



微单倍型遗传标记的研究进展及在动物保护上的应用潜力

谢大鑫¹, 许蔚^{2,3}, 沈富军^{2,3*}, 张修月^{1,3*}

1. 四川大学生命科学学院, 生物资源与生态环境教育部重点实验室, 成都, 610065;
2. 成都大熊猫繁育研究基地, 成都, 610081;
3. 四川省濒危野生动物保护生物学重点实验室, 成都, 610081)

稿件运行过程

收稿日期: 2023-03-12
修回日期: 2023-04-09



关键词: 微单倍型;
遗传标记;
发展演变;
动物学应用潜力

Key words: Microhaplotype;
Genetic markers;
Evolution;
Application potential of zoology

中图分类号: Q75

文献标识码: A

文章编号:

2310-1490(2024)-01-0193-11

DOI: 10.12375/ysdwx.20240124

摘要

新型遗传标记微单倍型, 凭借高稳定性、无“stutter”峰干扰、高多态性、短扩增片段和 SNP 位点需求量少等显著优势逐渐受到法医物证学、种群遗传学等领域研究人员的关注, 并在亲缘关系分析与亲子鉴定方面得到较好应用。本文逐一综述了微单倍型的发展、命名规则与筛选标准等, 回顾了微单倍型数据获取、分析与检测的研究进展, 首次对微卫星、SNP 和微单倍型 3 种遗传标记进行了全面比较, 并介绍了微单倍型遗传标记在动物学领域的应用潜力及面临的挑战与机遇, 期冀为相关研究提供参考。

Advances on Microhaplotype Genetic Markers and Application Potentials in Conservation of Animals

XIE Daxin¹, XU Wei^{2,3}, SHEN Fujun^{2,3*}, ZHANG Xiuyue^{1,3*}

1. Key Laboratory of Bio-resource and Eco-environment of Ministry of Education, College of Life Sciences, Sichuan University, Chengdu, 610065, China;
2. Chengdu Research Base of Giant Panda Breeding, Chengdu, 610081, China;
3. Sichuan Key Laboratory of Endangered Wildlife Conservation Biology, Chengdu, 610081, China)

基金项目: 四川省自然科学基金重点项目(2022NSFSC0009)

第一作者简介: 谢大鑫(1997—), 男, 硕士研究生; 主要从事野生动植物保护与资源利用研究。E-mail: 2581565668@qq.com

* 通信作者: 沈富军, E-mail: shenfj@panda.org.cn; 张修月, E-mail: zhangxiuyue@scu.edu.cn

Abstract: The microhaplotype, as a kind of new genetic marker, has gradually attracted the attention of researchers in the fields of forensic genetics and population genetics in capable of their high stability, no interference from “stutter” peaks, high polymorphisms, short amplification fragments and fewer SNP locus requirements, and has been well applied in kinship analysis and paternity testing. This paper reviews the developments, nomenclature and screening criteria of the microhaplotypes, retrospects the progress of data acquisition, analysis and detection, comprehensively compares the microhaplotypes with the microsatellites and SNPs at first, and introduces the application potentials, challenges and opportunities of the microhaplotype genetic markers in zoology, with the hope of providing a reference for related research.

微单倍型(microhaplotype, MH)作为一种新型遗传标记,指在较短DNA片段范围内,拥有2个及以上高度紧密连锁单核苷酸多态性(single nucleotide polymorphism, SNP)位点的基因座^[1]。在过去的5年间,由于微单倍型遗传标记优势显著,众多实验团队开展了相关研究,相关论文数量不断增加,也促进了高多态性微单倍型位点数量的日益增长^[2],并在法医物证学领域得到了较多应用。在法医物证学领域,随着许多高多态性微单倍型位点的开发,在个体识别、祖源推断和亲缘关系分析等方面的应用得到了突破性进展^[3-6]。在动物学领域,国内外研究人员利用微单倍型遗传标记也进行了部分研究^[7-9],但罕

有文章系统总结现有研究成果和存在的问题。本文基于微单倍型的发展演变、研究进展等,介绍微单倍型在动物学领域的应用潜力,并对微卫星(microsatellite)、SNP和微单倍型3种不同遗传标记进行系统总结,突出微单倍型标记的应用优势并对其面临的挑战与将来的发展进行了展望。

1 微单倍型的发展演变

伴随基因组测序技术的进步,微单倍型遗传标记由诞生走向成熟应用,数十年的发展史可总结为图1。

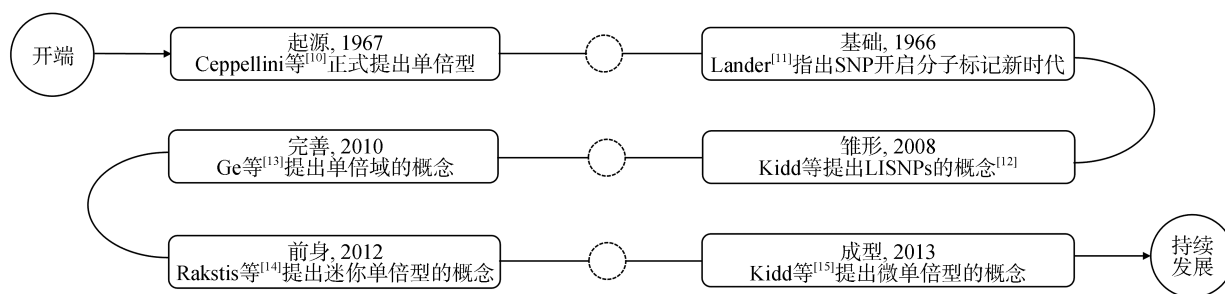


图1 微单倍型发展演变

Fig. 1 Development and evolution of microhaplotypes

1.1 单倍型

单倍型(haplotype)指的是同一染色体上共同遗传的多基因座上等位基因的组合,最早由Ceppellini等^[10]于1967年在第三届国际组织相容研讨会中正式提出,用以描述抗原抗体反应共同涉及的遗传基因位点,并由世界卫生组织将其命名为人类白细胞抗原(human leukocyte antigen, HLA)^[16]。发展至2001年,人类基因组全序列信息的公布揭露了人类基因组存在单倍型结构的事实^[17],研究人员对单倍

型应用的探索也日渐深入。

1.2 单核苷酸多态性

SNP是指基因组水平上由于单个核苷酸的变异所引起的DNA序列多态性,凭借遗传稳定性高、位点丰富且分布广泛等优点,在农学、医学等领域得到了众多应用^[18],Lander^[11]曾于1996年指出SNP标记开启了分子标记新时代,是继微卫星标记后的第三代分子标记技术。由于二等位基因SNP占据绝大多数,因此大多SNP标记多态性也较低^[19],但不可否

认, SNP的出现为微单倍型的诞生奠定了基础。

1.3 谱系信息 SNPs

Kidd在2007年的国际法医遗传学会(International Society of Forensic Genetics, ISFG)会议中提出LISNPs(lineage informative SNPs, LISNPs)的概念, 将一组高度连锁的SNPs作为一个具备多等位基因的遗传标记^[12]。该标记的提出是对高度连锁SNPs所包含遗传信息的进一步挖掘, 标志着SNP标记迈向微单倍型标记的第一步。

1.4 单倍型域

2010年, Ge等^[13]基于前人研究基础, 提出单倍型域(haplotype block), 它是一组包含强连锁不平衡(linkage disequilibrium, LD)且紧密相邻的SNP集, 并成功筛选出24个单倍型域作为法医学备用位点。经Ge等^[13]研究证明, 以单倍型域作为遗传标记较单个SNP的鉴别能力会得到进一步提升。

1.5 迷你单倍型

2012年, Kidd团队提议将一组片段长度小于10 kb且包含3个及以上SNP的单倍型遗传标记, 命名为迷你单倍型(mini-haplotype)^[14], 并强调该遗传标记应处于非重组热区且标记包含的SNPs应具备一定的多态性。作为微单倍型的前身, 该遗传标记在亲权鉴定、祖源推断方面具备一定的研究价值, 但其分型受检测条件的限制, 因此未能得到广泛应用^[20]。

1.6 微单倍型

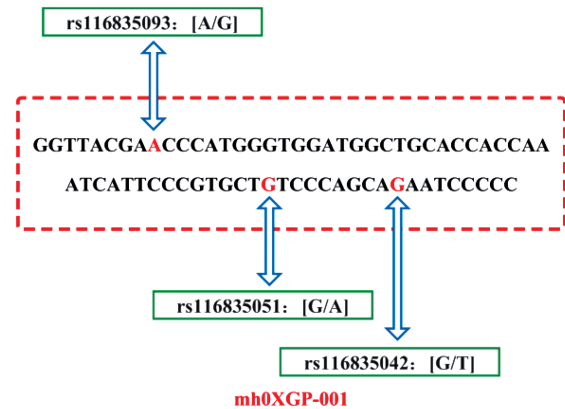
伴随高通量测序技术的发展与普及, 2013年Kidd团队^[15]基于前期研究基础, 将片段长度在200 bp内, 拥有两个及以上高度紧密连锁SNP位点的多态性基因座定义为微单倍型, 示例如图2。该遗传标记克服了微卫星标记扩增子会出现“stutter”峰、突变率较高和SNP标记多态性较低的显著缺点, 兼具扩增片段短、适用于高通量测序技术的优点, 首先受到了法医学领域研究人员的关注, 在个体识别、祖源推断等方面得到了相应应用^[3, 21], 同时在动物学领域也具备较佳的应用前景^[9, 22]。

2 微单倍型的命名规则与筛选标准

2.1 命名规则

作为一种新型遗传标记, 将命名规则标准化是非常必要的。科学文献的标准程序是在早期定义一个短字符, 如SNP代表单核苷酸多态性。不同实验室可能会采取不同的短字符定义同一个遗传标记,

这将给科研人员交互引用公共数据库中的遗传标记带来困扰。因此Kidd等^[23]建议以mh/MH(微单倍型的英文缩写)作为标记前缀, 紧跟染色体编号, 再拼接各研究团队自定义的标志符号, 以“-”作为分割, 最后连接各团队开发的微单倍型序列号。命名逻辑如图3A所示, 该命名规则因简洁明了的优点已被多数实验团队认可。



紧密连锁的3个SNP位点, rs116835093, rs116835051和rs116835042组合而成微单倍型基因座mh0XGP-001。

The microhaplotype locus mh0XGP-001 consists of three closely linked SNP loci rs116835093, rs116835051, and rs116835042.

图2 微单倍型基因座示例

Fig. 2 Example of a microhaplotype locus

Zhu等^[24]肯定了Kidd团队命名法的实用性, 并指出微单倍型标记中的SNPs可能在不同研究团队或深入研究中发生变化而出现无法区分不同微单倍型的缺陷, 如为获得更高多态性, 需增加新的SNPs到候选微单倍型中, 而更改前后的候选微单倍型, 无法凭借Kidd团队的命名法区分。Zhu等^[24]建议在Kidd团队命名法的基础上, 附加SNP集的版本号。该建议进一步提升了Kidd团体命名法的适用性, 命名逻辑如图3B所示。

此外, 陈鹏等^[25]设计出一套有别于Kidd团队的命名规则。命名方法为所处研究团队标志符号+染色体编号+染色体英文首字母+微单倍型编号+SNP集序列字母。该命名规则较Kidd命名方式稍显繁琐, 字符分割界限不清, 暂未得到大规模应用, 命名逻辑如图3C所示。

2.2 筛选标准

在群体遗传学中, 一个基因座上同等频率的中性等位基因数被定义为有效等位基因数(effective number of alleles, A_e)^[26-27], A_e 能够将每个微单倍型基

因座转化为相同“标准”从而进行量化比较,达到对微单倍型检测能力定性的目的。Kidd团队^[27]研究发现, $A_c > 3$ 的微单倍型具有法医学领域实践潜力,可用于DNA混合物分析。Zhu等^[28]也提出了3个标准用于微单倍型的筛选:(1)微单倍型内任一SNP的次要等位基因频率(minor allele frequency, MAF)应大于0.05;(2)一个微单倍型遗传标记至少应包含3个SNP,具有较多SNP的微单倍型的多态性通常较高;(3)所有SNPs必须位于200 bp的区域片段内,200 bp的序列长度是大规模并行测序(massively parallel sequencing, MPS)平台的常规测序长度。总体而言,针对法医学或动物学实践案例,提取的降解DNA样本

片段长度是有限的,同时还要考虑遗传标记的多态性,应利用含较多SNP且片段长度较短的微单倍型。

Oldoni等^[29]认为随着测序技术的进步,在排除基因座内具备更高交换频率的位点后,即排除重组热点后,某些微单倍型的读取长度可放宽至300 bp,以便在扩增子内获取更多SNPs。信息度(informativeness, I_n)量化了群体之间等位基因频率的变化^[29]。较高的信息度表明微单倍型具备受试群体常见的等位基因,因此Oldoni等^[29]强调理想的微单倍型同时应该具备较高的 A_c 与 I_n 值。该表述与刘聪等^[30]认为微单倍型用于受试样本群体中,等位基因频率分布应差异显著说法一致。



图3 不同实验团队定义的微单倍型命名规则^[23-25]

Fig. 3 Nomenclature for microhaplotypes as defined by different experimental teams^[23-25]

3 微单倍型的数据获取、分析方法与检测手段

3.1 微单倍型数据库

耶鲁大学医学院开发的等位基因频率数据库(ALFRED, <http://alfred.med.yale.edu>)是遗传学领域应用较为广泛的微单倍型数据库^[31],该数据库旨在为人类基因多样性及分子人类学相关的科学研究服务,收录数据来源广泛且均经过严谨审核。截至2023年1月,共计收录664 708个多态性标记、762个人群和66 726 252个频率表,其中来源于22对常染

色体的微单倍型标记共计198个。数据库中展示的微单倍型基础信息包括微单倍型名称、数据库编号、微单倍型基因座的SNP组成及rs编号、片段长度、SNP的染色体位置、等位基因的分型及相关参考文献信息^[32]。该数据库因检索便捷、信息详尽、数据全面且实时更新等优点深受科研人员的认可。

美国国家生物法医分析中心的Standage团队^[33]认为现有数据库信息不够详尽,无法快捷检索不同群体微单倍型的等位基因频率,因此开发并发布了名为MicroHapDB的数据库(<https://osf.io/gr7h6/>)。该数据库收录了2020年2月前发表的绝大多数微单

倍型数据,并展示了26个全球样本集群的微单倍型等位基因频率,旨在推进基于微单倍型的人类取证能力。该数据库除初始安装及版本更新外,无需再通过网络访问,操作便捷可自由调取至任何分析环境。同时,该数据库虽然支持用户自行审查错误数据,因缺少权威监管,无法动态更新。总体而言,这是一个使用便捷但难以与时俱进的数据库,使用较为受限。

Xue等^[34]认为目前数据库中所记载的微单倍型,均是源于特定目的或特殊人群。研究人员仅能选择已发表微单倍型标记用于实验,但所选标记往往不能满足实验需求,因此开发出名为D-SNPsDB的数据库(https://1drv.ms/u/s!AnbHu1NQuWvTgRH8sKil_CCIfy8J?e=ddhhsj)。库中以2-SNP微单倍型为主,旨在为混合DNA解析、亲缘关系分析和个体识别服务。该数据库中2-SNP微单倍型基础信息全面,安装完成后的数据库无需再通过网络访问,通过R软件提示可自动启动,检索界面简洁,使用方式简便,整体功能全面,但数据动态更新性同样较差,该数据库在完成与dbSNP数据库和人类在线孟德尔遗传数据库(online mendelian inheritance in man, OMIM)的动态连接后,能为各科学领域的研究人员提供更便捷的服务。

3.2 微单倍型数据调用及分析软件

微单倍型数据调用与分析软件不仅有助于研究人员在不同需求环境便捷调用不同微单倍型遗传标记,还有助于直观比较各遗传标记或体系的适用性。Baetscher等^[22]开发的MICROHAPLOT,是一款经由R包和相关Shiny应用实现的程序,可满足使用者对微单倍型读取深度的需求。预先过滤不达标微单倍型,再便捷导入含微单倍型的扩增子数据,从而合理调用微单倍型位点。Zhu等^[35]开发了一款名为FLfinder的软件用于微单倍型分型。经人工验证导出的基因分型结果准确无误,但该软件允许插入、缺失和突变带来的序列错配,允许错配的默认设置可能增加出错的概率,如何优化算法避免错配仍是该软件亟待解决的问题。Li等^[36]基于FLfinder的技术基础,开发出一款名为MHanalyser的软件。该软件克服了FLfinder只能对单个样本进行微单倍型数据分析的缺陷,可同时对多个样本和多个位点进行数据分析。

伴随二代测序技术(next generation sequencing,

NGS)的发展,Zhang等^[37]开发了MHTyper软件。这是一款用于微单倍型等位基因分型、测序深度统计且仅适用于Linux/Unix系统的流程化脚本。在针对150个微单倍型位点的测试实验中,MHTyper软件报告的基因型与生物信息学软件Integrative Genomics Viewer分析的可视化基因结果一致,证明了MHTyper软件可作为二代测序技术后续数据分析的有效工具。

3.3 微单倍型检测方法

由于微单倍型和SNP都属于序列多态性遗传标记,但微单倍型遗传标记包含的多个SNP难以转换为长度多态性,因此法医物证学领域常用的毛细管电泳法无法推广应用于微单倍型的检测分型。基于该科研难点,众多实验团队已开发出诸如一代测序、二代测序和核酸质谱法等检测手段,并逐渐得到应用。

3.3.1 一代测序

一代测序即Sanger测序,代表的是DNA测序的金标准^[38],又名DNA末端终止法测序技术。该技术利用了DNA复制的原理,由于双脱氧核苷三磷酸(dideoxy ribonucleoside triphosphate, ddNTP)可中止反应过程中DNA链的延伸,在电泳平台分离产物后,便可按顺序读取相应DNA序列。随着科技发展,该测序技术渐渐无法满足更精准的科研需要,测序时如遇两个及以上位点杂合时,则无法从基因组DNA中确定单个SNP等位基因之间的顺反位置关系,进而影响分型判断^[3]。但该技术读取的原始数据质量较好,在病毒的基因组测序以及PCR产物测序中仍发挥着重要作用^[39]。

3.3.2 二代测序

二代测序技术又称下一代测序或大规模并行测序,相较于传统的一代测序,该技术可以边合成边测序每条链的等位基因组合,从而克服了一代测序无法判断等位基因顺反位置关系的障碍,并能获取复合体系中SNP位点的全部基因分型,达到对微单倍型遗传标记准确分型的目的^[25,40]。虽然该技术具备高通量、集成化等显著优势,但较昂贵的检测设备与检测费用,不利于二代测序技术的推广使用。同时,大多数微单倍型的研究都需要使用到复杂的基因组分析工具如GATK、Ion Torrent等软件^[41],简单开源工具的开发仍是尚未解决的难题。

3.3.3 核酸质谱法

质谱技术具有高准确性、高灵敏度、高分辨率和高通量等优点,在近30年间发展迅速。基质辅助激光解吸电离飞行时间质谱(matrix assisted laser desorption ionization time of flight mass spectrometry, MALDI-TOF-MS)是20世纪末期发展起来的一种新型软电离生物质谱技术^[42]。该技术只需将样品与芯片基质共价结合形成结晶后经质谱仪的激光激发,核酸分子就会解吸附为单电荷离子,其飞行时间与离子的质量成反比,再通过检测核酸分子的飞行时间可判断样品分子质量,进而得到样品的基因型信息。质谱法分离速度和分辨率均远超于以凝胶电泳为基础的测序方法。同时,针对相同长度但序列不同的DNA片段,质谱法可以精准分辨碱基种类^[43],因此核酸质谱法常使用在微单倍型的SNP位点检测中。但较为昂贵的仪器设备、无法检测的未知突变、对检测样本及位点数的要求,一定程度上桎梏着核酸质谱技术的应用。

4 微卫星、单核苷酸多态性和微单倍型遗传标记的比较

微卫星是指以1~6个核苷酸为单元,多次串联重复的DNA序列,又称短串连重复序列(short tandem repeats, STRs)或简单重复序列(simple sequence repeats, SSRs)^[44],近20年来,该标记因多态性信息含量丰富、共显性遗传等优势,广泛应用于种群遗传多样性、种群遗传结构分析等领域^[45-46],为濒危物种的保护遗传学研究做出了突出贡献。SNP和MH的应用优势在前文已有简述,3种标记应用至不同领域也各有千秋。

STR在基因组中的分布约为6~10 kb/个^[47],单倍型和SNP数量则更为丰富,人类基因组中每1 000个核苷酸约有1个SNP^[48],李茜等^[49]研究发现平均每Mb分布约1 000个微单倍型(2 891 927/3 000 Mb)。

STRs标记在进行PCR扩增过程中,由于发生滑链效应常产生“stutter”峰,会产生比真实等位基因产物少或多一个重复单位的扩增子,影响后续分型分析及结果判断^[50],尤其是在不平衡混合DNA样本中,主要与次要贡献者等位基因峰高度一致时,难以进行准确分型。微单倍型的一个显著优势是不会产生“stutter”峰^[51],这明显降低了分型分析难度。与

STRs相比,微单倍型的另一个优点是一个基因座内的所有等位基因长度均一,而STRs基因座内的等位基因长度变化甚至会高达100 bp,如果样品DNA降解或其中含有PCR抑制剂,同一供体不同长度的等位基因表达程度不一,会导致混合物来源解析复杂化^[52]。

微单倍型拥有类同于STRs的高多态性,也保留了SNP突变率低的特点^[51],大多SNP标记仅有2个等位基因,如果 n 个二等位基因SNP构成一个微单倍型位点,理论上该标记最多有 2^n 种SNP的组合,即该标记甚至可拥有 2^n 个等位基因^[53]。遗传标记在某个群体中的等位基因数目越多,表示该标记多态性越好,个体识别能力也越高^[54]。Chen等^[4-5]基于千人基因组数据库筛选25个微单倍型,发现在汉族人群中平均杂合度达0.617 8,但累积的随机匹配概率(random matching probability, RMP)只约等于15个STRs基因座的累积识别能力,即为达到与STRs相同鉴别能力的目的,需要更多微单倍型的参与。

微卫星一般采用毛细管电泳(capillary electrophoresis, CE)技术检测。针对SNP的分型技术繁多,测序分型(genotyping-in-thousands by sequencing, GT-seq)是一种对多重PCR产物进行二代测序的技术,可同时对大量SNP位点测序,其读取深度与分型成本效益相对较高,近几年来发展迅速^[55]。经Hayward团队^[56]优化选择的324个SNP位点构成的GT-seq体系,即可用于区分加拿大北极熊(*Ursus maritimus*)个体并能表征该地区北极熊种群的遗传结构。Schmidt团队^[57]借助于362个SNP标记构建的GT-seq体系,明确了西部响尾蛇(*Crotalus oreganus*)的遗传多样性分布与不同地区的遗传分化程度,有利于人类从生物学的角度更好地维持西部响尾蛇种群的生存。此外,GT-seq技术在大眼狮鲈(*Sander vitreus*)遗传种群鉴定、辣椒(*Capsicum* spp.)遗传育种领域也有着突出表现^[58-59]。诚然,该技术的优势明显,但对SNP位点数的过高需求量,会导致多重PCR体系所需引物数量众多,过多的引物数反而会干扰基因分型效率。优化引物组合,减少引物组内引物间的扩增差异,也难以显著提升基因分型率。微单倍型的检测大多也采取对多重PCR产物进行二代测序的方式,该技术无论在单一还是混合DNA样本中均可获得微单倍型等位基因,相较于CE技术,二代测序的分型效率更高,但较昂贵的设备成本且复杂的分析过

程,在一定程度上阻碍了微单倍型的应用^[60]。相较于GT-seq技术在SNP上的应用,微单倍型位点多态性更高,对SNP位点及多重PCR体系混合引物组,引物的需求量更少,将GT-seq技术应用于微单倍型位点上,其基因分型效率与效益都更为优异。

从整体研究成本分析,常规SNP标记多态性低于STR标记,常规微单倍型多态性介于SNP与STR之间。为达到相同鉴别能力,SNP标记所需数量最多,微单倍型其次,STR标记需求数量最少,同种分型方式,检测更多基因座的成本也将更高。此外,SNP分型时构建的多重PCR体系所需引物数量庞大,引物合成及体系构建耗资不菲^[61],也是科学研究面临的巨大挑战。综上,利用STR标记从事相关研究的成本较低,效益较高;利用微单倍型从事相关研究的成本适中;利用SNP从事相关研究的成本较高。这也是在法医学领域,STR标记无法被完全取代的重要原因。

3种遗传标记的特点比较如表1所示,总体而言,微单倍型作为分布在整个基因组中的单拷贝多SNP基因座,较常规STR、SNP标记更具优势和发展潜力。

表1 3种遗传标记的特点^[4-5,47-61]

Tab. 1 Characteristics of three kinds of genetic markers^[4-5,47-61]

特点 Features	微卫星 STR	单核苷酸多态性 SNP	微单倍型 MH
基因组分布 Genomic distribution	6~10 kb/个	约1 000个/Mb	约1 000个/Mb
滑链效应 Slipped-strand mispairing	有	无	无
突变率 Mutation rate	较高	较低	较低
多态性 Polymorphism	高	低	较高
等位基因数 Number of alleles	一般较多	多为2	≥2且≤2 ⁹
主要分型技术 Major typing techniques	CE	种类繁多	MPS
研究成本 Research costs	较低	较高	介于两者之间

5 微单倍型在动物保护遗传学上的应用潜力

5.1 探究种群遗传结构

微单倍型包含紧密连锁位点的连锁不平衡信息,有助于提高遗传祖源推断研究的分辨率,结合对

微单倍型的分析,提高了种群遗传结构的分辨率。早有利用微卫星分析补充放养对鲑科(Salmonidae)鱼群遗传结构模式产生何种影响的文献报道^[62],但基因组工具很少用于测试放养对鲑科以外鱼类遗传结构的影响^[63]。Bootsma等^[64]创新性地使用微单倍型标记对来自23个种群的954只大眼狮鲈约20 000个位点进行了基因分型,并分别利用SNPs和MHs分析了大眼狮鲈的种群遗传结构,研究发现,微单倍型数据反应了不同种群的细微差异,并捕捉到了SNPs和微卫星没有发现的种群遗传结构差异信息,表明微单倍型对种群遗传结构分析更加精细。

5.2 亲缘关系分析与亲子鉴定

Baetscher等^[7]利用96个微单倍型基因座组成的MHs体系对超过14 500条鲈科(Scorpaenidae)鱼类进行了基因分型,根据参考数据库的等位基因频率区分了不同种的鲈科鱼类,再通过CKMRsim程序完成了6 000多条墨绿平鲈(*Sebastes atrovirens*)的亲缘关系分析与亲子鉴定。该研究还利用微卫星基因型结合CERVUS^[65]、COLONY^[66]软件验证谱系关系鉴定的有效性,结果表明,8对亲子关系置信度高于95%,成功识别25对全同胞关系,微单倍型分析结果与微卫星验证结果整体保持高度一致性,揭示了微单倍型应用于动物种群亲缘关系分析与亲子鉴定的巨大发展潜力。此外,高多态性微单倍型体系还能对低遗传分化的物种进行遗传资源鉴定。Baetscher等^[8]凭借构建的MHs体系成功区分来自4个不同繁殖地的暴风鹱(*Fulmarus glacialis rogersii*)种群,该结果表明,对于种群规模较大的物种来说,微单倍型较其他遗传标记鉴别能力更具优势。

夏雷等^[9]首次利用微单倍型对一个草鱼(*Ctenopharyngodon idellus*)群体进行了亲子鉴定,其研究采用HapCUT2软件高效完成了亲代微单倍型的组装、子代微单倍型的分型,最后使用CERVUS软件进行亲子鉴定分析。结果显示,3个和5个微单倍型标记亲子鉴定结果与微卫星鉴定结果一致性分别达97.08%、99.42%,证明其建立的微单倍型筛选标准与亲子鉴定分析流程的可行性,也为其他动物群体的亲子鉴定工作提供了新思路。

综上,由于微单倍型遗传标记具有多态信息含量(polymorphic information content, PIC)丰富的优势,能够在一次测序中对多个个体进行基因分型,对比从DNA序列数据中直接调用SNPs,需要更少的微

单倍型位点就可达到相同的目的。根据 Baetscher 等^[22]在墨绿平鲈样本数据集的模拟分析中,发现对于 96 个 MH 组成的 MHs 体系,假阴性率(false-negative rate, FNR)为 0.01,假阳性率(false-positive rates, FPR)为 0.065,这意味着在 20 个非同胞的鉴定比较中,有 1 个以上的结果为阳性。对比 96 个基因座中多态信息含量排名前 96 的 SNP 集,假阳性率为 0.01,假阴性率为 0.44,证实 96 个 SNP 组成的 SNP 集几乎完全缺乏区分同父异母的半同胞和无关个体的能力^[22],进而表明在更具挑战性的大规模动物群体亲缘关系分析与亲子鉴定中,微单倍型的鉴别优势更为突出。

6 微单倍型面临的挑战及展望

6.1 面临的挑战

微单倍型作为一种新型遗传标记正以其独有优势逐渐受到多领域研究人员的关注,并在混合 DNA 分析、亲缘关系判断和亲子鉴定等方面取得了一定成果,但微单倍型在动物学领域的相关应用仍面临着不小的挑战:(1)检测成本较高,学者们大多认可二代测序将会成为微单倍型检测的金标准,但构建二代测序平台所需费用昂贵,单次检测成本较高,不适用于少量样本多次检测的实验方式。(2)缺乏动物学领域微单倍型数据库,现有的 ALFRED、Micro-HapDB、D-SNPsDB 数据库均是为了便捷法医物证学领域的研究人员调用微单倍型而构建,难以用于动物群体的个体识别、亲子鉴定。完善动物学领域微单倍型数据库,实时收录高多态性微单倍型标记,更有利于微单倍型在动物学领域得到广泛应用。(3)遗传标记筛选条件有待优化,可供选择的微单倍型数量丰富,但如何构建更实用的复合微单倍型体系,排除位点检测不平衡的干扰,仍需不断优化筛选条件,调整微单倍型体系。

6.2 应用前景展望

目前,直接利用分子遗传标记对野生动物群体进行的分析鉴定仍旧匮乏,野外往往仅能获取野生动物的毛发、粪便等非损伤性实验材料,这些材料大多已降解,常规遗传标记难有完美鉴定效果,而高稳定性、无“stutter”峰干扰、高多态性、短扩增片段和 SNP 位点需求量少的微单倍型优势显著。相信随着研究的深入,作为传统遗传标记的深化研究,微单倍型标记定将为濒危动物的保护贡献更多力量。

参考文献:

- [1] 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 Science International: Genetics*, 2014, 12: 215–224.
- [2] KIDD K K, PAKSTIS A J. State of the art for microhaplotypes [J]. *Genes*, 2022, 13(8): 1322.
- [3] KIDD K K, SPEED W C, PAKSTIS A J, *et al.* Evaluating 130 microhaplotypes across a global set of 83 populations[J]. *Forensic Science International: Genetics*, 2017, 29: 29–37.
- [4] CHEN P, YIN C Y, LI Z, *et al.* Evaluation of the microhaplotypes panel for DNA mixture analyses[J]. *Forensic Science International: Genetics*, 2018, 35: 149–155.
- [5] CHEN P, DENG C W, LI Z, *et al.* A microhaplotypes panel for massively parallel sequencing analysis of DNA mixtures [J]. *Forensic Science International: Genetics*, 2019, 40: 140–149.
- [6] ZHU J, CHEN P, QU S Q, *et al.* Evaluation of the microhaplotype markers in kinship analysis [J]. *Electrophoresis*, 2019, 40 (7): 1091–1095.
- [7] BAETSCHER D S, ANDERSON E C, GILBERT-HORVATH E A, *et al.* Dispersal of a nearshore marine fish connects marine reserves and adjacent fished areas along an open coast[J]. *Molecular Ecology*, 2019, 28(7): 1611–1623.
- [8] BAETSCHER D S, BECK J, ANDERSON E C, *et al.* Genetic assignment of fisheries bycatch reveals disproportionate mortality among Alaska Northern Fulmar breeding colonies [J]. *Evolutionary Applications*, 2022, 15(3): 447–458.
- [9] 夏雷, 石米娟, 张婉婷, 等. 基于微单倍型分子标记的草鱼亲子鉴定方法[J]. *水生生物学报*, 2020, 44(3): 509–517.
XIA L, SHI M J, ZHANG W T, *et al.* A method for paternity testing of grass carp (*Ctenopharyngodon idellus*) using microhaplotypes[J]. *Acta Hydrobiologica Sinica*, 2020, 44(3): 509–517.
- [10] CEPPELLINI R, CURTONI E S, MATTIUIZ P L, *et al.* Genetics of leukocyte antigens: a family study of segregation and linkage[C]//CURTONI E S, MATTIUIZ P L, TOSI R M. *Histocompatibility testing*. Baltimore: Williams & Wilkins Company, 1967: 149–187.
- [11] LANDER E S. The new genomics: global views of biology [J]. *Science*, 1996, 274(5287): 536–539.
- [12] BUTLER J M, BUDOWLE B, GILL P, *et al.* Report on ISFG SNP panel discussion [J]. *Forensic Science International: Genetics Supplement Series*, 2008, 1(1): 471–472.
- [13] GE J Y, BUDOWLE B, PLANZ J V, *et al.* Haplotype block: a new type of forensic DNA markers [J]. *International Journal of Legal Medicine*, 2010, 124(5): 353–361.
- [14] PAKSTIS A J, FANG R X, FURTADO M R, *et al.* Mini-haplotypes as lineage informative SNPs and ancestry inference SNPs [J]. *European Journal of Human Genetics*, 2012, 20

- (11): 1148–1154.
- [15] KIDD K K, PAKSTIS A J, SPEED W C, *et al.* Microhaplotype loci are a powerful new type of forensic marker[J]. *Forensic Science International: Genetics Supplement Series*, 2013, 4(1): e123–e124.
- [16] TERASAKI P I. 器官移植体液免疫理论的100年历史回顾[J]. 周永昌,译. *中华细胞与干细胞杂志(电子版)*, 2012, 2(2): 82–89.
TERASAKI P I. A 100-year historical review of the theory of humoral immunity for organ transplantation[J]. ZHOU Y C, Trans. *Chinese Journal of Cell and Stem Cell (Electronic Edition)*, 2012, 2(2): 82–89.
- [17] DALY M J, RIOUX J D, SCHAFFNER S F, *et al.* High-resolution haplotype structure in the human genome [J]. *Nature Genetics*, 2001, 29(2): 229–232.
- [18] 唐立群, 肖层林, 王伟平. SNP分子标记的研究及其应用进展[J]. *中国农学通报*, 2012, 28(12): 154–158.
TANG L Q, XIAO C L, WANG W P. Research and application progress of SNP markers[J]. *Chinese Agricultural Science Bulletin*, 2012, 28(12): 154–158.
- [19] 阿丽耶·库热什. 20个微单倍型遗传标记的法医学应用研究[D]. 乌鲁木齐: 新疆医科大学, 2020.
KURESHI A. Research on forensic application of 20 microhaplotype genetic markers [D]. Urumqi: Xinjiang Medical University, 2020.
- [20] 周靖, 王艳, 徐恩萍. 微单倍型在法医遗传学中的研究进展[J]. *浙江大学学报(医学版)*, 2021, 50(6): 777–782.
ZHOU J, WANG Y, XU E P. Research progress on application of microhaplotype in forensic genetics [J]. *Journal of Zhejiang University (Medical Sciences)*, 2021, 50(6): 777–782.
- [21] WANG H, ZHU J, ZHOU N, *et al.* NGS technology makes microhaplotype a potential forensic marker [J]. *Forensic Science International: Genetics Supplement Series*, 2015, 5: e233–e234.
- [22] BAETSCHER D S, CLEMENTO A J, NG T C, *et al.* Microhaplotypes provide increased power from short-read DNA sequences for relationship inference [J]. *Molecular Ecology Resources*, 2018, 18(2): 296–305.
- [23] KIDD K K. Proposed nomenclature for microhaplotypes[J]. *Human Genomics*, 2016, 10(1): 16.
- [24] ZHU J, LONG B, QU S Q, *et al.* A new proposed nomenclature for microhaplotypes [J]. *Forensic Science International: Genetics Supplement Series*, 2019, 7(1): 813–815.
- [25] 陈鹏, 朱镜, 姜又菁, 等. 遗传标记微单倍型在法医学中的研究进展[J]. *中国法医学杂志*, 2017, 32(5): 484–487.
CHEN P, ZHU J, JIANG Y J, *et al.* Research progress of microhaplotype in forensics [J]. *Chinese Journal of Forensic Medicine*, 2017, 32(5): 484–487.
- [26] CROW J F, KIMURA M. An introduction to population genetics theory[M]. Caldwell: The Blackburn Press, 1970: 322–327.
- [27] KIDD K K, SPEED W C. Criteria for selecting microhaplotypes: mixture detection and deconvolution [J]. *Investigative Genetics*, 2015, 6: 1.
- [28] ZHU J, LÜ M L, ZHOU N, *et al.* Genotyping polymorphic microhaplotype markers through the Illumina® MiSeq platform for forensics [J]. *Forensic Science International: Genetics*, 2019, 39: 1–7.
- [29] OLDONI F, KIDD K K, PODINI D. Microhaplotypes in forensic genetics [J]. *Forensic Science International: Genetics*, 2019, 38: 54–69.
- [30] 刘聪, 肖超, 黄钰洁, 等. 新型连锁遗传标记的法医学研究进展[J]. *中国法医学杂志*, 2018, 33(3): 269–272.
LIU C, XIAO C, HUANG Y J, *et al.* Advances in studies on the new type of linkage genetic marker [J]. *Chinese Journal of Forensic Medicine*, 2018, 33(3): 269–272.
- [31] RAJEEVAN H, SOUNDARARAJAN U, KIDD J R, *et al.* ALFRED: an allele frequency resource for research and teaching [J]. *Nucleic Acids Research*, 2012, 40 (Database issue): D1010–D1015.
- [32] 于文帅, 康克莱, 张驰, 等. 微单倍型遗传标记研究新进展及其在法医混合DNA分析中的应用[J]. *沈阳药科大学学报*, 2023, 40(9): 1245–1252.
YU W S, KANG K L, ZHANG C, *et al.* Research progress in the application of microhaplotype genetic markers in forensic DNA mixture analysis [J]. *Journal of Shenyang Pharmaceutical University*, 2023, 40(9): 1245–1252.
- [33] STANDAGE D S, MITCHELL R N. MicroHapDB: a portable and extensible database of all published microhaplotype marker and frequency data [J]. *Frontiers in Genetics*, 2020, 11: 781.
- [34] XUE J M, QU S Q, TAN M Y, *et al.* An overview of SNP-SNP microhaplotypes in the 26 populations of the 1000 Genomes Project [J]. *International Journal of Legal Medicine*, 2022, 136(5): 1211–1226.
- [35] ZHU J, ZHOU N, JIANG Y J, *et al.* FLfinder: a novel software for the microhaplotype marker [J]. *Forensic Science International: Genetics Supplement Series*, 2015, 5: e622–e624.
- [36] LI Z, YANG J W, LIANG W B, *et al.* Application of MAnalyzer software in the study of microhaplotypes in forensics [J]. *Forensic Science International: Genetics Supplement Series*, 2019, 7(1): 271–273.
- [37] ZHANG C, CAO Y D, SONG J J, *et al.* MHTyper: a microhaplotype allele-calling pipeline for use with next generation sequencing data [J]. *Australian Journal of Forensic Sciences*, 2021, 53(3): 283–290.
- [38] BUTLER J M. Fundamentals of forensic DNA typing [M]. Amsterdam: Academic Press, 2009: 376–378.

- [39] 田李, 张颖, 赵云峰. 新一代测序技术的发展和应[J]. 生物技术通报, 2015, 31(11): 1-8.
TIAN L, ZHANG Y, ZHAO Y F. The next generation sequencing technology and its applications [J]. Biotechnology Bulletin, 2015, 31(11): 1-8.
- [40] 饶皎, 李彩霞, 赵钊, 等. 微单倍型遗传标记及其法医遗传学应用[J]. 刑事技术, 2017, 42(4): 324-328.
RAO M, LI C X, ZHAO Z, *et al.* Microhaplotypic genetic markers and their applications in forensic genetics [J]. Forensic Science and Technology, 2017, 42(4): 324-328.
- [41] KWON Y L, LEE E Y, KIM B M, *et al.* Application of a custom haplotype caller to analyze sequence-based data of 56 microhaplotypes [J]. Forensic Science International: Genetics, 2022, 61: 102778.
- [42] DEMIREV P A, HO Y P, RYZHOV V, *et al.* Microorganism identification by mass spectrometry and protein database searches [J]. Analytical Chemistry, 1999, 71(14): 2732-2738.
- [43] 李云峰, 马文晨, 方义亮, 等. 生物质谱技术在核酸检测中的应用[J]. 中国国境卫生检疫杂志, 2017, 40(2): 141-144.
LI Y F, MA W C, FANG Y L, *et al.* Application of mass spectrometry in detection of nucleic acids [J]. Chinese Journal of Frontier Health and Quarantine, 2017, 40(2): 141-144.
- [44] LITT M, LUTY J A. A hypervariable microsatellite revealed by *in vitro* amplification of a dinucleotide repeat within the cardiac muscle actin gene [J]. American Journal of Human Genetics, 1989, 44(3): 397-401.
- [45] BALLOUX F, LUGON-MOULIN N. The estimation of population differentiation with microsatellite markers [J]. Molecular Ecology, 2002, 11(2): 155-165.
- [46] COLLINS J R, STEPHENS R M, GOLD B, *et al.* An exhaustive DNA micro-satellite map of the human genome using high performance computing [J]. Genomics, 2003, 82(1): 10-19.
- [47] 李生斌, 孙瑞芳, 李欣, 等. 一种新的3个人类抗肌萎缩蛋白基因 STR 位点及其分型方法: CN101712997B[P/OL]. 2012-03-28 [2023-03-06]. https://kns.cnki.net/kcms2/article/abstract?v=ZPxNsP0ekIElMboj9cGV5q3X82tn5qjGO61GbEY8kj1Lzn1_ytQBpSs1vQaDozUUW8NKHhAuDqFskBZGEURcywQyc91dmsNn21u-iAh_JSl8ykAo4ae1RhIwr63jh0q&uniplatform=NZKPT&language=CHS.
LI S B, SUN R F, LI X, *et al.* A new three STR loci of the human anti-myosin gene and its typing methods: CN101712997B [P/OL]. 2012-03-28 [2023-03-06]. https://kns.cnki.net/kcms2/article/abstract?v=ZPxNsP0ekIElMboj9cGV5q3X82tn5qjGO61GbEY8kj1Lzn1_ytQBpSs1vQaDozUUW8NKHhAuDqFskBZGEURcywQyc91dmsNn21u-iAh_JSl8ykAo4ae1RhIwr63jh0q&uniplatform=NZKPT&language=CHS.
- [48] 杨泽, 孙亮, 屈彦纯. 一种鉴定 PGC-1 基因启动子-1436 位 SNP 分子标记的引物、方法及试剂盒: CN100368562C[P/OL]. 2008-02-13 [2023-03-06]. https://kns.cnki.net/kcms2/article/abstract?v=ZPxNsP0ekIHX_dKHhHhZ5m_ixZU_Pa4wTl-PZluuMrF7jvzM0yEoUJ7yuPLu1yAkMKwW336tfwRfz699zQbH-cjn32iYxtuHt3A07Jd6RceQZVMCRmqClhVvkqyXVRFZR4W0-wGNjX8c=&uniplatform=NZKPT&language=CHS.
- [49] 李茜, 王浩宇, 曹悦岩, 等. 微单倍型遗传标记的法医基因组学研究[J]. 遗传, 2021, 43(10): 962-971.
LI X, WANG H Y, CAO Y Y, *et al.* Forensic genomics research on microhaplotypes [J]. Hereditas (Beijing), 2021, 43(10): 962-971.
- [50] GILL P, HANED H, BLEKA O, *et al.* Genotyping and interpretation of STR-DNA: low-template, mixtures and database matches: twenty years of research and development [J]. Forensic Science International: Genetics, 2015, 18: 100-117.
- [51] OLDONI F, PODINI D. Forensic molecular biomarkers for mixture analysis [J]. Forensic Science International: Genetics, 2019, 41: 107-119.
- [52] 周晶. 法医个体识别微单倍型标记的筛选及NGS检测体系的建立[D]. 石家庄: 河北医科大学, 2019.
ZHOU J. Genome-wide screening for microhaplotype for forensic individual identification and development of NGS panel [D]. Shijiazhuang: Hebei Medical University, 2019.
- [53] 陈峰. DNA 微单倍型的研究现状、挑战与展望 [J]. 南京医科大学学报(自然科学版), 2020, 40(8): 1081-1084.
CHEN F. Research progress, challenge and prospect of DNA microhaplotype [J]. Journal of Nanjing Medical University (Natural Sciences), 2020, 40(8): 1081-1084.
- [54] 乔麦菊, 冉江洪, 张和民. 微卫星标记在大熊猫研究中的应用进展 [J]. 兽类学报, 2019, 39(1): 103-110.
QIAO M J, RAN J H, ZHANG H M. The application of microsatellite markers in giant panda research [J]. Acta Theriologica Sinica, 2019, 39(1): 103-110.
- [55] CAMPBELL N R, HARMON S A, NARUM S R. Genotyping-in-thousands by sequencing (GT-seq): a cost effective SNP genotyping method based on custom amplicon sequencing [J]. Molecular Ecology Resources, 2015, 15(4): 855-867.
- [56] HAYWARD K M, CLEMENTE-CARVALHO R B G, JENSEN E L, *et al.* Genotyping-in-thousands by sequencing (GT-seq) of noninvasive faecal and degraded samples: a new panel to enable ongoing monitoring of Canadian polar bear populations [J]. Mo-

- lecular Ecology Resources, 2022, 22(5): 1906–1918.
- [57] SCHMIDT D A, GOVINDARAJULU P, LARSEN K W, *et al.* Genotyping-in-thousands by sequencing reveals marked population structure in western rattlesnakes to inform conservation status [J]. Ecology and Evolution, 2020, 10(14): 7157–7172.
- [58] BOOTSMA M L, GRUENTHAL K M, MCKINNEY G J, *et al.* A GT-seq panel for walleye (*Sander vitreus*) provides important insights for efficient development and implementation of amplicon panels in non-model organisms [J]. Molecular Ecology Resources, 2020, 20(6): 1706–1722.
- [59] JO J, KIM Y, KIM G W, *et al.* Development of a panel of genotyping-in-thousands by sequencing in *Capsicum* [J]. Frontiers in Plant Science, 2021, 12: 769473.
- [60] ZHANG R R, TAN Y, JIAN H, *et al.* A new approach to detect a set of SNP-SNP markers: combining ARMS-PCR with SNaPshot technology [J]. Forensic Science International: Genetics Supplement Series, 2019, 7(1): 150–151.
- [61] BUTLER J M, COBLE M D, VALLONE P M. STRs vs. SNPs: thoughts on the future of forensic DNA testing [J]. Forensic Science, Medicine, and Pathology, 2007, 3(3): 200–205.
- [62] UTTER F. Population genetics, conservation and evolution in salmonids and other widely cultured fishes: some perspectives over six decades [J]. Reviews in Fish Biology and Fisheries, 2004, 14(1): 125–144.
- [63] ROUGEMONT Q, CARRIER A, LE LUYER J, *et al.* Combining population genomics and forward simulations to investigate stocking impacts: a case study of muskellunge (*Esox masquinongy*) from the St. Lawrence River Basin [J]. Evolutionary Applications, 2019, 12(5): 902–922.
- [64] BOOTSMA M L, MILLER L, SASS G G, *et al.* The ghosts of propagation past: haplotype information clarifies the relative influence of stocking history and phylogeographic processes on contemporary population structure of walleye (*Sander vitreus*) [J]. Evolutionary Applications, 2021, 14(4): 1124–1144.
- [65] KALINOWSKI S T, TAPER M L, MARSHALL T C. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment [J]. Molecular Ecology, 2007, 16(5): 1099–1106.
- [66] JONES O R, WANG J L. COLONY: a program for parentage and sibship inference from multilocus genotype data [J]. Molecular Ecology Resources, 2010, 10(3): 551–555.