

# 苗族上古起源的基因与文化综合解析

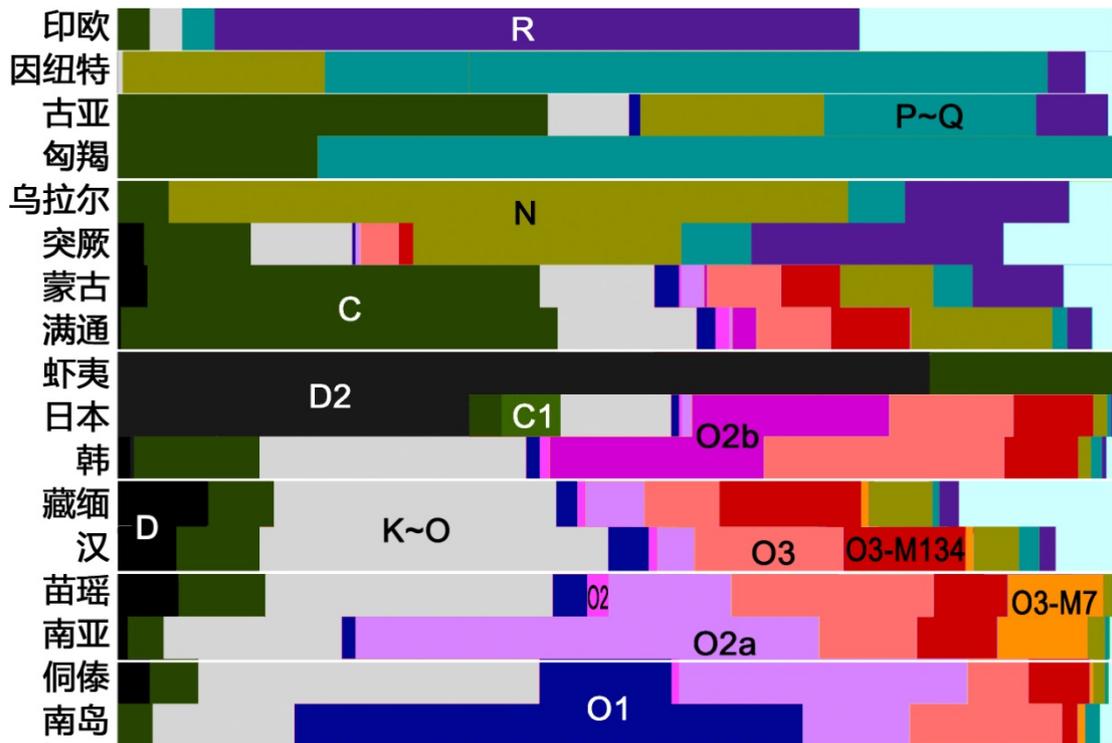
李 辉

复旦大学人类学与人类遗传学系

人类的演化历史中不断发生着群体分化的过程。人属约 230 万年前起源于东非，分化成若干物种，其中早期的主流物种为能人。约 190 万年前，匠人从能人中进化出来，到 140 万年前全部取代能人。约 180 万年前，部分匠人走出非洲，演化成亚洲的直立人亚种。约 120 万年前，非洲匠人中进化出智人物种。约 80 万年前，海德堡智人走出非洲，40 万年以后，在欧亚大陆西部演化成尼人亚种，东部演化成丹人亚种，留在非洲的演化成罗德西亚人亚种。约 20 万年前东非的罗人进化出现代人亚种。约 7 万年前，部分现代人走出非洲，扩散到全世界。由于旧石器时代的地理隔离，现代人长期适应不同区域的气候环境，演化成了八个地理种，包括东非种、中非种、西非种、地中海种（旧称高加索种）、东亚种、巽他种、萨胡种、美洲种【1】。近一万年内，人类进入新石器时代，孕育出丰富多样的文化，不同地域的人群产生了文化隔离，演化出民族。所以，民族是新石器时代以来文化隔离的产物。但是由于长期隔离，民族之间也积累了遗传差异，差异程度随隔离时间而加大。所以文化差异是民族形成的原因，遗传差异是民族形成的结果。民族不以遗传血缘来判定，但遗传差异可以反过来分析民族的演化历程。

## 从遗传学和考古学分析东亚诸语系演化的时空框架

民族文化的差异中，最核心的是语言差异。东亚语言间的演化关系，虽然比非洲和西亚要复杂一点，但是比美洲和澳洲的情况要简单得多。而且东亚的语言学调查和研究做得非常细致。但是东亚语系划分问题，至今国内外争议颇多。这可能是由于在具体语系划分时没有执行相对统一的标准。两个语言类群分化超过多久可以算作不同语系，这应该有一个相对统一的标准。目前较为一致的看法是：两个语言类群分化超过 8000 年，其间的相似性就难以判别了，如果超过一万年，可能就完全看不出关系了。这在遗传学上也有相似的表现。8000 年这个标准年代，应该是一个很有意义的年代，因为这是新石器时代和农业全面开始的年代。以此为语系划分的标准，也就是说语系是新石器时代人群文化区系集中的结果。那么，不同的语系势必要追溯到不同的新石器时代文化区系。反过来说，如果没有独特的新石器文化，语系就没有时空来源。所以语系必定有着承载它的人类群体以及造就它的新石器文化区系，并孕育出民族类群。遗传学的人群区分和年代计算、考古学的区系文化比较，在语言学的语系演化的研究中应该起到重要的作用。

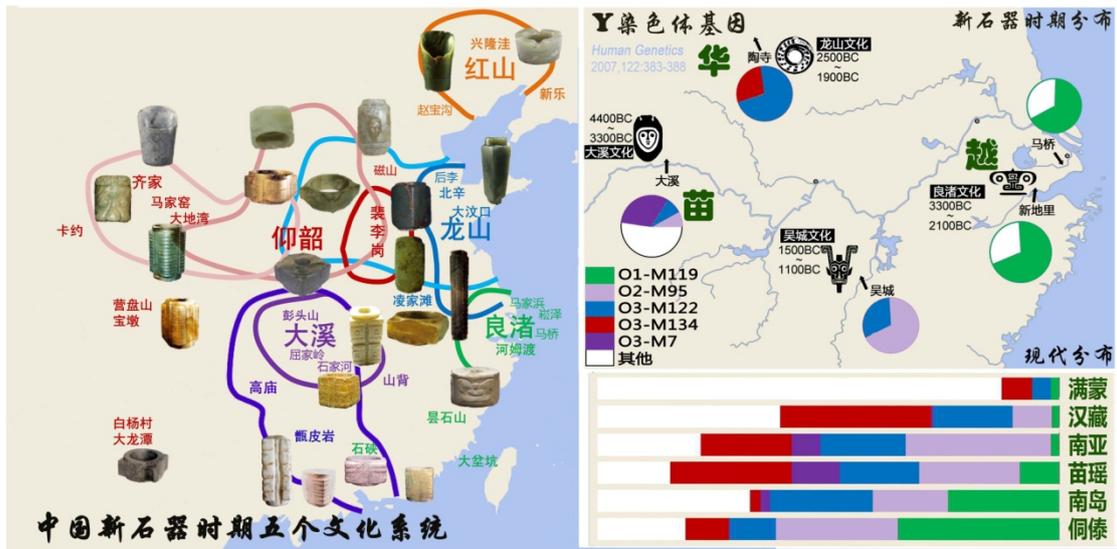


东亚各语言族群的 Y 染色体类型比例（苗瑶中比例最高的 O3 部分基本是 O3-F11）

由于大多数民族群体经历了长期的父系社会历史，所以父系遗传的 Y 染色体与民族群体的相关性最明显。从遗传学的分析，特别是 Y 染色体精细分型数据看来，东亚的各个语言类群的人群之间的分化年代已经比较清楚了。汉语族和藏缅语族的分化年代大约是 5~6 千年，也就是在 8 千年之内，所以支持汉藏语系的概念。汉语族与乌拉尔语系分化略超过 8 千年，毫无疑问是不同的语系，即便还有些许同源词。汉藏与苗瑶之间的原初分化在 1~1.2 万年，就支持把汉藏语系与苗瑶语系分开。而侗傣与前二者的分化年代超过了 1.5 万年，更加支持其独立的语系地位。后期的语言接触造成的语言相似性应该是被排除在语系划分所考虑的因素之外的。所以历史时期汉藏南迁造成的苗瑶、侗傣与汉藏之间的诸多相似性，不能用以支持三者的合并。侗傣与南岛的人群之间仅有大约 6 千年的分化，其语言间的同源词也比较清晰，按此标准可以作为一个语系，不过两者在类型学上差异过大，是分是合可以进一步讨论。

考古文化区系与语系起源的对应关系，也可通过遗传学的古 DNA 分析来确认。在良渚文化区系中的人骨 DNA 检测出高频的 Y 染色体 O1 单倍群，与侗傣和南岛语的族群高度一致。从 5900 年前的崧泽文化开始，长江下游的人群开始受到中华文明的影响，体现在八角星纹等符号的传播，侗傣族群和南岛族群的祖先可能就此起始分化。南岛族群的祖先或是分布于闽粤台的大坌坑文化人群。长江中游的大溪文化区系中检测出了高频的 Y 染色体 O3-F11 和 O3-M7 单倍群，这又与现代的苗瑶族群高度一致。黄河中下游地区有两个文化区系，西边的仰韶和东边的大汶口-龙山。而两个区系的边界在不断地往西推，从最早的河南-山东边界，渐渐到了陕西-河南，最后到了甘肃青海-陕西。仰韶文化（马家窑）退到了藏缅族群的分布区，龙山文化彻底占据了中原，融合了部分仰韶的因素。这可能体现了汉语族和藏缅语族先民的早期的冲突和互动。汉语族先民带着龙山文化在西进的过程中，人口上吸纳了仰韶的居民，文化上与之同化。而拒绝同化的部分仰韶

先民只能向西退到甘青地区，成为藏缅语族的人群。造成的效应是现代的藏缅语族人群和汉族之间，在语言的多样性上是西高东低，在遗传的多样性上是东高西低。这种相反的结构，能用前述的过程解释。龙山文化的遗骸检测出高频的 Y 染色体 O3-M122(F11)和 O3-M134，与现代汉族的遗传结构吻合。西辽河流域的红山文化区系也是一个重要的部分，其中的高频的 N 和 O3 单倍群显示，这个文化区系可能与乌拉尔语系人群和汉族的祖先有关，很可能是乌拉尔人的起源地。东北以龙为标记的红山文化和西南以凤为标记的高庙文化在大汶口文化的碰撞和融合，可能奠定了汉族最早的基础。这需要遗传学上的进一步证明。



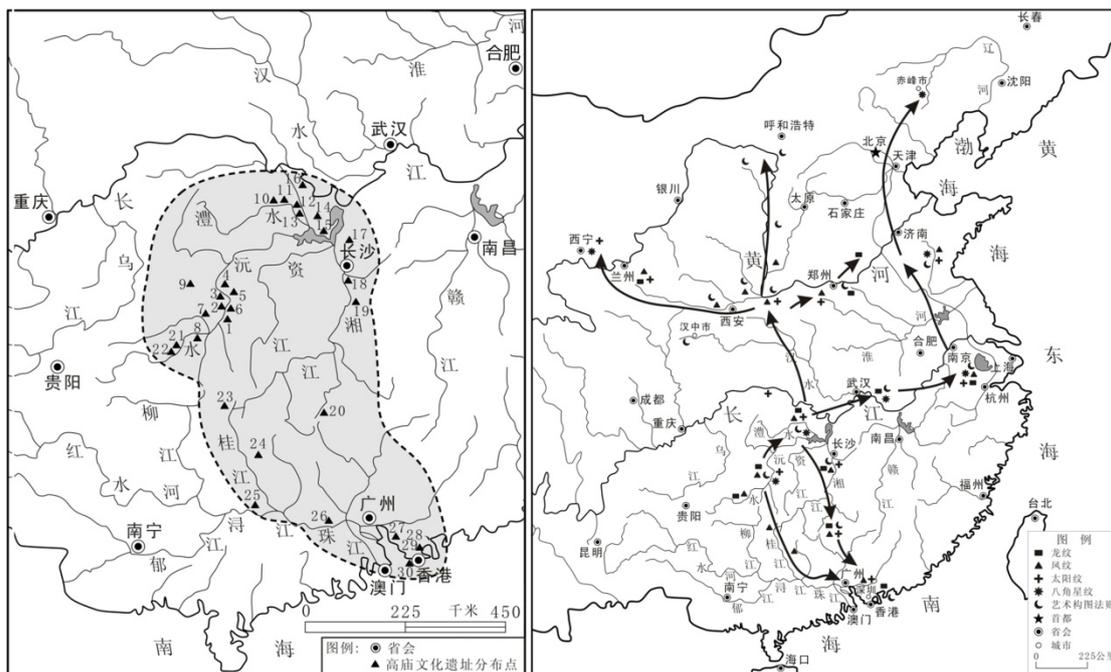
考古文化与民族基因的关联

所以，从整个东亚地区新石器时代与现代的族群分布看来，苗族最有可能是起源于湖广地区，与现在的分布区并不遥远。其源头可能与湖广地区的大溪文化或者更早的高庙文化有关。

### 苗族服饰中的高庙文化遗存

苗族的服饰有着其鲜明的特色，尤其是其服饰上的各种图案花纹，有着苗族族群的独特造型风格，也强烈地暗示着这个族群的起源。其中最具有代表性的纹样是凤鸟纹与八角星纹。

凤鸟图案在中国各个民族服饰上普遍出现，已经成为中华文化符号之一。但是苗瑶的凤鸟纹造型却有自己的特色，最典型的是苗瑶花纹中凤鸟的嘴喙是向下弯曲的，而其他民族的凤鸟嘴喙往往是直的。苗瑶的凤鸟往往夸大其眼睛，而并不重视体现其尾部，这一点也与汉族为代表的其他民族的凤鸟造型不同。在很多苗族的服装上，凤鸟甚至只剩下一只大眼睛和一只大弯嘴的组合。有观点认为这种纹样可能来自长江下游的良渚文化或者黄河流域的仰韶文化。虽然这两个新石器考古文化中有很多鸟的图案出现，但是飞翔的鸟类是很容易引起人们联想和崇拜的对象，仅仅喜鸟这一特征并不足以说明族群之间的关系。另外，良渚文化与仰韶文化中的鸟类造型与苗瑶凤鸟差异还是较大，很难建立必然联系。然而，在新石器考古文化中，有一个文化却大量出现与苗瑶凤鸟几乎完全一样的造型。那就是大约 7800 年前起源于湘西的高庙文化【2】。



高庙文化的分布和传播【2】



始建于 6800 年前的城头山古城是中国最早的城市遗址

高庙文化可能是东亚最早出现的文化面貌系统完整的文化类型，出现了大量稳定的刻画图案和符号，表现着当时的人群有着极其完善的信仰体系。高庙文化的分布区域覆盖湖北南部、湖南全境、贵州和广西东部、广东西部一直到香港。在大约 6800 年前建立了东亚第一个城——城头山古城，位于湖南常德境内，至

今还保留着部分护城河。高庙文化还不断向四周扩散影响，长江下游、黄河下游是其重点影响的区域。高庙文化中出土的陶器中，用点刻的方式绘制了大量凤鸟图，而其中最常见凤鸟造型就是大眼睛和大弯嘴的组合，与现代苗瑶服饰上的凤鸟造型完全一致。这种特殊的表现方式，并不容易用人类普遍的认知或者巧合来解释，非常有可能是直接的传承。何况高庙文化与苗瑶族群的起源与分布范围是几乎重合的。所以寻找苗瑶凤鸟纹饰的起源应该不必舍近求远、舍早求晚、舍似求非。高庙文化很可能是其最近、最早、最像的源头。汉族式的直嘴的凤鸟可能是高庙文化的凤鸟概念传入以后其自己演绎发展出来的。而苗瑶还保留着高庙最早的凤鸟造型。



高庙文化桂阳千家坪凤鸟纹与现代苗族服饰凤纹比较

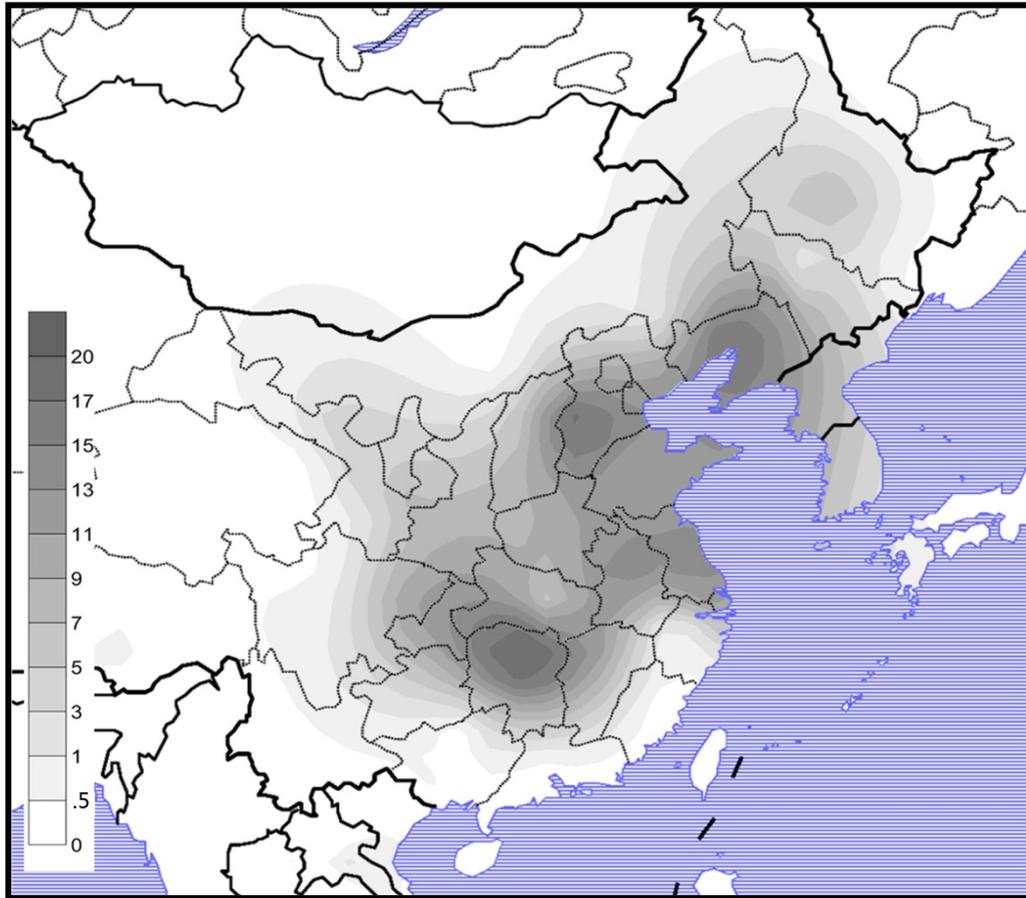
除了凤鸟图案，苗瑶服饰上还常常出现八角星图案，这种图案的源头恰恰也在高庙文化中。7800年前，高庙文化中就已经出现了八角星纹。6800年前的城头山古城中出土的白陶盆中绘制的八角星纹，已经几乎与现代苗瑶服饰上的八角星纹完全一致了。这种纹饰随着高庙文化的辐射传播到了周边的文化中，例如江浙的崧泽文化。而山东的大汶口文化是八角星纹最集中出现的第二个文化类型。大汶口文化非常可能直接传承自高庙文化，从考古遗存上看，高庙文化从长江中游传播至下游，再从苏北到山东的路径比较清晰。所以苗瑶的八角星纹应该也是起源于高庙文化的。虽然其他民族的服饰上偶尔也会出现八角星纹，比如侗族、水族、布依族等，但是这些民族的八角星纹有更多的变形和演绎，会添上花瓣状的光芒，与高庙的造型不同，可能是对八角星纹的内涵有着完全不同的理解。



高庙文化城头山遗址白陶盘八角星纹与现代苗族八角星纹比较

## 苗族古史传说与遗传和考古证据吻合

山东与湖广考古文化上的关联，与古史传说中这两地民族的关联是吻合的。传说中湖广是三苗族群的分布区，而山东是东夷九黎的分布区，而三苗与九黎有着密切的渊源关系，相对与神农和轩辕的炎黄族群，三苗九黎属于同类。如果山东的大汶口文化起源于湖南的高庙文化，是否也暗示着九黎族群本来就起源于原初的三苗族群，承载着文化的人群也从湖南扩张到了山东。有趣的是，这种假设在对考古遗址人骨和现代人群的遗传基因分析中也得到了证据支持。因为族群的父系主导和传承，分析族群源流最合适的遗传材料是父系遗传的 Y 染色体【1】。山东人群中最高频的 Y 染色体是 O3-F11，也是苗族的高频类型之一，这种类型还高频出现在河北东南部和湖南【3】。而其中湖南的 O3-F11 谱系分化程度最复杂，可追溯的历史最久，也就是说山东人中有很大一部分上古时期来自湖南。O3-F11 这个谱系是非常特殊的一个谱系，在目前的中国男性中占了大约 19% 的比例，在各同级谱系中最高。而这 19% 的男性从谱系上都只能追溯到约 6800 年前的同一个祖先【4】，也就是说很大一部分的汉族和苗族男人有着一位共同的祖先。这位共祖的年代与城头山古城的年代一致，这恐怕不是巧合。如前所述，在考古遗骸的分析中，我们还没有高庙文化的材料，但是我们分析了其后的大溪文化遗骸，其中有高频的 O3-F11 和 O3-M7，与现代苗瑶人群的类型是一样的。而我们在良渚文化遗骸中检测出高频的 O1 类型，这是现代的侗傣族群高频的类型，与大溪文化和苗瑶族群是不同的。在龙山文化陶寺遗址中，虽然有 O3-F11，但是没有 O3-M7。所以，从人群本身而言，我们倾向于认为苗瑶族群起源于湖广当地的新石器时代人群，而侗傣族群可能与江浙地区的新石器时代人群有关。



**Y 染色体 O3-F11 频率分布图**

（湖南的多样性最高，苏北山东的类型是湖南分出的一个支系）

现代苗族普遍认为他们的祖先是五千多年前的蚩尤。如果蚩尤真实存在过，那么由于他战败被杀，留下大量子孙的可能性很小。那么为何苗族对蚩尤有这么深的印象呢？因为各种巨大的灾难往往会给群体留下更深刻的印象。例如全世界流传的上古大洪水。很多时候灾难在群体记忆中会被夸大，例如南北朝的“五胡乱华”给北方汉族带来巨大灾难，而回顾这段历史的人大多会认为北方汉族的血缘几乎被胡人替代。实际上现代遗传学检测发现，历代混入北方汉族的外族血缘只是很少的比例。所以，蚩尤可能是苗族历史上的一位重要首领，但是不太可能是现代苗族血缘上的共祖。大多数苗族人的共同祖先从目前数据上应该追溯到更早的历史。我们注意到，苗族起源传说中的始祖“姜央”，其形象和故事版本与伏羲女娲完全一致，这或许更可能是苗族可以追溯的祖先。

传说中东夷九黎族群奉太昊、少昊为祖，东方与南方曾经一致的神话体系中太昊氏也是极为关键的。太昊伏羲演八卦，是否与八角星纹有关，或许是一个需要研究的课题。太昊也是汉族祖先的三皇之首，他是否就是汉族与苗族的 O3-F11 谱系的那位共祖？而传说中太昊氏以风为姓，是否就是“风”？苗族的神树“枫”与之又有何关系？高庙文化中出现的建木图案与枫树有何关系？通过民族学、考古学、遗传学、语言学的跨学科交叉研究，一幅美妙的史前人群演化图景正在慢慢展开。

【参考文献】

1. 李辉, 金力(2015)Y 染色体与东亚族群演化. 上海: 上海科技出版社.
2. 贺刚(2013)湘西史前遗存与中国古史传说. 长沙: 岳麓书社.
3. Wang CC, Yan S, Qin ZD, Lu Y, Ding QL, Wei LH, Li SL, Yang YJ, Jin L, Li H, the Genographic Consortium (2013) Late Neolithic expansion of ancient Chinese revealed by Y chromosome haplogroup O3a1c-002611. *J Systematics Evol* 51(3):280–286.
4. Yan S, Wang CC, Zheng HX, Wang W, Qin ZD, Wei LH, Wang Y, Pan XD, Fu WQ, He YG, Xiong LJ, Jin WF, Li SL, An Y, Li H, Jin L (2014) Y Chromosomes of 40% Chinese Descend from Three Neolithic Super-Grandfathers. *PLoS ONE* 9(8): e105691.