

文章编号: 0427-7104(2003)04-0621-09

遗传和体质分析草苗的起源

李 辉, 李 昕¹, 杨宁宁¹, 文 波¹, 侯井榕²,
黄兴球², 金建中¹, 金 力¹

(1. 复旦大学 生命科学学院 现代人类学研究中心, 上海 200433;

2. 广西民族学院 民族学人类学研究所, 南宁 530006)

摘 要: 我国湘桂黔交界处的草苗是苗族的一个特殊支系, 具有汉族、苗族和侗族的风俗, 操侗语, 是国内民族融合最复杂的群体之一, 对民族理论和群体遗传学理论研究有重要意义. 用主成分分析的方法研究了广西融水草苗的三类指标. 对草苗研究了 Y 染色体单核苷酸多态位点 (SNP) 单倍型, 发现其与汉族基本一致, 而与侗傣群体和苗瑶群体不同. 列出了草苗的体质特征, 发现其面貌特征接近侗傣群体. 对草苗的肤纹分析又发现其最接近苗瑶群体. 这说明草苗的父系祖先应该来自汉族; 而母系成分最早一段时期可能来自苗族, 使其心理上认同为苗族; 后来由于处于侗族的优势分布区中, 不断与侗族交流, 又吸收了侗族的遗传结构并接受了侗语. 草苗的复杂遗传结构说明中国的民族交流是频繁的, 研究民族源流时一定要充分认识到这种复杂性.

关键词: 草苗; Y 染色体单核苷酸多态 (Y-SNP); 体质; 肤纹; 族源; 主成分分析

中图分类号: Q 986; Q 983

文献标识码: A

草苗可能是中国最奇特的民族群体之一, 他们心理上认同为苗族, 又自称来源于汉族而唱汉歌, 却完全讲侗语^[1]. 而苗族、汉族、侗族分属于完全不同的苗瑶、汉藏 (氏羌)、侗傣 (百越) 族群, 3 种迥异的文化却在草苗人群中有机地结合了起来. 中国没有其他少数民族群体像草苗这样体现出如此复杂的民族融合现象. 所以研究草苗这一群体的成因, 对于民族交流融合的理论探讨有非常重大的意义. 在这类融合群体中研究人群交流融合造成的效应, 比如迁徙对受纳群体基因频率的影响, 对群体遗传学也有特殊的意义.

草苗现有人口约八万, 分布于广西三江和融水、贵州黎平、湖南通道一带 3 省交界地区^[1]. 文化人类学的研究推测其最早开始形成于湖南通道的戍边汉兵, 主要在贵州黎平苗汉通婚时壮大并成为苗族, 后来在迁到广西三江的侗族地区, 开始改用侗语^[1]. 2001 年 7 月, 我们组织对融水的各民族作调查时, 首先调查了草苗, 收集了体质人类学数据、肤纹和血样 DNA. 以期从遗传学和体质人类学方面探究这一独特群体的遗传结构和群体来源.

Y 染色体非重组区可以基本解决混血群体的分析问题, 其上的各类突变位点稳定地构成各种单倍型, 并且单倍型间存在严格的系统发生关系, 在不同系统的族群中有特异性的分布^[2]. 混合群体中的每一种 Y 染色体单倍型都可以通过各种分析方法找到它的来源, 而且由于 Y 染色体只有男性体内存在, 每个男子只有一份, 同一家族的个体都一样, 所以群体中的实际总体大小比其他遗传标记小得多, 采样时只要很小的样本量就可以了. 一般不相关样本量大于 6, 类型数据就有效, 大于 24, 频率数据有效. 因此这一遗传标记成了民族遗传研究的最佳材料^[3,4], 也是本项研究的主要手段.

体质特征除了受一定的环境影响外, 其遗传的部分同时取决于父母双方的贡献, 肤纹的遗传也同时来源于父母双方, 所以本项研究以此弥补 Y 染色体单纯父系遗传的不足.

* 收稿日期: 2002-10-11

基金项目: 国家自然科学基金九五计划重大项目 (39993420); 复旦大学学生科技创新活动基金 (2001B1-603)

作者简介: 李 辉 (1978—), 男, 博士研究生.

1 材料和方法

1.1 调查对象

在广西壮族自治区柳州地区融水苗族自治州四荣乡荣塘村选取苗族草苗支系成年人作为调查对象,调查个体对象都健康且同性别,个体间彼此无近亲关系.其可追溯的祖辈直系亲属都属于草苗支系.全部个体都操侗语北部方言的草苗话^[5].对其中 33 个男性分析 Y 染色体 SNP 单倍型.各人都按法规签署知情同意书.由于 Y 染色体是具有家族代表性的遗传标记,对于八万左右的草苗总人口来说已具充分的代表性.对男女各 50 人进行体质观测.由于手部伤残等原因而肤纹捺印不全,实际肤纹分析数据完全的有男 32 人,女 23 人,出于提供数据目的,还是把这项数据在文中列出.

1.2 Y 染色体 SNP 单倍型分析

用 FLINDERS/ FITZCO 公司生产的 FTA 血样 DNA 采集纸卡,野外收集每一被研究个体的血样 0.5 mL.常温干燥保存.实验室中,在 FTA 卡上切下带血样的 3 mm²小纸片,用 FLINDERS/ FITZCO 公司提供的 DNA 纯化试剂 200 mL 常温轻摇浸泡小纸片 5 min,吸去反应过的试剂,再更换新试剂 2 次,同样反应.吸去纯化试剂,加入 pH 8.0 的 Tris-EDTA 试剂,同样反应 2 次.60 °C 烘干待用.将烘干的小纸片直接放入反应体系中进行多聚酶链式扩增(PCR)反应.采取二步扩增的策略.先设计特异性强的引物进行增长 PCR,在基因组中找到研究位点所在区段,扩增出 800 bp 左右片段较少份数.再设计巢式扩增引物,引入因 SNP 位点突变与否而异的酶切位点,用增长扩增的产物作模板进行反应,大量扩增 100 ~ 250 bp 的片段.PCR 产物分别用相应的核酸限制性内切酶及其体系进行酶切反应.各位点所用酶和突变、酶切状况见表 1.

表 1 10 个位点的酶切反应信息

Tab. 1 Enzyme reaction information of 10 SNP loci

酶及其反 应状况	位点									
	M119	M110	M95	M88	M122	M134	M7	M9	M89	M45
酶	<i>Bst</i>	<i>Nla</i>	<i>Hha</i>	<i>Hha</i>	<i>Nla</i>	<i>Nla</i>	<i>Bfa</i>	<i>Bam</i> H	<i>Nla</i>	<i>Bfa</i>
原始型	A	T	C	A	C	G	C	C	C	G
突变型	C	C	T	G	T	G 缺失	G	G	T	A
切动型	C	C	C	G	C	G	C	C	C	G

产物用琼脂糖凝胶电泳检测.

根据实验结果确定各个体的 Y 染色体单倍型,单倍型的编号依据宿兵的系统^[6].草苗的单倍型频率与文献报道的其他有关民族群体的数据进行比较.用美国芝加哥 SPSS 公司的 SPSS10.0 统计分析通用软件对各群体数据进行主成分分析(Principle Component Analysis)^[7],观察各群体的亲缘关系远近.由于草苗语言上属于侗台语族侗水语支^[5],所以我们选取了同属侗水语支的侗族^[8]、水族^[9]及同语族壮傣语支的壮族^[8]的数据做比较.又由于其与汉族和苗瑶语族也关系密切,所以有选取最近的汉族支系六甲人^[10]和南北方的汉族总群体^[11]及苗族花苗支系、瑶族(盘瑶和布努支系,盘瑶和广西隆林花苗为实验室未发表数据)、畲族^[8]的数据也参加分析做对照.

1.3 体质观测方法

按《人体测量方法》^[12]的规范,用眼色表、弯脚规、直角卡尺和马丁测高仪观察和测量 13 个群体的头部各项指标以及体部测量.对原始数据按规范转化为指数.

侗傣族群中较近的毛南族、仫佬族、水族^[9]、壮族五色支系,汉藏族群中上海松江的汉族^[13]、云南的白族和彝族^[14]、上海崇明的汉族吴人支系^[13](可能来源于苗瑶族群^[15]),苗瑶族群中的广西大瑶山盘瑶^[14]和云南禄劝苗族^[16]的体质特征已见诸文献,与我们采集的同族群体直接比较,观察其差异.头部观察性特征作为定性的指标有较强的群体特异性,由于文献中报道的数据不够全面,我们选取了蒙古褶、眼裂开

度和倾斜度、中部水平断面、鼻根高度、鼻背硬骨部和软骨部形状、鼻尖方向、鼻基底形状、鼻孔形状、鼻翼突度、上唇皮肤部突度等 12 项指标^[17]与我们采集的群体进行比较分析。同样用 SPSS10.0 作主成分分析。

1.4 肤纹的捺印和分析

用油墨法捺印受测者的指掌肤纹于普通白纸上,用肤纹放大镜观察。计数指纹类型(分为筒弓 A^s, 帐弓 A^t, 桡箕 L^r, 尺箕 L^u, 斗纹 W, 双箕斗 D), 指纹嵴数 FRC(十指总嵴数 TFRC), 掌轴三叉参数(轴三叉掌面纵向百分距 tPD, 轴三角 atd), 指基三叉 a, b 间嵴数(αbRC), 指间各区(, , ,)和大鱼际 T, 小鱼际 H 的真实花纹出现频率, 掌褶线的类型(正常, 桥贯, 通贯和悉尼线)。与文献报道的有关群体比较^[18], 用 SPSS10.0 直接作聚类分析, 以肯定无关的中国俄罗斯族数据作根位置画出聚类系统树。对各项数据都以公式 $(X - Min) / (Max - Min)$ 转化成 0~1 之间的数值, 其中 X 为特定未转化数值, Min 为该项数据各比较民族中最小值, Max 为最大值。以转化后的数据作主成分分析, 画出三维散点图。对肤纹数据的这种处理是本课题组首次应用的^[19]。

2 结果

2.1 草苗的 Y-SNP 单倍型

对 33 个草苗男子的遗传分析检出 5 种 Y-SNP 单倍型, 其中绝大多数是 H6 (M122), 其次是 H8 (M134), 这与汉族比较一致, 但 H6 的比例却是现已发现的群体中最高的。没有发现苗瑶群体特征型 H7, 而发现 H9 (M119), 这种类型可能源于百越^[20], 各地汉族中吸纳百越成分时混进了一定比例^[21]。草苗和参照群体的各单倍型的具体频率见表 2。

表 2 草苗和其他参考群体的 Y 染色体 SNP 单倍型频率

Tab. 2 Percentages of Y-SNP haplotypes of Grass Hmong and other populations compared

群体		样本量/人	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12	H13
草苗		33					8.2	66.7		12.1	10.0		3.0		
布努		10	50.0				20.0		30.0						
苗瑶	花苗	17	7.7		7.7	7.7	38.5	15.4	7.7				15.4		
	盘瑶	14	7.1			7.1	21.4	14.3		35.7			14.3		
	畲	11	18.2				9.1	18.2	27.3	18.2			9.1		
侗		20	15.0					15.0		15.0	25.0	10.0	20.0		
侗傣	壮	29	3.6		3.6	7.1	3.6	3.6		25.0	17.9		25.0	10.7	
	水	50				8.0	10.0			20.0	18.0		44.0		
汉北		82	8.5			2.4	22.0	29.3		23.2	9.8				4.9
汉族	汉南	283	7.8	0.4		1.4	12.9	25.4	1.8	27.9	16.8		3.5	0.7	1.4
	六甲	32	11.1				7.4	40.7		37.0			3.7		

根据表 2 中的数据, 进行主成分分析, 利用前三主成分构建立体直角坐标系, 各群体在坐标系中的分布见图 1 (第 625 页)。

从图 1 中可见, 汉族群体、侗傣群体和苗瑶群体是不相混的, 前二者内部各群体比较集中, 苗瑶内部的群体比较分散。草苗明显处于汉族群体中。

2.2 草苗的体质特征

草苗的头面部观察性特征的各类型频率见表 3 和表 4。

表 3 中可见草苗的眼睛虹膜颜色为浅褐色, 双眼皮较多, 蒙古褶不发达, 面型多上宽下窄, 其他特征不突出。

表 3 草苗的眼部、面形和口部观察特征频率

Tab. 3 Ocular facial and buccal observations of Grass Hmong

		Martin &Schultz 眼色号						上眼睑褶			蒙古褶				眼裂开度			眼裂倾斜			
性别		9	10	11	12	13	14	无	>2 mm	1~2 mm	达睫毛	无	微显	中等	发达	狭窄	中等	较阔	外倾	水平	内倾
		男		2	32	18	12	28	8	4	42	24	30	18	52	24	6	24	64	12	32
女		4	74	4	10	4	4	4	48	12	36	22	52	18	8	22	56	22	56	44	

		正面形					颧弓			颈部侧面观			人中突度		人中高度		上红唇厚度							
性别		椭	卵	圆	长	方	菱	倒	梯	突出	中等	紧收	后斜	直	前凸	凸唇	正唇	低	中等	高	薄唇	中唇	厚唇	肿胀
		男		12	34		14	6	22	12	38	58	4	20	58	22	32	68	12	74	14	28	42	18
女		12	44	14		4	22	4	52	48		22	56	22	48	52	18	82		16	66	18		

注:Martin & Schultz 眼色表各号代表虹膜色泽(9 浅黄色,10 黄色加灰紫色边,11 深黄色,12 黄褐色,13 浅褐色,14 深褐色,15 深褐色加黑色边,16 黑褐色)。

表 4 草苗的耳部和鼻部观察特征频率

Tab. 4 Aural and nasal observations of Grass Hmong

		达尔文结节					耳壳外形				耳垂		耳壳外展度			鼻翼高			鼻翼宽			盯眈					
性别		长耳	锐	带	痕	无	方	倒	椭	倒	梨	梨	菱	圆	方	三角	紧	中	外	高	中	低	宽	中	窄	干	湿
		男		4	12	14	54	16	10	26	46	12	4	4	32	34	22	58	20	62	34	6	6	28	66	22	78
女		4	14	12	48	22	30	52	4		14		34	48	18	38	58	4	70	22	8	4	38	58	16	84	

性别		鼻根		鼻背硬骨部			鼻背软骨部			鼻背整体形			鼻尖方向		鼻基底线		鼻孔形状			鼻翼突						
		微	中	高	凹	直	凸	凹	直	凸	凹	直	凸	柔顺	波折	翘	前	垂	翘	平	垂	圆	圆	圆	突	突
男		8	76	16	10	42	48	36	48	16	12	38	22	28	6	76	18	18	76	6	4	36	60	44	56	
女		26	74		16	44	40	38	58	4	18	48	12	22	8	84	8	26	66	8		4	96	8	40	52

草苗的鼻翼较高突,其他特征不突出。

草苗的体质测量数据和相应指数见表 5。

表 5 草苗的体质测量数据(mm)和指数

Tab. 5 Physical metrage and indexes of Grass Hmong

性别	头长	头宽	头高	额宽	面宽	下颌角宽	耳屏间宽	容貌面高	形态面高	口裂宽	上唇厚	两唇厚	人中高	鼻宽	鼻高
男	184.2	148.7	219.4	116.9	134.5	108.6	138.5	182.4	109.1	49.8	6.9	16.2	14.6	41.8	44.7
女	175.8	145.0	215.2	112.4	127.1	102.2	131.1	172.3	104.6	48.9	6.1	14.1	13.1	36.9	44.6

性别	身高	指距	上臂长	前臂长	手长	上肢长	大转子点	胫骨上点	内踝下点	坐高	躯干后高
男	1 570.2	1 565.6	305.1	246.0	167.9	719.0	679.5	457.2	59.3	858.6	649.7
女	1 452.7	1 478.3	289.8	217.4	156.3	663.6	653.8	431.1	57.5	816.7	618.4

性别	头指数			颌宽指数			面宽指数		鼻指数	身高指数			
	宽高	长宽	长高	头宽	额宽	颧宽	口宽	容貌面		形态面	指距	上肢长	坐高
男	83.9	80.7	67.8	136.9	107.6	123.9	45.9	135.6	81.1	93.6	99.7	45.8	54.7
女	81.7	82.4	67.4	141.9	110.0	124.4	47.8	135.6	82.3	82.7	101.8	45.7	56.2

从表 5 可见,草苗的各项指标处于南方民族的中等水平。

利用草苗和相关参照群体的面貌特征数据作主成分分析,根据前三主成分构建立体直角坐标系,各群体在坐标系中的分布见图 2。

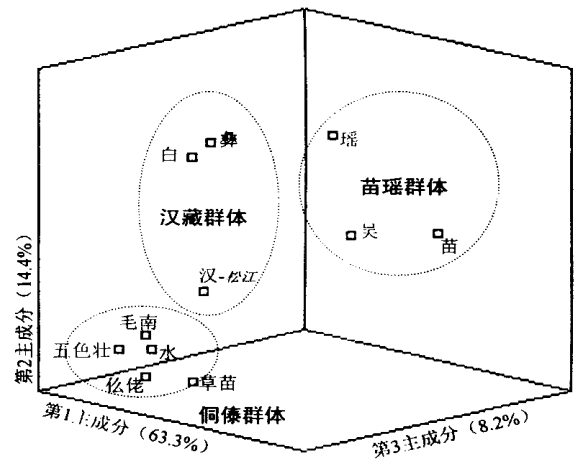
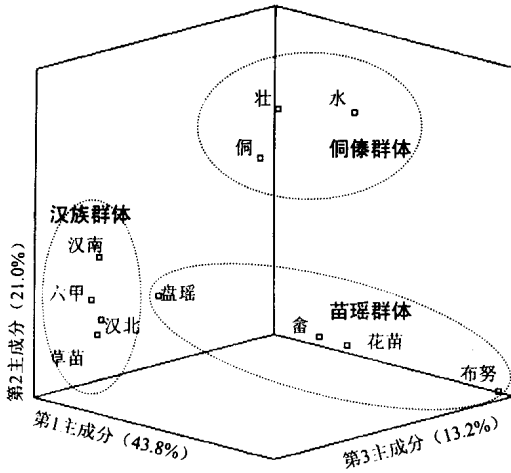


图 1 草苗和其他群体的 Y 染色体单倍型主成分分析散点图

图 2 草苗和其他群体的面貌特征主成分分析散点图

Fig. 1 Spot plot of principle component analysis with Y-SNP

Fig. 2 Spot plot of principle component analysis with facial features

图 2 中三类群体也区分得较清晰。吴人落入了苗瑶群体中。而草苗虽然第 1 主成分与苗瑶群体较一致,但第 2 和第 3 主成分都把草苗拉向了侗傣群体。

2.3 草苗的指纹和掌纹

草苗的各指端指纹类型频率见表 6。

表 6 草苗指纹各纹型出现频率

Tab. 6 Percentages of finger print types of Grass Hmong

											%	
性别	纹型	5L	4L	3L	2L	1L	5R	4R	3R	2R	1R	全手
男	A ^s				6.25					6.25		1.25
	A ^r				3.13		3.13			6.25		1.25
	L ^r	6.25			18.75		6.24	3.13	3.13	18.75		5.63
	L ^u	50.00	40.62	53.12	25.00	21.88	50.00	34.38	50.00	34.37	15.63	37.50
	W	40.62	59.38	37.50	31.25	40.62	37.50	62.50	40.62	25.00	43.75	41.88
女	D	3.13		9.38	15.62	37.50	3.13		6.25	9.38	40.63	12.50
	A ^s					4.35			4.35	4.35		1.30
	A ^r				4.35	8.70						1.30
	L ^r					4.35				8.70		1.30
	L ^u	65.22	30.43	47.83	43.48	17.39	60.87	30.43	52.17	34.78	34.78	41.74
平均	W	30.43	56.52	39.13	43.48	43.48	34.78	65.22	39.13	43.48	34.78	43.04
	D	4.35	13.05	13.04	8.70	21.73	4.35	4.35	4.35	8.70	30.43	11.30
	A ^s				3.64	1.82			1.82	5.45		1.27
	A ^r				3.64	3.64	1.82			3.64		1.27
	L ^r	3.64			10.91	1.82	3.64	1.82	1.82	14.55		3.82
平均	L ^u	56.36	36.36	50.91	32.73	20.00	54.55	32.73	50.91	34.55	23.64	39.27
	W	36.36	58.18	38.18	36.36	41.82	36.36	63.64	40.00	32.73	40.00	42.36
	D	3.64	5.45	10.91	12.73	30.91	3.64	1.82	5.45	9.09	36.36	12.00

注:栏目中数字 1~5 表示拇指到小指,L 和 R 是左右手,下同。

男性的食指指纹类型最丰富,女性的左手拇指比食指类型更丰富。男女间的全手纹型频率较一致。草苗的各指指纹峰数见表 7。

表 7 草苗的指纹嵴数

Tab. 7 Finger ridge counts of Grass Hmong

性别	5L	4L	3L	2L	1L	5R	4R	3R	2R	1R	TFRC
男	11.72	14.34	14.44	11.63	16.13	12.78	14.13	12.25	12.06	16.41	135.88
女	11.74	14.13	11.43	10.78	13.30	11.35	13.57	11.04	10.61	14.39	122.35
平均	11.73	14.24	12.94	11.20	14.71	12.06	13.85	11.65	11.34	15.40	129.11

女性的拇指嵴纹数显著少于男性,其他手指也略少,总指纹嵴数少于男性。
草苗的掌面各三叉点的有关参数见表 8。

表 8 草苗的掌面三叉点各参数

Tab. 8 Palm triradius parameters of Grass Hmong

性别	手	$\alpha b RC$	$atd/(\circ)$	tP/cm	DP/cm	tPD
男	左	37.34	38.39	1.85	10.26	18.04
	右	36.81	38.63	1.81	10.18	17.78
	双手	37.08	38.51	1.83	10.22	17.91
女	左	35.43	40.26	1.98	9.56	20.68
	右	35.83	41.10	2.00	9.36	21.61
	双手	35.63	40.68	1.99	9.46	21.15
平均		36.35	39.59	1.91	9.84	19.53

女性的 DP 值偏小,表明其掌面偏短,所以 atd 和 tPD 偏大,但没有显著差异。
草苗掌面各区域的真实花纹出现频率见表 9。

表 9 草苗的掌面各区真实花纹出现频率

Tab. 9 Percentages of true patterns in palmer areas of Grass Hmong

性别	手	%					
		T					H
男	左	0.00	3.23	0.00	37.50	87.50	3.13
	右	3.23	0.00	3.33	40.63	75.00	6.45
	双手	1.61	1.61	1.67	39.06	81.25	4.79
女	左	9.09	4.55	4.76	13.04	82.61	22.73
	右	0.00	0.00	0.00	34.78	78.26	22.73
	双手	4.55	2.27	2.38	23.91	80.43	22.73
平均		3.08	1.94	2.02	31.49	80.84	13.76

女性大鱼际和小鱼际真实花纹出现频率极显著多于男性,而指间三区花纹显著少于男性。
草苗各种掌褶线类型的频率见表 10。

表 10 草苗掌褶线各类型出现频率

Tab. 10 Percentages of palm pleat line types of Grass Hmong

性别	手别	掌褶线型			性别	手别	掌褶线型			性别	手别	掌褶线型		
		正常	桥贯	通贯			正常	桥贯	通贯			正常	桥贯	通贯
男	左	70.00	26.67	3.33	女	左	80.95	14.29	4.76	两性	左	74.51	21.57	3.92
	右	75.86	10.34	13.79		右	85.00	15.00	0.00		右	79.59	12.24	8.16
	平均	72.93	18.51	8.56		平均	82.98	14.64	2.38		平均	77.05	16.91	6.04

没有发现悉尼线 (Sydney line). 男性右手的通贯型 (Simian line) 极显著多于女性, 桥贯型在男女间分布的手别也不同. 根据指纹类型总频率, $TFRC$, αbRC , 掌面真实花纹频率等 11 项数据与文献比较^[19], 直接画出图 3 聚类系统树.

图 3 中不相关的俄罗斯族处于根位置上, 广西贵州等的各南方民族都较接近, 而草苗和四川苗族与其他南方民族较远. 从树中看不出草苗与其他民族的关系. 图 4 是根据标准化的数据画出的主成分分布图.

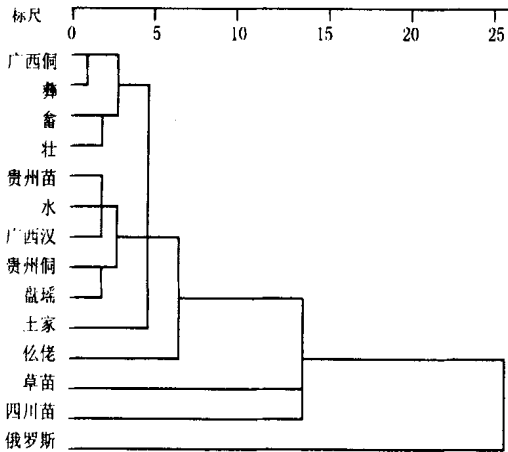


图 3 草苗和相关对照群体的肤纹聚类系统树
Fig. 3 Dendrogram using single linkage hierarchical cluster analysis of dermatoglyphics

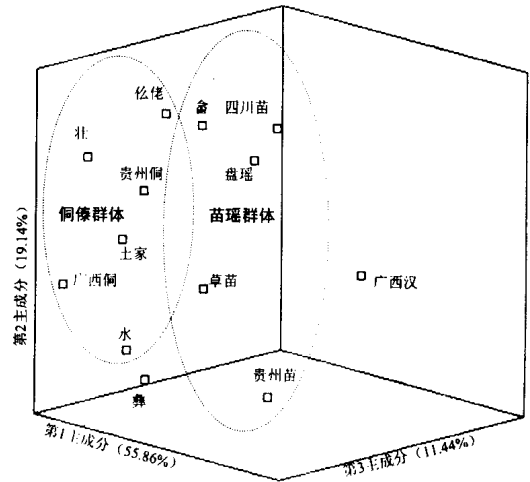


图 4 草苗和相关对照群体的肤纹主成分分析散点图
Fig. 4 Spot plot of principle component analysis with dermatoglyphics

从图 4 可以明显地看出, 第 1 主成分区分了苗瑶群体和侗僚群体, 草苗处于苗瑶群体中, 但也略偏侗僚群体. 广西汉族、土家族和彝族这 3 个汉藏群体没有聚成一类. 第 2 主成分上草苗与广西侗族和广西汉族基本一致. 整体效应上, 草苗与广西侗族、广西汉族、贵州苗族和贵州侗族的距离几乎相等.

3 讨 论

3.1 群体系统分类的材料与方法的取舍

对群体的系统分类分析首先要求选用的项目尽可能遗传率高, 因为系统发生完全是遗传体系. DNA 分子标记是完全遗传性的标记, 肤纹的遗传率也较高, 但是体质特征的遗传率就很不一致. 体质特征中头部定性特征平均遗传率较高^[13], 本文就是用了这些项目分析.

本文中肤纹的分析同时用到了聚类分析系统树和主成分分析散点图. 聚类分析对数据的选用是平均对待的, 得到的分类结果是总体的混合效应. 而主成分分析则可以把数据结构进行分类筛选, 把主要的数据结构提取出来, 得到的分类结果区分为各种不同的效应. 本项研究中的肤纹分析明显可见, 聚类分析的分类结果是不具系统性的, 而主成分分析的结果却明显把苗瑶系统和侗僚系统分了开来, 系统性较强. 另外两类材料的主成分分析对汉藏、苗瑶和侗僚的系统分类也相当清晰. 这说明在研究族源等系统分类型的分析时, 主成分分析是种更好的方法, 而研究民族的总体成分时聚类分析仍不失为一种好方法.

3.2 草苗的早期父系来源可能是汉族

Y 染色体遗传标记与父系遗传是完全对应的, 所以 Y 染色体的研究结果就是父系遗传结构. 草苗的 Y 染色体类型基本符合汉族的普遍类型, 而与侗僚、苗瑶相差较远, 没有侗僚的特征型 H10 和 H12 及苗瑶的特征型 H7. 可以认为其早期的父系来源应该是汉族, 此后这一父系结构一直保持了下来, 而女性则一直与其他民族交流.

草苗还有部分 H9, 而南方汉族的 H9 最有可能是来源于侗僚系统^[21]. 与草苗进入贵州大致同期进入

广西的汉族六甲人中原来没有 H9,近年才由侗族的融入带进了一定比例^[10]。草苗的 H9 可能是汉族进来前就已从其他百越支系中吸收了,也可能是后来来自侗族的成分。虽然草苗称不与侗族通婚,但因为他们使用侗语,所以与侗族有过交流是很可能的。

草苗的 H6 比例极高,比汉族平均水平高得多。南方其他一些汉族支系的此比例也偏高,包括六甲人和客家人。这可能是迁徙过程中的遗传漂变。

3.3 苗族和侗族应是不同过程融入草苗的

草苗的面貌接近侗傣系统,而肤纹主要接近苗瑶系统,这体现出草苗的遗传结构的复杂性,苗族和侗族都应该是其来源成分。而不同的材料体现出了其结构的不同方面。但是苗族和侗族的成分可能是不同的过程中进入草苗的。从草苗的民族心理看来,苗族可能更早融入这个群体。或许正如文化人类学的研究结果,草苗产生于元代到清代间汉父苗母的结合^[11]。之后在民族迁徙过程中,草苗的村落散布于侗族主要分布区内,与侗族不断地接触使他们的语言渐渐变成了侗语,遗传结构上也接纳了部分侗族的成分。

草苗复杂的民族融合结构说明中国的民族群体交流是很密切的,不同的民族系统的成分也可以融合在同一个人群中。同一系统中的民族交流可能更频繁。所以在调查民族渊源时千万不可片面武断,要充分认识到这种复杂性。

广西壮族自治区民族事务委员会和融水苗族自治县民族事务局安排和协助了调查工作,在此表示感谢。

参考文献:

- [1] 朱慧珍. 草苗的历史与风俗考析[J]. 广西民族学院学报(哲学社会科学版), 1998, 20(1): 48-53.
- [2] Jin L, Su B. Natives or immigrants: Modern human origin in East Asia[J]. *Nature Reviews Genetics*, 2000, 1(2): 126-133.
- [3] 李辉, 宋秀峰, 金力. 人类谱系的基因解读[J]. 二十一世纪(香港), 2002, 71(6): 98-108.
- [4] 李辉. 东亚人的遗传系统初识[J]. 国立国父纪念馆馆刊(台北), 2002, 10: 123-136.
- [5] 梁敏, 张均如. 侗台语族概论[M]. 北京: 中国社会科学出版社, 1996.
- [6] Su B, Xiao J H, Underhill P, et al. Y-Chromosome evidence for a northward migration of modern humans into Eastern Asia during the last Ice Age[J]. *Am J Hum Genet*, 1999, 65(6): 1718-1724.
- [7] Cavalli-Sforza L L, Menozzi P, Piazza A. The history and geography of human genes[M]. Princeton: Princeton Univ Press, 1994.
- [8] Su B, Jin L, Underhill P, et al. Polynesian origins: Insights from the Y chromosome[J]. *Proc Natl Acad Sci USA*, 2000, 97: 8225-8228.
- [9] 李辉, 杨宁宁, 黄兴球, 等. 广西融水北高村人的遗传学和体质人类学特征[J]. 中央民族大学学报(自然科学版), 2002, 11(2): 108-116.
- [10] 李辉, 侯井榕, 杨宁宁, 等. 广西六甲人来源的分子人类学分析[J]. 广西民族学院学报(哲学社会科学版), 2002, 24(5): 38-43.
- [11] Su B, Xiao C J, Deka R, et al. Y chromosome haplotypes reveal prehistorical migrations to the Himalayas[J]. *Hum Genet*, 2000, 107: 582-590.
- [12] 吴汝康, 吴新智, 张振标. 人体测量方法[M]. 北京: 科学出版社, 1984.
- [13] 林凌, 李辉, 张海国, 等. 上海郊区人群的体质特征和遗传关系[J]. 人类学学报, 2002, 21(4): 293-306.
- [14] 中国人类学会. 中国八个民族体质调查报告[M]. 昆明: 云南人民出版社, 1982.
- [15] 李辉, 周振鹤, 徐立群, 等. 上海本地人源流的主成分分析[J]. 中央民族大学学报(自然科学版), 2003, 12(2): 167-173.
- [16] 余发昌, 李明, 刘冠豪. 云南苗族的体质特征研究[J]. 人类学学报, 1994, 13(4): 321-326.

- [17] 李 辉,潘方芳,张敏华,等.滇西北8个民族群体面貌特征观察的聚类分析[J].复旦学报(自然科学版),2001,40(5):568-576.
- [18] 张海国,丁明,焦云萍,等.中国人肤纹研究.中国52个民族的肤纹聚类[J].遗传学报,1998,25(5):381-391.
- [19] 柳扬帆,李 辉,金建中,等.白族勒墨人的肤纹分析及其统计方法探索[J].中央民族大学学报(自然科学版).2003,12(1):45-52.
- [20] 李 辉.百越族群遗传结构分析的初步思路[A].现代人类学国际研讨会论文集[C].上海:复旦大学现代人类学研究中心,2002,89-94.
- [21] 李 辉.百越遗传结构的一元二分迹象[J].广西民族研究,2002,70(4):26-31.

Origin of Grass Hmong as Revealed by Genetics and Physical Anthropology

LI Hui¹, LI Xin¹, YANG Ning-ning¹, WEN Bo¹, HOU Jing-rong²,
HUANG Xing-qiu², JIN Jian-zhong¹, JIN Li¹

(1. Center for Anthropological Studies, School of Life Sciences, Fudan University, Shanghai 200433, China;
2. The Ethnology and Anthropology Institute, Guangxi University for Nationalities, Nanning 530006, China)

Abstract: Grass Hmong (Cao Miao) on the boundary of three provinces, Hunan, Guangxi and Guizhou, is a special branch of official Hmong Nationality. They have the customs of Han, Hmong and Kam, and speak Kam, so this population shall be one of the most mixed populations in China, and will be quite valuable to academic studies of ethnology and population genetics. Three kinds of characters of Grass Hmong in Rongshui Guangxi were studied with principle component analysis. Y-chromosome single nucleotide acid polymorphism haplotypes show Grass Hmong is consistent to Han Chinese, and different from Hmong-Mien and Kam Tai. Physical characters of Grass Hmong are also listed in this paper. The facial features of Grass Hmong is found fairly close to Kam Tai populations. Grass Hmong's dermatoglyphy is close to Hmong-Mien populations. This reveals that the paternal ancestors of Grass Hmong shall come from Han, and maternal members may be from Hmong at the beginning which made them identify with Hmong, when they lived in the surrounding of Kam later they accepted some Kam members and Kam language. The complex genetic structure of Grass Hmong clues on the continual intercourse between ethnic groups in China, and this complexity must be noticed during ethnic headstream studies.

Key words: Grass Hmong (Cao Miao); Y-chromosome single nucleotide acid polymorphism; physical; dermatoglyphy; ethnic origin; principle component analysis