

语言与基因:论南岛语族的起源与扩散

范志泉 邓晓华 王传超

【摘要】南岛语族的起源与扩散问题在语言学、考古学等相关领域一直是个具有争议性的重大理论问题。语言和基因演变的相关性,是语言学和分子人类学的研究领域之一。本文从理论、方法以及材料的角度,梳理了历史语言学视角下南岛语族起源与扩散问题的不同假说,并评述了前人时贤的研究成果及局限性。并在此基础上尝试从跨学科协作角度,结合遗传学、分子人类学最新的研究成果,提出对南岛语族起源与扩散问题的新思考。侗台-南岛语人群的先民可追溯到大陆华南地区,在距今5000-6000年前,开始对外迁移与扩散,其中一支向东迁徙至台湾,经菲律宾、印度尼西亚东部后扩散至远大洋洲,产生了早期Lapita文化;另一支向西,经由中南半岛、东南亚岛屿走廊后迁徙至印度尼西亚东部。现今的东南亚诸多族群是以上两组人群与操巴布亚语的土著人群混合而成的。而留在华南地区的演化为侗台语族等,侗台人群又与南下的北方汉族混合形成了南方汉族。

【关键词】语言;基因;南岛语族;演化

【作者简介】范志泉,厦门大学人类学与民族学系博士研究生;邓晓华,福建工程学院特聘教授;王传超,厦门大学人类学与民族学系教授(福建 厦门 361005)。

【原文出处】《学术月刊》(沪),2018.10.175~184

【基金项目】本文为国家社科基金重点项目“台湾原住民问题研究”(11AZD069)的阶段性成果;受厦门大学南强青年拔尖人才支持计划(X2123302)、中央高校基本科研业务费(ZK1144)与国家自然科学基金青年项目“中国东南各族群的遗传混合”(31801040)支持。

一、引言

南岛语系(Austronesian),亦称马来-波利尼西亚语(Malay-Polynesian)遍布于太平洋和印度洋中的岛屿以及东南亚的中南半岛和马来群岛等地,约有4亿人口,语言总数达1200多种;北起台湾、夏威夷岛,南至新西兰,东达复活节岛,西抵马达加斯加。南岛语族一般以语言学分类法为主要依据,分为东、西两大区。即东部大洋洲族群(Oceanic),包括美拉尼西亚、密克罗尼西亚、波利尼西亚三区;西部赫斯佩拉尼西亚族群(Hesperonesian),旧称印度尼西亚(Indonesian),包括中国台湾地区^①、菲律宾、印度尼西亚、马来西亚等地(见图1)。^②

自南岛语族共同体确认以来,学术界对南岛语族的研究已有百年历史。在众多研究课题中,最引人注目也备受争议的焦点之一,莫过于南岛语族的起源与扩散问题。有关南岛语族的起源与扩散问题,起初便是来自语言学的实践。但是由于语言学者们囿于各自理论、方法以及材料的限制,在该问题的研究方面处于各自



图1 南岛语族的地理分布图(Diamond 1988)^③

为政、各说各话的状态而始终未达成共识。

至20世纪前半期,一些学者如奥地利史前学者 Geldern^④,考古学者 Beyer^⑤,我国民族学者林惠祥^⑥、凌纯声^⑦,日本学者鹿野忠熊^⑧等,分别从不同角度将南岛语人群的起源与扩散研究都联系到了中国大陆东南部地区。随着考古资料的增多,澳洲学者 Bellwood 依据更新的考古资料,并参考其他学科的资料,提出了一个南岛语族扩散理论,他认为南岛语族的祖先是居住在大陆东南沿海的新石器时代的农民。^⑨然而,令人遗憾的是,目前在南岛语族起源问题的研究上,考古学所能提供的证据及材料,其实还是相当薄弱的。^⑩再者,关于南岛语民族的起源和迁移历史,台湾南岛语学者李壬癸认为我们可以使用几种不同学科的资料来推测:遗传学(基因研究)、语言学、考古学、文化人类学等,其中,基因的稳定性最高也最可靠,语言次之,文化又次之,因为文化最容易移借。所以,本文认为采用语言学与分子人类学相结合的研究方法旨在通过更为可靠的、客观的研究证据来阐释南岛语族的起源与扩散。为此,本文将从理论、方法以及材料的角度评述前人时贤的研究,并在此基础上尝试从跨学科协作角度,结合遗传学、分子人类学最新的研究成果,提出对南岛语族起源与扩散模式的新思考。

二、南岛语族起源与扩散假说

尽管学者们依据不同的研究理论、方法以及材料存在着许多不同看法,但大抵可归纳为四类假说:一类是认为起源于中南半岛沿海一带;一类是认为起源于中国大陆华南地区;一类是起源于南洋群岛的西新几内亚;最后一类是认为独立起源于台湾,并由台湾开始向外扩散。

(一)中南半岛沿海一带起源说

荷兰学者柯恩(Hendrik Kern)主要是运用构拟的动植物群的古语材料来推断南岛语族的起源地,尤其是只在某种特定气候条件下才能找到的动植物群。他于1889年在《断定马来波利尼西亚人群起源地的语言学证据》(Linguistic evidence for the determination of the original homeland of the Malayo-Polynesian peoples)一文中,比较了遍布各地区的一百种南岛语言,并构拟出一些原始南岛语词汇。如:甘蔗、椰子、香蕉、竹子、芦苇;鲨鱼、乌贼、虾、鰂鱼、海龟;船、帆、船桨等热带动植物名以及航海词汇,这些词汇证据表明了古南岛语族起源于热带地区且靠海的地方。再从如:稻、米、旱田、水牛、鳄鱼、黑毛猴、向海“PAN*laHud”、内陆/高地“PAN*Daya”等词汇的分布区域来看,则进一步将古南岛语族的起源地指向中南半岛沿海一带。^⑪日本学者土田滋与柯恩的看法比较一致,并从“穿山甲”一词的分布地域更进一步证实了柯恩的看法。^⑫

(二)中国大陆华南起源说

最早提出大陆华南土著民族侗台语诸语与南岛语具有同源关系的是美国学者本尼迪克特(P.K.Benedict)。他于1942年发表《台语、加岱语和印度尼西亚语-东南亚的一个新联盟》(Thai, Kadai and Indonesian: A new Alignment in Southeastern Asia)一文,首次将南岛语与侗台语进行比较,列出了近30个南岛、侗台语同源词,论述了南岛语与侗台语的同源关系,并结合民族学、考古学知识认为它们共同起源地在华南地区。

法国学者欧远古(A. G. Haudricourt)也持相同看法,他认为南岛语族的起源地应该在亚洲大陆东南沿海一带,介于海南岛和台湾之间;因为这个区域比起南洋群岛的南岛语言不仅最分歧和存古现象最多,而且还和华南、越南北部、海南等地的各种语言有亲属或接触关系。^⑬

在20世纪40年代,本尼迪克特提出侗台-南岛同源说,并于70年代进一步加以论述从而引起国内外学者重视。倪大白(1988、1994),蒙斯牧(1992、1995),龙耀宏(1993),陈保亚(1996),欧亚觉亚、郑贻青(2004),邓晓华、王士元(2009),邓晓华、邓晓玲(2011)等学者也纷纷发表论著讨论并肯定侗台-南岛语的亲缘关系。此外,邓晓华(1992)列举了11个侗台语和南岛语有同源关系的,反映原始南岛人文化特征的基本核心词,从这些文化特征也可以窥见侗台语人群与南岛语人群史前早期的密切关系。^⑭另外,他还通过对比南方汉语方言与南岛语中有关基本生活、生态、身体部位的一些核心词,发现南方汉语方言中仍保存着原始南岛语“底层”成分,如:飞

廉、雷、蛟、肉、火尾、杀、穿山甲、哨笛、坂、“多”(此字有音无字)、舌头、蟑螂、蛇、砍。可以说南方汉语的形成是随汉人南迁的中原汉语与南方土著语言,即与原始南岛语长期相互融合的结果,留在大陆的古南岛人可以分成两部分:一种是已经汉化的闽、粤、吴人;另一种是受中原文化影响,但尚未汉化的苗瑶族群、侗台族群。^⑤

(三)西新几内亚起源说

萨丕尔(Edward Sapir)曾经提过一个这样的观念:我们可以经由有亲属关系的语言的地理分布来推论民族迁徙的历史与方向,语言最分歧的地区也就是这个民族的古代居留中心。^⑥

这一观点在戴恩(Isidore Dyen)1956年的《语言分布与民族迁徙理论》(Language distribution and migration theory)一文中得到了全面阐释。并在此基础上提出了语言分布与民族迁徙学说,即语言最分歧地区最可能是该语族的起源地。^⑦1965年戴恩在《南岛语言的词汇统计分类》(The Lexicostatistical Classification of the Austronesian Languages)一文中,运用语言分布与民族迁徙学说来推测南岛语族的起源问题。他以斯瓦迪士200基本核心词表来收集南岛语词汇,发现有三个地区语言同源百分比率最低、最分歧:

- (1)新几内亚-美拉尼西亚区;
- (2)中国台湾地区;
- (3)苏门答腊及其西岸岛屿。

其中,尤以新几内亚-美拉尼西亚区所包含的语言总数最多、最分歧。由于东新几内亚与美拉尼西亚的语言属于东部南岛语大洋洲支系的一支,它们有一些共同的语言演变。那么南岛语言最分歧的地区也就只有西新几内亚及其附近,因此这一地区最有可能是古南岛语族的起源地。^⑧

(四)台湾起源说

白乐思(Robert Blust)根据南岛语言的音韵、词汇等的演变,认为台湾南岛语言的高度歧义性,代表着原始南岛语最早的几个主要分支。而东区的波利尼西亚、美拉尼西亚、密克罗尼亚广大区域中的上千种语言,仅由剩余一支马来波利尼西亚语群演化出来(见图2)。^{⑨⑩}

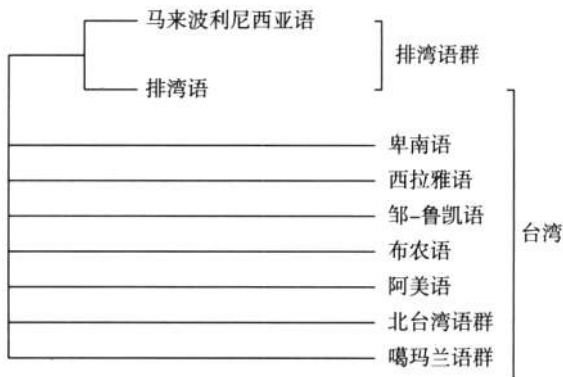


图2 《南岛语基本词汇数据库》的南岛语谱系树形图^⑪

承接戴恩、白乐思等人的研究,李壬癸(1979、1999、2004、2006、2008),何大安、杨秀芳等(2000)等台湾学者从不同角度继续论证了台湾南岛语言的多样性、分歧性。他们的观点基本一致,认为台湾的南岛语言除了保存最多原始南岛语特征外,其语言最分歧,语言之间的差异也最大。一个地区的语言越早分裂,表示语言分化的年代越久远,彼此之间的分歧也越大,也越接近该语系的起源地;越晚分裂的语言,彼此之间的差异也越小,离起源也就越远。南岛语言越分歧的台湾,显示着台湾的南岛语言时代纵深越长。^⑫再者,不同的语言创新演变,可以串成结果互异的语言群的现象,这种现象也显示出台湾的南岛语言不属于一个单一的语群,而且彼此之间接触、影响程度很深。^⑬

澳洲学者贝尔伍德(Peter A. Bellwood)则利用白乐思等语言学家建立起的一个南岛语言演化的大略时空架构,提出南岛语族在太平洋岛屿上的迁徙与扩散路径,即“Express-Train”(快车)模型。^④简单来说,“Express-Train”模型认为南岛语族起源于台湾,从台湾开始,经菲律宾、印度尼西亚东部,而后快速向波利尼西亚扩散。

与传统的历史比较语言学的谱系分类方法不同,不少学者开始利用生物学种系发生学理论与方法,并结合语言学中的词源统计分析法,使用计算机算法程序,对南岛语言的谱系分类及谱系树构建进行了研究。

格雷(Gray)、乔丹(Jordan)对77种南岛语言5185个词项进行词源统计,并运用进化树分析软件PAUP*产生了一棵最简洁的南岛语言谱系树,研究结果发现:南岛语谱系树形图与“Express-Train”模型高度兼容,“Express-Train”模型完全可以被转换成一个有序的地理字符并映射到语言谱系树上(见图3);另外,在谱系树上很明显地可以看出台湾的南岛语(阿美、排湾)的确处于根部位置(见图4)。^⑤之后,格雷(Gray)、德拉蒙德(Drummond)和格林希尔(Greenhill)等人基于统计推断系统发育谱系树的贝叶斯法,该法通过后验概率来评估各谱系树拓扑结构的可靠性,来构建南岛语谱系树。他们构建的南岛语谱系树支持南岛语族在太平洋群岛上定居停留与对外扩张的设想,该设想认为南岛语族大约距今5200年前起源于台湾,并在一系列的定居停留和对外扩张中迅速地穿越太平洋。^{⑥⑦}

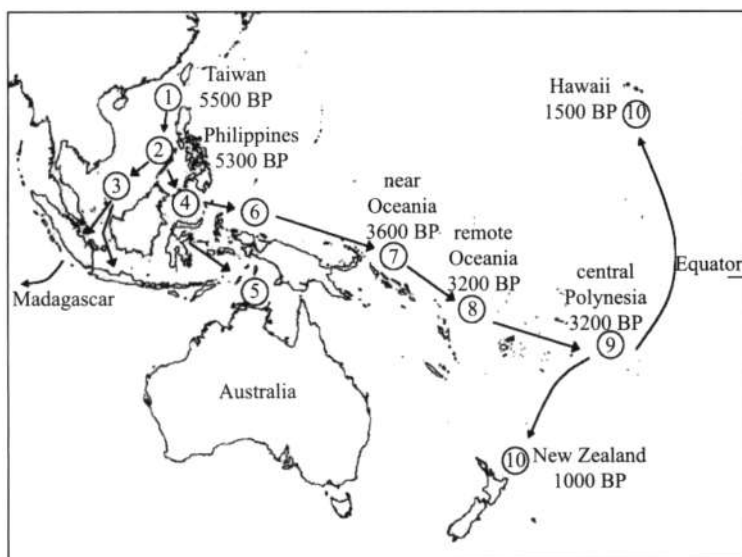


图3 “Express-Train”模型(Gray, et al. 2000)

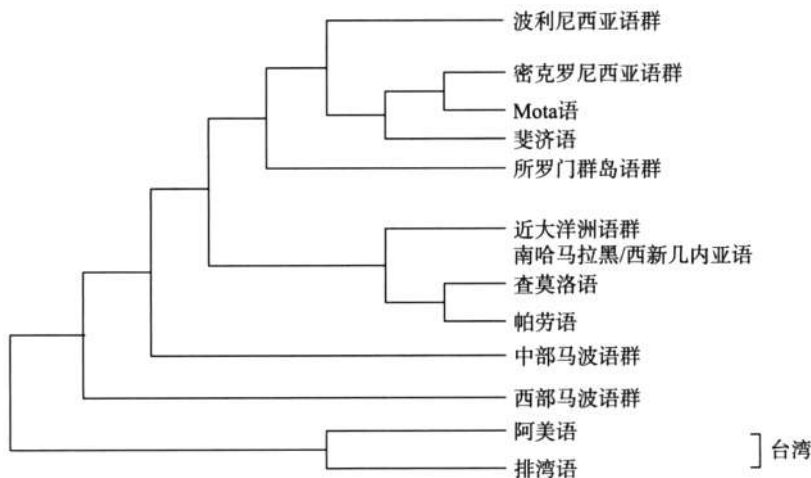


图4 南岛语言谱系树形图(改自 Gray, et al. 2000)^⑧

从上面的论述,我们可以看到前人及时贤在探讨南岛语族起源与扩散研究方面取得了诸多成果,但结果仍旧存在分歧。这主要是因为学者们所采用的理论、方法及材料方面存在某些问题。具体表现在:

理论方面,所谓的“语言最分歧的地区就最可能是该语族的起源地”,此观点基于遗传学理论。近几年国际顶级学术期刊《科学》(Science)上连续发表了数篇文章,全面讨论了世界各种语言的语音多样性分布与语言起源问题。如:爱金逊(Atkinson)(2011、2012a、2012b);^②西苏(Cysouw)(Cysouw M., et al. 2012);^③王传超(Wang)(Wang C. C., et al. 2012);^④朱丽叶(Tuyl)(Tuyl R. V., et al. 2012);^⑤耶格尔(Jaeger)(Jaeger T. F., et al. 2012)。^⑥其中,尤其西苏(Cysouw)等人的研究最为彻底,他们认为:在缺乏来自外源的基因突变的情况下,遗传漂变可使人口趋于同质性而导致多样性的减少,但语言无法像生物基因那样,不能简单地根据分歧的数量来定义语言的多样性,因为世界各地都有可能是语言的起源地。换言之,导致语言多样性的原因或许是由于不同族群的迁徙所致。因此,我们无法简单地根据分歧的数量来定义和阐释语言多样性与语族起源地之间的直接关系。

方法方面,无论是采用历史比较语言学的谱系分类方法,还是采用基于词源统计的定量分析研究方法,都能产生相对一致的谱系树形图。树形图简单明了,不但能够清晰地表现出语言之间的亲缘关系,还能直观的表现整个语言纵向的演化历史。^⑦然而树形图只是一种理想化的状态,其忽视了语言间的接触与交流,因为语言的接触必然会有更多的瓜葛,不可能干净整齐地表现为不同的树枝。^{⑧⑨}总的来说,语言的谱系分类,实际上也只能反映语言自身纵向地演变、分化,却无法传递出因语言由于横向接触而在内部结构和核心词中所造成的深刻影响以及因语言横向接触产生的语言融合、语言转换进而转入到另一个语言集团中参与发展变化这些现象。^{⑩⑪}

材料方面,本尼迪克特、邓晓华等诸多学者发现南岛语与侗台语、汉语方言这两类语言结构类型不同构的语言存在着一批具有同源关系或底层关系的词汇,这些同源词填补了长久以来基于以欧美学者主导的原始南岛语拟测研究中“今南岛语”重建的原始南岛语所缺乏的华南地区的语言材料。同时,这一语言材料的充实也弥补了部分台湾语言学者所认为的今日中国大陆上已无任何南岛语言遗留的空白。

在族群的起源与扩散问题上,起初大多是来自语言学领域的实践。时至今日单纯的语言框架已难以解释复杂多样的语族起源与扩散问题。且单纯从语言的表面现象下断言,有时会得出与事实相悖的结论。因此,不少学者业开始转向尝试论证跨学科、跨领域的可行性。

三、南岛语族起源的分子人类学分析

分子人类学是一门利用人类遗传信息来分析人类起源、族群演化、古代社会文化结构等多方面多层次的问题的学科,其研究的材料是人类基因组。^⑫人类的基因组由细胞核中的染色体和细胞质中的线粒体DNA组成,其物质基础是DNA大分子;在遗传过程中,DNA由于复制错误会积累突变,在遗传漂变、瓶颈效应或自然选择的作用下,突变类型在群体之间会形成一定比例的差异。其中,线粒体DNA遗传标记属于严格的母系遗传,而Y染色体遗传标记则属于严格的父系遗传。

(一)Y染色体与“侗台-南岛”

自20世纪80年代起,线粒体DNA就被用来研究世界上不同族群之间的遗传演化历史关系。而与南岛语族起源问题相关的线粒体DNA研究则始于1992年至1996年台湾“中研院”的主题计划《台湾与东南亚南岛民族的生物与文化类缘关系之研究》,英国遗传人类学家麦尔登(Melton)也参与其中。^⑬麦尔登(Melton, et al. 1995)等的线粒体DNA研究结果显示波利尼西亚人群起源于台湾,^{⑭⑮}而另一些线粒体DNA研究对此提出不同看法,认为波利尼西亚人群其实起源于印度尼西亚东部。^{⑯⑰}除线粒体DNA研究结果具有较大争议性外,其本身与语言分类的相关性并不好;相反地,大量的证据表明Y染色体类型与语言分类的相关性更为密切。^⑱另外,Y染色体有效群体小,突变率低、位点丰富以及有群体的特异性的类型分布,因而一直被广泛应用于追溯群体的父系源流、揭示性别间的迁徙及群体的差异。^⑲

利用Y染色体分析南岛语族起源问题的是2000年宿兵(Su Bing)、金力(Jin Li)等人发表于美国国家科学院刊上的《从Y染色体看波利尼西亚人群的来源》(Polynesian origins: Insights from the Y chromosome)一文。宿兵和金力等人分析了华南、东南亚、台湾(布农、泰雅、雅美、排湾、阿美)、美拉尼西亚、密克罗尼西亚、波利尼西亚的36个族群,共计551位男性的Y染色体样本。研究发现:台湾南岛人群的Y染色体单倍型不存在于波利尼西亚人群中,在密克罗尼西亚人群中有且少量。在密克罗尼西亚和波利尼西亚人群中高频率出现的Y染色体单倍型却不存在于台湾南岛人群,说明Y染色体不太支持波利尼西亚人群起源于台湾。^⑧

2008年,李辉(Li Hui)、文波(Wen Bo)等人用Y染色体SNP单倍群频率数据进行主成分分析,结果显示:侗台人群比其他东亚、东南亚人群和南岛西群(即马来人群,台湾原住民人群)具有强烈的遗传亲缘性。其中,O1a-M119是台湾南岛人群的主要Y染色体单倍型,平均为77%,这一单倍型在侗台人群和马来人群中分别约占20.5%和21.2%,远高于其他东亚人群;其次,通过Median Joining法构建了三个种群之间的Y染色体单倍群O1a*的网络结构图(见图5),侗台人群处于网络图的中心,分别与马来人群、台湾南岛人群直接或间接共享单倍型,而马来人群与台湾原住民人群之间几乎没有共享单倍型。此外,在遗传分化距离上,马来人群与台湾南岛人群间的分化差异远远大于与侗台人群的分化差异,马来人群和台湾南岛人群分别与侗台人群的亲缘性比起他们相互之间的还要近。换言之,从父系遗传结构上看马来人群并非直接起源于台湾南岛人群,两者分别独立起源于大陆的侗台人群。^⑨

承接宿兵等(Su Bing, et al.)和李辉等(Li Hui, et al.)等人研究,韦兰海(Wei Lan-hai)、严实(Yan Shi)等人选用17个STR位点来分析了Y染色体单倍型O3a2b*-P164(×M134),并在310个具有谱系分辨意义的Y染色体多态性的基础上,生成修正系统发育树。研究结果显示:所有可得的Y染色体单倍型O3a2b*-P164(×M134)属于新定义的O3a2b2-N6单倍型,O3a2b2-N6(×F706)单倍型来自中国北方或东亚东南沿海地区,O3a2b2a2-F706(×B451)单倍型出现在中国大陆东部沿海地区。而来自南岛语人群的样本属于O3a2b2a2-F706单倍型。其中,O3a2b2a2b-B451单倍型均来自南岛语人群,是南岛语族特有的Y染色体单倍型,属于O3a2b2a2-F706(×B451)单倍型的分支。可见,以Y染色体O3a2b2a2b-B451单倍型为代表的南岛语人群与亚洲大陆东南族群之间具有遗传亲缘关系。^⑩

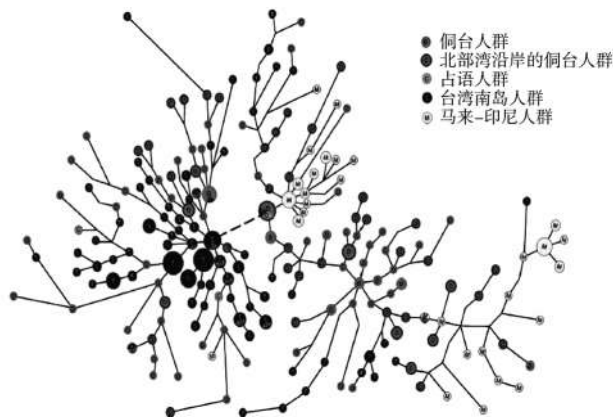


图5 Y染色体单倍群O1a*的网络结构图(引自Li Hui, et al. 2008)

(二)古人类DNA与全基因组

2011年至2012年台湾“中研院”在中国大陆福建沿海马祖列岛上发现了两具距今7500-8000年左右的亮岛古人类遗骸。葛明轩(Ko Min-Shan)、陈仲玉(Chen Chung-yu)等人通过分析亮岛人的线粒体DNA序列,并将其与8个台湾南岛人群和4个其他族群的550个样本进行比较,研究结果显示:亮岛人的线粒体DNA序列与台湾南岛人群最为接近,其线粒体DNA类型属于现存南岛语人群中最古老的单倍群E的早期分支之一,是由M9

单倍群分化而来,而M9单倍群主要分布在东亚大陆上,可以看出亮岛人与东亚大陆人群早期密切关系。^⑤此外,线粒体DNA单倍群E在台湾以及东南亚等南岛语族群中的频率分布如图6所示,饼图中标红部分为单倍群E所占之比例。因此,由图6可知,线粒体DNA单倍群E不仅在东南亚,乃至非洲的马达加斯加都有一定频率分布,所以亮岛人不但和东亚大陆人群有关,而且也能说明和东南亚、大洋洲居民的紧密关系。

另外,2016年至2018年,斯科格勒德等(Skoglund, et al.)、利普森等(Lipson, et al.)、波什等(Posth, et al.)等在国际学术刊物《自然》(Nature)、《当代生物学》(Current Biology)、《自然生态与进化》(Nature Ecology & Evolution)等上陆续发表了数篇与太平洋岛屿古人类DNA相关的文章,系统全面展示了南岛语族在太平洋岛屿上的迁徙与扩散。

斯科格勒德等(Skoglund, et al.)、利普森等(Lipson, et al.)、波什等(Posth, et al.)等通过分析距今3100–2300年前,位于波利尼西亚的瓦努阿图、汤加等地的古DNA全基因组,并将其与收集到的东亚、大洋洲现代人群样本数据进行比较,研究结果发现与早期Lapita文化有关的远大洋洲人群的祖先,在距今约5000年前时,从台湾扩散至菲律宾,经西美拉尼西亚,最后达到远大洋洲的。但是在近、远大洋洲的现代人群中发现有一定量的土著巴布亚人群有关的遗传贡献,这些土著人群是在南岛语人群未到来之前便已生活在新几内亚高地了;因此,他们再进一步分析了距今2900–150年前的瓦努阿图古DNA全基因组,发现巴布亚人群的相关血统是在距今2300年前左右到达瓦努阿图、汤加的,这很可能反映了几百年前Lapita末期文化的人群迁徙情况(见下页图7),图7反映了以汤加为例的人群历史模型。^{⑥⑦}然而尽管远大洋洲人群的遗传结构发生了巨大变化,但作为外来的巴布亚语言并没有取代南岛语言。^⑧

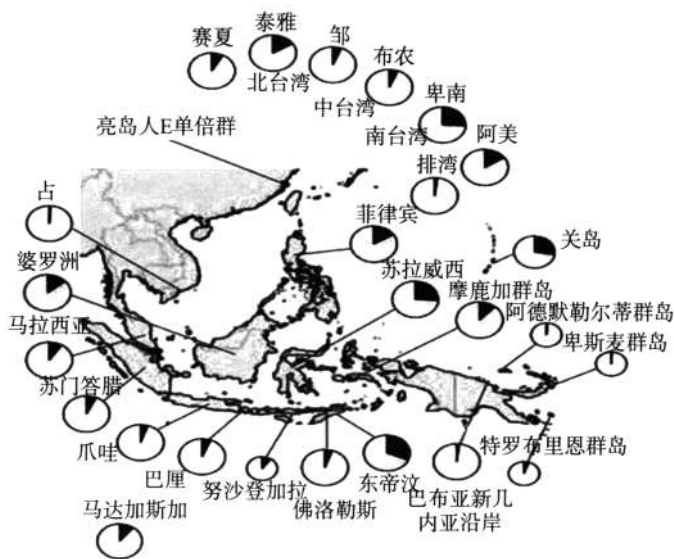


图6 线粒体DNA单倍群E在台湾和东南亚各族群里的分布频率(改自Ko, et al. 2014)

此外,利普森(Lipson, et al.)等则通过分析56个东南亚岛屿人群的全基因组芯片数据,所有东南亚岛屿的南岛语人群与台湾南岛人群具有强烈的亲缘关系,同时也接受了来自大陆中南半岛操南亚语人群的遗传贡献,这种南亚语相关的遗传贡献有可能是因为在南岛语人群迁徙之前,东南亚岛屿就已有操南亚语人群的存在;又或是南岛语人群经由中南半岛走廊而后再扩散至印度尼西亚东部地区。^⑨

总的来说,虽然Y染色体结果不直接支持马来人群、波利尼西亚人群起源于台湾,但它表明包括台湾原住民在内的南岛语族和侗台语族具有遗传亲缘关系,他们的起源地在华南地区。因此,不论是语言学研究领域,还是分子人类学研究领域,要研究南岛语族的起源问题,将包括侗台人群在内的华南族群涵括在内是十分有

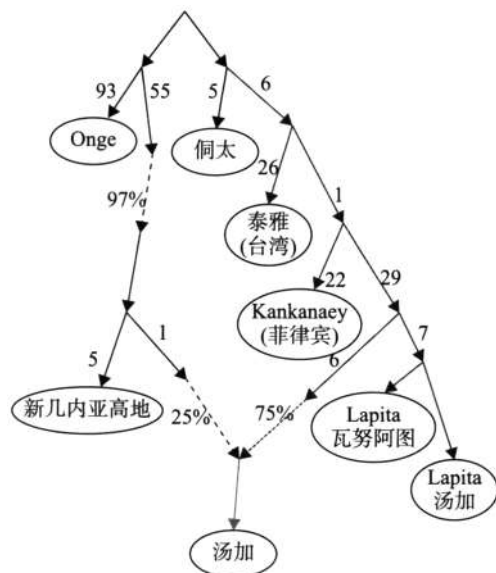


图7 汤加人群的历史模式
(改自 Skoglund, et al. 2016)

必要的。但是从以上的Y染色体、古人类DNA和全基因组芯片数据显示:南岛语族在太平洋岛屿上的迁徙与扩散过程中呈现出遗传谱系的性别差异以及不同历史层次的问题,或许是因为自然选择、瓶颈效应或遗传漂变导致部分线粒体DNA和Y染色体的丢失。

四、问题与结论

本文从理论、方法以及材料的角度,梳理了历史语言学视角下南岛语族起源与扩散问题的不同假说,并评述了前人时贤的研究成果及局限性。由于学者们囿于各自理论、方法以及材料的限制,在该问题的研究方面上始终未达成共识。分子人类学是一门新兴学科。语言和基因演变的相关性,是语言学和分子人类学的一个研究领域。分子人类学家在研究人类起源的同时便已注意到语言的演变与生物的遗传演变有着类同的关系。语言分类与遗传谱系总体来说大致是平行演化的,^⑤这为语言学利用和借鉴分子人类学的研究成果提供了一个新视角。^⑥

为此,本文尝试从跨学科协作角度,结合考古学、分子人类学最新的研究成果,提出对南岛语族起源与扩散问题的新思考。

侗台-南岛语人群的先民是居住在亚洲大陆东南部的新石器时期居民,他们共同组成了以环南中国海为中心的“亚洲地中海文化圈”,在其他异文化相继移植之前,构筑了一个巨大的土著文化共同体体系。^⑦距今约5000-6000年前时,开始向外迁徙与扩散。留在大陆华南地区的先民,在东亚漫长而又复杂的历史演变中,一部分受华夏中原文化影响,与南迁汉人长期相互融合,形成南方汉人与南方汉语,南方汉语底层仍保留着一批南岛语“底层”;另外一部分则逐渐演化为现在的侗台人群,侗台语由于长期受到周边单音节、有声调语言影响,经历了“语言的类型转换”^⑧。离开华南地区的:1)一支向东迁徙至台湾,到达台湾后,在岛内独立演化发展后,才经菲律宾南下至印度尼西亚东部,产生了早期的Lapita文化,而后犹如“Express-Train”般快速扩散至远大洋洲,成为了波利尼西亚人群的早期祖先;2)向西的一支经由中南半岛、东南亚岛屿走廊南下,迁徙到印度尼西亚东部。在距今2300年前左右,来自东亚的南岛语人群逐渐开始和新几内亚岛的操巴布亚诸语人群接触交流,之后也向波利尼西亚扩散与迁徙。

注释:

①有学者将中国台湾地区纳入西部南岛语群,也有学者将其独立为北区南岛语(即台湾),至今尚无定论。具体请参见李壬癸:《台湾原住民族:语言篇》,南投:台湾省文献会,1999年。

②李壬癸:《台湾南岛语言论文选集》,台北:“中央研究院”语言学研究所,2004年。

③Diamond Jared M, The express train to Polynesia. *Nature*. 1988(336). pp. 307-308.

④Geldorn R., “Ueheimat und Früheste Wanderungen dez Austronesiar,” *Anthropology* 27, 1932(3-4), pp. 543-619.

⑤Beyer H. O., “Philippine and East Asian archaeology and its relation to the origin of the Pacific islands population,” *National Research Council of the Philippine Bulletin* 29, 1948.

⑥林惠祥:《人类学论著》,福州:福建人民出版社,1981年。

⑦凌纯声:《东南亚古文化研究发凡》,《新生报(民族学研究专刊)》,1950年。

- ⑧鹿野忠熊:《台湾考古学民族学概观》,宋文薰译,台中:台湾省文献委员会,1955年。
- ⑨Bellwood P., “Austronesian prehistory in Southeast Asia: Homeland, expansion and transformation,” *The Austronesians*, ed. by Peter Bellwood, James J. Fox and Darrell Tryon, Canberra: The Department of Anthropology, Research School of Pacific and Asian Studies, The Australian National University, 1995, pp. 96–106.
- ⑩臧振华:《再论南岛语族的起源与扩散问题》,《南岛研究杂志》(*Journal of Austronesian Studies*)2012年第1期。
- ⑪Kern H., “Taalkundige gegevens terbepaling van het stamland der Maleisch-Polynesische volken,” 1889, pp. 270–287; McFarland Curtis D, Shigeru Tsuchida, “Linguistic evidence for the determination of the original homeland of the Malayo-Polynesian peoples,” *Oceanic Studies: Linguistics, Anthropology and Sociology*, Tokyo, 1976.
- ⑫土田滋:「高砂族の言語」,『国際文化』1972年。
- ⑬Haudricourt A. G., “Les origines asiatiques des langue malayo-polynésiennes,” *JSO10*, 1954, pp. 1–7.
- ⑭邓晓华:《从语言推论壮侗语族与南岛语系的史前文化关系——谨以此文悼念恩师严学窘教授》,《语言研究》1992年第1期。
- ⑮邓晓华:《南方汉语中的古南岛语成分》,《民族语文》1994年第3期。
- ⑯转引自李壬癸:《从语言的证据推论台湾土著民族的来源》,《大陆杂志》1979年第1期。
- ⑰Dyen I., “Language distribution and migration theory,” *Language*, 1956(32–4), pp. 30–35.
- ⑱Dyen I., “The Lexicostatistical Classification of the Austronesian Languages,” *Supplement to IJAL*, 1965(31), pp. 20–26.
- ⑲Blust R., “The pro-Austronesian pronouns and Austronesian subgrouping: A preliminary report,” *working papers in linguistics*, 9.2 U. H. 1977, pp. 1–15.
- ⑳Blust R., “Subgrouping, circularity and extinction: Some issues in Austronesian comparative linguistics,” In Elizabeth Zeitoun and Paul Jen-kuei Li, eds, *Selected Papers from the 8th International Conference on Austronesian Linguistics*, Symposium Series of the Institute of Linguistics (Preparatory Office), Taipei: Institute of History and Philology, Academia Sinica, 1999, pp. 31–94.
- ㉑Greenhill S. J., Blust R., Gray R. D., “The Austronesian Basic Vocabulary Database: From Bioinformatics to Lexomics,” *Evolutionary Bioinformatics*, 2008(4), pp. 271–283.
- ㉒李壬癸:《珍惜台湾南岛语言》,台北:前卫出版社,2010年。
- ㉓何大安、杨秀芳:《南岛语与台湾南岛语》,《台湾南岛语言丛书导论》,台北:远流出版事业股份有限公司,2000年。
- ㉔Diamond J. M., “The express train to Polynesia,” *Nature*, 1988(336), pp. 307–308.
- ㉕Gray R. D., Jordan F. M., “Language trees support the express-train sequence of Austronesian expansion,” *Nature*, 2000(405), pp. 1052–1055.
- ㉖Gray R. D., Drummond A. J., Greenhill S. J., “Language Phylogenies Reveal Expansion Pulses and Pauses in Pacific Settlement,” *Science*, 2009(323), p. 479.
- ㉗Greenhill S. J., Drummond A. J., Gray R. D., “How Accurate and Robust Are the Phylogenetic Estimates of Austronesian Language Relationships?,” *PLoS ONE* 5, 2010(3), pp. 1–6.
- ㉘由于篇幅原因,本图为原图的简化图。详细树形图请参看原文。
- ㉙Atkinson Q. D., “Phonemic diversity supports a serial founder effect model of language expansion from Africa,” *Science*, 2011(332), pp. 346–349; Atkinson Q. D., “Response to Comments on ‘Phonemic Diversity Supports a Serial Founder Effect Model of Language Expansion from Africa’,” *Science*, 2012a(335), p. 657; Atkinson Q. D., “Response to Comment on ‘Phonemic Diversity Supports a Serial Founder Effect Model of Language Expansion from Africa’,” *Science*, 2012b(335), p. 1042.
- ㉚Cysouw M., et al., “Comment on ‘Phonemic Diversity Supports a Serial Founder Effect Model of Language Expansion from Africa’,” *Science*, 2012(335), p. 657.
- ㉛Wang C. C., et al., “Comment on ‘Phonemic Diversity Supports a Serial Founder Effect Model of Language Expansion from Africa’,” *Science*, 2012(335), p. 657.
- ㉜Tuyl R. V., et al., “Comment on ‘Phonemic Diversity Supports a Serial Founder Effect Model of Language Expansion from Africa’,” *Science*, 2012 (335), p. 657.

- ③Jaeger T. F., et al, "Comment on 'Phonemic diversity supports a serial founder effect model of language expansion from Africa'," *Science*, 2012(335), p. 657.
- ④邓晓华,高天俊:《演化语言学的理论、方法与实践》,《山西大学学报(哲学社会科学版)》2014年第2期。
- ⑤臧振华:《再论南岛语族的起源与扩散问题》,《南岛研究杂志》(Journal of Austronesian Studies)2012年第1期。
- ⑥徐丹:《中国境内的混合语及语言混合的机制》,《语言战略研究》2018年第2期。
- ⑦罗美珍:《试论台语的系属问题》,《民族语文》1983年第2期。
- ⑧陈保亚:《论语言接触与语言联盟——汉越(侗台)语源关系的解释》,北京:语文出版社,1996年。
- ⑨李辉、金力:《Y染色体与东亚族群演化》,上海:上海科学技术出版社,2015年。
- ⑩陈叔倬、许木柱:《台湾原乡论的震撼——族群遗传基因资料的评析》,《语言暨语言学》2001年第1期。
- ⑪Melton T., et al, "Polynesian genetic affinities with Southeast Asian populations as identified by mtDNA analysis," *American Journal of Human Genetics*, 1995(57), pp. 403-414.
- ⑫Sykes B., et al, "The origins of the Polynesians: an interpretation from mitochondrial lineage analysis," *American Journal of Human Genetics*, 1995(57), pp. 1463-1475.
- ⑬Trejaut J. A., et al, "Traces of Archaic Mitochondrial Lineages Persist in Austronesian-Speaking Formosan Populations," *PLoS Biol.* 2005(3.8), pp. 1362-1372.
- ⑭Richards M., Oppenheimer S., Sykes B., "mtDNA suggests Polynesian origins in Eastern Indonesia," *American Journal of Human Genetics*, 1998(63), pp. 1234-1236.
- ⑮Lum K. J., Cann R. L., "mtDNA lineage analyses: origins and migrations of Micronesians and Polynesians," *American Journal of Physic Anthropology*, 2000(113), pp. 151-168.
- ⑯Forster P., Renfrew C., "Mother Tongue and Y Chromosomes," *Science*, 2011(333), pp. 1390-1391.
- ⑰李辉、金力:《Y染色体与东亚族群演化》,上海:上海科学技术出版社,2015年。
- ⑱Su B., Jin L., et al, "Polynesian origins: Insights from the Y chromosome," *Proceedings National Academy Sciences of USA*. 2000 (97), pp. 8225-8228.
- ⑲Li H., Wen B., et al, "Paternal Genetic Affinity between Western Austronesians and Daic Populations," *BMC Evolutionary Biology*, 2008(8), pp. 146-151.
- ⑳Wei L. H., Yan S., et al, "Phylogeography of Y-chromosome haplogroup O3a2b2-N6 reveals patrilineal traces of Austronesian populations on the eastern coastal regions of Asia," *PLoS ONE* 12(4), 2017, pp. 1-12.
- ㉑Ko Albert M. S., Chen C. Y., et al, "Early Austronesians: Into and Out Of Taiwan," *American Journal of Human Genetics*, 2014(3), pp. 426-436.
- ㉒Skoglund P., et al, "Genomic insights into the peopling of the Southwest Pacific," *Nature*, 2016(538), pp. 510-513.
- ㉓Lipson M., et al, "Population Turnover in Remote Oceania Shortly after Initial Settlement," *Current Biology*, 2018(28), pp. 1-9.
- ㉔Posth C., et al, "Language continuity despite population replacement in Remote Oceania," *Nature Ecology & Evolution*, 2018(2), pp. 731-740.
- ㉕Lipson M., et al, "Reconstructing Austronesian population history in Island Southeast Asia," *Nature Communications*, 2014, pp. 1-7.
- ㉖Cavalli-Sforza L. L.:《追踪亚当与夏娃——从演化历史看基因、民族和语言的关系》,吴一丰、郑谷苑、杨晓珮等译,台北:远流出版事业股份有限公司,2003年。
- ㉗徐丹:《研究语言的新视角:语言和基因的平行演变》,《当代语言学》2015年第2期。
- ㉘吴春明:《“南岛语族”起源研究述评》,《广西民族研究》2004年第2期。
- ㉙倪大白:《中国的壮侗语与南岛语》,《中央民族学院学报》1988年第3期。