

# 后基因组时代的生物语言学研究

李 慧

(北京师范大学珠海分校, 珠海 519085)

**提 要:** 近年来, 生物语言学再次成为国外语言学的研究热点。人类语言起源和进化是生物语言学研究的 5 个基本问题之一。*FOXP2* 的发现及过去 10 年来分子遗传学技术的迅猛发展宣告人类语言起源和进化研究的新时代已经到来。对语言相关基因进行鉴定并开展种间、种内进化研究, 将为理解人类语言起源和进化提供新的实证证据。文章综述生物语言学的研究历程以及后基因组背景下言语和语言障碍的遗传学研究进展, 分析利用语言相关基因进行语言起源和进化研究的思路, 讨论语言与基因的关系。

**关键词:** 生物语言学; 语言起源和进化; 言语和语言障碍; 语言相关基因

中图分类号: H05

文献标识码: A

文章编号: 1000-0100(2013)01-0008-8

## Biolinguistics in Post-genome Era

Li Hui

(Beijing Normal University <Zhuhai> Zhuhai 519085 ,China)

Biolinguistics has become the research hot topic in linguistic circles abroad in recent years. The origin and evolution of human language is one of the five fundamental questions that biolinguistics intends to answer. The discovery of *FOXP2* gene and advances in molecular genetic technology in the last decade have heralded a new era for human language evolutionary studies. The identification of the language-associated genes and a close examination of between — and within-species alteration of the genes in question will provide new empirical data in understanding the origin and evolution of human language. This paper provides a review of the research development in biolinguistics and the molecular genetic studies of speech and language disorders in post-genome era. The application of language-associated genes to the study of human language evolution is analyzed and the relationship between language and genes is also discussed about.

**Key words:** biolinguistics; origin and evolution of human language; speech and language disorders; language-associated gene

### 1 引言

“‘生物语言学’已经成为语言学领域甚至认知科学领域内出现最频繁的术语之一。”(Boeckx 2011: 449) 早在 1959 年, “生物语言学”这一术语已出现在 Meader 和 Muyskens 合著的 *Handbook of Biolinguistics* 一书中。该书“首次将生物学的研究成果与语言学结合起来界定‘生物语言学’”(吴文 2012: 36)。它讨论了言语产生参与组织的发育、言语器官的胚胎发育、言语过程的器官功能整合、神经肌肉系统整合、语言的决定因素等问题 (Meader & Muyskens 1950), 但当时所受关注度不高。上世纪 60 年代, Lenneberg(1967) 的著作 *Biological Foundation of Language* 《语言的生物学基础》是这一领域的基础性文献。他从脑神经

科学的角度研究语言产生的物质基础, 把儿童的语言发育看成受发音器官和大脑等神经机能制约的自然成熟过程, 并且提出了“关键期”假说。在该书中, Lenneberg 预见了许多目前备受关注的议题, 如语言习得的基因学、语言异常的基因学、双生子研究、失语症、语言演化等。1974 年第一届生物语言学大会由 Massimo Piattelli-Palmarini 组织, 与会者有语言学家、发育生物学家、神经生物学家和其他对语言的习得、起源、进化感兴趣的个人和团体。“在会上, Jenkins 极力倡导发行生物语言学研究的期刊, MIT 的科学团体也已成立了生物语言学的研究组织。”(Boeckx 2011: 449) Chomsky 从 20 世纪 50 年代以来一直采用生物语言学的研究方法研究内在语言, 他提出了语言获

得装置( language acquisition device , LAD)、生成语法( Generative Grammar)、标杠理论( X-bar theory)、管辖约束理论( government and binding theory)、原则和参数( principles and parameters)、最简方案( minimalist program) 等学说,以说明语言是以规则为基础的复杂交际系统。“普遍语法( universal grammar) 是一切人类语言所必须具有的原则、条件和规则系统,代表了人类语言最基本的东西。”( 刘润清 2002: 161) “一切语法规则、一切心理运算,最终都要表现为相应的人脑的物质机制。”( 刘润清 2002: 163) 因此对普遍语法的研究最终要对人脑物质进行科学探索。S. Pinker 和 L. Jenkins 是继 Chomsky 之后两位坚定的语言天赋论支持者。S. Pinker 认为,语言是一种本能( instinct) 而不是文化产物( cultural artifact) ( Pinker 1994)。“语言是大脑独特的生物组件,是儿童自发发展的复杂专门技能,没有有意识的训练和正规的指导,不理解其潜在的逻辑……一些认知科学家把语言描述成心理机能、心智器官、神经系统、计算机模块。在这里,我更愿意把它称为‘本能’,正如蜘蛛知道如何织网一样。”( Pinker 1994: 12) 他多次引用 Chomsky 的“普遍语法”概念,然而他认为,“Chomsky 对于语言机能的自然属性的论断基于词句结构的技术分析,流于深奥的形式主义( abuse formalism) ……一个结论的产生需要有各种证据的汇聚支撑”( Pinker 1994: 18)。在认知神经科学、语言学和生物学等领域都有建树的 S. Pinker 致力于通过心理语言学和神经语言学的研究来论证“语法基因”的存在,阐述语言和大脑神经网络的关系。“语法基因在大脑的某个时间、某个部位编码蛋白质或引发蛋白质的转录,从而指导、吸引或粘合神经元形成神经网络,在学习过程中配合突触协调( synaptic tuning),从而形成解决语法问题的必要的运算机制。”( Pinker 1994: 316) 同时,他不赞成 Chomsky 怀疑达尔文进化学说对语言本能的解释。L. Jenkins 对语言学与生物学的统一以及 Chomsky 提出的有关生物语言学的 5 个基本问题( 语言知识的构成、获得、使用、相关的大脑机制以及发展进化) 进行了有益的探索( Jenkins 2000)。进入 21 世纪,分子神经生物学和分子遗传学将被广泛地用来研究语言功能与疾病。王士元总结了语言研究和基因研究的 3 点联系:( 1) 遗传学与语言学都在研究自然信息系统,这两个系统都是层级性与适应性的。达尔文早就讨论过它们在演化上的相似性,当今的科学家还是继续为此着迷;( 2)

关于从基因到大脑再到语言的演化过程。遗传学与语言学都想了解从基因到神经回路再到语言的发展过程。FOXP2 的研究是个好的开始,让我们得知哪些基因是语言的基础;( 3) 基因和语言都能为我们提供研究、了解人类历史的一个窗口”( 王士元 2011: 115)。

## 2 关于语言起源问题的争论

人类语言的起源和进化是生物语言学研究的 5 个基本问题之一。目前,进化心理学家和语言学家已经对语言是一种自然发生体系达成了共识,然而对语言的起源和进化的理论框架存在分歧。Aaron 等列举了目前 4 种代表性的理论框架:( 1) 突变论( macro-mutation): 认为促使语言产生的神经结构突然改变,从而造成了人类语言与动物交流之间无法逾越的鸿沟。语言是进化的偶然事件,它的发生可能是因为突变导致了大的神经隆起,此事件堪比为宇宙起源的“大爆炸”理论。突变可能为一次性的大突变或长期沉寂的小突变的累积爆发;( 2) 渐变论( microevolutionary): 认为几千年来在自然选择的压力下产生了无数微小的基因变异,从而造成了表型的逐渐改变,这一过程包含随机突变、适应、选择、遗传传递;( 3) 新表型发生学说( neophenogenesis): 认为语言是基因和基因组以外的体系( 包括社会文化环境) 相互作用的产物,社会需求和文化因素可能在内在语言能力的表达以及推进语言进化到一个空前复杂的体系水平方面起了决定性作用;( 4) 功能变异理论( exaptation): 认为基因组内有无数潜在的遗传基因,只有具有环境适应优势的基因才被选择性表达。行使初始功能的基因可能被征用行使其他功能,如控制动作手势的神经基因也可以控制语言产生。( Aaron *et al.* 2006: 266) 总结起来,关于人类语言起源争论的焦点在“语言能力是否为人类独具; 人类语言能力是突变还是渐变的; 人类语言能力是自然选择的结果还是其他能力进化的副产品”( 董粤章 张韧 2009: 358)。

语言起源的实证理据有 3 方面的来源: 考古学来源,中间形式语言的语言重构( 历史语言学范畴) 以及语言进化限制因素的计算机模拟。这些研究取得了许多重要的结果,然而由于取证困难等因素导致了关于语言起源的研究长期没有突破性进展。生物研究领域内,在公元 2000 年前后的几年中,人和一些模式生物的基因组全序列相继被测定,积累了大量的基因组学数据。人类进入后基因组时代,分子遗传学技术的发展,为人类

语言的起源和进化研究不断提供新的实证数据。语言起源的分子遗传学研究手段有:比较灵长类基因组序列、比较灵长类基因表达调控、发育性交流障碍相关基因研究。本文着重综述近年来言语和语言障碍相关基因的研究进展以及对人类语言起源和进化的启示。

### 3 言语和语言障碍相关基因研究

虽然儿童具有天生的语言学习能力,然而有相当比例的儿童患有原因不明的语言相关缺陷,如智障、脑瘫、孤独症和听力损伤等。从病理学的角度研究语言,观察患有语言疾患的儿童,有助于认识儿童语言系统的内部结构,逆向寻找到语言功能相关基因。

#### 3.1 言语和语言障碍

美国精神障碍评估手册第四版(DSM-IV)(American Psychiatric Association 1994: 225)把言语和语言障碍分为5大类,即表达型语言障碍、表达-理解混合型语言障碍、构音障碍、口吃和待分类的交流障碍。“其中,言语障碍指口语中的发音、发声及言语节律性障碍,包含口吃(即语畅障碍)、构音障碍。语言障碍相比不易察觉,且原因复杂,它包括构词、造句、语义理解、语用方面的缺陷,可能影响非言语性语言能力(例如,发育性阅读和/或书写障碍,即阅读和/或书写能力被选择性地破坏,而说话及其他认知功能不受伤害的语言障碍),也可能影响语言的理解和表达。特殊性语言障碍(specific language impairment, SLI)是概括性术语,通常指表达型语言障碍和表达-理解混合型语言障碍,有时也包含构音障碍。”(Newbury 2010: 309)

虽然许多言语和语言障碍在临床上有区分,然而它们彼此高度共存。很多患儿不会纯粹属于某种诊断类别,且在语言发育过程中诊断类别可能会相互转换,如“约15%患永久性言语障碍的儿童同时患有语言障碍,约5%的SLI患者同时也有言语障碍”(Shriberg 1999: 1465),这些都导致了鉴定病因的困难。言语障碍和语言障碍之间、言语和语言障碍与其他发育问题(如孤独症、学习障碍、听力损伤等)之间的确切关系也颇有争议。目前,虽然言语和语言障碍的确切发病机制尚不清楚,但诸多研究显示遗传因素起主要作用,如“言语和语言障碍显示明显的家族聚集性;同卵双生子的同病一致率高于异卵双生子;语言的不同方面有中等以上的遗传度;而各种言语和语言障碍(如口吃和SLI)的遗传度通常都很高”

(Dediu 2011: 281)。鉴定言语和语言障碍相关基因对预测和治疗言语和语言障碍,探索语言习得和进化过程无疑具有重大价值。

#### 3.2 孟德尔遗传言语和语言障碍相关基因——FOXP2 基因

上世纪90年代初,媒体广泛报道了1个大型3代家族(称为KE家族)的许多成员患有由单基因突变引起的所谓的“语法特异障碍”(grammar-specific disorder)(Fisher 2006: 285)。这些报道源自1篇发表在Nature上的短讯,短讯称“该患病家族成员不能判断、理解和运用语法(1种被称为‘语言特征盲’的缺陷),此缺陷在患者自发性言语、书写和复述中都有表现,而根源可能在于潜在的语法程序发生了错误,且表现出单基因遗传模式”(Gopnik 1990: 715)。该短讯被认为是人类具有“语法基因”的难得的一手证据,似乎“普遍语法”这一语言学概念已获得生物学研究提供的实证理据。尽管当时还没有鉴定出导致这一障碍的基因,但媒体已高呼发现了“语法基因”。

后续研究发现“KE家族所患的言语失用症(developmental verbal dyspraxia (DVD)或childhood apraxia of speech(CAS))是由于言语肌肉运动的位置、范围、协调性和运动序列的编程能力受损产生的运动性言语障碍。除了言语问题,KE家族中患病成员还有表达-理解混合型语言障碍,某些病例还有语言书写缺陷以及非言语性认知损伤”(Watkins 2002: 458)。Newbury简述了DVD相关基因——FOXP2的鉴定及功能研究过程:“Fisher通过全基因组连锁分析(genome-wide linkage analysis, GWLA)将该基因初步定位于7q31处,Lai等通过基因精确定位发现所有KE家庭患者的FOXP2基因都有突变,而未患病家庭成员没有。紧接着,研究者又发现一位患有类似言语障碍的非KE家族儿童,他的染色体重排导致了FOXP2基因的断裂,FOXP2基因和言语失用症的相关性从而得到支持。此后,又有一些独立的FOXP2基因筛查研究鉴定出了另外几例FOXP2断裂的病例,他们都有言语失用症状。KE家族携带FOXP2的杂合性错义突变,从而改变了FOXP2蛋白的DNA结合区域。另有研究表明,FOXP2的另一处杂合性无义突变也和言语和语言障碍有关,这种无义突变严重地截短了蛋白,剪除了必需的功能模序,包括蛋白质相互作用区、DNA结合区和可能的核定位信号。目前,虽然FOXP2在言语失用症中的真正作用有待明晰,但可以确定的是该基因对负责精细动作控制的大脑

区域(运动皮层、纹状体、小脑)的发育非常重要,它的断裂对言语发育会产生极其严重的后果”(Newbury 2010: 310-311)。

*FOXP2* 的鉴定开启了覆盖语言学、神经影像学、动物模型学(主要是小鼠和鸣禽)、分子生物学、人类学和进化学等领域的研究。*FOXP2* 基因是否就是所谓的“语言基因”或“语法基因”呢? Fisher 对此进行了驳斥,且给出了充分的论据:“首先,*FOXP2* 基因并非人类特有,在哺乳动物世界中它以高度保守的形式存在,如鼠的 *FOXP2* 蛋白版本与人类只有 3 个氨基酸位点的差别,这与 *FOXP2* 仅仅是为了人类的言语和语言能力而存在的观点不符;其次,对 *FOXP2* 的表达研究还发现,其功能的重要性不仅仅体现在中枢神经系统,它在胚胎发育过程中对肺、心脏、肠的发育也发挥着重要的调控作用。作为转录因子,*FOXP2* 在不同的发育背景下调控多重生物学通路,这也是复杂生物学系统的普遍特征;最后,*FOXP2* 在脑中的表达也极为复杂,在人类和啮齿类动物的早期发育中,*FOXP2* 在脑区广泛表达,包括皮质、丘脑、下丘脑、纹状体、小脑和髓质,表达一直持续到成年期。因此,人类的 *FOXP2* 基因不仅影响传统认为的与语言相关皮层区域的发育和功能,也影响其他的皮层区域。所以,“*FOXP2* 只局限于 Broca 区”(Fisher 2006: 288)。

### 3.3 FOXP 家族其他基因以及 FOXP2 蛋白调控基因

#### 3.3.1 FOXP1

在 *FOXP2* 数据的基础上,研究者认为其他的 *FOX* 基因也是言语和语言损伤的候选基因。“人类 *FOX* 基因家族有超过 40 个成员,根据 DNA 结合区域的特定模序分成 19 个亚族(命名为 *FOXA* 到 *FOXS*)。*FOXP* 亚族包含 4 种功能各异的基因(*FOXP1-4*)。这些基因编码的蛋白彼此结合形成活跃的 DNA 异源二聚体”(Hannenhalli 2009: 236)。“*FOXP1* 和 *FOXP2* 彼此联系紧密,在调节组织发育中彼此配合”(Shu *et al.* 2007: 1992)。Pariani 等描述了一位由于 3 号染色体缺失导致 *FOXP1* 基因断裂和其他 3 个转录本(*ELF4E3*, *GPR27* 和 *PROK2*) 缺失的病人,最初的临床诊断是先天性睑裂狭窄综合症和先天性多发性关节挛缩症。然而,该患者还呈现发育迟缓及言语和语言缺陷的症状,被认为可能由 *FOXP1* 断裂引起。(Pariani *et al.* 2009: 123-127) 后来, Vernes 等对临床诊断有发育性言语失用症的 49 位患者的 *FOXP1* 编码区进行测序,仅发现在 1 例病例中

*FOXP1* 编码区的错义改变,且这种改变在未患病的对照组中也存在(Vernes *et al.* 2009: 1354-1358)。另有研究描述了一位患有小脑扁桃体异常和癫痫的病人。这个患者最初由于言语发育迟缓而就诊,结果检查出 *FOXP1* 的缺失(Carr 2010: 1216-1220)。“虽然此病例与 Pariani 等所观察的病人的临床表现大相径庭,研究者普遍认为,Pariani 观察的病例所表现出来的先天性睑裂狭窄综合症,先天性多发性关节挛缩症可能是因为其他转录本的缺失引起的。他们继续推测 *FOXP1* 断裂可以解释患者临床表现相似的那一部分,称为运动发育缺陷和言语迟缓缺陷。”(Newbury 2010: 313)

#### 3.3.2 CNTNAP2

“*FOXP2* 蛋白由锌指结构、叉头 DNA 结合区(叉头框)和多聚谷氨酸富集区 3 部分组成,充当转录因子,有 4 种选择性亚型,调控 300 至 400 个神经系统基因,其中许多在功能上被视为言语和语言障碍的候选基因。”(Newbury 2010: 311) 受 *FOXP2* 调控的基因包含 1 个 7 号染色体上的 *CNTNAP2* 基因。最近研究显示,该基因与语言损伤易感性有关。该研究发现“*CNTNAP2* 的 9 种常见变异和 1 组语言损伤家庭在语言测试(包括语言理解和表达测试、语音短时记忆测试等)中的得分低有显著相关性”(Vernes 2008: 2339)。另外,有人发现,“该基因的选择性变异(包含常见变异和罕见断裂或突变)也与一系列的神经发育障碍,如自闭症、精神分裂症、癫痫、注意力缺陷多动障碍(ADHD)和学习障碍有关”(杨曹骅 2012: 351)。

“*CNTNAP2* 基因编码轴突蛋白,负责发育神经元钾离子通道的定位,为轴突-胶质细胞相互作用提供便利。”(杨曹骅 2012: 352) “大脑表达研究显示,该基因在啮齿类动物大脑中均匀表达,然而在雄性鸣禽发声控制核团中呈现特异性表达,在人脑前额叶(与智力密切相关的脑区)富集。”(Panaitof 2010: 2001) “MRI 研究表明,携带两个拷贝的 *CNTNAP2* 变异基因的孤独症病人,在多个脑区域(如前额叶、梭状脑回、额枕皮质、小脑)的灰白质显著减少,这些均与自闭症状有密切关系,同时也佐证了 *CNTNAP2* 基因的遗传突变在孤独症核心症状之一的语言发育障碍方面可能占到主导作用。”(杨曹骅 2012: 353)

#### 3.4 多基因遗传言语和语言障碍相关基因

单基因突变导致的语言遗传疾病非常罕见。研究者通常认为,言语和语言障碍疾病都是多基

因遗传疾病。遗传性状的形成,不仅取决于两个以上微效基因的累加作用,同时还受环境因子的影响,因此相关基因的鉴定非常困难。然而,随着复杂表型鉴定技术、基因分型技术以及统计方法的不断进步,已经取得了丰硕的成果。

### 3.41 SLI 相关基因

“SLI 是指在正常语言学习环境中,在没有听力缺陷、智力迟钝、神经或精神损伤的情况下言语能力得不到正常发展。”(Law *et al.* 2000: 165)

运用 GWLA 分析,研究者已发现和 SLI 相关的 3 个主要位点,它们分别位于 13 号染色体 (*SLI3*)、16 号染色体 (*SLI1*) 和 19 号染色体 (*SLI2*)。目前,16 号染色体的 *SLI1* 区是研究得最为深入的。“在此区域进行关联分析发现, *ATP2C2* 和 *CMIP* 两个基因的常见变异与 *SLI1* 有显著而独立的相关性,它们都与语音短期记忆相关。”(Newbury 2009: 267)

Dash 等发现,“*ATP2C2* 编码钙离子 ATP 酶,负责细胞钙离子和镁离子的运输。钙离子是重要的胞间信使,参与许多包括短期记忆的神经过程。镁离子在高浓度时对神经细胞有致毒性,镁离子缺失与自闭症相关”(Dash *et al.* 2007: 555)。“*CMIP* 蛋白是细胞骨架的适配蛋白,在大脑中有大量表达,而细胞骨架重建在突触形成和神经迁移中起关键作用;*CMIP* 蛋白和 *FilaminA* 相互作用,后者的突变与神经元移行障碍的脑室周围灰质异位有关。”(Heng 2010: 43) 其他研究发现,“*CMIP* 也和 *RelA* (转录因子 *NF- $\kappa$ B* 的亚单位)、*P13* 激酶蛋白相互作用”(Kamal 2009: 993)。这说明 *CMIP* 在多种生物途径中发挥作用。然而,“*CMIP* 和 *ATP2C2* 都非 *FOXP2* 的调控基因,两基因之间以及它们和 *CNTNAP2* 基因之间也没有明显的关系”(Vernes 2008: 2342),这可能预示着语言发育的新的生物学途径。

### 3.42 口吃相关基因

“口吃是语畅障碍,其特征是连续语流的打断和不自主地重复或延长音节。”(Prasse 2008: 1271) GWLA 分析和连锁分析发现,染色体 12q 处与口吃有稳定而显著的相关性 (Riaz *et al.* 2005: 650)。Kang 等通过对此位点的 45 个基因测序显示口吃患者的 *GNPTAB* 基因的 4 个编码区突变,这些突变在对照组中未被发现 (Kang *et al.* 2010: 677 - 685)。之前,Paik 等有研究显示,“*GNPTAB* 和 *GNPTG* 基因突变分别会导致代谢疾病、粘脂糖症 II 型和粘脂糖症 III 型。它们是一种溶酶体贮积病,是溶酶体酶 N-乙酰基-1-磷

酸转移酶 (GlcNAc-phosphotransferase, GlcNAc-PT) 活性缺失使得溶酶体酶不能正常进入溶酶体的常染色体隐性遗传病。*GlcNAc-PT* 蛋白以复合体形式存在,其中  $\alpha/\beta$  亚基由 *GNPTAB* 编码, $\gamma$  亚基由 *GNPTAG* 编码。位于染色体 16q 位点的 *NAGPA* 编码 N-乙酰氨基葡萄糖-1-磷酸二脂  $\alpha$ -N-乙酰葡萄糖胺糖苷酶,为 *GlcNAc-PT* 的下游酶”(Paik *et al.* 2005: 310)。“对口吃患者的 *GNPTG* 和 *NAGPA* 编码区测序发现有突变位点,而在对照组中没有。这些研究暗示了言语障碍的另一种生物学机制——溶酶体信号通路。”(Kang 2010: 683)

## 4 从言语和语言障碍相关基因研究语言起源和进化

首先确定人类发育性言语和语言障碍相关基因,然后以这些基因为目标,进行种间及种内对比,研究其进化机制和进化历程。

人类大脑容量在灵长类进化中显著增长,其内在分子机制在 *MCPH* 疾病相关基因的研究中得到了启示。“*MCPH* 疾病的主要特征是病人的头比正常人小得多,相当于早期古猿的脑容量大小。”(Tang 2006: 911) “*MCPH* 疾病有多基因基础,已知其中至少 4 个(包括 *MCPH1*, *ASPM*, *CDK5RAP2* 和 *CENPJ*) 在从猿到人的灵长类动物中经历了显著的正选择。”(Tang 2006: 911) 如 Evans 等把人类 *ASPM* 和其他 6 种灵长类动物的 *ASPM* 进行了比较研究,发现 *ASPM* 在大猿系列进化明显加剧,而从猿类到人类尤为明显。此外,从人类最近的祖先到现代人类,其 *ASPM* 基因的  $Ka/Ks > 1$  说明受到显著的正选择作用。运用 McDonald-Kreitman 检测种间多态性和种内分歧进一步证实了 *ASPM* 在最近人类进化中经历了强烈的正选择。该项测试还显示自从 5 到 6 百万年前人猿分离后, *ASPM* 平均每 300,000 到 400,000 年固定会有一个优势氨基酸变化,这符合 *ASPM* 基因在人类进化中调控大脑容量的推断。(Evans *et al.* 2004: 489 - 494) “*MCPH1* 在 DNA 受损反应和调控细胞周期检测点中发挥作用,而其他基因编码微管蛋白相关的中心体蛋白,从而调控神经元前体细胞的对称细胞分裂。”(Tang 2006) 2005 年 *ASPM* 和 *MCPH* 基因由于自然选择导致的“最新型演化基因型(简称为 *ASPM-D* 和 *MCPH-D*)”分别被报道,研究者发现“它们呈现出倾向性的地理分布”(Evans 2005: 1717 - 1720)。Dediu 等提出声调语言分布倾向与 *ASPM* 和 *MCPH1* 相关,

即持原始型(ancestral forms) *ASPM* 和 *MCPH* 基因的现代人倾向说声调语言,而持 *ASPM-D* 和 *MCPH-D* 基因型的人倾向说非声调语言(Dediu *et al.* 2007: 10944 – 10949)。

此思路在 Enard 等对 *FOXP2* 基因的进化研究中也得到证实: *FOXP2* 基因序列在脊椎动物中是高度保守的,表明其功能在进化过程中受到强烈限制(Enard *et al.* 2002: 869 – 872)。尽管这种存在很保守,人类世系中 *FOXP2* 氨基酸替换率依然有显著增长。对人类替换氨基酸序列两侧的内含子序列进行研究发现在过去 20 万年间可能发生了选择性清除,即替换氨基酸序列或者其两侧内含子序列在人类进化史中经历了正选择。人类 *FOXP2* 与鼠的 *FOXP2* 间有 3 个氨基酸的不同,与黑猩猩、大猩猩和恒河猴等灵长类动物的 *FOXP2* 之间有两个不同的氨基酸。数据分析显示后者差异的“出现和在人类世系中快速传播约发生在过去 20 万年内的某个时间点,这似乎与之前对人类语言和言语出现的时间点(约 5 万年前到 10 万年前之间)的推断相吻合”(Boyd & Silk 2000: 135)。这些数据再次催生了媒体报道中“语言基因”标签论的复苏,有学者认为“人类 *FOXP2* 氨基酸上的微小替换导致了人类 *FOXP2* 功能的根本性变化,从而成为人类语言进化的主要驱动力”(Fisher 2006: 289)。然而,生物学研究再次对此观点提出了质疑。“首先,人类 *FOXP2* 蛋白中这两个氨基酸的功能重要性尚未明确,因为它们位于该蛋白有明显功能特征的结构域之外。其次,由于 *FOXP2* 对除脑以外人类其他组织的发育也是必需的,所以该基因的任何改变都应该很大程度地保持其原始功能。”(Fisher 2006: 289)

## 5 思考与展望

“探索意识的生物学基础是 2005 年到 2030 年亟待解决的 25 个重大科学课题之一,作为意识最易管窥的反映物 and 最重要的载体之一,语言的生物机制是意识生物基础的关键一环。”(董粤章 张韧 2009: 355) 从言语和语言障碍相关基因探索语言起源和进化是分子遗传学研究语言机制的视角之一。目前,普遍认为常见的言语和语言障碍是 1 种复杂的多基因遗传疾病,由许多微效基因相互作用,并由环境致病因子诱发,共同导致言语和语言障碍的发生。过去科学家通过 GWLA 和候选基因关联分析寻找到了如上所述的言语和语言障碍相关基因。然而,“连锁分析的成功依赖

于几个重要条件:遗传因素完全或接近完全决定疾病的发生,致病基因具有很高的外显率等”(唐劲松等 2010: 57)。在单基因遗传疾病的研究中,GWLA 是非常实用的方法,而且成功地鉴定了很多单基因疾病的基因,如 *FOXP2*。然而单基因语言疾病相对比较罕见,而常见的语言疾病基本都是复杂的多基因疾病。因而连锁分析在寻找复杂疾病的致病基因的过程中收效甚微。“关联分析可以检测一个比较短的基因组区域(约 100 Kb, 千碱基)是否包含致病基因,但研究的基因非常有限。”(唐劲松 2010: 184) “近几年来,随着人类基因组测序计划和人类基因组单体型图(HapMap)计划的完成为全基因组关联分析(genome-wide association study, GWAS)提供了条件。”(韩建文 张学军 2011: 26) 然而,“目前基于 HapMap 设计的 GWAS 的研究对象主要是 SNPs,对 SNPs 以外的其他变异(如拷贝数变异、基因表达、表观遗传修饰)检出效能非常微弱”(权晟 张学军 2011: 101),容易造成遗传遗失(missing heredity),须要采取深度分析策略,发现更多的言语和语言障碍易感基因。

此外,前述言语和语言障碍相关基因均参与不同的生物学发展和变异,在不同发育过程和疾病中起作用。因此须要对上述基因进一步进行功能上的鉴定,如采取在细胞或动物模型中进行基因操作,以梳理出这些基因对语言的确切功能。关于动物模型的选择,“除了来源有限的黑猩猩外,在啮齿类动物、鸣禽以及其他发声学习者身上也可以开展分子、解剖、发育和行为等多水平的研究”(Fisher & Marcus 2006: 18)。同时,以鉴定的语言相关基因为研究对象,进行进化分析,从不同层面探索种间-种内的相似性及差异性以及人类语言功能的进化限制因素,阐述语言起源、进化机制以及进化历程。

“图书馆有大量讨论有关人类语言的文献以及各种关于语言进化的版本,与之形成鲜明对比的是,很少有人关注蜜蜂交流系统的进化。”(Chomsky 2007: 14) 语言学家着眼于人类语言的发声学、形态学以及语义、语用等语言原则领域的研究,很少涉及其他物种的发声生物学和心理学。神经科学家只关注人类语言的神经机制,专注于研究大脑皮层负责处理语言的两个区域(Broca 区和 Wernicke 区)。这些努力对认清语言的本质,探测句子产生或理解时大脑中发生的电生理活动是非常有效的。然而,没有哪一种表型是脱离进化历程的。如有充分的分子生物学和发育生

物学证据证明,“脊椎动物的 4 肢是由鱼类的鳍循序渐进演化而来,而且由 *Hox* 基因群掌控整个过程”(Nicolas 2003: 589)。越来越多的遗传学研究证明,现代人类非凡的语言功能同样也不是一项突然出现的创举。目前,已经清楚的是“人类语言的出现一定经历了漫长的灵长类动物祖先形态上的量变和质变,如脑容量的改变、神经关联强度的改变、新的神经通路的产生、声音感觉机制的形态学改变;而形成这些改变的分子机制有基因型改变、基因表达时空改变和基因表达量的改变”(Fisher 2006: 270)。然而,基因表达产物为调控因子、信号分子、受体、酶等,它们之间形成高度复杂的作用网络且受环境调节,从而构建和维系大脑功能,包括大脑的高级功能——语言。因此,基因和语言的关系是间接而复杂的。生物语言学研究语言功能的生物学机制,已成为语言研究的另一种范式和视角,充满生命力。“把生物学引入语言学中,或许有助于对语言本质的揭示,促进语言学进一步发展。”(袁晓红 刘桂玲 2008: 106)而关于语言起源的问题,由于古人类学、考古学、分子生物学等不同专业学者所依赖的证据不同,至今没有定论,呈现出百花争鸣的局面。

#### 参考文献

- 董粤章 张 韧. 语言生物机制研究的新视野: *FOXP2* 与人类语言能力 [J]. 东北大学学报(社会科学版), 2009(4).
- 韩建文 张学军. 全基因组关联研究现状 [J]. 遗传, 2011(1).
- 刘润清. 西方语言学流派 [M]. 北京: 外语教学与研究出版社, 2002.
- 权 晟 张学军. 全基因组关联研究的深度分析策略 [J]. 遗传, 2011(2).
- 唐劲松 陈晓岗 刘春宇. 精神疾病的全基因组关联分析 [J]. 中国神经精神疾病杂志, 2010(3).
- 王士元. 语言与基因 [OL]. 2011. <http://comonca.org.cn/Abs/2011/019.htm>.
- 吴 文. “生物语言学”及术语考究 [J]. 中国科技术语, 2012(2).
- 杨曹骅. CNTNAP2 基因在孤独症发病机制中的作用 [J]. 中国儿童保健杂志, 2012(4).
- 袁晓红 刘桂玲. 语言研究中的生物观 [J]. 外语学刊, 2008(4).
- Aaron, P. G., & Joshi, R. M. Written Language is as Natural as Spoken Language: A Biolinguistic Perspective [J]. *Reading Psychology*, 2006(4).

- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, 4th ed (DSM-IV)* [Z]. Washington, DC: American Psychiatric press, 1994.
- Boeckx, C. Biolinguistics: A Brief Guide for the Perplexed [J]. 语言科学, 2011(5).
- Boyd, R. & Silk, J. B. *How Humans Evolved* [M]. New York: W. W. Norton, 2000.
- Carr, C. W., Moreno-De-Luca, D., Parker, C., Zimmerman, H. H., Ledbetter, N., Martin, C. L., Dobyns, W. B. & Abdul-Rahman, O. A. Chiari I Malformation, Delayed Gross Motor Skills, Severe Speech Delay, and Epileptiform Discharges in a Child with FOXP1 Haploinsufficiency [J]. *Eur. J. Hum. Genet*, 2010(18).
- Chomsky, N. Of Minds and Language [J]. *Biolinguistics*, 2007(1).
- Dash, P. K., Moore, A. N., Kobori, N. & Runyan, J. D. Molecular Activity Underlying Working Memory [J]. *Learn Mem*, 2007(3).
- Dediu, D. & L. R. Ladd. Linguistic Tone is Related to the Population Frequency of the Adaptive Haplogroups of Two Brain Size Genes, ASPM and Microcephalin [J]. *Proc Natl Acad Sci U S A*, 2007(26).
- Dediu, D. Are Languages Really Independent from Genes? If Not, What Would a Genetic Bias Affecting Language Diversity Look Like? [J]. *Human Biology*, 2011(2).
- Enard, W., Przeworski, M., Fisher, S. E., Lai, C. S., Wiebe, V., Kitano, T., Monaco, A. P., & Pääbo, S. Molecular Evolution of FOXP2, A Gene Involved in Speech and Language [J]. *Nature*, 2002(6900).
- Evans, P. D., Anderson, J. R., Vallender, E. J., Gilbert, S. L., Malcom, C. M., Dorus, S., & Lahn, B. T. Adaptive Evolution of ASPM, A Major Determinant of Cerebral Cortical Size in Humans [J]. *Hum Mol Genet*, 2004(5).
- Evans, P. D., Gilbert, S. L., Mekel-Bobrov, N., Vallender, E. J., Anderson, J. R., Vaez-Azizi, L. M., Tishkoff, S. A., Hudson, R. R., Lahn, B. T. *Microcephalin*, A Gene Regulating Brain Size, Continues to Evolve Adaptively in Humans [J]. *Science*, 2005, (309).
- Fisher, S. E., & Marcus, G. F. The Eloquent Ape: Genes, Brains and the Evolution of Language [J]. *Nat Rev Genet*, 2006(1).
- Fisher, S. E. Tangled Webs: Tracing the Connections between Genes and Cognition [J]. *Cognition*, 2006(2).
- Gopnik, M. Feature-blind Grammar and Dysphasia (correspondence) [J]. *Nature*, 1990(6288).

- Hannenhalli, S. & Kaestner, K. H. The Evolution of Fox Genes and Their Role in Development and Disease [J]. *Nat Rev Genet*, 2009(4).
- Heng, J. I., Chariot, A., Nguyen, L. Molecular Layers Underlying Cytoskeletal Remodelling during Cortical Development [J]. *Trends Neurosci*, 2010 (1).
- Jenkins, L. *Biolinguistics: Exploring the Biology of Language* [M]. Cambridge: Cambridge University Press, 2000.
- Kamal, M., Valanciute, A., Dahan, K., Ory, V., Pawlak, A., Lang, P., Guellaen, G. & Sahali, D. C-mip Interacts Physically with RelA and Inhibits Nuclear Factor Kappa B Activity [J]. *Mol Immunol*, 2009(5).
- Kang, C., Riazuddin, S., Mundorff, J., Krasnewich, D., Friedman, P., Mullikin, J. C. & Drayna, D. Mutations in the Lysosomal Enzyme-targeting Pathway and Persistent Stuttering [J]. *N Engl J Med*, 2010(8).
- Law, J., Boyle, J., Harris, F., Harkness, A. & Nye, C. Prevalence and Natural History of Primary Speech and Language Delay: Findings from a Systematic Review of the Literature [J]. *Int J Lang Commun Disord*, 2000 (2).
- Lenneberg, E. H. *Biological Foundations of Language* [M]. New York: John Wiley & Sons, Inc., 1967.
- Meader, C. L. & Muyskens, J. H. *Handbook of Biolinguistics* [M]. Toledo: Herbert C Weller, 1959.
- Newbury, D. F., Winchester, L., Addis, L., Paracchini, S., Buckingham, L. L. & Clark, A. CMIP and ATP2C2 Modulate Phonological Short-term Memory in Language Impairment [J]. *Am J Hum Genet*, 2009 (2).
- Newbury, D. F. & Monaco, A. P. Genetic Advances in the Study of Speech and Language Disorders [J]. *Neuron*, 2010(2).
- Nicolas, S., Papillon, D., Perez, Y., Caubit, X. & Le Parco Y. The Spatial Restrictions of 5' HoxC genes Expression are Maintained in Adult New Spinal Cord [J]. *Biol Cell*, 2003 (9).
- Paik, K. H., Song, S. M., Ki, C. S., Yu, H. W., Kim, J. S., Min, K. H. & Chang, S. H. Identification of Mutations in the GNPTA (MGC4170) Gene Coding for GlcNAc-phosphotransferase Alpha/beta Subunits in Korean Patients with Mucopolidiosis Type II or Type IIIA [J]. *Hum Mutat*, 2005(4).
- Panaitof, S. C., Abrahams, B. S., Dong, H., Geschwind, D. H. & White, S. A. Language-related Cntnap2 Gene is Differentially Expressed in Sexually Dimorphic Song Nuclei Essential for Vocal Learning in Songbirds [J]. *J Comp Neurol*, 2010 (11).
- Pariani, M. J., Spencer, A., Graham, J. M. & Rimoin, D. L. A 785kb Deletion of p14.1p13, Including the FOXP1 Gene, Associated with Speech Delay, Contractures, Hypertonia and Blepharophimosis [J]. *Eur J Med Genet*, 2009 (2-3).
- Pinker, S. *The Language Instinct: How the Mind Creates Language* [M]. New York: Harper Collins, 1994.
- Prasse, J. E. & Kikano, G. E. Stuttering: An Overview [J]. *Am Fam Physician*, 2008(9).
- Riaz, N., Steinberg, S., Ahmad, J., Pluzhnikov, A., Riazuddin, S., Cox, N. J., Drayna, D. Genomewide Significant Linkage to Stuttering on Chromosome 12 [J]. *Am J Hum Genet*, 2005 (4).
- Shriberg, L. D., Tomblin, J. B., McSweeney, J. L. Prevalence of Speech Delay in 6-year-old Children and Comorbidity with Language Impairment [J]. *J Speech Lang Hear Res*, 1999(6).
- Shu, W., Lu, M. M., Zhang, Y., Tucker, P. W., Zhou, D. & Morrissey, E. E. Foxp2 and Foxp1 Cooperatively Regulate Lung and Esophagus Development [J]. *Development*, 2007 (10).
- Tang, B. L. Molecular Genetic Determinants of Human Brain Size [J]. *Biochem Biophys Res Comm*, 2006 (3).
- Vernes, S. C., Newbury, D. F., Abrahams, B. S., Winchester, L., Nicod, J., Groszer, M., Alarcón, M., Oliver, P. L., Davies, K. E., Geschwind, D. H., Monaco, A. P. & Fisher, S. E. A Functional Genetic Link between Distinct Developmental Language Disorders [J]. *N Engl J Med*, 2008 (22).
- Vernes, S. C., MacDermot, K. D., Monaco, A. P. & Fisher, S. E. Assessing the Impact of FOXP1 Mutations on Developmental Verbal Dyspraxia [J]. *Eur J Hum Gene*, 2009(10).
- Watkins, K. E., Dronkers, N. F. & Vargha-Khadem, F. Behavioural Analysis of an Inherited Speech and Language Disorder: Comparison with Acquired Aphasia [J]. *Brain*, 2002(3).