

基因家谱学在中国的发展历程

文少卿¹ 韩 昇² 李 辉¹

(1. 复旦大学 现代人类学教育部重点实验室, 上海 200433; 2. 复旦大学 历史系, 上海 200433)

摘 要: 基因家谱学是将遗传学运用于传统谱牒学的一门应用学科, 它正日益受到公众的关注。首先, 我们追溯了基因家谱学在全球, 特别是国内的发展历程, 包括源头、萌芽和最新进展。此外, 我们集中回顾了五个方面: 家族历史、群体结构、突变速率、考古学应用和法医学应用, 其中, 基因家谱学和其他学科的跨学科研究产生了许多新的成果。最后, 我们对于方法和案例的介绍能够促进从业者和爱好者共同推进基因家谱学在中国的发展。值得注意的是, 基因家谱学的热潮也暗含风险, 主要涉及 DNA 测试者及其家人, 以及同名人士的隐私问题。

关键词: 基因家谱; Y 染色体 DNA 谱系; 名人家族史; 深度家系; 法医学应用

中图分类号: K207 **文献标志码:** A **文章编号:** 1674-6627(2018)01-0011-10

谱牒又称家谱、家乘、宗谱、族谱和世谱, 是家族的历史书。谱牒、方志和正史是我国史学研究的三大基石, 故有“家乘犹国史”的美誉。谱牒反映了中国宗法血缘文化, 在我国有着悠久的历史。最早出现的谱牒为“甲骨家谱”, 出土于殷墟的甲骨文中就有关于氏族的骨片^[1]。商代有专门的谱牒档案, 并设有掌管谱牒的官员。战国时期, 史官编著了我国第一部谱牒专著——《世本》。战国、秦汉至魏晋, 受门阀制度影响, 谱牒学渐盛。宋以后, 科举盛行, 重视门第出身, 家谱渐渐不示外姓, 自此, 谱牒学由盛转衰。至明清, 修谱又达鼎盛期, 遍及城乡。据统计, 我国现存家谱至少有6 000多种^[2], 这些谱牒是中华民族传统文化的有机组成部分, 有着宝贵的文献价值^[3]。梁启超曾说: “族姓之谱六朝唐极盛, 宋后寝微。然此实重要史料之一: 例如欲考族制组织法, 欲考各时代各地地方婚姻平均年龄、平均寿命, 欲考据父母两系遗传, 欲考男女产生比例, 欲考出生率与死亡率比较……无数问题, 恐除族谱家谱外, 更无它途可以得资料。我国乡乡家家皆有谱, 实可谓史界瑰宝。将来我国立大图书馆, 能尽集天下之家谱俾学者分科研究, 实不朽之盛业也。”^[4]

谱牒学是一项专门研究谱牒的历史科学。从两汉鼎盛时期的谱学到清代章学诚的谱学, 曾积累了一大批丰厚的文献, 供后人研究。谱牒学研究一直面临着三大困难: 一是谱牒整部造假。唐以前, 为了出仕和地位, 多以郡望相标榜。二是谱牒真真假假。主要是为了抬高家族的身份, “如周姓则祖后稷, 吴姓则祖太伯, 姜姓则祖太公望, 袁姓则祖袁绍”^[5]。三是材料损失严重^[6]。如唐末农民大起义

收稿日期: 2017-10-10

基金项目: 国家自然科学基金项目“中国西北突厥语族人群混合起源的遗传学解析”(31671297); 上海市“曙光计划”项目“欧亚大陆东西方人群在突厥族群中的混合研究”(14SG05)

作者简介: 文少卿(1982-), 男, 湖北荆门人, 复旦大学现代人类学教育部重点实验室博士后研究人员, 主要从事历史人类学研究; 韩昇(1957-), 男, 江苏海安人, 复旦大学历史系教授, 博士, 博士生导师, 主要从事中国古代史研究; 李辉(1978-), 男, 上海人, 复旦大学现代人类学教育部重点实验室教授, 博士, 博士生导师, 主要从事历史人类学研究。

时期、抗日战争时期、“土改”和“文革”时期。虽然谱牒学研究可以运用版本学、校勘学等方面的知识鉴别真伪,考订谬误^[6],但很多时候,传统谱牒学研究受限于材料本身,很难有突破性进展。

遗传学究其本质是一门历史科学。遗传突变世代相传,并因为随机性或者选择作用在群体和物种间/内积累下来^[7]。人体内有 23 对染色体,其中 22 对常染色体中,每一对染色体都有一条来自于父亲,一条来自于母亲,两条染色体在传代过程中会发生部分重组。另一对性染色体,包括 X 染色体和 Y 染色体。在女性体内, X 染色体成对出现,也会发生重组现象。但是在男性体内, X 染色体来自母亲, Y 染色体只能来自父亲。因此, Y 染色体具有父子相传的特点,严格遵从父系遗传的方式,传递过程中不受社会文化和自然因素的影响(如图 1 所示)。

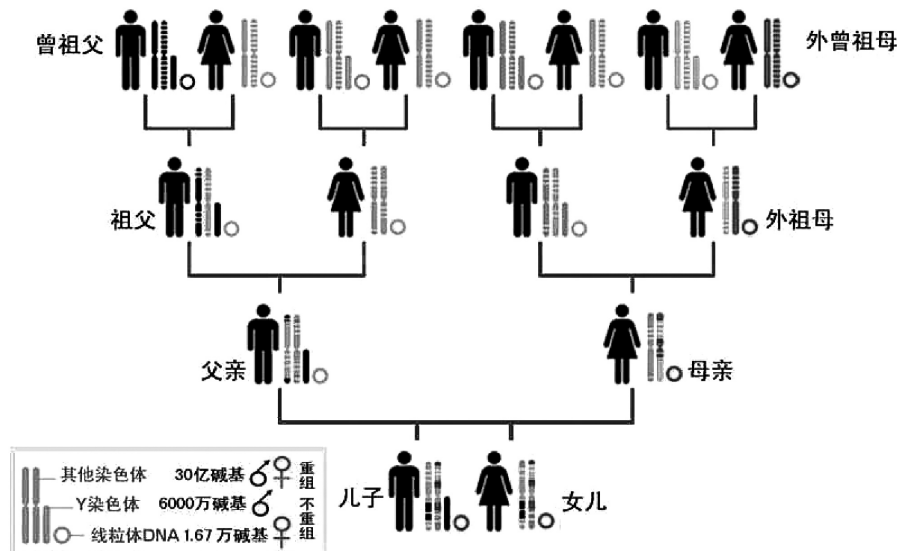


图 1: 遗传物质的传递规律

基因家谱研究是遗传学在传统谱牒学上的应用,它试图通过 DNA 检测,特别是 Y 染色体标记体系,弄清家族内部个体间的遗传关系^[8]。基因家谱研究极大地弥补了传统谱牒学的不足。在我国,家族、宗族都是指父系制度的同姓亲族,是以婚姻和血缘关系结合成的社会单位,因此,相比其他遗传标记体系, Y 染色体是研究基因家谱的绝佳材料。 Y 染色体上有着充足的拓扑学关系明晰的遗传标记,可以满足父系遗传谱系的构建;另外, Y 染色体上有着稳定的“分子钟”,可以估算不同 Y 染色体间的分化时间^[9]。因此,当遗失的或者不忠实记载的谱牒无法作为追寻祖先的可靠依据时,以现代分子生物为基础的基因家谱研究是直接追溯宗族成员之间亲缘关系的最佳方法,是验证祖先与后代的父系关联,补全和重构家族谱牒的唯一手段。

一、世界范围内基因家谱的研究动态

对于人类根源的探索缘于人类本能的好奇心。20 世纪末期,来自数个国家的数百万人开始热衷于寻找自己的祖先来源。这一狂热的爱好催生了应用科学一个新的分支——基因家谱学,在该学科中, DNA 分子起到了承载家族历史的作用^[10]。

(一) 源头

基因家谱研究的出现离不开 DNA 分型技术。该技术的突破源于英国莱斯特大学的 Alan Jeffreys 发现了名为微卫星的高变遗传标记^[11]。20 世纪 80 年代中期, Alan Jeffreys 等基于这些序列,开发了多位点探针,这一检测是鉴定生物亲缘关系最为有效的工具^[12]。该方法主要基于限制性片段长度多态性,被称为 DNA 指纹技术(DNA fingerprinting)。到 80 年代末, DNA 指纹法被法医界广泛接受和应用。

法医遗传学不仅为基因家谱学提供了技术支持,同时也提供了方法上的参考。法医遗传学主要涉及两个核心问题:个体鉴定和亲缘关系评估^[13]。后者的目标和基因家谱学的目标一致。因此,首个基因家谱研究实际上是一次法医学调查。这些90年代早期的案例中,最有名的是罗曼诺夫沙皇家族^[14]和纳粹罪犯约瑟夫·门格勒^[15]的遗骸鉴定。通过各种亲缘关系的比较,前者确定出土于叶卡捷琳堡的一座墓葬中的9块遗骨来自俄罗斯末代沙皇——尼古拉斯二世及其家族成员;后者证明1985年出土于巴西的遗骸属于约瑟夫·门格勒本人。

(二) 萌芽

1996年6月,旅居美国的斯诺伐克籍谱牒学、民俗学者Helen Cincebeaux在明尼阿波利斯举行的东欧家族历史学会联盟会议上首次提出了基因家谱学(Genetic Genealogy)概念。1998年11月,一位名为Alan Savin的业余谱牒学者在其个人网站上再次谈到了这个名称,他也是《基因家谱学导论》(*An Introduction to Genetic Genealogy*)的作者。

基因家谱学真正的萌芽是在20世纪90年代末,主要标志为一系列高质量文献的发表^{[16][17][18][19]}。从文献发表的先后顺序和学术意义来看,排在首位的是1997年Scorecki等关于犹太祭祀Y染色体的研究^[16]。《圣经》中记载犹太人中的祭司是由犹太教的第一祭司长Aaron开始按血缘代代相传,作者们发现德系和西班牙系犹太祭司们与非祭司的犹太人相比,有较近的亲缘关系,也就是说,祭司们可跨越3300年追溯到一个共同的父系祖先。Y染色体的分析与《圣经》故事的完美契合着实让人吃惊。这项研究开创了通过Y染色体标记进行基因家谱研究的先河。

1998年,两篇文章的问世为基因家谱学的建立添砖加瓦。这两篇文章都与历史名人有关,因此引发了极大的关注。第一项研究与美国第三任总统Thomas Jefferson和他的黑人女仆Sally Hemings有无私生子的鉴定有关^[18]。Foster比较了Jefferson的叔叔、Sally的大儿子和最小的儿子的男性后代Y染色体,认为Jefferson是Sally的最小儿子的生父。这项研究理清了Sally Hamming后代错综复杂的亲缘关系,为随后通过Y染色体标记重塑父系谱系关系提供了范例。同年,Decorte发表了一项疑似法国国王Louis XVI和Marie-Antoinette的儿子的线粒体DNA鉴定^[19]。作者们对Carl Wilhelm Naundorf(声称是这对皇室夫妇的后代)、Marie-Antoinette和她的两个姐妹,以及现存的母系亲属的线粒体DNA进行了分型。结果表明,虽然Naundorf被冠以Louis XVII的名号,但他的线粒体DNA遗传类型与其他母系亲属测试者的完全不同,因此他是个伪冒者。这项研究是最早通过线粒体DNA进行基因家谱学的工作,为通过母系遗传研究基因家谱学提供了参考。

(三) 发展

近20年,受公众对名人家族历史的浓厚兴趣的驱动,基因家谱学中探索历史名人的父系类型一直是最为活跃的研究方向,例如前面提到的美国总统Thomas Jefferson^[18]、法国皇帝拿破仑^[20]和波旁家族^[21]、俄罗斯的罗曼诺夫家族^[22],以及中国的爱新觉罗家族^[23]、成吉思汗^[24]和曹操^[25]。其中,成吉思汗和爱新觉罗·觉昌安的父系类型最早是通过特殊的Y染色体单倍型,即所谓的星簇(Star Cluster)和满族簇(Manchu cluster)推测而来的,它们分别高频出现于蒙古语和满通古斯语人群^{[24][26]}。这种通过分析Y染色体单倍型数据寻找晚近人群扩张信号的方法还运用在其他课题中。例如,2015年Balaresque分析了欧亚大陆上127个人群的5321个Y染色体数据,发现了11个星簇状的扩张信号,其扩张年代为公元前2100年至公元1100年^[27]。排除所用位点过少、可能存在的遗传类型错判等问题,这项研究还是有些有趣的结论,比如揭示了晚近扩张人群与生活方式的关系——阿尔泰游牧人群Y染色体的大范围传播可能与马的使用有关。

基因家谱研究受DNA检测技术发展的推动。现如今,基于现代DNA和古代DNA高通量测序的手段也进入到基因家谱学之中。最好的例子是2014年Olalde I等测试了一个疑似含法国国王Louis XVI血液的手帕上的全基因组,他们发现,根据所获得的数据来推断的祖先和表型与历史记载不符^[21],否定了葫芦内所含血样属于Louis XVI,该研究为破解历史悬案提供了有力的证据。

在1999~2000年,第一家提供基因家谱分析的基因测试公司成立。随后,基因家谱研究被这些商业公司陆续推出的项目所主导^[28]。其中,第一家也是最成功的公司即Family Tree DNA。该公司由一位美国的业余谱牒专家Beenett Greenspan于1999年创立。迄今,有大约40家^[29]甚至更多的实验室和公司可以提供祖源检测和重构家族谱牒的服务。截至2009年,接受这类服务的用户已经超过了100万人。

2005年,首个基因家谱学的专业刊物在线出版。同年,国际基因家谱学会(ISOOG)成立。这个非商业机构的目标是通过研讨会、网站、演讲、论坛和会议,教育大众如何在谱牒研究中正确使用遗传学知识。2008年,俄罗斯DNA家谱学院出版了杂志*Herald of the Russian Academy of DNA Genealogy*。同年,另一个基因家谱学的杂志*Russian Journal of Genetic Genealogy*创刊。此外,也有一些专业的博客介绍基因家谱学的动态,甚至有电视频道专门播放基因家谱相关内容的节目。

二、国内基因家谱的研究动态

国内基因家谱的研究要从曹操家族说起。这项研究通过对曹操后代家系的遗传学梳理,确定了曹操的遗传类型和身世,证实曹操的父亲是家族内过继,而不是异姓(夏侯家)收养,这解决了历史学遗留了2000年的问题^{[25][30][31]}。这项研究是国内通过分子生物学手段研究历史疑难问题的新尝试,它开创了历史人类学这一新领域。该研究通过媒体传播在全国引起了较大反响,并引发了基因家谱研究的新潮流。这项研究催生了后续一系列基于家族水平的探索性课题,包括历史谜案探索、突变速率估算、东亚遗传结构、考古学和法医学的应用,以及国内首个分子人类学专业刊物——《现代人类学通讯》的问世。

(一) 历史谜案

曹操的祖父是东汉时期的大宦官曹腾,曹操的生父曹嵩是其养子,因此,曹操的身世是一大史学谜题。曹操在作《家传》时自称为周朝曹叔振铎之后,后来再度改为出自帝舜。据《曹瞒传》记载,曹操之父曹嵩出自夏侯氏。宋代以后,曹操出自夏侯氏的说法流传甚广。而后世制作的家谱几乎都追溯到西汉开国功臣曹参。传统谱牒学方法仅从文献资料无法确定曹操的真实身世,使用最新的科学技术成为解决这个问题的新思路。

复旦大学现代人类学教育部重点实验室从2009年开始对曹操的遗传类型归属展开了系统的调查,分别从曹操可能存在的后人家系^[25]和曹氏宗族墓群中曹操的叔祖父——曹鼎遗骸^[30]两方面入手,论证曹操的父系Y染色体类型极有可能为O2*-M268。根据曹操后人家系的测试结果,反推单倍群O2*-M268为曹操所属类型的可能性是92.71%;根据古代样本Y染色体12个STR位点的测试结果,利用基于贝叶斯等位基因频率法的数据库,推测曹鼎属于O2*-M268+PK4-,M176-的概率为60.18%,属于单倍群C3*-M217的概率为13.97%,属于其他单倍群的概率均低于11%。2016年,课题组又对该样本进行了Y染色体单核苷酸多态位点(SNP)分型测试,明确其遗传类型属于O2-M268+,F1462+,PK4-,弥补了之前通过现代人家系推测曹操遗传类型和Y-STR推测Y-SNP的缺陷^[31]。至此,曹操及其后人的遗传类型为F1462+,PK4-的结论得以盖棺定论。

如前文所述,一些研究者根据蒙古语和满通古斯语人群中特有的高频单倍型(星簇和满族簇)以及对应的共祖时间,推测其父系祖先分别为成吉思汗和爱新觉罗·觉昌安^{[24][26]}。但是,由于缺乏明确的家系样本和古代遗骸的支持,这两个推测结果真实性一直饱受争议^[32]。2015年,我们采集了7个爱新觉罗氏的家系(3个家系有完整谱牒)的男性样本,进行了Y染色体短串联重复片段(STR)分型,结果显示,其中3家(2家有完整家谱,其最近共祖为清太祖努尔哈赤)的父系遗传类型非常接近,且极为罕见,对这3个样本的Y-SNP测试结果表明,此类型属于C3b2b1*-M401单倍群,即爱新觉罗家族的父亲遗传类型^[23]。2016年,我们又对爱新觉罗家族的男性个体进行了Y染色体测序,发现这

些个体属于更为精细的遗传类型 C3b1a3a2-F8951,与星簇为兄弟支系,与之前推测的 C3b2a-M48 满族簇毫无关系^[33]。因此,之前推断的成吉思汗与星簇之间的关系可能也不牢靠。

2017年,韦兰海等探讨了星簇的起源,以及它与成吉思汗和各个蒙古语人群的关系^[34]。通过对34个属于星簇的Y染色体全序及与其较近支系的Y序列的研究,我们发现这些星簇样本属于单倍群 C3b1a3a1-F3796,其分布于中国东北到北高加索的广大区域。这些人群的源头被追溯到古代尼伦蒙古部落或者普通的蒙古部落,比如乌兹别克人和诺盖人中的曼格特部(Manghit tribe)、乌兹别克人中的柯纳格斯部(Keneges tribe)、阿富汗的哈扎拉人、中国的达斡尔人,以及哈萨克人群中的拉特部(Dulat tribe)、乌孙部(Uysun tribe)和克烈部(Kerey tribe)。而他们的晚近共祖时间为2576年前,与整个蒙古语人群扩张的时间一致,这显然与成吉思汗或其他男性亲属无关。Zerjal等认为哈扎拉人是成吉思汗的直系子孙,他们认为这是支持成吉思汗与星簇之间关系的有力证据。但事实上,哈扎拉人来源于成吉思汗派遣的蒙古大军。而这两万战士来自于不同的蒙古部落,并没有直接证据表明哈扎拉人都是成吉思汗的后代。

事实上,寻找成吉思汗的父系类型困难重重。一方面,由于蒙古人保持祖先神秘感的葬俗,导致蒙古贵族(特别是皇室)的墓地极少被发现^[35]。尽管如此,基于有限墓葬的古DNA研究给出了极为矛盾的结论。根据蒙古国东部塔班陶勒盖的疑似黄金家族墓地遗骸的DNA鉴定结果,Lkhagvasuren等认为成吉思汗和他的家族属于遗传类型 R1b-M343^[36],然而另一项来自中国河北的高等级贵族墓的遗传调查表明,成吉思汗的遗传类型为单倍群 Q 也不能完全排除^[37]。另一方面,疑似成吉思汗后代的现代人的DNA分型结果表明,不同家族分支的父系类型也不一致^[38]。比如说,3个现存的成吉思汗家族分支属于遗传类型 C3 下的3个早期分化的不同亚型,此外,成吉思汗亲兄弟 Khasar 的后人属于遗传类型 O3,成吉思汗父亲的同父异母兄弟 Belgutei 的后人属于遗传类型 R1a1a。

近年来,本课题组对西北地区永登县的鲁土司家族进行了父系遗传调查,根据鲁氏家谱的记载,他们的祖先为脱欢(成吉思汗的第六子阔列坚的后代)。在元末明初,脱欢和他的部众归降于明廷。后因作战勇猛,自第三代起赐汉姓鲁。从此,鲁土司家族驻守西北永登县长达19代,历时561年。Y染色体分型结果表明,鲁土司家族的遗传类型为 C3b1a1a1a-F1756(又称 C3b3a2a-448del),常见于蒙古语和突厥语人群^[39]。进一步的遗传网络分析发现,鲁土司家族后人与哈萨克斯坦DNA计划(The Kazakhstan DNA Project)中的 Tore 支系共享单倍型^[40]。Tore 支系为成吉思汗大儿子术赤的后代。至此,成吉思汗的遗传类型又有了一个新的可能——C3b1a1a1a-F1756。

(二) 突变速率

Y染色体短串联重复片段(STR)常常用于估算单核苷酸多态位点(SNP)支系和历史名人的年代。但是由于采用不同的速率和估年方法,所产生的结果会有数倍的偏差,选择何种突变速率和估年方法仍有争议。王传超等使用两个有着完整、可靠谱牒记录的深度家系直接评估Y染色体的突变速率和估年方法。结果表明,“BATWING法加上家系突变率”能非常准确地对历史人物进行年代估算^[41]。这项研究的两个家族的祖先分别为史诏(南宋名士,翼国公史简之子,生活于公元885~957年前)和操妙荣(任职饶湖镇总,生活于公元603~673年前)。

司马光家族延续了上千年,家谱记载较为完整,为突变速率的估计提供了较好的材料。杜盼新等对11个声称为司马光后代的家族进行了Y-STR分型,结果表明,有5个家族的STR单倍型彼此之间十分接近,同属下游单倍群 O1a1a1a1a1a-F492, F656。同时,我们使用“BATWING法加上家系突变率”逐层计算所有支系的最近共祖时间,其结果与根据谱牒资料构建的家族谱系图非常吻合^[42]。

这两项研究表明,深度家系对于计算父系遗传的男性Y染色体在家族中的变化速度非常有价值。

(三) 遗传结构

正如前文所提到的,Balaresque通过5000多个Y-STR数据,筛选出了欧亚大陆上包括星簇和满

族簇的 11 个晚近扩张信号^[27]。相比之下,二代测序技术应用于 Y 染色体全序列,能够构建更多信息量的谱系树,它的枝长与时间成比例,因此,可以直接评估每个节点的最近共祖时间(TMRCA)。此外,根据基于高通量测序的谱系树,我们可以系统地分析不同的因素如何影响人口大小和结构。比如,这些有着较长枝长的支系暗示了较低的人口增长和频繁的瓶颈,而那些被视为星状结构的有着较短枝长的支系被解释为人口快速扩张的强烈信号。除了星状结构,有时在谱系树中存在一种有规律的分叉结构,它也与人口扩张有关。但是,这两种不同的结构所反映的人口历史是完全不同的。比如,Sikora 等通过溯祖模拟发现,R1b 星状结构和 E1b1a 的分叉结构分别与西欧亚和撒哈拉以南非洲的新石器时代扩张有关^[43]。具体而言,R1b 的星状结构表明,当它进入这块大陆时,就在很短的时间内发生了成功的扩张,而 E1b1a 有着非常有规律的分叉结构,显示 E1b1a 扩张始于较大的人口大小,并持续了数千年。

关于东亚的父系人口历史,2014 年,严实等应用二代测序技术发现了 Y 染色体上约 4 000 个新的遗传标记,并定义了大量新的分支,发现东亚 3 个父系星状扩张发生在约 5 000 ~ 6 000 年前的新石器时代,现代约 40% 的中国人来自当时这 3 个子嗣繁多的祖先^[44]。我们最近的研究进一步观察到这次新石器时代的强烈的瓶颈效应和人口扩张^[45]。东亚特有的父系谱系中,有 5 个呈星状结构——O3a2c1a-F5(O α)、O3a2c1-F46(O β)、O3a1c1-F11(O γ)、C3a1-F2613(C α)和 Q1a1-M120(Q α),以及一个呈复杂分叉结构——O1a1a1-F78(O σ)。值得注意的是,我们发现 65% 的现代中国人都是这 6 个超级祖父的后代,他们的扩张时间分别为 5 400 年前(O α)、6 500 年前(O β)、6 800 年前(O γ)、6 400 年前(C α)、5 200 年前(Q α)和 5 000 年前(O σ)^[46]。正如前文所述,我们揭示了汉族主体来自于新石器时代的少数几个父系祖先,但他们是谁?每个超级祖父与古史传说中的传奇领袖是否有关?不久的将来,联合古 DNA 和现代人 DNA 的研究将有助于解决这些问题,它对于理解中国人群的起源非常关键。

(四) 考古应用

考古发掘往往面临一个问题,汉之前的遗址极少有墓砖、墓志等表明身份信息的出土物,再加上墓地被盗严重(墓志等有经济价值),绝大部分遗骨缺少身份信息,但是对于墓地发掘来说,确定墓主人的身份是考古学的一个核心问题。

山东省临沂市洗砚池晋墓位于洗砚池街北侧、王羲之故居公园东北部。此处共发现两座墓葬,其中,M1 为双室墓葬,保存完整,同一墓葬埋葬 3 个未成年人,为考古发掘中极为罕见的现象;M2 为夫妻合葬墓,墓室结构完整,在一号墓之西,与一号墓相距不远。该遗址一经发现,便引起了社会和学术界的广泛关注,被评为 2003 年度十大考古发现之一。但是,对于墓主人的身份一直存在两种猜测:一种认为是琅琊王氏的后代,另一种认为是西晋司马氏的后代。2016 年,杜盼新等人对该遗址进行了基因家谱调查:(1) 排除了 M1 西室 6 岁女童与 M2 夫妇间的亲缘关系;(2) 鉴定了 M2 墓葬中男性遗骨的父亲遗传类型为 C3 南支-F948 + ,F3880 + ,该遗传类型在北方汉族中最为常见^[47]。该研究为后续琅琊王氏和司马家族的父亲遗传类型研究奠定了基础。

还有一个案例是复旦大学和吉林大学古 DNA 课题组的一次合作。两个课题组对内蒙古阿鲁科尔沁旗辽代耶律羽之墓地的两个男性人骨进行了古 DNA 鉴定。一个遗骸为 M6:3(耶律道清,耶律羽之的三代子孙),一个为 M10(无墓志)。实验结果表明,两个个体的遗传类型分属 N 南支和 N 北支,这两个支系的分化时间已有上万年。通过父系亲缘关系鉴定,排除了 M10 是耶律羽之家族男性后代的可能,其具体身份有待进一步比较验证。

基于上述两个案例,我们可以看到,对于大量已经发掘的身份未定的墓地,DNA 鉴定并将其结果纳入一个可以相互对比的古 DNA 数据库的重要性。依据这个数据库,我们可以基于亲缘关系,推测未知遗骸的可能身份。另外,进一步结合碳十四测年技术和文献材料,多重证据锁定未知墓主的身份,真正发挥科技考古的威力。

(五) 法医应用

一方面,法医学的DNA指纹法为基因家谱学奠定了基础,另一方面,族群、氏族、家族水平的谱系研究也促进了法医遗传学的发展。

甘肃白银市连环杀人案的破获是法医学运用基因家谱学原理的一个成功案例^[48]。白银连环杀人案是中华人民共和国成立以来发生的重大恶性案件之一。从1988年到2002年,11位女性(包括一名8岁的女童)遭到强奸和杀害。嫌疑犯逍遥法外长达30年,警方做出了各种努力,如指纹比对、DNA比对、按口述画像、重金悬赏等,但均无功而返。但是前期的积累是有意义的,尤其是Y染色体数据库的建立,成为后续破案的关键。当嫌疑人的远方堂叔犯案,警方发现其Y染色体单倍型与当年疑犯的DNA类型一致时,整个高氏家族男性都被纳入了排查范围,最终,白银市连环杀人案得以快速破获。该案件的破获是Y染色体基因家谱学的成功运用,这意味着法医数据库不仅只能做同一认定,每个男性数据背后代表着一个大的同姓家族。Y染色体DNA分型和基因家谱学原理正在成为法医遗传学的常规手段。

近年来,本课题组开发了一个专门用于群体遗传学和法医学应用的Y染色体数据库^{[49][50]},通过该数据库,我们推测了入缅远征军^[50]、淮海战役士兵^[51],以及一些刑事案件中的疑犯和失踪人口的可能地理来源和父系亲属。

2007年,我国首个历史人类学专业刊物——《现代人类学通讯》问世了。该杂志旨在从社会、语言、民俗、历史、考古、体质等方面回答人类起源和演化的问题。基因家谱学属于其中的一个重要板块,该刊物上发表了两项相关研究。操姓主要有鄱阳郡操姓和重庆长寿操姓两大分支。据传,鄱阳郡操姓源自逃难的曹操后人。2012年,王传超等的Y染色体调查显示,鄱阳郡操姓与曹操家族,乃至其他曹姓均无关系^[52]。2014年,王传超等又依据谱牒材料,对赛典赤及郑和的后裔——云南的纳姓和马姓进行父系Y染色体分型,发现他们属于南亚西部常见的单倍群L1a-M76,揭示了赛典赤与郑和的波斯祖源^[53]。

三、展 望

姓氏最早出现于中国,其历史可追溯到5000年前。在其他国家中,西欧人姓氏最早的记录不足千年(爱尔兰人约900年),日本人的姓氏也才150余年^[54]。我国现有姓氏4100个,但大多数姓氏的起源并不单一。周朝的姓氏大多是以封国为氏,后改为姓,封国内的百姓来源本来就是多种多样的。所以中国的姓氏总体上内部父系血缘并不完全一致。根据我们对超过10000例带有姓氏信息的Y-SNP/STR数据的分析,在同一个姓氏中,几乎可以观察到中国人中常见的所有父系类型,同时,在同一个父系类型中,也几乎可以看到中国人群中所有常见的姓氏。这意味着姓氏和Y染色体不是简单地一一对应关系,基于较小样本量的推测,Y染色体类型与姓氏之间的关联没有意义。但是,修著家谱一直是中国人的传统,一些家谱甚至跨越3000年。这些谱牒材料和深度家系都是我国的人文资源。研究族群/姓氏内部和深度家系的遗传动态,可能是厘清族群/姓氏来源和演化的一个事半功倍方案。

现如今,国外的基因家谱学研究正被商业公司所引领,主要原因是:(1)该学科的原理、测试和分析手段日渐成熟,商业公司可以以标准化流程提供更好的服务;(2)在不考虑科研目的的情况下,商业公司更能满足普通大众寻根问祖的个性化需求;(3)嫁接互联网技术,消费者可以在线下单,通过唾液采集器收集口腔上皮细胞,邮递至公司,最后的报告也会通过移动端反馈结果,方便快捷;(4)有专业的人员进行项目计划、组织和结果的个性化解读,客户体验感好。因此,基因家谱学正成为娱乐性遗传检测的一部分,这些检测包括通过全基因组标记去评估个人祖源、亲缘关系和疾病易感性等。可以预见的是,我国的基因家谱学研究也会朝着这个方向发展。

国外的业余爱好者们展示着令人惊讶的关于分子进化、群体遗传学和统计学的知识^[55]。比

如,他们时常在 *Journal of Genetic Genealogy* 杂志上发表文章。该杂志虽然缺少标准的同行评议审稿,但是有一些遗传学工作者也乐在其中。这可能也是让公众参与科学研究的一个有趣的例子。我们相信,随着基因家谱学的进一步推广,国内业余爱好者的水平也会显著提高。《现代人类学通讯》更关注东亚的族群/家族历史,反对任何形式对权威的盲从,同时面向专业和业余作者,促进人类学的各个领域百花齐放、百家争鸣。

基因家谱学的推广也暗含风险,需要格外警惕。首先,可能面临的风险是意外地发现了一次非父事件或者一段珍贵的口述历史被证明是不可靠的。在声称是美国总统 Thomas Jefferson 后代的家系的鉴定案例中,上述两种情况同时发生,这些家属长期不能接受检测结果,造成不良影响^[56]。其次,尽管 Y 染色体被认为和疾病没有太大关系,但是一些 Y 染色体的缺失被发现与男性不育有关^[57],这意味着一些特殊的 Y-STRs 和 SNPs 的缺失^[58],这些结果涉及检测者的个人隐私。最后,Y 染色体谱系常常被误导性地对应地理起源和族群分布^[59],引起不必要的纠纷。但是,如果前期知情同意工作做得细致,后期解释和沟通工作做到位的话,这些风险完全可以避免。

再次回想梁启超的“大图书馆梦”,或许借助互联网的力量收集、整理谱牒,借助 DNA 的力量鉴别谱牒,并集聚基因家谱学从业者和爱好者的共同力量,“能尽集天下之家谱俾学者分科研究,实不朽之盛业也”。

参考文献:

- [1] 陈梦家. 殷墟卜辞综述 [M]. 北京: 科学出版社, 1956.
- [2] 常建华. 家族谱研究概况 [J]. 中国史研究动态, 1985(2).
- [3] 汪俊. 略论谱牒学在文史研究中的意义 [J]. 扬州大学学报(人文社会科学版), 1997(3).
- [4] 梁启超. 中国近三百年学术史·谱牒学 [M]. 上海: 东方出版社, 2004.
- [5] 刘黎明, 樊雄, 李鉴踪, 等. 祠堂·灵牌·家谱: 中国传统血缘亲族习俗 [M]. 成都: 四川人民出版社, 1993.
- [6] 王云度. 谱牒学概述 [J]. 江苏师范大学学报(哲学社会科学版), 1991(2).
- [7] Krause J, Paabo S. Genetic Time Travel [J]. *Genetics*, 2016(1).
- [8] Calafell F, Mhd L. The Y Chromosome as the Most Popular Marker in Genetic Genealogy Benefits Interdisciplinary Research [J]. *Human Genetics*, 2017(5).
- [9] 李辉. Y 染色体与基因家谱 [J]. 世界科学, 2013(2).
- [10] Tëtushkin E Y. Genetic Genealogy: History and Methodology [J]. *Russian Journal of Genetics*, 2011(5).
- [11] Jeffreys A J, Wilson V, Thein S L. Hypervariable 'minisatellite' Regions in Human DNA [J]. *Nature*, 1985, 314(6006).
- [12] A. J. Jeffreys, V. Wilson, S. L. Thein. Individual-specific 'fingerprints' of human DNA [J]. *Nature*, 1985, 316(6023).
- [13] Ivanov P L. Human Individualization and Personal Identification: Molecular Biology in Forensic Medicine [J]. *Herald of the Russian Academy of Sciences*, 2003(12).
- [14] Gill P, Ivanov P L, Kimpton C, et al. Identification of the Remains of the Romanov Family by DNA Analysis [J]. *Nature Genetics*, 1994(2).
- [15] Jeffreys A J, Allen M J, Hagelberg E, et al. Identification of the Skeletal Remains of Josef Mengele by DNA Analysis [J]. *Forensic Science International*, 1992(1).
- [16] Skorecki K, Selig S, Blazer S, et al. Y Chromosomes of Jewish Priests [J]. *Nature*, 1997, 385(6611).
- [17] Thomas M G, Skorecki K, Ben-Ami H, et al. Origins of Old Testament Priests [J]. *Nature*, 1998, 394(6689).
- [18] Foster E A, Jobling M A, Taylor P G, et al. Jefferson Fathered Slave's Last Child [J]. *Nature*, 1998, 396(6706).
- [19] Decorte R. Mitochondrial DNA Analysis on Remains of a Putative Son of Louis XVI, King of France and Marie-Antoinette [J]. *European Journal of Human Genetics*, 1998(4).
- [20] Lucotte G, Thomasset T, Hrechdakian P. Haplogroup of the Y Chromosome of Napoléon the First [J]. *Journal of Mo-*

- molecular Biology Research 2011(1) .
- [21] Larmuseau M H D ,Delorme P ,Germain P ,et al. Genetic Genealogy Reveals True Y Haplogroup of House of Bourbon Contradicting Recent Identification of the Presumed Remains of Two French Kings [J]. European Journal of Human Genetics 2014(5) .
- [22] Coble M D ,Loreille O M ,Wadhams M J ,et al. Mystery Solved: The Identification of the Two Missing Romanov Children Using DNA Analysis [J]. Plos One 2009(3) .
- [23] Yan S ,Tachibana H ,Wei L H ,et al. Y Chromosome of Aisin Gioro ,the Imperial House of the Qing Dynasty [J]. Journal of Human Genetics 2014(6) .
- [24] Zerjal T ,Xue Y ,Bertorelle G ,et al. The Genetic Legacy of the Mongols [J]. American Journal of Human Genetics , 2003(3) .
- [25] Wang C ,Yan S ,Hou Z ,et al. Present Y Chromosomes Reveal the Ancestry of Emperor CAO Cao of 1 800 Years Ago [J]. Journal of Human Genetics 2012(3) .
- [26] Xue Y ,Zerjal T ,Bao W ,et al. Recent Spread of a Y-Chromosomal Lineage in Northern China and Mongolia [J]. American Journal of Human Genetics 2005(6) .
- [27] Balaesque P ,Poulet N ,Cussatblanc S ,et al. Y-chromosome Descent Clusters and Male Differential Reproductive Success: Young Lineage Expansions Dominate Asian Pastoral Nomadic Populations [J]. European Journal of Human Genetics 2015(10) .
- [28] Tëtushkin E Y. Genetic Genealogy: History and Methodology [J]. Russian Journal of Genetics 2011(5) .
- [29] Royal C D ,Novembre J ,Fullerton S M ,et al. Inferring Genetic Ancestry: Opportunities ,Challenges and Implications [J]. American Journal of Human Genetics 2010(5) .
- [30] Wang C C ,Yan S ,Yao C ,et al. Ancient DNA of Emperor CAO Cao ' s Granduncle Matches Those of His Present Descendants: a Commentary on Present Y Chromosomes Reveal the Ancestry of Emperor CAO Cao of 1 800 Years Ago [J]. Journal of Human Genetics 2013(4) .
- [31] 文少卿 ,王传超 ,敖雪 等. 古 DNA 证据支持曹操的父系遗传类型属于单倍群 O2 [J]. 人类学学报 2016(4) .
- [32] Abilev S ,Malyarchuk B ,Derenko M ,et al. The Y-chromosome C3* Star-cluster Attributed to Genghis Khan ' s Descendants is Present at High Frequency in the Kerey Clan from Kazakhstan [J]. Human Biology 2012(1) .
- [33] Wei L H ,Yan S ,Yu G ,et al. Genetic Trail for the Early Migrations of Aisin Gioro ,the Imperial House of the Qing Dynasty [J]. Journal of Human Genetics 2016 (3) .
- [34] Wei L H ,Yan S ,Lu Y ,et al. Whole Sequence Analysis Indicates that the Y Chromosome C2* -Star Cluster Traces Back to Ordinary Mongols ,rather than Genghis Khan [J]. European Journal of Human Genetics 2017(in pressing) .
- [35] 张晓东. 蒙元时期的蒙古人墓葬 [D]. 吉林大学硕士学位论文 2006.
- [36] Lkhagvasuren G ,Shin H ,Si E L ,et al. Molecular Genealogy of a Mongol Queen ' s Family and Her Possible Kinship with Genghis Khan [J]. Plos One 2016(9) .
- [37] Cui Y ,Song L ,Wei D ,et al. Identification of Kinship and Occupant Status in Mongolian Noble Burials of the Yuan Dynasty Through a Multidisciplinary Approach [J]. Philosophical Transactions of the Royal Society of London , 2015 370(1660) .
- [38] Batbayar K ,Sabitov Z M. The Genetic Origin of the Turko-Mongols and Review of The Genetic Legacy of the Mongols. Part 1: The Y-chromosomal Lineages of Chinggis Khan [J]. The Russian Journal of Genetic Genealogy 2012(2) .
- [39] Wei L H ,Huang Y Z ,Yan S ,et al. Phylogeny of Y-chromosome Haplogroup C3b-F1756 ,an Important Paternal Lineage in Altaic-speaking Populations [J]. Journal of Human Genetics 2017(10) .
- [40] Zh S ,Turuspekov Y ,Daulet B ,et al. The Kazakhstan DNA Project Hits First Hundred Y-profiles for Ethnic Kazakhs [J]. Russian Journal of Genetic Genealogy 2012(1) .
- [41] Wang C C ,Li H. Evaluating the Y chromosomal STR Dating in Deep-rooting Pedigrees [J]. Investigative Genetics , 2015(1) .
- [42] 杜盼新 ,文少卿 ,王凌翔 等. 重构司马光家族基因家谱 [J]. 人类学学报(出版中) .
- [43] Sikora M J ,Colonna V ,Xue Y ,et al. Modeling the Contrasting Neolithic Male Lineage Expansions in Europe and Af-

- rica [J]. *Investigative Genetics* 2013(1) .
- [44] Yan S ,Wang C C ,Zheng H X ,et al. Y Chromosomes of 40% Chinese Descend from Three Neolithic Super-Grandfathers [J]. *Plos One* 2014(8) .
- [45] Wang C C ,Huang Y ,Yu X ,et al. Agriculture Driving Male Expansion in Neolithic Time [J]. *Science China (Life Sciences)* 2016(6) .
- [46] Wen S Q ,Tong X Z ,Li H. Y-chromosome-based Genetic Pattern in East Asia Affected by Neolithic Transition [J]. *Quaternary International* 2016(426) .
- [47] 杜盼新 ,骆潇沁 ,文少卿 ,等. 临沂洗砚池晋墓遗骸 DNA 研究报告 [A]. 山东省文物考古研究所 ,临沂市文化广电新闻出版局. 临沂洗砚池晋墓 [C]. 北京: 文物出版社 2016.
- [48] Yao H ,Wen S ,Tong X ,et al. Y Chromosomal Clue Successfully Facilitated the Arrest of Baiyin Serial Killer [J]. *Science Bulletin* 2016(22) .
- [49] Wang C C ,Wang L X ,Shrestha R ,et al. Convergence of Y Chromosome STR Haplotypes from Different SNP Haplogroups Compromises Accuracy of Haplogroup Prediction [J]. *Journal of Genetics and Genomics* 2015(7) .
- [50] Wen S Q ,Tong X Z ,Wang C Z ,et al. Y-chromosomes from Skeletal Remains of Chinese Expeditionary Force offer a Clue to Their Paternal Relatives [J]. *Science Bulletin* 2016(6) .
- [51] 王迟早 ,文少卿 ,石美森 ,等. 淮海战役士兵遗骸的 Y 染色体分子鉴定 [J]. *法医学杂志* 2017(4) .
- [52] 王传超 ,严实 ,韩昇 ,等. 鄱阳操姓血缘上并非出自曹操 [J]. *现代人类学通讯* 2012(6) .
- [53] 王传超 ,王凌翔 ,张曼菲 ,等. Y 染色体揭示赛典赤·赡思丁和郑和的波斯祖源 [J]. *现代人类学通讯* 2014(8) .
- [54] King T E ,Jobling M A. What's in a name? Y Chromosomes ,Surnames and the Genetic Genealogy Revolution [J]. *Trends in Genetics* 2009(8) .
- [55] Balanovsky O ,Gurianov V ,Zaporozhchenko V ,et al. Phylogeography of Human Y-chromosome Haplogroup Q3-L275 from an Academic/citizen Science Collaboration [J]. *BMC Evolutionary Biology* 2017 ,17(Suppl 1) .
- [56] Williams S R. Genetic Genealogy: the Woodson family's Experience [J]. *Culture Medicine & Psychiatry* 2005(2) .
- [57] Vogt P H. AZF Deletions and Y Chromosomal Haplogroups: History and Update based on Sequence [J]. *Human Reproduction Update* 2005(4) .
- [58] King T E ,Bosch E ,Adams S M ,et al. Inadvertent Diagnosis of Male Infertility Through Genealogical DNA testing [J]. *Journal of Medical Genetics* 2005(4) .
- [59] King T E ,Bowden G R ,Balaresque P L ,et al. Thomas Jefferson's Y Chromosome Belongs to a Rare European Lineage [J]. *American Journal of Physical Anthropology* 2010(4) .

【责任编辑 海晓红】

The Development History of Genetic Genealogy in China

WEN Shao-qing¹ ,HAN Sheng² ,LI Hui¹

(1. Ministry of Education Key Laboratory of Contemporary Anthropology ,Fudan University ,Shanghai 200433 ,China;

2. History Department ,Fudan University ,Shanghai 200433 ,China)

Abstract: Genetic genealogy is an applied science which adopts genetics to traditional genealogy ,raising lots of concerns daily. At first ,this article traces the development history of genetic genealogy including its source ,infancy and latest advance in worldwide especially in China. Next ,the author mainly focuses on the aspects of family history ,group structure , mutation rate ,archaeological application and forensic application. New achievements have been made in interdisciplinary studies which combines genetic genealogy with other disciplines. Finally ,the author hopes that the descriptions of principles and case studies will help practitioner and amateur to promote genetic genealogy in China in the future. It is worthwhile to note that the popularity of genetic genealogy also follows with some drawbacks mainly because of privacy issues related to the DNA donor ,his close family and far-related namesakes.

Key words: Genetic Genealogy; DNA Genealogy of Y Chromosome; Celebrity Family History; Deep-rooting Pedigrees; Forensic Application