

中国汉族群体 5 个 STR 分子遗传标记 多态性研究

贾振军¹, 吴 谨¹, 李 慧², 侯一平¹, 张伟娟¹, 周雪平¹, 邓建强¹,
石美森¹, 张 霽¹, 李英碧¹

(1. 四川大学华西医学中心法医物证教研室, 成都 610041; 2. 解放军四十七医院内三科, 成都 610500)

摘要:为了解中国人 5 个 STR 基因座等位片段结构特征, 获得汉族群体 *D2S2955*、*D3S4014*、*D20S604*、*D22S689* 和 *GATA198B05* 基因座的群体遗传学数据。采取成都地区无血缘关系汉族个体血样 EDTA 抗凝血, Chelex 法提取 DNA, PCR 扩增, 非变性聚丙烯酰胺凝胶不连续缓冲系统水平电泳分型, 自动激光荧光测序仪测定 DNA 序列。序列分析显示, 中国人 *D2S2955*、*D3S4014*、*D20S604* 基因座具有简单重复序列, 而 *D22S689*、*GATA198B05* 基因座具有复杂重复序列。5 个 STR 基因座在成都汉族群体中均具有遗传多态性。揭示了我国汉族人群 5 个 STR 基因座的等位基因片段结构特征, 为人类群体遗传研究提供了数据, 建立的不连续缓冲系统水平电泳分型方法为检测这 5 个 STR 基因座提供了简便技术。

关键词:聚合酶链反应; 遗传多态性; 短串联重复序列; 中国汉族群体

中图分类号: Q347; D919.2 文献标识码: A

文章编号: 0253-9772(2005)03-0343-06

Study of Five STR Loci in A Chinese Han Population

JIA Zhen-Jun¹, WU Jin¹, LI Hui², HOU Yi-Ping¹, ZHANG Wei-Juan¹, ZHOU Xue-Ping¹,
DENG Jian-Qiang¹, SHI Mei-Sen¹, ZHANG Ji¹, LI Ying-Bi¹

(1. School of Preclinical and Forensic Medicine, Sichuan University, Chengdu 610041, China;

2. No. 47 Hospital of PLA, Chengdu 610500, China)

Abstract: To understand the allele structure and genetic polymorphisms at five STR loci in Chinese Han population, and construct a preliminary database, EDTA-blood specimens were collected from unrelated individuals. DNA samples were extracted with Chelex method and were amplified by PCR technique. The PCR products were analyzed using both PAGE horizontal electrophoresis with discontinuous buffer system and automated fluorescence detection approach. As a result, three STRs consist of simple repeat motifs, while two STR contain a complex repeat structure. The STR polymorphisms at all of the five loci have been observed in Chinese Han population. In a word, the obtained data are beneficial to understanding the population genetics of the five STR loci in Chinese Han population. As a simple approach, the PAGE horizontal electrophoresis can be employed for typing the five STR markers.

Key words: polymerase chain reaction; genetic polymorphism; short tandem repeats; Chinese Han population

收稿日期: 2004-06-08; 修回日期: 2004-12-30

基金项目: 四川省科技厅应用基础项目基金(编号: 03JY029-074-1) 及纽约中华医学基金(CMB, 2000-636)资助 [Supported by the Sichuan Provincial Funds(No: 03JY029-074-1), and China Medical Board, N. Y. (No: 2000-636)]

作者简介: 贾振军(1977—), 男, 硕士研究生, 研究方向: 法医遗传学。

通讯作者: 李英碧(1963—), 男, 副教授, 研究方向: 法医遗传学。E-mail: liyingbiscu@163.com

短串联重复序列(short tandem repeat, STR)不仅是绘制基因组物理图谱和遗传连锁分析的理想标记,也是群体遗传学和法医遗传学高信息基因座的丰富来源。近年来,国外学者报道了一些新的法医STR遗传标记^[1],但缺少中国大陆人群的群体遗传数据。为了解中国人群5个法医STR基因座等位基因片段结构特征,获得中国汉族群体D2S2955、D3S4014、D20S604、D22S689和GATA198B05基因座基因频率。我们对成都汉族群体5个STR基

因座的遗传多态性进行了研究。

1 材料和方法

1.1 样本 DNA

120份EDTA抗凝血采自中国成都地区无血缘关系汉族个体。Chelex法提取DNA^[2]。

1.2 PCR扩增

5个基因座基本参数及引物序列见表1。每一扩增样本含人类基因组DNA 2~40ng, 1×PCR缓

表1 5个STR基因座及PCR条件

Table 1 Five STR loci and the PCR condition

基因座 Locus	基因库号 GenBank accession	PCR 引物序列 PCR primer sequences	变性 Denature	退火 Annealing	延伸 Extension
D2S2955	M36981	Forward primer TAATCAAGATAATGTGGTACTGGC	94 °C	57 °C	72 °C
		Reverse primer GGTAAACCTTTATCCTGGACTTCA	30 s	30 s	30 s
D3S4014	M34865	Forward primer TCTCCTGGAAGGAAAGAGCT	94 °C	56 °C	72 °C
		Reverse primer TGCAGAACCCC ATACTGTTAT	30 s	30 s	30 s
D20S604	M51615	Forward primer TGAGTGAATTTCCTAATAAAATCCC	94 °C	58 °C	72 °C
		Reverse primer AAGCAATOTCATTTATTTTACACA	30 s	45 s	30 s
D22S689	M40668	Forward primer TATGTACAGACCTGCAACTTG	94 °C	58 °C	72 °C
		Reverse primer CCTGCCTGCCTATCTATCTG	30 s	30 s	45 s
GATA198 B05	M90030	Forward primer GGGAGCCAATTCAACCTAAT	94 °C	58 °C	72 °C
		Reverse primer GAGAGACAGAACCTATAGCATGC	30 s	30 s	45 s

冲液, 1.5 mmol/L MgCl₂, 150 μmol/L dNTP, 1 U Taq聚合酶(深圳博彩公司产品), 每一种引物0.25 μmol/L, 反应体积37.5 μL, 在热循环仪(PE9600)中, 按表1参数循环32次。

1.3 PCR产物检测和等位基因的确定

PCR产物用非变性聚丙烯酰胺凝胶不连续缓冲系统水平电泳分离, 银染显示电泳谱带^[3]。样本PCR产物与等位基因分型标准物同步电泳, 比较样本PCR产物的电泳谱带对应分型标准物所处的位置, 确定样本表型。等位基因分型标准物由不同基因型个体的DNA混合后按1.2方法进行PCR扩增制备, 制备的分型标准物与测定序列的样本同步电泳, 获得命名^[4]。

1.4 DNA序列分析

每基因座每一种等位片段随机选择一个样本, 从谱带处切下凝胶洗脱出DNA, 采用PCR循环测序法, 用自动激光荧光测序仪(ABI310)测定序列。

1.5 数据处理

按文献[5,6]的方法。

2 结果

2.1 等位基因结构及命名

采用自动激光荧光测序仪对所选择样本进行测序, 结果显示: 中国人D2S2955、D3S4014和D20S604基因座具有简单重复序列, 其中D2S2955重复序列为[AGAT]_n, D3S4014重复序列为[CTAT]_n, D20S604重复序列为[ATCT]_n。按国际法医血液遗传学会DNA委员会推荐原则^[5], 以核心序列重复次数命名等位基因。例如, 具有简单重复序列的D20S604基因座共有8个等位片段, 其核心序列ATCT的重复次数n为11~18, 因此, 这8个等位基因分别命名为: 11~18。另一方面, D22S689、GATA198B05基因座具有复杂重复序列, D22S689的重复序列为[TAGA]_n[CAGA]_n, 共9个等位基因, 命名为10~18; GATA198B05基因座重复序列为[ATCT]_n[CTAC]_n, 共10个等位基因, 命名为11~20。

2.2 等位基因分型标准物构建及STR分型

采用混合不同基因型个体DNA进行PCR扩增,PCR扩增产物与测定序列样本同步电泳的方法,制备了5个STR基因座分型标准物。图1、图2、图3、图4分别给出了在人类等位基因分型标准物对照下,电泳测定不同个体表型的实例。

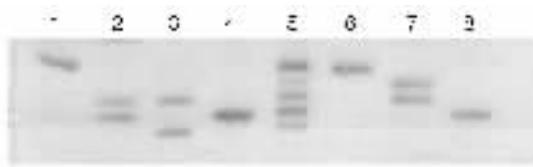


图1 自制等位基因标准物对照下不同个体

D3S4014 电泳分型

1:12-12; 2:9-10; 3:8-10; 4:9-9; 5:等位基因标准物8~12; 6:12-12; 7:10-11; 8:9-9。

Fig.1 D3S4014 typing with the home made allele ladder

1:12-12; 2: 9-10; 3: 8-10; 4: 9-9; 5: human allele ladder 8~12; 6: 12-12; 7: 10-11; 8: 9-9.



图2 自制等位基因标准物对照下不同个体

D20S604 电泳分型

1:11-14; 2:14-15; 3:15-15; 4:等位基因标准物11~18; 5:13-16; 6:12-16; 7:13-16; 8:14-15; 9:17-18; 10:11-16; 11:14-18; 12:15-17。

Fig.2 D20S604 typing with the home made allele ladder

1:11-14; 2:14-15; 3:15-15; 4: human allele ladder 11~18; 5:13-16; 6:12-16; 7:13-16; 8:14-15; 9:17-18; 10:11-16; 11:14-18; 12:15-17.

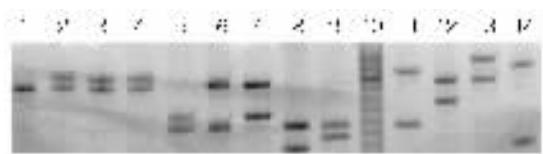


图3 自制等位基因标准物对照下不同个体

D22S689 电泳分型

1:16-16; 2:16-17; 3:16-17; 4:16-17; 5:12-13; 6:12-16; 7:13-16; 8:10-12; 9:11-12; 10:等位基因标准物10~18; 11:12-17; 12:14-16; 13:16-18; 14:10-17。

Fig.3 D22S689 typing with the home made allele ladder

1:16-16; 2:16-17; 3:16-17; 4:16-17; 5:12-13; 6:12-16; 7:13-16; 8:10-12; 9:11-12; 10: human allele ladder 10~18; 11:12-17; 12:14-16; 13:16-18; 14:10-17.

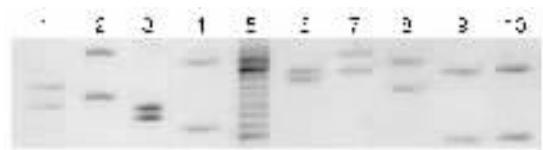


图4 自制等位基因标准物对照下不同个体

GATA198B05 电泳分型

1:14-16; 2:15-20; 3:13-14; 4:12-19; 5:等位基因标准物11~20; 6:17-18; 7:18-20; 8:16-19; 9:11-18; 10:11-18。

Fig.4 GATA198B05 typing with the home made allele ladder

1:14-16; 2:15-20; 3:13-14; 4:12-19; 5: human allele ladder 11~20; 6:17-18; 7:18-20; 8:16-19; 9:11-18; 10:11-18.

2.3 STR基因座等位基因的群体分布

表2及表3列出了5个STR基因座在成都地区汉族群体的基因型观察值及等位基因频率,按文献[5]的方法进行卡方检验,结果表明各基因座基因型分布符合Hardy-Weinberg平衡。

表2 成都汉族群体5个STR基因型分布

Table 2 The distribution of genotypes at five STR loci in Han population in Chengdu

GATA198B05				D2S2955		D3S4014		D20S604		D22S689	
基因型	观察数	基因型	观察数	基因型	观察数	基因型	观察数	基因型	观察数	基因型	观察数
Genotype	Obs.	Genotype	Obs.	Genotype	Obs.	Genotype	Obs.	Genotype	Obs.	Genotype	Obs.
11-18	1	17-18	8	10-10	6	8-10	3	11-15	2	10-13	1
12-14	1	17-19	1	10-11	2	9-9	32	12-12	1	10-14	1
12-17	1	17-20	1	10-12	16	9-10	66	12-13	1	10-15	7
12-18	2	18-18	16	10-13	4	9-11	4	12-15	3	10-16	3
12-19	1	18-19	10	11-11	12	9-12	1	12-16	4	11-13	1

续表 2

GATA198B05				D2S2955		D3S4014		D20S604		D22S689	
基因型	观察数	基因型	观察数	基因型	观察数	基因型	观察数	基因型	观察数	基因型	观察数
Genotype	Obs.	Genotype	Obs.	Genotype	Obs.	Genotype	Obs.	Genotype	Obs.	Genotype	Obs.
13-14	3	18-20	5	11-12	20	10-10	9	12-17	4	11-15	2
13-16	3	19-19	1	11-13	13	10-11	4	12-18	1	11-16	2
13-17	5	19-20	1	12-12	25	10-12	3	13-13	2	12-14	1
13-18	3	20-20	1	12-13	10			13-15	3	12-15	3
13-19	6			13-13	6			13-16	16	12-16	2
14-14	3							13-17	1	13-13	3
14-15	5							14-14	2	13-14	4
14-16	2							14-15	6	13-15	12
14-17	4							14-16	3	13-16	2
14-18	11							14-17	1	13-17	1
14-19	1							14-18	2	14-14	3
14-20	2							15-15	18	14-15	18
15-15	1							15-16	23	14-16	11
15-16	2							15-17	2	14-17	1
15-18	6							15-18	1	15-15	14
15-19	2							16-16	19	15-16	11
15-20	1							16-17	1	15-17	1
16-16	1							17-17	2	15-18	4
16-18	2							17-18	1	16-16	5
16-19	2									16-17	4
16-20	2									16-18	1
17-17	2									17-17	3
Total	119			114				119		120	119

表 3 成都汉族群体 5 个 STR 等位基因频率

Table 3 The distribution of genotypes at five STR loci in Han population in Chengdu

GATA198B05		D2S2955		D3S4014		D20S604		D22S689	
基因	频率	基因	频率	基因	频率	基因	频率	基因	频率
Allele	Fre. (%)	Allele	Fre. (%)	Allele	Fre. (%)	Allele	Fre. (%)	Allele	Fre. (%)
11	0.4	10	14.9	8	1.3	11	0.8	10	5.0
12	2.1	11	25.9	9	56.7	12	6.3	11	1.3
13	8.4	12	42.1	10	38.2	13	10.5	12	2.5
14	14.7	13	17.1	11	2.1	14	6.7	13	11.3
15	7.6			12	1.7	15	31.9	14	17.6
16	6.3					16	35.7	15	36.1
17	10.1					17	5.9	16	18.5
18	33.6					18	2.1	17	5.5
19	10.9							18	2.1
20	5.9								
HWE		HWE		HWE		HWE		HWE	
$\chi^2 = 6.0984$		$\chi^2 = 3.6891$		$\chi^2 = 4.3067$		$\chi^2 = 5.193$		$\chi^2 = 5.0368$	
$df = 9$		$df = 6$		$df = 5$		$df = 7$		$df = 10$	
$P > 0.05$		$P > 0.05$		$P > 0.05$		$P > 0.05$		$P > 0.05$	

Note: Fre:frequency(频率); HWE: Test for Hardy-Weinberg equilibrium(Hardy-Weinberg 平衡检验).

2.4 STR基因座的法医遗传学理论值

表4显示由5个STR基因座汉族群体等位基因频率所求得的法医遗传学理论值。*D2S2955*、*D20S604*、*D22S689*和*GATA198B05*基因座在我国汉族人群中等位基因频率分布均匀,5个STR基因累积非父排除概率及累积个人识别几率分别为0.9373301和0.9997456。

表4 5个STR基因座的杂合度、标准误、排除概率及个人识别机率

Table 4 The heterozygosity, standard errors, exclusion probability and discrimination power at

five STR loci in Chinese population

基因座 Locus	观察 杂合度 H_o	期望 杂合度 H_e	标准误 SE	非父排除 概率 EP	个人识别 机率 DP
<i>GATA198B05</i>	0.790	0.823	0.0151	0.580	0.944
<i>D2S2955</i>	0.570	0.704	0.0150	0.257	0.863
<i>D3S4014</i>	0.655	0.5317	0.0117	0.363	0.612
<i>D20S604</i>	0.625	0.751	0.0164	0.322	0.889
<i>D22S689</i>	0.765	0.783	0.0155	0.535	0.923

Note: H_o (heterozygosity observed), H_e (heterozygosity expected), SE (standard error), EP(power of exclusion), DP(power of discrimination).

3 讨论

STR作为DNA标记一经运用就显示了它的优越性,杂合度平均达到0.7,容易分析。所用的DNA样品较少、便于取材,一次PCR只需要10~15 ng基因组DNA。1998年Litt、Luty、刘雅诚和侯一平等连续报道了STR基因座的应用^[8,9]。STR基因座在人类学、法医学、民族学和多基因遗传病的研究等领域不断显示出重要的应用价值,目前STR作为遗传连锁标记已经得到了广泛的应用^[10~13]。但由于各实验室条件不同,使用的遗传标记,非父排除能力可有明显差异。因而,不断选择稳定、易于检测、非父排除能力高的遗传标记,合理取舍遗传标记系统,优化并不断提高检测体系的非父排除能力始终是各实验室努力追求的目标。所以我们选择了5个STR位点进行了研究。*D2S2955*、*D3S4014*和*GATA198B05*基因座,是新发现的STR遗传标记,缺乏国内外的群体遗传学数据。其中*D2S2955*、*D3S4014*基因座具有简单重复序列,等位基因相对

较少,*D22S689*、*GATA198B05*基因座具有复杂重复序列,等位基因较多且分布均匀。*D20S604*基因座是近年来各国学者用于人类群体遗传学研究和法医遗传学鉴定的新STR遗传标记,在国外对*D20S604*基因座的报道中是复杂重复序列ATCT×TATC,而我们研究显示是简单重复序列ATCT。上述分析结果显示:*D22S689*、*D20S604*和*GATA198B05*等3个基因座对成都汉族人群来讲,其基因频率分布较好,提供的信息量较大,是个人识别率高、非父排除能力较强且能稳定遗传的基因座,具有较高的应用价值。

参考文献(References):

- [1] http://www.chlc.org.
- [2] Walsh P S, Metzger D A, Higuchi R. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques*, 1991, 10(4): 506~513.
- [3] Allen C R, Graves G, Budowle B. Polymerase chain reaction amplification products separated on rehydratable polyacrylamide gels and stained with silver. *Biotechniques*, 1989, 7(7): 736~744.
- [4] HOU Yi-Ping, WU Jin, LI Ying-Bi, Tang Jian-Pin, ZHANG Si-Zhong, CHU Jia-You, M Prinz. Genetic polymorphisms of five STR loci in a Chinese Han population. *China J Med Gene*, 1999, 16(4): 228~232.
侯一平, 吴 谦, 李英碧, 唐剑频, 张思仲, 褚嘉祐, M Prinz 成都汉族群体五个STR基因座遗传多态性研究. 中华医学遗传学杂志, 1999, 16(4): 228~232.
- [5] Hou Y, Prinz M, Staak M. Comparison of different tests for deviation from Hardy-Weinberg equilibrium of AMPFLP population data. In: Barw, Fiori A, Rossi U. *Advances in Forensic Haemogenetics 5*, Springer-Verlag: Berlin Heidelberg New York, 1994, 511~514.
- [6] HOU Yi-Ping, GOU Qing, WU Mei-Yun, M Staak, M Prinz. Population genetics of short tandem repeat locus VWA. *China J Med Gene*, 1996, 13(2): 65~69.
侯一平, 苟 清, 吴梅筠, M Staak, M Prinz. 人类短串联重复序列VWA基因座的群体遗传学研究. 中华医学遗传学杂志, 1996, 13(2): 65~69.
- [7] Bar W, Brinkmann B, Budowle B, Carracedo A, Gill P, Lincoln P, Mayr W, Olaisen B. DNA recommendations. Further report of the DNA Commission of the ISFH regarding the use of short tandem repeat systems. International Society for Forensic Haemogenetics. *Int J Legal Med*, 1997, 110(4): 175~176.
- [8] HOU Yi-Ping, LI Ying-Bi, TANG Jian-Pin, WU Jin, ZHANG Si-Zhong. Genetic polymorphisms of eight STR loci in a Chinese Han population. *China J Med Gene*, 2000, 17(4): 236~240.

- 侯一平, 李英碧, 唐剑频, 吴 谨, 张思仲. 成都汉族群体八个STR基因座遗传多态性研究. 中华医学遗传学杂志, 2000, 17(4): 236~240.
- [9] LIU Ya-Cheng, JIAO Zhang-Ping, TANG Hui, WANG Jing, YAN Jiang-Wei. Genetic polymorphisms of Penta E, Penta D in a Chinese Han population in Beijing. *China J Med Gene*, 2001, 16(2): 95~96.
刘雅诚, 焦章平, 唐 晖, 王 静, 严江伟. 北京地区汉族人群 Penta E、Penta D 遗传多态性. 中国法医学杂志, 2001, 16(2): 95~96.
- [10] ZHOU Qiang, WU Si-Kun, YU Fang, HE Rong-Yue. Genetic polymorphisms of STR loci TH01, TPOX, CSF1PO in Guizhou Han population. *Hereditas (Beijing)*, 2004, 26(1): 31~34.
周 强, 吴思鹃, 喻 芳, 何荣跃. 贵州地区汉族人群 TH01、TPOX、CSF1PO 基因座的遗传多态性. 遗传, 2004, 26(1): 31~34.
- [11] ZHANG Hong-Bo, ZHAO Jun-Hai, LAI Jiang-Hua, CHEN Teng, LI Sheng-Bin, LAI Shu-Ping. The study of HLA-Cw polymorphism in Uygur population. *Hereditas (Beijing)*, 2003, 25(5): 549~551.
张洪波, 赵钧海, 赖江华, 陈 腾, 李生斌, 赖淑萍. 新疆维吾尔族 HLA-Cw 基因座遗传多态性调查. 遗传, 2003, 25(5): 549~551.
- [12] LI Yi, HAO Bing-Tao, YANG Yan-Li, ZHU Wen-Yu, SI Yan-Mei, WANG Ying-Tai. Genetic polymorphisms of human short tandem repeat vWA. *Hereditas (Beijing)*, 2002, 24(6): 639~642.
李 怡, 郝冰涛, 杨艳丽, 朱文玉, 司艳梅, 王应太. 河南汉人群体短串联重复 vWA 遗传多态性研究. 遗传, 2002, 24(6): 639~642.
- [13] YU Lei, WU Xin-Yao, HUANG Yan-Mei, CAI Gui-Qing, XU Chuan-Chao, TONG Da-Yue. Studying the polymorphism at DYF155S1 locus in Chinese Han population by MVR-PCR maked with fluorescene. *Acta Genetids Sinica*, 2003, 30(1): 15~19.
庾 蕤, 伍新尧, 黄艳梅, 蔡贵庆, 许传超, 童大跃. 用荧光标记 MVR-PCR 方法研究中国汉族人群 DYF155S1 基因座多态性. 遗传学报, 2003, 30(1): 15~19.

第三届全国动植物数量遗传学术讨论会 第一轮通知

第三届全国动植物数量遗传学术讨论会拟定于 2005 年 10 月 28 至 30 日在南京召开。在包括分子数量遗传学在内的生物信息学迅速发展的今天, 动物、植物和人类数量遗传学已取得了较大的进展, 有必要将从事动植物数量遗传与育种应用的同行组织在一起, 共同交流我国动植物数量遗传研究领域几年所取得的重要成果与进展, 并探讨数量遗传学的相关问题。本次会议将邀请知名专家回顾我国数量遗传学领域取得的成果、展望本学科的发展趋势和推广数量遗传学的新理论与新方法, 包括动物、植物和人类数量遗传学的进展与展望、QTL 作图方法、标记辅助选择、动植物数量性状的遗传与育种以及遗传多样性等。本届会议除大会报告, 还欢迎科研成果展示和企业产品介绍。现将有关事宜通知如下:

一、会议主题

基因组时代的动植物数量遗传与育种

二、征文范围

· 数量遗传学的新思想、新理论和新方法 · QTL 作图的理论和方法 · 分子标记辅助选择 · 动植物数量性状的遗传与育种 · 动植物遗传多样性

三、征文要求

符合本届会议征文范围的研究论文和文献综述的中文文摘, A4 纸上 1 页, 约 1 000~1 200 字。包括论文题目、作者姓名、单位和邮编、E-mail 地址。论文摘要请作者用 word 文件。论文摘要截止日期 2005 年 8 月 31 日。会议将论文摘要汇编成册。

四、会议注册

正式代表 600 元, 学生 300 元。参会费用将在第二轮通知后交纳。

五、组织单位和主办单位

组织单位: 中国遗传学会、江苏省遗传学会; 主办单位: 南京林业大学、南京农业大学; 组委会主席: 施季森 教授

六、联系方式

联系人: 季孔庶 (E-mail: ksji@njfu.edu.cn) 地 址: 南京市龙蟠路新庄 9 号南京林业大学

邮 编 210037 电 话: 025-85427308 传 真: 025-85427412