

线粒体 DNA 多态性揭示丝绸之路上游地区少数民族的基因融合

杨留启, 谭思杰, 俞海菁, 郑冰蓉, 乔恩发, 董永利, 咎瑞光, 肖春杰*

云南省生物资源保护和利用重点实验室, 云南大学人类遗传学研究中心, 昆明 650091

* 联系人, E-mail: cjxiao@public.km.yn.cn

收稿日期: 2007-10-12; 接受日期: 2008-01-07

国家重点基础研究发展计划(批准号: 2006CB708502)和国家自然科学基金(批准号: 30660076)资助项目

摘要 河西走廊位于中国古丝绸之路的上游地区, 是连接东亚和中亚地区的主要通道. 为了在遗传背景上评价该地区的基因融合, 研究了该地区 5 个少数民族人群的线粒体 DNA 多态性. 采用测序和 PCR-限制性内切酶片段长度多态分析方法对采自河西走廊地区的三个人群和邻近的新疆两个人群总共 242 个无关个体进行单倍型分型, 并且把得到的数据同以前报道的东亚、中亚、西亚以及欧洲人群数据进行比较. 结果发现欧洲特异单倍型和东亚特异单倍型同时存在于河西走廊地区人群中, 从而揭示了从欧洲经中亚向河西走廊的基因流动. 河西走廊地区人群在遗传上分布于中亚和东亚人群之间, 这意味着中亚和东亚人群对该地区的母系遗传结构做出了重要贡献. 本研究更加完整地展示了沿着丝绸之路地区从欧洲经中亚到丝绸之路上游的河西走廊的基因流动和基因融合事件.

关键词

多态性
线粒体 DNA
基因融合
河西走廊

公元前 2 世纪, 我们的先人建立了一条东起中国长安(今陕西西安)经由河西走廊、新疆、中亚、西亚, 西至地中海沿岸的古老的贸易路线, 后人称之为“丝绸之路”. 丝绸之路的出现不仅对东西方的文化发展产生了重要影响, 而且对东亚人和欧洲人之间的基因交流起到了重要的作用.

对丝绸之路沿途人群的遗传结构研究为东方和西方人群在这一地区的基因融合研究提供了许多资料. 其中 Quintana-Murci 等人^[1]、Comas 等人^[2,3]和姚永刚等人^[4,5]对生活在西亚、中亚和我国新疆维吾尔自治区的民族人群的母系遗传结构进行了研究, 认为东亚人和欧洲人的基因流到了中亚并且在这一地区发生了东西方人群间广泛的基因融合.

河西走廊是丝绸之路的上游地区, 地域范围包括甘肃省和青海省部分地区, 是连接东亚和中亚的一条狭长通路(图 1). 历史上曾经发生诸如回鹘(生活在东北亚的一个古老民族)和蒙古民族沿着丝绸之路从东向西的迁徙事件^[6], 因此, 我国东部地区典型的东亚人群和中亚人群之间的基因交流事件就很可能发生. 尽管河西走廊地处连接东亚和中亚的重要的地理位置, 这一地区人群的基因融合程度仍没有被评估. 我们猜测在这一地区存在着东亚人群和中亚人群之间的基因交流, 这一想法推动我们对这一地区人群的遗传结构展开研究.

我们在河西走廊地区采集了三个少数民族人群样本, 分析了他们的线粒体 DNA 序列, 并把他们同我

们采集的两个新疆维吾尔自治区的少数民族人群样本作比较(图 1)。除此之外,姚永刚等人^[7]、Quintana-Murci等人^[11]、Torrioni等人^[8]、Francalacci等人^[9]、Malyarchuk等人^[10]和Comas等人^[2]的数据也被引用作为参照人群。我们的研究旨在阐明居住在河西走廊地区的少数民族的母系遗传结构,同时结合前人对邻近地区人群的研究,使整个丝绸之路沿途地区人群的母系遗传结构框架更加完整。

1 材料与方法

1.1 样本的采集

在知情同意原则下,采集了甘肃、青海和新疆维吾尔自治区的 5 个少数民族人群无关个体血样 242 份,他们是:甘肃省积石山县的 50 名保安人(Baoan)和东乡县的 64 名东乡人(Dongxiang)、青海省土族自治县的 48 名土族人(Tu)以及新疆维吾尔自治区疏附县的 50 名维吾尔人(Uygur)和伊宁市的 30 名乌兹别克人(Uzbek)(图 1)。样本处理经过云南大学道德与伦理委员会核准。

为使我们的数据具有可比性,4 个汉族人群(广东汉族人guangdonghan、云南汉族人yunnanhan、山东汉族人shandonghan和辽宁汉族人liaoninghan)^[2]、一个维吾尔族人群(Uygur)、一个哈萨克族人群(Kazak)和两个吉尔吉斯人群(Talas Kirghiz和Sary-Tash Kir-

ghiz)^[2]、一个托斯卡纳人群(Tuscan)^[8]、一个西西里人群(Sicilian)^[9]和一个斯洛文尼亚人群(Slovnican)^[10]也被纳入我们的分析中作为比较人群。

1.2 DNA 提取及线粒体 DNA 单倍型的划分

常规酚-氯仿方法提取基因组DNA。选取线粒体DNA高变区 1(HVS-1)进行PCR-测序分析,编码区 16 个多态位点进行PCR-RFLP分析。引物L15996 和 H16498^[2]用于高变区 1(HVS-1)PCR扩增,纯化的PCR扩增产物用ABI PRISM BigDye Terminator V3.1 试剂盒进行测序反应,ABI 377 测序仪(Applied Biosystems)完成该区的序列测定。用Kivisild等人^[11]、Macaulay等人^[12]和姚永刚等人^[7]使用的引物,对线粒体编码区的 16 个多态位点进行PCR-RFLP分析,使用的引物和内切酶见表 1。用于欧、亚特异单倍型识别的内切酶及多态位点见表 2。为了更准确地识别欧洲人群特异的单倍型U在我们所研究的样本中的有无,我们还用引物L12188(5'-CTTACGACCCCTT-ATTTACCG-3')和H12472(5'-ATAAAGGTGGATGC-GACAATG-3')针对性地扩增 12308 多态位点。对于存在于高变区 1(HVS-1)和编码区的序列变异但尚不能确定其单倍型的样本,我们进一步用引物L29 和 H408^[2]对其进行了高变区 2 (HVS-2)的PCR-序列测定,以确定其单倍型。



图 1 丝绸之路上游路线及样本来源地示意图

表 1 线粒体多态位点引物序列、PCR 退火温度、限制性内切酶及多态性判别^{a)}

位点	引物序列	退火温度/°C	限制性内切酶及多态性判别
10394/10397	AAACAACCTAACCTGCCACTA GAAGTGAGATGGTAAATGCT	60	10394Dde I ⁺ 10397Alu I ⁺
5176	AACTTAAACTCCAGCACAC TAGGTAGGAGTAGCGTGGTA	55	Alu I ⁻
4831	AACATGCTAGCTTTTATTCCAG AGAAGAAGCAGTCCGGATGT	55	Hha I ⁺
3391	CCAACCTCCTACTCCTCATTG TGGGACCTTTGCGTAGTTGTA	53	HaeIII ⁺
4715	CAAGTATTTCCCTCACGAAGCA GGTAGTATTGGTTATGGTTCATGG	57	HaeIII ⁺
9824	CATTTCCGACGGCATCTAC GGTGGATTTTTCTATGTAGCC	56	HinfI ⁺
7598	AAAACCATTTTCATAACTTTGTC GTGTTAGGAAAAGGGCATAAC	56	Hha I ⁻
663	AATTTTATCTTTTGGCGGTATG TTGATGCTTGTCCCTTTTGA	53	HaeIII ⁺
5417	ATCTCGCACCTGAAACAAGCTAAC TGGGGTGGGTTTTGTATGTTCCGTA	56	Rsa I ⁻
12406	AACCACCTAACCCTGACTT AGTATGGTAATTAGGAAGATGAG	55	Hpa I ⁻
10646	AGGAATAATACTATCGCTGTT TAGTCTAGGCCATATGTGTTG	55	Rsa I ⁺
10310	TGAGCCCTACAAACAACCTAACAT ATACCAATTCGGTTCAGTCTAATC	54	NlaIII ⁻
9bp	ATGCTAAGTTAGCTTTACAG ACAGTTTCATGCCATCGTC	57	
14766	CATTATTCTCGCACGGACTAC TGGTAACTGATTTTATTAGGG	55	Mse I ⁻
7025	ACTCCACGGAAGCAATATGA CTATGATGGCAAATACTGCTC	55	Alu I ⁻
4216	CTACTACAACCCTTCGCTGAC GATAGGTGGCACGGAGAATT	55	NlaIII ⁺

a) 限制性内切酶及多态位点基于Macauley等人^[12], Kivisild等人^[11]和姚永刚等人^[2]的研究

表 2 欧亚特异单倍型识别的内切酶及多态位点^{a)}

单倍型	限制性酶切位点
东亚类型	
M	10394Dde I ⁺ , 10397Alu I ⁺
A	663HaeIII ⁺
B	9bp deletion
D	10394Dde I ⁺ , 10397Alu I ⁺ , 5176Alu I ⁻
G2a	10394Dde I ⁺ , 10397Alu I ⁺ , 7598Hha I ⁻
F	10310NlaIII ⁻ , 12406Hpa I ⁻
G	10394Dde I ⁺ , 10397Alu I ⁺ , 4831Hha I ⁺
M7	10394Dde I ⁺ , 10397Alu I ⁺ , 9824Hinf I ⁺
M8	10394Dde I ⁺ , 10397Alu I ⁺ , 4715HaeIII ⁺
M9	10394Dde I ⁺ , 10397Alu I ⁺ , 3391HaeIII ⁺
M10	10394Dde I ⁺ , 10397Alu I ⁺ , 10646Rsa I ⁺
R9	10310NlaIII ⁻
N9	5417Rsa I ⁻
欧洲类型	
HV	14766Mse I ⁻
H	14766Mse I ⁻ , 7025Alu I ⁻
JT	4216NlaIII ⁺

a) 多态识别位点基于姚永刚等人^[2]和Malyarchuk等人^[10]的研究

1.3 数据分析

用DNASTAR软件(DNASTAR, Inc.)对DNA序列进行分析, 结果和剑桥序列(CRS) [13]进行比较. 用SPSS11.5 软件进行主成分分析, 用ARLEQUIN 2.000 软件 [14]按照Tamura和Nei [15]的方法计算 F_{ST} 遗传距离. 基于 F_{ST} 值, 用Phylip 3.62 软件构建Neighbour-Joining 树. 参照Roberts 和 Hiorns [16]的方法, 用Admix 2.0 软件 [17]估算东亚人和中亚人在保安、东乡、土、维吾尔和乌兹别克族人群中的基因混合度, 由于基因混合度估算中亲本人群的选择至关重要 [18,19], 我们选择了从东亚到中亚的多个民族人群作为参照人群以便尽量避免结果的误差, 参照的广东汉、云南汉、山东汉和辽

宁汉共四个汉族人群 [7]作为东亚亲本人群, 参照的维吾尔、哈萨克和吉尔吉斯族人群作为中亚亲本人群 [2,5].

2 结果

2.1 单倍型频率分布

在检验的 5 个少数民族人群的 242 分样本中共发现了 27 种单倍型, 大多数单倍型可以归入欧洲特异性单倍型 [20,21]和东亚特异性单倍型 [7,12,22] (表 3, 图 2), 而且每个民族样本中均发现了欧洲和东亚特异性单倍型. 在河西走廊地区的保安、东乡和土族人群中, 东亚特异单倍型 A, B, D, F 和 M8 等占主导地位, 而在两个新疆人群(维吾尔和乌兹别克)中, 除东亚特异

表 3 研究及参照样本的来源地及单倍型频率分布^{a)}

人群	BA	Tu	DX	Uzb	Uy	GH ^a	YH ^a	SH ^a	LH ^a	Sic ^b	Tus ^c	Slo ^d	Uyg1 ^e	Kaz1 ^e	Kaz2 ^f	Uyg2 ^f	Kit2 ^f	Kir2 ^f
单倍型	(50)	(48)	(64)	(30)	(50)	(30)	(43)	(50)	(51)	(49)	(48)	(104)	(47)	(53)	(55)	(55)	(48)	(47)
A	8.0	43.8	17.2	3.3	6.0	0.0	4.7	4.0	5.9	0.0	0.0	0.0	4.3	3.8	9.1	7.3	6.3	2.1
B	14.0	2.1	6.3	0.0	2.0	20.0	18.6	12.0	17.6	0.0	0.0	0.0	2.1	3.8	5.4	7.3	4.2	10.6
D	20.0	14.6	20.3	13.3	10.0	16.7	14.0	36.0	25.5	0.0	0.0	0.0	10.6	13.2	18.2	16.4	14.6	25.5
F3	4.0	0.0	0.0	3.3	0.0	0.0	2.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
F	4.0	4.2	7.8	0.0	0.0	23.3	23.3	12.0	5.9	0.0	0.0	0.0	6.4	7.6	3.6	7.2	4.2	2.1
G2a	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
G	4.0	10.4	3.1	10.0	0.0	0.0	0.0	6.0	7.8	0.0	0.0	0.0	12.7	5.7	5.5	1.8	14.6	2.1
M7	14.0	0.0	4.7	0.0	2.0	3.3	18.6	4.0	7.8	0.0	0.0	0.0	6.4	1.9	0.0	5.4	0.0	0.0
M8	4.0	8.3	12.5	3.3	8.0	3.3	4.7	8.0	11.8	0.0	0.0	0.0	6.4	26.4	9.1	1.8	12.5	14.9
M9	4.0	6.3	3.1	0.0	2.0	3.3	0.0	4.0	3.9	0.0	0.0	0.0	0.0	1.9	1.8	1.8	4.2	0.0
M10	0.0	0.0	0.0	0.0	2.0	0.0	2.3	2.0	5.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
M11	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
N9a	2.0	0.0	6.3	0.0	2.0	6.7	7.0	6.0	2.0	0.0	0.0	0.0	0.0	1.9	0.0	1.8	6.3	0.0
M*	0.0	4.2	3.1	0.0	0.0	23.3	2.3	2.0	0.0	0.0	0.0	0.0	2.1	3.8	5.5	3.6	6.3	4.3
Y	0.0	0.0	0.0	0.0	2.0	0.0	0.0	2.0	2.0	0.0	0.0	0.0	0.0	0.0	1.8	0.0	0.0	0.0
R9b	0.0	0.0	0.0	6.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
R	0.0	0.0	1.6	3.3	0.0	0.0	0.0	2.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	12.7	4.2	2.1
HV	0.0	0.0	0.0	20.0	12.0	0.0	2.3	0.0	0.0	0.0	0.0	6.7	4.2	5.7	21.8	0.0	12.5	25.5
H	8.0	2.1	6.3	13.3	6.0	0.0	0.0	0.0	0.0	52.0	41.7	47.1	10.6	7.5	0.0	10.9	0.0	0.0
I	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	4.2	1.9	0.0	1.9	0.0	0.0	0.0	0.0
J	0.0	2.1	1.6	6.7	2.0	0.0	0.0	0.0	0.0	14.0	14.6	9.6	4.2	1.9	0.0	0.0	4.2	6.4
T	0.0	0.0	0.0	0.0	24.0	0.0	0.0	0.0	2.0	14.0	10.4	5.8	2.1	7.6	7.3	1.8	2.1	0.0
U	6.0	2.1	1.6	16.7	14.0	0.0	0.0	0.0	0.0	8.0	16.7	23.1	14.9	3.8	7.3	20.1	4.2	4.3
V	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	1.9	0.0	0.0	0.0	0.0
W	0.0	0.0	0.0	0.0	6.0	0.0	0.0	0.0	0.0	2.0	2.1	4.8	6.4	0.0	3.6	0.0	0.0	0.0
X	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.0	8.3	1.0	0.0	0.0	0.0	0.0	0.0	0.0
N*	4.0	0.0	4.7	0.0	0.0	0.0	0.0	0.0	0.0	2.0	2.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0

a) 单倍型D包括D4和D5亚型; M8包括M8a和Z亚型; U包括U1~U8和K. 圆括号内数字为样本个数. *表示此单倍型不能归入任何一种亚型, 民族人群保安, 土, 东乡, 乌兹别克, 维吾尔, 广东汉, 云南汉, 山东汉, 辽宁汉, 西西里, 托斯卡纳, 斯洛文尼亚, 维吾尔 1, 哈萨克 1, 哈萨克 2, 维吾尔 2, 吉尔吉斯 1, 吉尔吉斯 2 依次由BA, Tu, DX, Uzb, Uy, GH, YH, SH, LH, Sic, Tus, Slo, Uyg1, Kaz1, Kaz2, Uyg2, Kit2, Kir2 表示. ^a 姚永刚等人 [7], ^b Francalacci 等人 [9], ^c Torroni 等人 [8], ^d Malyarchuk 等人 [10], ^e 姚永刚等人 [5], ^f Comas 等人 [2]

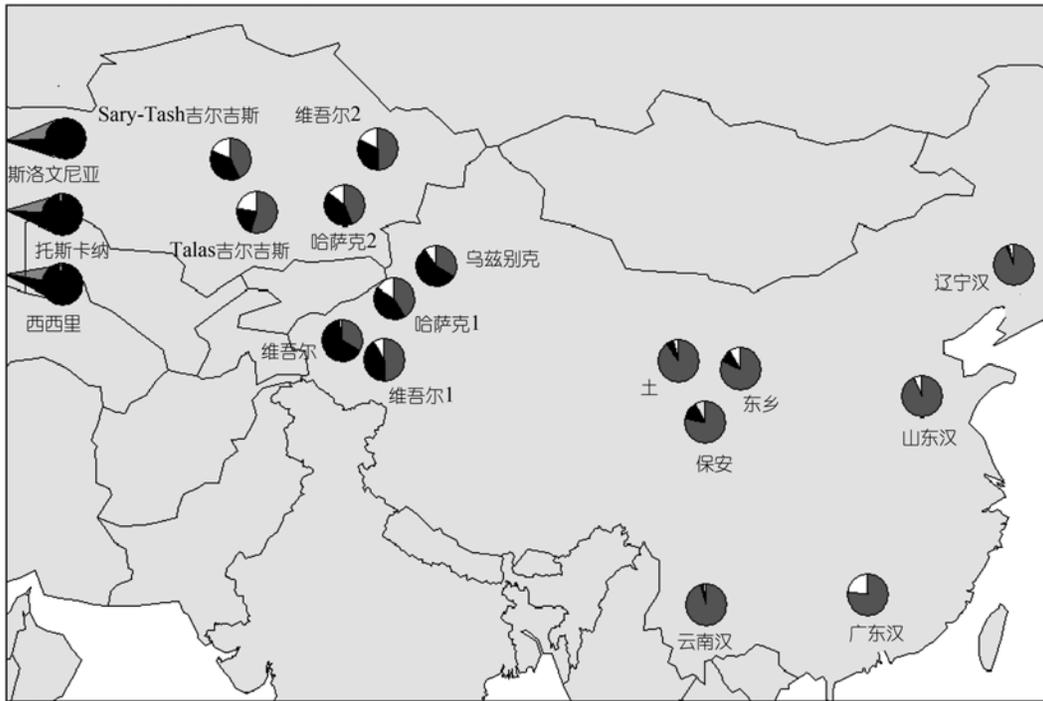


图 2 研究及参照样本来源地及单倍型比例示意图

圆饼图示研究样本的线粒体单倍型频率分布, 黑色部分表示欧洲特异单倍型(HV, H, V, J, T, U, I, W, X); 灰色部分表示东亚特异单倍型(M7, M8, M9, M10, A, B, D, G, F, F3 和 N9a); 白色部分表示其他单倍型(R, R9b, M*, N*)

的 A, D 和 M8 外, 欧洲特异的 H, J 和 U 单倍型也均有大量分布. 在参照人群当中, 4 个汉族人群只有少量欧洲特异单倍型分布, 3 个欧洲人群(托斯卡纳、西西里和斯洛文尼亚)中没有东亚特异单倍型分布. 河西走廊地区的保安、东乡和土族人群的欧洲和东亚特异单倍型频率处在东亚人群(4 个汉族)和中亚人群(维吾尔和乌兹别克)之间. 从单个单倍型频率看, A 在土族中有 43.8% 的高频分布, T 在维吾尔人中有 24.0% 的高频分布.

2.2 主成分(principal component)分析

为了说明各人群间的遗传亲缘关系, 我们对研究及参照的 18 个群体 mtDNA 单倍型频率进行了主成分分析, 结果见图 3. 由图 3 可见, 2 个主成分累计贡献率占总方差的 69.28%. 代表东亚人群的 4 个汉族人群聚集在图的右下角, 2 个南方汉族(广东汉和云南汉)和 2 个北方汉族(山东汉和辽宁汉)聚在一起; 3 个欧洲人群(托斯卡纳、西西里和斯洛文尼亚)分布在图的左上角, 而代表中亚人群的两个维吾尔、一个乌

兹别克和两个吉尔吉斯却分布在东亚人群和欧洲人群之间的位置. 河西走廊地区的保安、东乡和土族分布在中亚人群和东亚人群之间. 保安和东乡族更接近北方汉族, 土族分布比较独立, 处在汉族、哈萨克和吉尔吉斯人之间.

2.3 邻接系统树(Neighbor-Joining tree)分析

基于 ARLEQUIN 2.000 软件计算出 18 个群体之间的 F_{ST} 遗传距离, 用 Phylip 3.62 软件构建邻接树, 结果见图 4. 保安和东乡靠近由 4 个东亚汉族人群形成的一枝, 土族单独形成了较长的一枝与东乡和保安相连. 3 个欧洲人群形成了和东亚人群距离最远的一枝. 我们研究的维吾尔和乌兹别克人群同参照的 2 个维吾尔、2 个哈萨克和 2 个吉尔吉斯人群分布在东亚-河西走廊人群和欧洲人群之间的位置. 这一结果表明在遗传关系上河西走廊人群更接近东亚和中亚人群而远离欧洲人群. 这一结果和主成分分析的结果相一致, 在一定程度上也和这些人群分布的地理位置相对应.

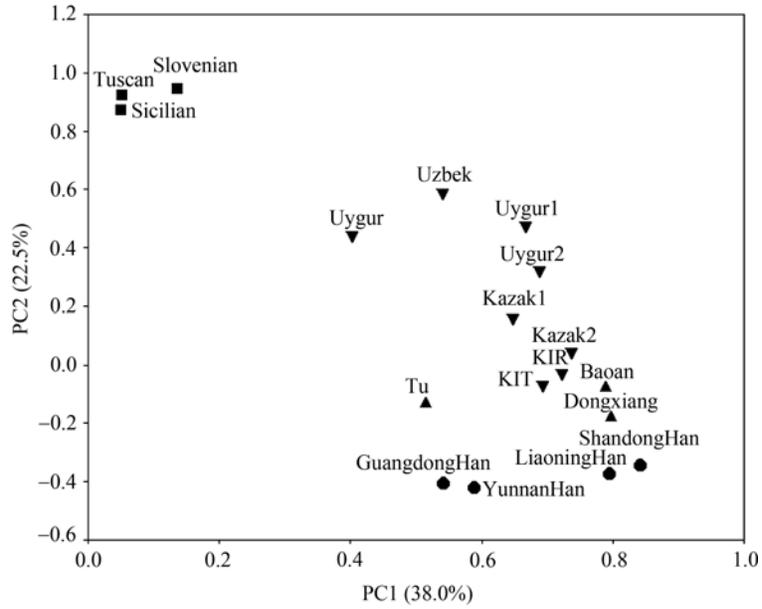


图 3 18 个群体 mtDNA 单倍群频率主成分坐标图

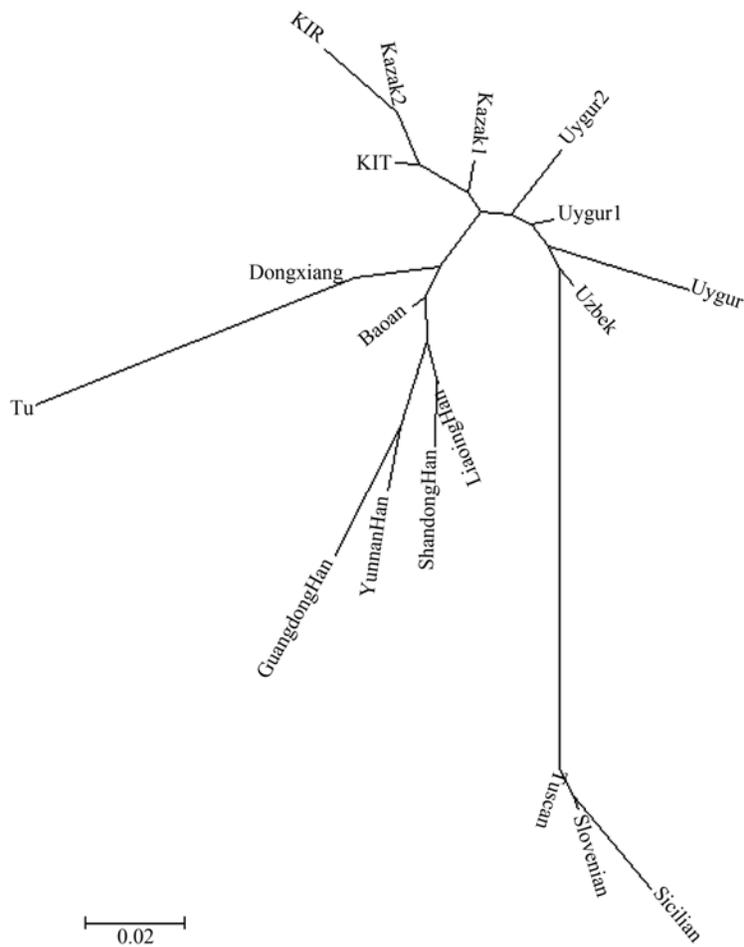


图 4 18 群体 mtDNA 邻接系统树

2.4 基因混合程度估算

用参照的东亚汉族人群(广东汉、云南汉、山东汉和辽宁汉)和中亚人群(维吾尔、哈萨克和吉尔吉斯)作为 2 个亲本人群, 采用Roberts和Hiorns^[16]的方法对所研究的 5 个民族人群进行了基因混合程度估算, 结果见表 4. 我们可以看出, 东亚和欧洲的mtDNA基因库对这 5 个人群的贡献形成两类, 正如预期, 中亚

人群对于维吾尔族和乌兹别克族的基因贡献是 100%, 而东亚人群和中亚人群对于 3 个河西走廊地区人群均有基因贡献. 对于土族和保安族来说, 更大的基因贡献来自于东亚人群; 而对于东乡族, 更大的基因贡献来自于中亚人群. 我们还将相关的 26 个欧亚人群中东亚和欧洲 mtDNA 特异单倍型频率以及不能划入东亚和欧洲特异的单倍型频率列于表 5.

表 4 5 个民族人群东亚和中亚基因贡献程度估算

人群	基因贡献程度/ M_{RH}^a	
	东亚人群	中亚人群
土族人	0.6696	0.3304
东乡人	0.2868	0.7132
保安人	0.6695	0.3305
维吾尔人	0.0000	1.0000
乌兹别克人	0.0000	1.0000

a) 东亚和中亚人基因贡献程度估算方法源自Roberts和Hiorns^[16]

表 5 26 个欧洲、西亚、中亚、河西走廊和东亚人群中东亚和欧洲特异单倍型频率以及不能划分的单倍型频率分布^{a)}

区域	人群	东亚特异单倍型/%	欧洲特异单倍型/%	不能划分的单倍型/%
欧洲	托斯卡纳 ^a	0	97.9	2.1
	斯洛文尼亚 ^b	0	100	0
	西西里 ^c	0	98	2
	土耳其 ^d	8	92	0
西亚	土耳其(阿塞拜疆) ^d	2.5	97.5	0
	Mazandarian ^d	0	95.2	4.8
	Gilaki ^d	2.7	94.6	2.7
	伊朗 ^e	15	80	5
	土库曼 ^d	36.5	58.4	5.1
	乌兹别克 ^d	30.9	54.9	14.2
	哈萨克 ^f	56.4	40.0	3.6
	Talas 吉尔吉斯 ^f	72.9	25.0	2.1
中亚	Sary-Tash 吉尔吉斯 ^f	59.6	31.9	8.5
	维吾尔族 ^f	54.5	34.5	11
	哈萨克族 ^g	66.2	30.3	3.5
	维吾尔族 ^g	55.3	42.4	2.3
	乌兹别克族 ^g	51.3	37.6	11.1
	乌兹别克族	33.3	56.7	10
	维吾尔族	36	64	0
	保安族	82	14	4
河西走廊	东乡族	81.1	9.5	9.4
	土族	89.5	6.3	4.2
	广东汉族 ^h	76.7	0	23.3
东亚	云南汉族 ^h	95.4	2.3	2.3
	山东汉族 ^h	96	0	4
	辽宁汉族 ^h	96	2	2

a) mtDNA单倍型的划分基于Macaulay等人^[12], Kivisild等人^[11]和姚永刚等人^[7], ^a Torroni等人^[8], ^b Malyarchuk等人^[10], ^c Francalacci等人^[9], ^d Quintana-Murci等人^[1], ^e Comas等人^[3], ^f Comas等人^[2], ^g 姚永刚等人^[5], ^h 姚永刚等人^[7]

3 讨论

东亚和欧洲mtDNA特异单倍型在研究的维吾尔和乌兹别克人群中均有高频分布(表 3, 图 2), 这一结果支持姚永刚等人 [4.5]和Comas等人 [2.3]所认为的在中亚地区存在着广泛的东、西方民族间基因融合的观点. 东亚和欧洲mtDNA特异单倍型同时存在于保安族、土族和东乡族人群中说明基因融合也存在于河西走廊地区. 河西走廊人群中较低的欧洲mtDNA特异单倍型频率分布表明这是欧洲人的血脉沿丝绸之路流向中亚过程中再向东延续的结果. 伴随着各人群沿丝绸之路分布的地理位置的不同, 一个明显的趋势是: 由西向东, 欧洲特异单倍型频率逐渐降低而东亚特异单倍型频率逐渐升高. 河西走廊人群中低频欧洲特异单倍型和高频东亚特异单倍型的分布暗示相对于中亚, 较少的欧洲人的血脉流到了河西走廊地区人群.

主成分和邻接树分析表明, 在遗传关系上, 河西走廊人群位于中亚和东亚人群之间, 比研究的维吾尔和乌兹别克两个新疆人群更接近东亚人, 这一结果也和单倍型频率分布(表 5)以及混合程度分析(表 4)相对应.

历史研究表明, 从公元前 3 世纪到公元 17 世纪, 东亚人和欧洲人之间沿丝绸之路的经济和文化的交流就断断续续地盛行. 几次大的历史事件可能造成了丝绸之路沿途人群现在的母系遗传结构 [2.3]. 保安和东乡民族主要来源自蒙古而土族源于东亚人 [2.4], 这几个民族人群从公元 7~8 世纪就定居在丝绸之路上游的河西走廊地区. 作为保安族和东乡族主要祖先的蒙古人在公元 13 世纪曾征战欧亚北部大陆 [2.5],

当地妇女的母系遗传成分注入蒙古民族已经不可避免地发生. 与此同时, 居住在河西走廊地区的几个少数在其民族形成的过程中, 来自西亚的阿拉伯人和波斯人和东方的民族人群也融入其中 [2.4], 数千年的东方和西方人之间的联姻、部落沿丝绸之路的迁徙和扩张也会将东方和西方人群的母系血统引入丝绸之路沿途的民族人群 [2.3].

前面已经提到, Quintana-Murci等人 [1], Comas等人 [2.3]和姚永刚等人 [4.5]对丝绸之路沿途的西亚、中亚和我国新疆地区人群的母系遗传结构研究之后认为东亚和欧洲的母系血统流到了中亚. 现在我们对丝绸之路上游部分的河西走廊地区的民族人群的母系遗传结构进行了研究, 发现相对于中亚人群, 较少的欧洲母系遗传成分流到了河西走廊地区, 这就造成了欧洲母系血统在这一地区人群内存在但含量较少. 河西走廊地区人群所处的中亚和东亚人群之间的遗传位置提示除了丝绸之路中游和下游地区发生的广泛的东亚和欧洲人群间的基因交流之外, 在丝绸之路上游的河西走廊地区也发生了中亚和东亚人群间的基因融合.

4 结论

我们对丝绸之路上游河西走廊地区的三个民族人群(保安族、东乡族和土族)进行了母系遗传分析, 发现在这一地区存在着中亚和东亚人群间的基因融合. 这一研究揭示了基因从中亚到东亚的流动, 使沿着丝绸之路的基因流动和基因融合路线图更加完整. 至此, 欧洲人群基因自欧洲经中亚再到河西走廊的流动路线显得更为清晰.

致谢 感谢为我们研究提供样品的献血者. 感谢宿兵博士对文稿给予指导性建议.

参考文献

- 1 Quintana-Murci L, Chaix R, Wells R S, et al. Where west meets east: the complex mtDNA landscape of the southwest and Central Asian corridor. *Am J Hum Genet*, 2004, 74: 827—845 [DOI](#)
- 2 Comas D, Calafell F, Mateu E, et al. Trading genes along the Silk Road: mtDNA sequences and the origin of Central Asian populations. *Am J Hum Genet*, 1998, 63: 1824—1838 [DOI](#)
- 3 Comas D, Plaza S, Wells R S, et al. Admixture, migrations, and dispersals in Central Asia: evidence from maternal DNA lineages. *Eur J Hum Genet*, 2004, 12: 495—504 [DOI](#)
- 4 Yao Y G, Lu X M, Luo H R, et al. Gene admixture in the silk road region of China: evidence from mtDNA and melanocortin 1

- receptor polymorphism. *Genes Genet Syst*, 2000, 75: 173—178 [\[DOI\]](#)
- 5 Yao Y G, Kong Q P, Wang C Y, et al. Different matrilineal contributions to genetic structure of ethnic groups in the silk road region in china. *Mol Biol Evol*, 2004, 21: 2265—2280 [\[DOI\]](#)
 - 6 宋蜀华, 陈克进. 中国民族概论. 北京: 中央民族大学出版社, 2001. 553—560
 - 7 Yao Y G, Kong Q P, Bandelt H J, et al. Phylogeographic differentiation of mitochondrial DNA in Han Chinese. *Am J Hum Genet*, 2002, 70: 635—651 [\[DOI\]](#)
 - 8 Torroni A, Huoponen K, Francalacci P, et al. Classification of European mtDNAs from an analysis of three European populations. *Genetics*, 1996, 144: 1835—1850
 - 9 Francalacci P, Morelli L, Underhill P A, et al. Peopling of three Mediterranean islands (Corsica, Sardinia, and Sicily) Inferred by Y-Chromosome biallelic variability. *Am J Phys Anthropol*, 2003, 121: 270—279 [\[DOI\]](#)
 - 10 Malyarchuk B A, Grzybowski T, Derenko M V, et al. Mitochondrial DNA variability in Bosnians and Slovenians. *Ann of Hum Genet*, 2003, 67: 412—425 [\[DOI\]](#)
 - 11 Kivisild T, Tolk HV, Parik J, et al. The emerging limbs and twigs of the East Asian mtDNA tree. *Mol Bio Evol*, 2002, 19: 1737—1751
 - 12 Macaulay V, Richards M, Hickey E, et al. The emerging tree of west Eurasian mtDNA: a synthesis of control-region sequence and RFLPs. *Am J Hum Genet*, 1999, 64: 232—249 [\[DOI\]](#)
 - 13 Andrews R M, Kubacka I, Chinnery P F, et al. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet*, 1999, 23: 147 [\[DOI\]](#)
 - 14 Schneider S, Roessli D, Excoffier L. Arlequin version 2.000: A software for population genetics data analysis. *Genetics and Biometry Laboratory*. Geneva: University of Geneva, 2000
 - 15 Tamura K, Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol*, 1993, 10: 512—526
 - 16 Roberts D F, Hiorns R W. Methods of analysis of the genetic composition of a hybrid population. *Hum Biol*, 1965, 37: 38—43
 - 17 Dupanloup I, Bertorelle G. Inferring admixture proportions from molecular data: extension to any number of parental populations. *Mol Biol Evol*, 2001, 18: 672—675
 - 18 Chakraborty R. Gene admixture in human populations: models and predictions. *Yearb Phys Anthropol*, 1986, 29: 1—43 [\[DOI\]](#)
 - 19 Sans M, Salzano F M, Chakraborty R. Historical genetics in Uruguay: estimates of biological origins and their problems. *Hum Biol*, 1997, 69: 161—170
 - 20 Finnila S, Lehtonen M S, Majamaa K. Phylogenetic network for European mtDNA. *Am J Hum Genet*, 2001, 68: 1475—1484 [\[DOI\]](#)
 - 21 Herrnstadt C, Elson J L, Fahy E, et al. Reduced-median-network analysis of complete mitochondrial DNA coding-region sequences for the major African, Asian, and European haplogroups. *Am J Hum Genet*, 2002, 70: 1152—1171 [\[DOI\]](#)
 - 22 Kong Q P, Yao Y G, Sun C, et al. Phylogeny of East Asian mitochondrial DNA lineages inferred from complete sequences. *Am J Hum Genet*, 2003, 73: 671—676 [\[DOI\]](#)
 - 23 Cavalli-Sforza L L, Menozzi P, Piazza A. *The history and geography of human genes*. Princeton, NJ: Princeton University Press, 1994.195—252
 - 24 包尔汉, 江云, 牙含章. 等. 中国大百科全书-民族卷. 北京: 中国大百科全书出版社, 1986. 42: 97
 - 25 张帆, 曹永年, 叶新民, 等. 蒙古民族通史. 呼和浩特: 内蒙古大学出版社, 2002. 1: 224—251