

· 述 评 ·

## DNA 微单倍型的研究现状、挑战与展望

陈 峰\*

南京医科大学基础医学院法医学系,江苏 南京 211166

**[摘要]** 法医DNA在刑事侦查中发挥不可替代的作用,面对日益复杂的作案方式,法医DNA面临多重挑战,寻找有效的新型法医DNA遗传标记具有重要的应用价值和科学意义。微单倍型(microhaplotype)是一种新型法医遗传标记,它是长度较短、包含2个以上单核苷酸多态性(single nucleotide polymorphism, SNP)位点的多态性基因座,兼有短串联重复序列(short tandem repeat, STR)和SNP的优点,具备较高多态性和极低突变率。最新研究发现其在祖源推断、个人识别、亲权鉴定等方面具有独特优势,尤其是在混合DNA分析、降解或微量检材检测中具备巨大潜力。文章对微单倍型的概念、特点、研究进展、挑战及展望等进行介绍,希望其将来在法医学应用中发挥更大的作用。

**[关键词]** 微单倍型;祖源推断;个人识别;亲权鉴定

**[中图分类号]** Q523

**[文献标志码]** A

**[文章编号]** 1007-4368(2020)08-1081-04

**doi:** 10.7655/NYDXBNS20200801

### Research progress, challenge and prospect of DNA microhaplotype

CHEN Feng\*

Department of Forensic Medicine, Nanjing Medical University, Nanjing 211166, China

**[Abstract]** The analysis of forensic DNA plays an irreplaceable role in criminal investigation. Forensic DNA identification is facing challenges while dealing with complex cases. The discovery of novel genetic markers for multiple forensic purposes exhibits important application value and scientific significance. Microhaplotype is a type of novel genetic marker, which is defined by two or more closely linked single nucleotide polymorphisms (SNPs) within a short length of DNA fragment. Thus, the multiple allelic combination of SNPs leads to high polymorphism of microhaplotype. It combines strengths of short tandem repeat (STR) and SNP, with relatively high polymorphism and extremely low mutation rate, while applied in forensic identification. Recent researches reveal that microhaplotype possesses unique advantages in ancestry inference, personal identification and paternity testing, as well as great potential in unbalanced DNA mixture analysis and degraded or trace DNA detection. This review summarized the concepts and characteristics, retrospected the research achievements and discussed the challenges in the application of microhaplotype, aiming at exploring a brighter future of microhaplotype in forensic application.

**[Key words]** microhaplotype; ancestry inference; personal identification; paternity testing

[J Nanjing Med Univ, 2020, 40(08): 1081-1084]

2020年,南京医科大学(原南京医学院)学生林某遇害案时隔28年告破,警方通过DNA比对确定犯罪分子的近亲,继而顺藤摸瓜,成功锁定犯罪分子本

人。回想2016年,另一桩世纪悬案“白银案”也以同样方式解决。在这些案件中,法庭科学DNA遗传标记的成功应用发挥了核心作用。

DNA遗传标记在刑事侦查、遇难者及失踪人员身份认定、民事亲权关系鉴定等领域均扮演重要角色。目前应用的遗传标记包括短串联重复序列(short tandem repeat, STR)、单核苷酸多态性(single nucleotide polymorphism, SNP)、插入/缺失多态性

**[基金项目]** 国家自然科学基金(81922041, 81801879);江苏省自然科学基金杰出青年基金项目(BK20170048);江苏省特聘教授计划

\*通信作者(Corresponding author), E-mail: fchen@njmu.edu.cn

(insertion/deletion polymorphism, InDel)等,它们各具特点,适用于不同法医学场景,但在实际应用中面临更加复杂的难题,如复杂混合DNA分析,降解或微量检材检测等,现有DNA遗传标记面临巨大挑战。

2013年,美国耶鲁大学Kidd教授研究团队从现有基因组数据库中,筛选得到300 bp以内、包含至少2个SNP位点的序列片段,将其命名为微单倍型(microhaplotype, MH)<sup>[1]</sup>。本文介绍微单倍型的概念、研究进展,阐述微单倍型的法医学应用优势,并展望其在法医鉴定中的应用前景。

## 1 微单倍型简介

微单倍型最先由Kidd教授团队提出,是指在小于300 bp的DNA区段内,包含至少2个SNP的多态性位点。在300 bp DNA片段长度范围内,微单倍型内部SNP间的重组率几乎为0<sup>[2]</sup>。一个微单倍型位点可视作一个独立基因座,其多态性由SNP的特定组合所构成。

微单倍型兼具STR和SNP的优势:①高度多态性。通常SNP位点仅有2个等位基因,若 $n$ 个二等位基因的SNP构成一个微单倍型位点,那么该微单倍型位点理论上最多有 $2^n$ 种SNP的组合,即最多有 $2^n$ 个等位基因。②低突变率。微单倍型的突变率相当于SNP的突变率,为 $10^{-9}$ ~ $10^{-8}$ /代,是STR突变率的百万分之一到十万分之一<sup>[2]</sup>,在亲权鉴定中有独特优势。③检测无阴影带。基于电泳技术分型的STR会产生阴影带,不利于复杂混合DNA样本的分析。微单倍型通过测序手段检测,无阴影带,且二代测序具有高通量、高灵敏度的优点,在定量分析复杂混合DNA中有很大的潜力。④长度优势。STR基因座的等位基因长度跨度很大,因此会产生扩增不平衡问题,而较长的等位基因在降解检材中极可能被破坏,得不到准确的分型结果,微单倍型长度相对均一,可减少因长度差异带来的扩增不平衡问题。

## 2 微单倍型研究现状

### 2.1 祖源推断

从生物地理的角度,对洲际间及洲际内人群进行祖源推断,是法医遗传学研究热点之一。Chen等<sup>[3]</sup>基于千人基因组数据筛选出10个微单倍型,其信息度(informativeness,  $I_n$ 值)均大于0.15,群体结构分析(structure analysis)可成功区分非洲、欧洲、东亚和南亚人群。Kidd等<sup>[4]</sup>从83个不同人群中筛选130个微单倍型位点,其中94个位点的平均有效等位基

因数(effective allele number,  $A_e$ 值)大于2,28个位点平均 $A_e$ 值大于3,7个位点平均 $A_e$ 值大于4,在进行群体结构分析时,选用其中50个 $I_n$ 大于0.2的位点即可将这83个人群划分出六大祖先成分。Bulbul等<sup>[5]</sup>在先前研究的基础上,根据 $I_n$ 值和 $A_e$ 值挑选出65个微单倍型位点,平均 $A_e$ 值达2.849,平均 $I_n$ 值达到0.243,对96个人群进行祖源推断,发现这65个位点可进一步进行亚人群划分,说明微单倍型具有更细致的人群划分和祖源推断的能力。Turchi等<sup>[6]</sup>对常用的先祖推断标记进行了比较分析,从既往研究报告中根据 $I_n$ 、固定指数(fixation index,  $F_{st}$ )、 $A_e$ 等指标挑选出272个二等位基因SNP位点、36个三等位基因SNP位点和128个微单倍型位点,并基于千人基因组数据评估它们区分洲际间人群的能力,结果发现微单倍型效果最优。

### 2.2 个人识别

个人识别要求遗传标记在人群中具有较高的多态性。Kidd等<sup>[7]</sup>在早期研究中,基于54个不同人群构建了31个微单倍型复合检测体系,其杂合度平均为0.54,累积随机匹配概率相当于15个STR基因座。Chen等<sup>[8-9]</sup>基于千人基因组数据库先后筛选了25个微单倍型,在汉族人群中的平均杂合度为0.6178,其累积随机匹配概率相当于15个STR基因座在汉族人群中的识别能力。研究发现一些较大的微单倍型扩增体系,如Oldoni等<sup>[10]</sup>构建的74个微单倍型复合体系,及近期由De la Puente等<sup>[11]</sup>构建的118个微单倍型复合体系,其个人识别能力远超前有STR扩增体系。

复杂混合DNA分析和微量或降解检材检测是个人识别的难点和重点。Oldoni等<sup>[10]</sup>将微单倍型用于对混合DNA的分析,发现对于总量10 ng、混合比例80:1及总量1 ng、混合比例40:1的样本,其DNA含量较少的个体成分在Ion平台上也能被成功检测。Turchi等<sup>[6]</sup>挑选了18个长度小于180 bp的微单倍型位点用于降解检材检测,发现分型质量主要取决于DNA含量,Ion平台上50 pg的DNA分型准确率大于90%。Chen等<sup>[8]</sup>研究则指出,可以根据阅读深度预测DNA混合比例,表明基于二代测序的微单倍型可能实现混合DNA的定量分析。

### 2.3 亲权鉴定

复杂亲缘关系鉴定是亲权鉴定中的难点问题。Zhu等<sup>[12]</sup>从千人基因组数据库中筛选出11个微单倍型位点,平均 $A_e$ 值为3.3,对20对实际亲子样本进行实测分析,得到似然率(likelihood ratio, LR)平

均达  $4.392\ 796\ 8 \times 10^7$ , 表明微单倍型可被用于亲子鉴定。Sun 等<sup>[13]</sup>筛选出 36 个微单倍型位点, 平均 A<sub>v</sub> 值 3.8。对近亲属样本中实测发现, 在 55 对叔侄或祖孙样本中, 使用 STR 位点错误鉴定 2 例, 无法排除 7 例; 而使用 36 个微单倍型位点, 无法排除 4 例, 无错误鉴定; 联合使用微单倍型位点和 STR 位点, 可得到全部正确的鉴定结果。这表明微单倍型在复杂亲权关系的鉴定中, 可以起到有效补充传统 STR 的作用, 提高鉴定准确性。

### 3 目前已报道的微单倍型位点及应用优势

目前公开报道的微单倍型位点共有 400 多个(表 1), 构成微单倍型的 SNP 位点均可从相应文献获得。部分微单倍型的信息及群体遗传数据可从 Allele FREquency Database (<https://alfred.med.yale.edu/>) 中搜索获得。此外, 一些基于二代测序平台的微单倍型复合扩增体系也已经建立并报道(表 2)。

表 1 文献报道的微单倍型位点

Table 1 The locus information of previously reported microhaplotypes

文献	数量(个)	长度
Hiroaki 等 <sup>[14]</sup>	8	52~98 bp
Voskoboinik 等 <sup>[15]</sup>	10	189~349 bp, 5 个 > 300 bp
van der Gaag 等 <sup>[16]</sup>	16	79~111 bp
Kidd 等 <sup>[4]</sup>	130	12~319 bp, 1 个 > 300 bp
Chen 等 <sup>[3,8,9]</sup>	46	4~76 bp
Sun 等 <sup>[13]</sup>	36	63~423 bp, 11 个 > 300 bp
De la Puente 等 <sup>[11]</sup>	118	3~96 bp
Zhu 等 <sup>[17]</sup>	13	164~250 bp
Gandotra 等 <sup>[18]</sup>	44	300~350 bp
Jin 等 <sup>[19]</sup>	23	<150 bp
Kureshi 等 <sup>[20]</sup>	15	8~177 bp

### 4 微单倍型的挑战与展望

微单倍型在溯源推断、个人识别、亲子鉴定中

表 2 文献报道的微单倍型复合扩增体系

Table 2 The multiplex panel of microhaplotypes

文献	微单倍型体系(个)	测序平台	应用优势
Cheney 等 <sup>[8]</sup>	25	Illumina <sup>TM</sup> HiSeq X-10	混合 DNA, 降解检材
De la Puente 等 <sup>[11]</sup>	118	Illumina <sup>TM</sup> MiSeq、ABI <sup>TM</sup> Ion S5	个人识别, 亲子鉴定, 混合 DNA, 降解检材
Sun 等 <sup>[13]</sup>	36	Illumina <sup>TM</sup> MiSeq	复杂亲权鉴定
Oldoni 等 <sup>[10]</sup>	74	ABI <sup>TM</sup> Ion S5	混合 DNA, 微量检材
Turchi 等 <sup>[21]</sup>	87	ABI <sup>TM</sup> Ion	混合 DNA, 降解检材
Gandotra 等 <sup>[18]</sup>	90	Illumina <sup>TM</sup> MiSeq	个人识别, 溯源推断
Jin 等 <sup>[19]</sup>	18	Illumina <sup>TM</sup> HiSeq X-10	个人识别
Pang 等 <sup>[22]</sup>	124	Illumina <sup>TM</sup> MiSeq	个人识别

均可有效补充传统 STR 的不足, 对复杂混合 DNA 分析、降解或微量检材检测以及复杂亲缘关系鉴定具有独特优势。然而目前微单倍型的法医学应用仍面临诸多挑战: ①缺乏群体遗传数据。法医学遗传标记的应用必须基于可靠的群体遗传数据, 研究证明微单倍型的多态性在不同人群中具有差异, 其群体遗传数据目前仍旧匮乏, 实际检案的法医学参数计算需今后逐步完善相关群体遗传学数据。②检测成本较高。二代测序是微单倍型的主流检测方法, 但二代测序平台在法医 DNA 实验室尚不普及, 单次检测的成本高于传统 STR 自动分型技术, 且无相应技术标准。③位点筛选条件有待优化。目前有较多微单倍型被筛选, 但构建复合扩增体系时还需考虑位点检测不平衡问题, 需要不断优化微单倍型的筛选策略, 调整扩增体系。

微单倍型作为新型遗传标记具有独特优势, 相信随着以上三方面问题逐渐被解决, 微单倍型能更好发挥其独特优势, 在法医 DNA 鉴定中得以广泛应用, 并在复杂混合 DNA 分析、降解或微量检材检测等疑难复杂问题中体现应用价值, 为疑难案件的侦破提供线索。

#### [参考文献]

- [1] KIDD K K, PAKSTIS A J, SPEED W C, et al. Microhaplotype loci are a powerful new type of forensic marker[J]. Forensic Sci Int Genet Supplement Series, 2013, 4(1): 123-124
- [2] OLDONI F, KIDD K K, PODINI D. Microhaplotypes in forensic genetics[J]. Forensic Sci Int Genet, 2019, 38: 54-69
- [3] CHEN P, ZHU W, TONG F, et al. Identifying novel micro-

- haplotypes for ancestry inference [J]. *Int J Legal Med*, 2019, 133(4):983-988
- [4] KIDD K K, SPEED W C, PAKSTIS A J, et al. Evaluating 130 microhaplotypes across a global set of 83 populations [J]. *Forensic Sci Int Genet*, 2017, 29:29-37
- [5] BULBUL O, PAKSTIS A J, SOUNDARARAJAN U, et al. Ancestry inference of 96 population samples using microhaplotypes [J]. *Int J Legal Med*, 2018, 132(3):703-711
- [6] TURCHI C, MELCHIONDA F, PESARESI M, et al. Performance of a massive parallel sequencing microhaplotypes assay on degraded DNA [J]. *Forensic Sci Int Genet Supplement Series*, 2019, 7(1):782-783
- [7] KIDD K K, PAKSTIS A J, SPEED W C, et al. Current sequencing technology makes microhaplotypes a powerful new type of genetic marker for forensics [J]. *Forensic Sci Int Genet*, 2014, 12:215-224
- [8] CHEN P, DENG C, LI Z, et al. A microhaplotypes panel for massively parallel sequencing analysis of DNA mixtures [J]. *Forensic Sci Int Genet*, 2019, 40:140-149
- [9] CHEN P, YIN C, LI Z, et al. Evaluation of the Microhaplotypes panel for DNA mixture analyses [J]. *Forensic Sci Int Genet*, 2018, 35:149-155
- [10] OLDONI F, BADER D, FANTINATO C, et al. A massively parallel sequencing assay of microhaplotypes for mixture deconvolution [J]. *Forensic Sci Int Genet Supplement Series*, 2019, 7(1):522-524
- [11] DE LA PUENTE M, PHILLIPS C, XAVIER C, et al. Building a custom large-scale panel of novel microhaplotypes for forensic identification using MiSeq and Ion S5 massively parallel sequencing systems [J]. *Forensic Sci Int Genet*, 2020, 45:102213
- [12] ZHU J, CHEN P, QU S, et al. Evaluation of the microhaplotype markers in kinship analysis [J]. *Electrophoresis*, 2019, 40(7):1091-1095
- [13] SUN S, LIU Y, LI J, et al. Development and application of a nonbinary SNP-based microhaplotype panel for paternity testing involving close relatives [J]. *Forensic Sci Int Genet*, 2020, 46:102255
- [14] HIROAKI N, KOJI F, TETSUSHI K, et al. Approaches for identifying multiple-SNP haplotype blocks for use in human identification [J]. *Leg Med (Tokyo)*, 2015, 17(5):415-420
- [15] VOSKOBOINIK L, MOTRO U, DARVASI A. Facilitating complex DNA mixture interpretation by sequencing highly polymorphic haplotypes [J]. *Forensic Sci Int Genet*, 2018, 35:136-140
- [16] VAN DER GAAG K J, DE LEEUW R H, LAROS J F J, et al. Short hypervariable microhaplotypes: a novel set of very short high discriminating power loci without stutter artefacts [J]. *Forensic Sci Int Genet*, 2018, 35:169-175
- [17] ZHU J, LV M, ZHOU N, et al. Genotyping polymorphic microhaplotype markers through the Illumina (R) MiSeq platform for forensics [J]. *Forensic Sci Int Genet*, 2019, 39:1-7
- [18] GANDOTRA N, SPEED W C, QIN W, et al. Validation of novel forensic DNA markers using multiplex microhaplotype sequencing [J]. *Forensic Sci Int Genet*, 2020, 47:102275
- [19] JIN X Y, CUI W, CHEN C, et al. Developing and population analysis of a new multiplex panel of 18 microhaplotypes and compound markers using next generation sequencing and its application in the Shaanxi Han population [J]. *Electrophoresis*, 2020. DOI:10.1002/elps.201900451
- [20] KURESHI A, LI J, WEN D, et al. Construction and forensic application of 20 highly polymorphic microhaplotypes [J]. *R Soc Open Sci*, 2020, 7(5):191937
- [21] TURCHI C, MELCHIONDA F, PESARESI M, et al. Evaluation of a microhaplotypes panel for forensic genetics using massive parallel sequencing technology [J]. *Forensic Sci Int Genet*, 2019, 41:120-127
- [22] PANG J B, RAO M, CHEN Q F, et al. A 124-plex microhaplotype panel based on next-generation sequencing developed for forensic applications [J]. *Sci Rep*, 2020, 10(1):1945

[收稿日期] 2020-02-29