



The enigma of speech and language acquisition in human - A genetic approach

Jayanta kumar Kundu, Suman Pratihar* and Rudra Prasad Nath

Department of Zoology, Genetics and Molecular biology Lab., Vidyasagar University, Midnapore - 721102 (West Bengal), INDIA

*Corresponding author. E-mail: pratihar_vu@yahoo.co.in

Abstract: The use of language is one of the most vital traits that distinguish human from other species. The gene FOXP2 (fork head box P2) is relevant to the human ability to develop language. FOXP2 is situated on human chromosome 7q31, and its major splice form encodes a protein of 715 amino acids. It contains a glutamine rich region consisting of two adjacent polyglutamine tracts, encoded by mixtures of CAG and CAA repeats. Two functional copy of FOXP2 seem to be required for acquisition of normal language in human. The chimpanzee, gorilla and the rhesus macaque FOXP2 proteins are all identical to each other and carry only one difference from the mouse and two differences from the human protein, whereas the orangutan two differences from the mouse and three from humans. This gene is also important on particular aspects of cognition awareness or grammatical processing.

Keywords: Language evolution, Human language, Primate language, FOXP2 gene, Language disorder

INTRODUCTION

Language is a uniquely human trait likely to have been a prerequisite for the development of human culture. The ability to develop speech requires fine control of the larynx and mouth, that are absent in chimpanzees and other great apes. FOXP2 (fork head box P2) is the first gene relevant to the human ability to develop language. A point mutation in FOXP2 co-segregates with a disorder in a family in which half of the members have severe articulation difficulties accompanied by linguistic and grammatical impairment. This gene is disrupted by translocation in an unrelated individual who has a similar disorder (Smith, 2007). Thus, two functional copy of FOXP2 seem to be required for acquisition of normal language. Sequencing of the complementary DNAs that encode the FOXP2 protein in the chimpanzee, gorilla, orangutan, rhesus macaque and mouse have been done, and compared with the human cDNA. Scientists also have investigated intraspecific variation of the human FOXP2 gene. They showed that human FOXP2 contain changes in amino acid coding and a pattern of nucleotide polymorphism which strongly suggest that this gene has been the target of selection during recent human evolution. FOXP2 is situated on human chromosome 7q31, and its major splice form encodes a protein of 715 amino acids. It contains a glutamine rich region consisting of two adjacent polyglutamine tracts, encoded by mixtures of CAG and CAA repeats (Bruce and Margolis, 2002). In the case of FOXP2, the lengths of the polyglutamine stretches differed for all taxa studied. Variation in all second polyglutamine tracts has been observed in a small

family affected with speech and language impairment, but this did not co-segregate with disorder, suggesting that minor changes in length may not significantly alter the function of the protein. If the polyglutamine stretches are disregarded, the human FOXP2 protein differs at only three amino- acids position from its orthologue in the mouse. The chimpanzee, gorilla and the rhesus macaque FOXP2 proteins are all identical to each other and carry only one difference from the mouse and two differences from the human protein, whereas the orangutan two differences from the mouse and three from humans. Two amino acid differences are both found in exon 7 of the FOXP2 gene and are a threonine to asparagine and an asparagine to serine change at position 303 and 325, respectively.

ORIGIN OF LANGUAGE

The origin of language is a topic that has attracted considerable speculations throughout human history. The use of language is one of the most vital traits that distinguish human from other species. Linguists agree that there are no existing primitive languages. All modern human population speaks languages of comparable complexity. While existing language differ in the size and subjects covered in their lexicons, all possess the grammar and syntax needed and can invent, translate and borrow the vocabulary necessary too express the full range of their speaker's concepts. All human possess similar linguistic abilities, no child is born with a biological pre disposition favoring any one type of language (Hauser *et al.*, 2002).

SPEECH VERSUS LANGUAGE

It is necessary to make a minor distinction between speech and language. Speech involves producing sounds from the voice box. Talking birds like parrots are able to imitate human speech with varying ability. However this ability to mimic human sounds is very different from the acquisition of syntax. On the other hand the deaf generally use speech but are able to communicate effectively with sign language which is considered as fully developed, complex, modern language. So the evolution of modern language required both development of anatomical apparatus and neurological changes in the brain.

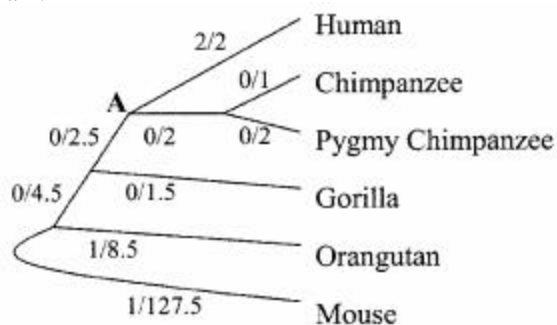


Fig. 1. Nucleotide substitution in the coding region of *foxp2* gene for different lineages of primate evolution.

LANGUAGE IN PRIMATES

Great ape language: The anatomical structure of their larynx does not enable them to make many of sounds like human. But in captivity some apes like kanzi have been reported to able to learn several hundred words. In wild vervet monkeys have been studied to make up about ten different vocalizations.

Chimpanzees: Wild chimpanzees have been talking in time of danger to warn others. If one chimpanzee sees any snake he makes low, rumbling noise to signal other to climb on the tree. However, single feature alone do not qualify such instances of communication as being true language.

Modern humans: Modern human first appeared 200000 years ago in Ethiopia. The greatest step in language evolution would have been the progression from primitive to modern languages. According to out of Africa hypothesis around 50000 years ago a group of human left Africa and proceeded to rest of the world. Some scientists believe that human did not live Africa before that because of modern cognition and language.

Monogenesis (Proto World language): Linguistic monogenesis or mother tongue theory is the hypothesis that there was one single proto language from which all other languages originated. All human alive today are descended from mitochondrial eve, a woman estimated to have lived in Africa some 150000 years ago. Possibility is that the proto world language could originate from that time.

DISCOVERY OF FOXP2 GENE

Search for gene was initially started as a result of investigation into KE or K family. Further investigation revealed that a point mutation in chromosome 7, sequenced and analyzed and referred as FOXP2 gene. The first protein that was discovered was the fork head transcription factor in *Drosophila* sp. Originally FOX family members were given vastly different names but in 2000 a unified nomenclature was introduced that grouped the FOX proteins into subclasses (FOXA-FOXS) based on sequence conservation.

GENES OF THE FOX FAMILY

FOXA1, FOXA2, FOXA3
FOXB1, FOXB2
FOXC1, FOXC2
FOXD1, FOXD2, FOXD3, FOXD4, FOXD5, FOXD6
FOXE1, FOXE2, FOXE3
FOXF1, FOXF2
FOXG1
FOXH1
FOXJ1, FOXJ2, FOXJ3
FOXK1, FOXK2

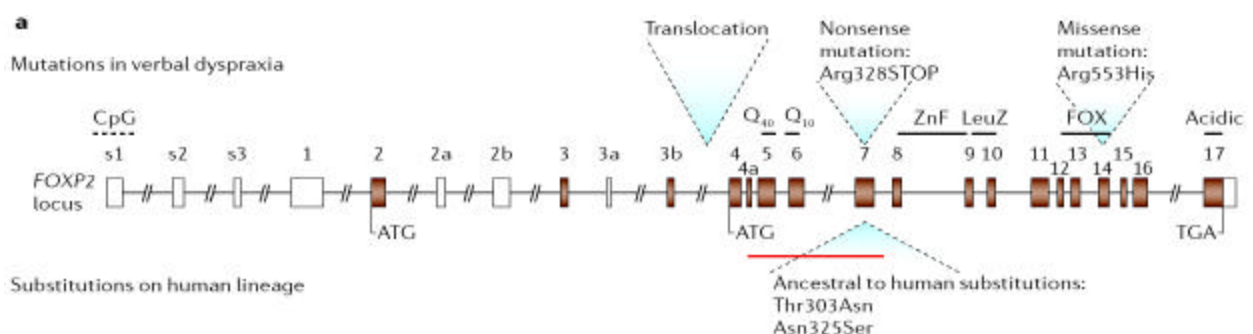


Fig. 2. Mutation in *foxp2* locus on chromosome 7.

FOXL1, FOXL2

FOXM1

FOXN1, FOXN2, FOXN3, FOXN4

FOXO1, FOXO2, FOXO3, FOXO4

FOXP1, FOXP2, FOXP3, FOXP4

FOXQ1

FOXR1, FOXR2 (Shu *et al.*, 2001).

GENE RESPONSIBLE FOR SPEECH DEVELOPMENT IN HUMAN

FOXP2 is a gene that is implicated in the skills including grammatical competence. FOXP2 is a member of large FOX families of transcription factors. Information from known human mutations and mouse studies suggests that FOXP2 regulates gene involved in the development of tissue like brain, lung and gut (Shu *et al.*, 2001). The exact identity of genes FOXP2 regulates is still not known. However in 2001 a point mutation in FOXP2 coding sequence was identified as the basis of an inherited speech and language disorder (Takahashi *et al.*, 2003).

EVOLUTION OF THE GENE FOXP 2

The FOXP2 protein sequence is highly conserved. Similar FOXP2 proteins are found in fish, reptiles and song birds. Human FOXP2 differs from chimpanzee FOXP2 by 2 amino acids, mouse FOXP2 by 3 amino acids, zebra finch FOXP2 by 7 amino acids. Great interest has centered on the evolution of FOXP2, given the abnormal speech and language development observed in the members of KE family who have mutation in one copy of gene. FOXP2 sequence is highly fixed into modern humans (Enard *et al.*, 2002).

MOLECULAR GENETICS OF THE FOXP 2 GENE

The gene that is responsible for the speech and language disorder in KE family was originally localized to the long arm of chromosome 7 (7q31) (Lai *et al.*, 2000). Ultimately after different analysis of the chromosomes of C.S. that led to the identification of the FOXP2 gene as the cause of the speech and language disorder (Lai *et al.*, 2001). Importantly the mutation substituted a histidine for an arginine at site 553 in the FOXP2 sequence. This arginine is invariant among FOX genes suggesting that it has a crucial functional role and lies adjacent to a histidine in the third helix of a fork head domain, where the protein contacts the DNA during transcriptional control. Loss of function of one copy of FOXP2 gene and that the one copy that remains is insufficient for normal brain development leading to the speech and language disorder (Lai *et al.*, 2003)

MOLECULAR BIOLOGY OF THE FOXP 2 GENE

A different point mutation is identified in multiple affected

members of a newly identified family with speech and language disorder. In this case (Fig. 2), the mutation produces a “stop” signal in the middle of the FOXP2 protein, severely truncating it, and presumably leading to loss of function. Again these individuals carry one normal copy of FOXP2 and one mutated copy. Disruption of only one copy of FOXP2 is enough to lead to speech and language impairment (Lehmann *et al.*, 2000).

Reduced amount of functional FOXP2 protein can lead to disordered brain development or function in a manner that primarily interferes with speech and language abilities. However, it is essential to realize that this is emphatically not the same as saying that FOXP2 is a “gene for speech” or a “gene for language”.

FOXP2 is not unique to humans and is found in extremely similar form throughout mammalian species. The protein encoded by mouse version of this gene differs from the human protein at only 3 amino acid positions out of more than 700. Since mice clearly lack in linguistic capabilities, the remarkably high conservation of FOXP2 across the species is incompatible with the idea that the gene exists solely to, sub serve the capacity for speech and language. Functional FOXP2 protein was already present in the common nonlinguistic ancestor of humans and rodents over 70 million years ago in a very similar form to that found in modern humans (Marcus *et al.*, 2003).

The expression of FOXP2 indicates that its functional importance is not restricted to the Central Nervous System (CNS) (Shu *et al.*, 2001). FOXP2 regulates key pathways in the developing lung, heart and gut. Detailed studies of FOXP2 in the brain indicate complexities of expression in both humans and rodents; the gene is switched on in a range of brain regions during early development including cortex, thalamus, hypothalamus, striatum, cerebellum and medulla, and expression persist into adulthood (Lieberman *et al.*, 2002).

A naïve view of a gene that exists to provide us with linguistic capabilities would predict an expression pattern that is restricted to a region of the brain such as Broca’s area (Marcus *et al.*, 2003). FOXP2 is a well conserved regulatory gene with multiple roles during development and is likely to influence patterning or function of the several regions of the CNS in all mammals (MacDermot *et al.*, 2005; Bosman *et al.*, 2004). The concordance between the rodent and human expression pattern suggests that FOXP2 was already playing such a role in the common ancestors of these species (Carlsson and Mahlapuu, 2002 and Zhang *et al.*, 2002).

IS FOXP 2 RELATED TO SCHIZOPHRENIA ?

Schizophrenia is a disorder of language may cause by mutation in FOXP2. Schizophrenia is a reflection of selective factors operating in the evolution of language. It is possible that FOXP2 could be responsible for

disorganization of brain. FOXP2 could be responsible for vulnerability to mental disorders, auditory hallucinations with language impairment such as schizophrenia and autism (Sanjuan *et al.*, 2006; Wassink *et al.*, 2002 and Tager-Flusberg *et al.*, 2001).

EVOLUTIONARY SIGNIFICANCE OF FOXP2 AND ITS FUNCTION

Despite the high conservation of FOXP2 coding sequence among mammals and the gene was involved in patterning corticostriatal or olivocerebellar circuitry many millions of years before the emergence of speech and language. Of the three amino acid changes that distinguish the mouse and human version of FOXP2 protein, two occur on the human lineage after the split between human and chimpanzee. At present it is unclear about the functional significance of the two amino acid changes that occur on the human lineage (Fig. 3 and Table 1). One of the changes has the potential to create a new site for the regulation of proteins by other proteins. Human specific version of FOXP2 is still essential for development of other tissues outside the brains. So, any modifications should have preserved the original function of gene to a large extent. When vocal communication was emerging in humans, changes in FOXP2 which improved motor sequencing and procedural learning abilities might have given them selective advantage

(Saleem *et al.*, 2003) Finally, the complexity of expression pattern may help to account for the fact that FOXP2 disruption leads to impairment in multiple aspects of CNS function (Bosman *et al.*, 2004). The FOXP2 protein probably act as a transcription factor binding to the regulatory region in the genomic DNA of target genes, modulating the rate at which the coding region of this genes are transcribed into messenger RNA (mRNA), which in turn serve as intermediate template that are then translated into proteins.

Conclusion

Common use of phrases like “Gene for Dyslexia” and “Language Gene” has helped many of those outside of molecular fields to come to erroneous conclusions (Grigorenko *et al.*, 1997). But it does not mean that such phrases have no value and should be outlawed. A basic knowledge of what genes can and cannot do is necessary to constrain an inform hypotheses about how genetic variability may relate to cognitive variability (whether in the normal range or with respect to disorder). Crucially, such hypotheses can actually play a fundamental role in our gene searches. For example if researchers had adopted a human specific “language gene” view when searching for the molecular basis of disorder in the KE family, then they may well have discounted FOXP2 at any early stage, since the gene is so highly conserved in

Table 1. Uniqueness of foxp2 amino acid sequence in human.

ORDERS	SPECIES	SITE	
		303	325
Galliformes	Chicken (<i>Gallus gallus</i>)	Thr	Asn
Artiodactyla	Cow (<i>Bos taurus</i>)	Thr	Asn
	Cat (<i>Felis catus</i>)	Thr	Ser
Carnivora	Dog (<i>Canis familiaris</i>)	Thr	Ser
	Bear (<i>Ursus maritimus</i>)	Thr	Ser
	Fox (<i>Alopex lagopus</i>)	Thr	Ser
Chiroptera	Bat (<i>Tadarida sp.</i>)	Thr	Asn
Rodentia	Mouse (<i>Mus musculus</i>)	Thr	Asn
Lagomorpha	Rabbit (<i>Sylvilagus floridanus</i>)	Thr	Asn
	Lemur (<i>Lemur catta</i>)	Thr	Asn
	Tamarin (<i>Saguinus oedipus</i>)	Thr	Asn
Primates	Monkey (<i>Macaca mulatta</i>)	Thr	Asn
	Orangutan (<i>Pongo pygmaeus</i>)	Thr	Asn
	Gorilla (<i>Gorilla gorilla</i>)	Thr	Asn
	Common chimp (<i>Pan troglodytes</i>)	Thr	Asn
	Pygmy chimp (<i>Pan paniscus</i>)	Thr	Asn
	Human (<i>Homo sapiens</i>)	Asn	Ser

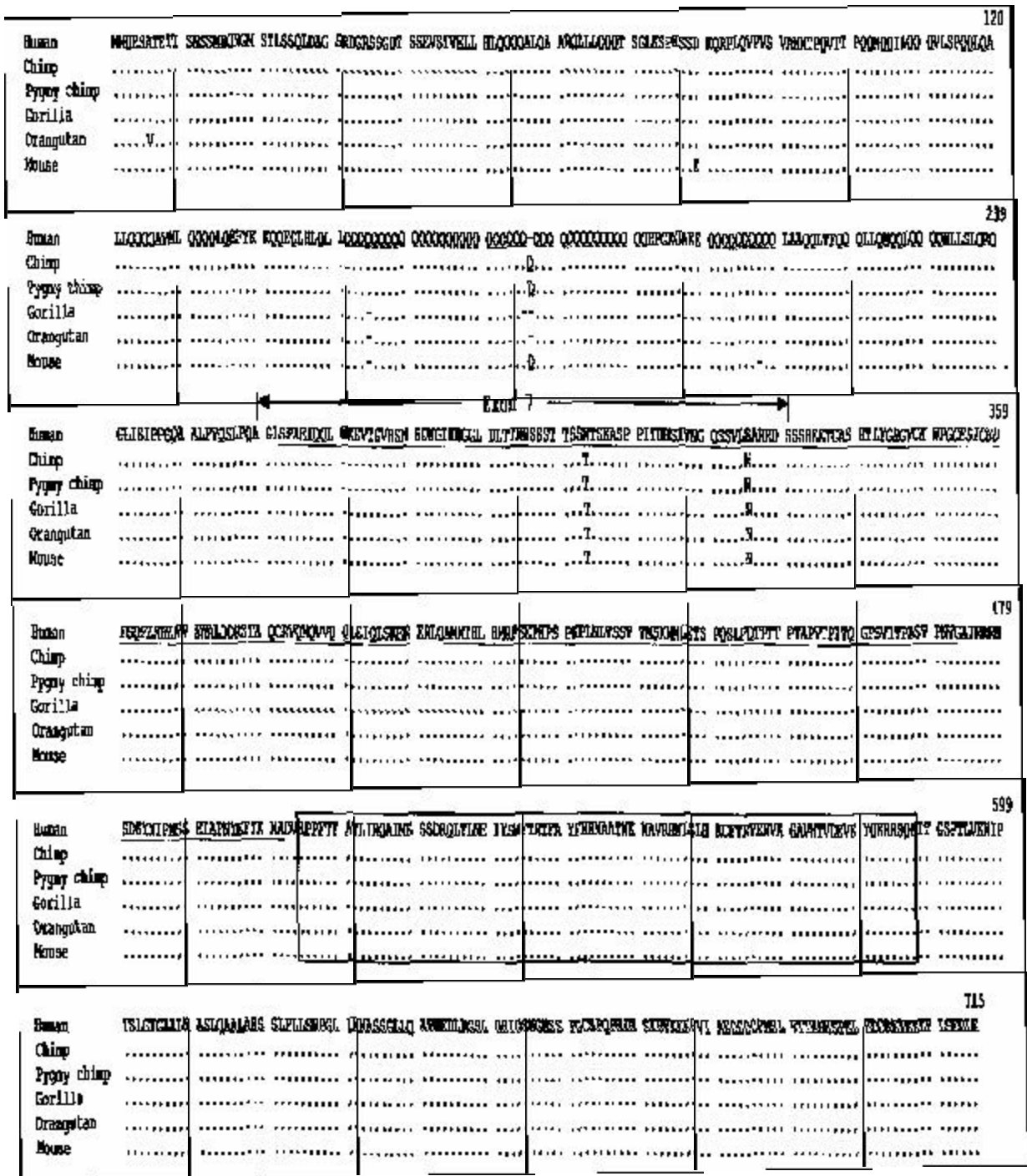


Fig. 3. A comparison of amino acid sequence of foxp2 in human, chimps, gorilla, orangutan and mouse. Exon 7 In figure showst the variation in amino acid sequence in human.

other nonlinguistic species. Moreover it is not obvious that cognitive scientists do become interested in the molecular pathways once the relevant genes are identified; even now discussion regarding the role of FOXP2 in human speech and language do not tend to exploit the relevant emerging data from molecular studies, but remain focused on approaches that where possible prior to isolation of the gene (Corballis *et al.* 2004). Putting

all this issues aside it remains to be seen whether or not we will ever identify a gene which appears to have specific influences on particular aspects of cognition such as phoneme awareness or grammatical processing. Considering of everything we have learned thus far from genetic studies of neurodevelopmental disorders leads to one inescapable conclusion. If we are able to place findings into a framework of system biology that

acknowledges molecular and developmental complexity, we will greatly improve our chances of untangling the webs that link gene to cognition (Consort, 2002)

Not much is known about details molecular action of the FOXP2 gene. Greater scientific research is needed to uncover the truth of “the language story”. Since now we cannot call the gene “the grammar gene”. We can only assume that with the help of new techniques such as neuroimaging, gene expression profiling, DNA sequencing etc. Scientists will be able to split the mystery open in near future.

REFERENCES

- Bosman, C., Garcia, R. and Aboitiz, F. (2004). FOXP2 and the language working memory system. *Trends in cognitive sciences*, 8: 251 – 252.
- Bruce, H. A. and Margolis, R. L. (2002). FOXP2: novel exons, splice variants and CAG repeat length stability. *Hum. Genet.*, 111: 136 - 144.
- Carlsson, P. and Mahlapuu, M. (2002). Forkhead transcription factors: key players in development and metabolism. *Dev. Bio.*, 250: 1- 23.
- Consort, S. L. I. (2002). A genomewide scan identifies two novel loci involved in specific language impairment *Am. J. Hum. Genet.*, 70: 384 – 398.
- Corballis, M. C. (2004). FOXP2 and the mirror system. *Trends in cognitive sciences*, 8: 95-96.
- Enard, W., Przeworski, M., Fisher, S.E., Lai, C. S. L., Wiebe, V., Kitano, T., Monaco, A. P. and Paabo, S. (2002). Molecular evolution of FOXP2, a gene involved in speech and language. *Nature*, 418: 869 – 872.
- Grigorenko, F. B., Wood, M. S., Meyer, L. A., Hart, W. C. and Shuster, A. (1997). Susceptibility loci for distinct components of developmental dyslexia on chromosomes 6 and 15. *Am. J. Hum. Genet.*, 60: 27 – 39.
- Hauser, M. D., Chomsky, N. and Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolved? *Science*, 298: 1569 – 1579.
- Lai, C. S., Fisher, S.E., Hurst, J.A., Vargna-khadem, F., Monaco, A.P. (2001). A forkhead – domain gene is mutated in a severe speech and language disorder. *Nature*, 413: 519 – 523.
- Lai, C. S. L., Fisher, S.E., Hurst, J.A., Levy, E.R., Hodgson, S., Fox, M., Jeremiah, S., Jamison, D.C., Gran, E.D., Vargna-khadem, F., Monaco, A.P. (2000). 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.
- Lai, C. S. L., Gerrelli, D., Monaco, A. P., Fisher, S. E. and Copp, A. J. (2003). FOXP2 expression during brain development coincides with sites of pathology in a severe speech and language disorder. *Brain.*, 126: 2455 – 2462.
- Lehmann, O. J. (2000). Chromosomal duplication involving the forkhead transcription factor gene FOXC1 causes Irish hypoplasia and glaucoma. *Am. J. Hum. Genet.*, 67: 1129 – 1135.
- Lieberman, P. (2002). On the nature and evolution of the neural bases of human language. *Year book of Physical Anthropology*, 45: 36 – 62.
- MacDermot, K.D., Bonora, E., Skyes, N., Coupe, A. and Fisher, E.S. (2005). Identification of FOXP2 truncation as a novel cause of developmental speech and language deficits. *Am. J. Hum. Genet.*, 76: 1074 - 1080.
- Marcus, G. F. and Fisher, S. E. (2003). FOXP2 in focus: what can genes tell us about speech and language? *Trends in Cognitive Sciences.*, 7: 257 – 262.
- Marcus, G. F., Vouloumanos, A. and Sag, I. A. (2003). Does Broca’s play by the rules? *Nature Neuroscience*, 6: 651 – 652.
- Saleem, R. A., Banerjee- Basu, S., Berry, F. B., Baxevanis, A. D. and Walter, M. A. (2003). Structural and functional analysis of disease causing missense mutation in the forkhead domain of FOXC1. *Journal of Human Molecular Genetics*, 12: 2993- 3005.
- Sanjuan, J., Tolosa, A., Gonzalez, J. C., Aguilar, E. J., Perez-Tur, J., Najera, C., Molto, M. D. and DeFrutos, R. (2006). Association between FOXP2 polymorphisms and schizophrenia with auditory hallucinations. *Psychiatric Genetics*, 16: 67 – 72.
- Shu, W., Yang, H., Zhang, L., Lu, M. M. and Morrissey, E. E. (2001). Characterization of 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.
- Smith, D. S. (2007). Genes, language development, and language disorder. *Journal of Mental Retardation and Developmental Disabilities*. 13: 96 – 105.
- Tager- Flusberg, H., Joseph, R. and Folstein, S. (2001). Current directions in research on autism. *Ment. Retard. Dev. Disabil. Res. Rev.*, 7: 21 – 29.
- Takahashi, K., Liu, F. C., Hirokawa, K. and Takahashi, H. (2003). Expression of FOXP2, a gene involved in speech and language in the developing and adult striatum. *J. Neurosci. Res.*, 73: 61– 72.
- Wassink, T. H., Piven, J., Vieland, V. J., Pietila, J. and Goedken, R. J. (2002). Evaluation of FOXP2 as an autism susceptibility gene. *Am. J. Med. Genet.*, 114: 566 - 569.
- Zhang, J., Webb, D. M. and Podlaha, O. (2002). Accelerated protein evolution and origins of human specific features: FOXP2 as an example. *Genetics*, 162: 1825 -1835.