impairment

Specific language impairment (SLI) is diagnosed in children who experience an unexpected difficulty in the acquisition of language skills, despite otherwise normal development and adequate intelligence. A diagnosis of SLI relies on the absence of other neurologic conditions (eg, cerebral palsy, autism) [29].

It has been recognized for some time that SLI has a strong genetic component. However, unlike the speech and language difficulties associated with the KE family, the genetics underlying this disorder are expected to be complex, involving several loci that interact with each other and the environment to produce an overall susceptibility to disease onset [30–32•,33]. This complexity, until recently, precluded the application of traditional techniques in the genetic study of SLI. However, advances within the genetic, technologic and bioinformation fields mean that it is now possible to screen for susceptibility genes for complex disorders in much the same way as for those of monogenic disorders. Despite reduced power and the need for large sample sizes, these linkage studies have been proved to offer an effective method of gene identification [34].

Molecular genetic studies

To date, two genome screens have been performed for SLI. The first reported linkage to chromosomes 16q and 19q [35], and the second implicated loci on chromosomes 13q and 2p [36]. Interestingly, neither screen showed any evidence for linkage to the chromosome 7q region. The apparent lack of consensus between the results from these investigations may be attributed to differences in study design. Although both screens were related to SLI, each took a completely different approach to their subject selection, phenotyping, and genetic strategies. The SLI Consortium [35] used small nuclear families and quantitative measures of language abilities (expressive language, receptive language, and nonword repetition) alongside nonparametric analyses. Bartlett et al. [36] used large extended pedigrees, a selection of binary affection statuses (language impairment, reading impairment and clinical impairment), and parametric genetic analyses (under both a recessive and a dominant model of inheritance). Each study design has its own strengths: the use of quantitative traits [35] circumvents the need for the derivation of a consistent affection status, whereas the use of large pedigrees [36] can act to increase the homogeneity of the sample. It is possible that both have revealed loci that are of general importance to the SLI phenotype. As with other complex disorders, only by independent replication of these studies can the importance of each of the identified chromosomal regions be evaluated.

Comparisons between the results of these SLI scans and those reported for autism and dyslexia screens highlight the complexity of the phenotypic and, possibly, genetic overlaps between these neurodevelopmental conditions. Although current diagnostic scales treat these three disorders as distinct entities, they are highly comorbid and it has been suggested that they may represent different manifestations of the same underlying deficits [24,37]. The chromosome 13q locus identified by Bartlett et al. was found to be linked to a reading discrepancy measure within a language-impaired sample [36]. Intriguingly, although previous studies of reading and language impairments had not found linkage to this locus, it has been connected to autism by both an independent genome screen [16,22] and a recent meta-analysis [38]. Conversely, the chromosome 2p locus, which was found linked to the language-impaired subgroup [36], is close to a region frequently connected to dyslexia (DYX3) [39–41]. In addition, chromosome 19q has also been suggested as an autism locus by a single group [42]. However, this result was obtained using a narrow diagnostic category that excluded any children who may overlap with the SLI spectrum.

FOXP2 in specific language impairment

Although the orofacial dyspraxia of the KE family would exclude them from a traditional diagnosis of SLI, the FOXP2 gene remains the principal candidate for more general forms of language impairments. Newbury et al. screened the entire FOXP2 coding sequence within 43 language-impaired probands and found only a single mutation [25]. Analogous to those cases described by Wassink et al. [26], this change involved an addition of two residues to the shorter of the two FOXP2 polyglutamine tracts. However, this insertion was found to occur within a stretch of CAG codons on the intron–exon border of exon 6 and, thus, the exact position of the mutation remains unclear and may not actually form part of the transcribed unit. Furthermore, in the language-impaired family carrying this change, the mutation did not cosegregate with the SLI phenotype. Thus, the authors concluded that the change probably represented a rare polymorphism [25]. Again, association analyses of markers spread across the FOXP2 region (Fig. 3) were all nonsignificant [25]. It, therefore, would appear that FOXP2 does not play a major role in the onset of SLI. However, it should be noted that the probands used in the Newbury et al. study [25] were selected to have general expressive or receptive language impairments and did not show any evidence of verbal or orofacial dyspraxia. Thus, although it is unlikely that FOXP2 repre-
sents a major gene locus in the onset of language impairment, it possibly plays a role in a small number of specific cases of language impairments that are characterized by dyspraxic features. Furthermore, it is feasible that another gene within the FOXP2 pathway will be involved in the SLI phenotype.

Summary

Although a year has passed since the FOXP2 gene was identified, the genetics of language remain as complex as ever. It is believed that FOXP2 does not contribute a major gene effect to either autism or SLI, but it remains possible that it is involved in a specific subclass of language impairments. It is hoped that further studies of the FOXP2 pathway can highlight additional genes (e.g., downstream target genes) that may play a role in the mechanism of more general forms of language impairments.

This year has seen the initiation of molecular genetic studies of SLI. As yet, these studies do not concur on the importance of any single locus in the cause of language impairments and further investigations are required to assess the significance of the putative loci on chromosomes 2, 13, 16, and 19.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:
• Of special interest
•• Of outstanding interest


This paper presents the characterization of the FOXP2 mutations in both the KE family and CS.
36 Bartlett CW, Flax JF, Logue MW, et al.: A major susceptibility locus for spe-


