

庞春英, 梁莎莎, 文崇利, 等. 基于 KASP 技术的水牛 SNP 标记开发及其分子身份证构建 [J]. 畜牧与兽医, 2025, 57 (3): 1-10.

PANG C Y, LIANG S S, WEN C L, et al. Development of SNP markers and establishment of molecular identity cards for water buffalo based on KASP technology [J]. Animal Husbandry & Veterinary Medicine, 2025, 57 (3): 1-10.

基于 KASP 技术的水牛 SNP 标记开发及其分子身份证构建

庞春英[#], 梁莎莎[#], 文崇利, 钟华配, 韦科龙, 梁淦,
梁宇辰, 陈笑寒, 黄斌, 冯玲^{*}, 梁贤威^{*}

(广西水牛研究所/农业农村部(广西)水牛遗传繁育重点实验室, 广西 南宁 530001)

摘要: 旨在利用竞争性等位基因特异性聚合酶链式反应 (competitive allele specific PCR, KASP) 技术, 探讨水牛分子身份证构建的方法, 为今后水牛品种精准鉴定及其品种资源的保护等提供技术支撑。基于水牛基因组重测序获得单核苷酸多态性 (SNP) 位点信息, 筛选合适 SNP 位点, 通过转化标记获得可用于分子身份证构建的 SNP 位点, 进行遗传多样性、群体结构分析, 并分别使用二进制和十进制构建 14 个水牛品种的 DNA 分子身份证。结果: 基于前期研究的 10 头不同水牛品种基因组重测序数据, 筛选获得 284 个在不同种间有差异且品种内一致的 SNPs 位点, 在 40 个水牛样本中测序并结合桑格 (Sanger) 测序验证获得的 60 个 SNPs 位点, 6 个位点未能正常分型, 转化成功率为 90%, 有效等位基因 (N_e) 平均值为 1.522, 香农信息指数 (SHA) 平均值为 0.503, 观察杂合度 (H_o) 和期望杂合度 (H_e) 平均值分别为 0.182 和 0.328, 多态性信息含量值 (PIC) 平均值为 0.269; 除去 20 个在部分水牛品种中位点缺失的 SNPs 位点, 应用 34 个 SNPs 位点进行群体遗传结构分析表明 14 个水牛品种分为 2 个类群, 其中槟榔江水牛、尼里/拉菲水牛和摩拉水牛的遗传距离与其他 11 种水牛的较远, 同为 1 个类群; 最终根据 34 个 SNPs 位点构建了 14 个水牛品种的分子身份证。综上, 应用 KASP 技术并结合遗传多样性和群体结构分析, 成功开发出能精准区分全部供试水牛品种的 34 个 SNPs 位点, 并构建了 14 个水牛品种具有唯一性的二进制和十进制分子身份证, 说明此方法可有效构建水牛分子身份证。

关键词: 水牛; 竞争性等位基因特异性技术; SNP 标记; 分子身份证

中图分类号: S823 文献标志码: A 文章编号: 0529-5130(2025)03-0001-10

Development of SNP markers and establishment of molecular identity cards for water buffalo based on KASP technology

PANG Chunying[#], LIANG Shasha[#], WEN Chongli, ZHONG Huapei, WEI Kelong, LIANG Gan,
LIANG Yuchen, CHEN Xiaohan, HUANG Bin, FENG Ling^{*}, LIANG Xianwei^{*}
(Guangxi Buffalo Research Institute/Guangxi Key Laboratory of Buffalo Genetics, Breeding and
Reproduction Technology, Ministry of Agriculture and Rural Affairs, Nanning 530001, China)

Abstract: The competitive allele specific PCR (KASP) technology was used in this study to establish a method of constructing molecular identity cards for water buffaloes, in order to offer technical support for precise identification of water buffalo breeds and protection of breed resources in the future. Based on water buffalo genome resequencing, SNP site information was obtained, suitable SNP sites were screened, and SNP sites that can be used for molecular identity card construction were obtained through transformation markers. Genetic diversity and population structure analysis were conducted, and DNA molecular identity cards for 14 water buffalo breeds were constructed using binary and decimal systems, respectively. Based on genome resequencing data from 10 different breeds of water buffalo, 284 SNPs with differences between different breeds and consistency within the breed were screened. 60 SNPs were obtained by sequencing 40 water buffalo samples and by combining with Sanger sequencing verification. Six SNPs failed to be classified normally, resulting in a conversion power of 90%. The average effective allele (N_e) value was 1.522, and the average Shannon information index (SHA) value was 0.503. The average values of observed heterozygosity (H_o) and expected heterozygosity (H_e) were 0.182 and 0.328, respectively, and the polymorphism information content (PIC) was 0.269. Excluding 20 SNPs that were missing in some water buffalo breeds, a population genetic structure analysis using 34 SNPs

收稿日期: 2024-03-07; 修回日期: 2025-01-15

基金项目: 农业农村部畜禽种质资源精准鉴定项目 (202103); 广西科技重大专项 (桂科 AA16450002)

第一作者: 庞春英, 女, 本科, 研究员; 梁莎莎, 女, 助理研究员。[#]共同第一作者

^{*}通信作者: 冯玲, 助理研究员, 研究方向: 水牛繁育, E-mail: 29498987@qq.com; 梁贤威, 研究员, 研究方向: 水牛繁殖, E-mail: liangbri@126.com。

showed that 14 water buffalo breeds were divided into 2 groups. Among them, the genetic distance of the Binglangjiang buffalo, Nili/Lafite buffalo and Mora buffalo was far from that of the other 11 water buffalo breeds, and they belong to the same group. Finally, molecular identification cards for 14 water buffalo breeds were constructed based on 34 SNPs. By applying the KASP technology combined with genetic diversity and population structure analysis, 34 SNPs that can accurately distinguish all 14 tested water buffalo breeds were successfully developed, and 14 unique binary and decimal molecular identification cards for water buffalo breeds were constructed. This study indicated that the present method was able to effectively construct water buffalo molecular identification cards.

Keywords: water buffalo; KASP technology; SNP marking; molecular identity card

水牛是偶蹄目牛科水牛属 (*Bubalus*) 哺乳动物, 大多数分布于亚洲、非洲、拉丁美洲、大洋洲及东南欧的一些国家, 分为沼泽型和河流型^[1]。中国是水牛养殖大国, 主要分布于黄河流域以南的地区, 地方品种均为沼泽型, 以役用为主^[2]。中国水牛按不同的地域命名, 共有 27 个地方品种, 是世界上沼泽型水牛资源最丰富的国家。近几十年来, 为了提高中国水牛的生产性能, 先后有多个河流型水牛品种被引入国内并与沼泽型水牛进行杂交改良, 导致一些拥有优良性状的地方品种资源丢失^[3]。近年来, 随着生物学以及生物技术的发展, 可以直接利用个体的基因组信息对个体进行亲缘鉴定, 由于地方品种在外貌上相似性很高, 单从外貌上很难作区分, 从基因组水平对个体进行快速、精准的品种鉴定, 构建其分子身份证, 以解决水牛“同种异名”或“同名异种”等问题成为研究热点。

竞争性等位基因特异性聚合酶链式反应 (competitive allelespecific PCR, KASP) 因其灵活、便宜、准确等特点, 已被广泛应用于牛、羊等家畜的生产性状 SNP 位点挖掘^[4-6] 及遗传疾病鉴定等方面的研究^[7-8]。在牛育种方面, 使用 KASP 方法鉴定到了多个与产乳性状相关的候选基因如催乳素受体 (PRLR) 基因、K 酪蛋白 (CSN3) 基因等相关的单

核苷酸多态性 (SNP) 位点, 并发现 PRLR 基因的 20: g. 39071965C>T 位点与产乳量、乳脂量等性状显著相关^[9], 20: g11188A>G 位点 (AA 型) 与牛乳脂肪酸等乳脂性状有显著相关^[10]。而采用 KASP 分型确定了中国荷斯坦牛群最高的脊椎畸形综合征携带率为 10.5%^[11], 并用同样方法验证出牛胚胎时期 CYP26C1 错义突变导致其犍牛下颌面部骨质疏松^[12]。但是在牛或水牛的品种鉴定、SNP 指纹图谱构建及分子身份证构建方面的研究报道较少。本研究基于前期对国内一些地方水牛基因组重测序数据的研究基础上, 采用 KASP 技术等分析方法, 探讨构建水牛分子身份证的可能性, 为今后在基因组水平上进行水牛品种精准鉴定、品种资源保护、水牛育种及其种业的发展起到奠定基础的作用, 具有重要的现实意义。

1 材料与方法

1.1 试验材料

采集来自广西、云南、福建、江苏、江西等地 14 个水牛品种原产地区养殖户或核心种牛场共 40 头健康水牛的外周血液样本 (表 1), 用冰袋运输迅速带回实验室储存备用。

表 1 水牛样本信息

编号	品种名称	采集地点	编号	品种名称	采集地点
S00032479	盱眙山区水牛	江苏省淮安市盱眙县石顶后山	S06706892	富钟水牛	
S00031613	盱眙山区水牛	天泉湖镇	CW59412667	富钟水牛	广西壮族自治区贺州市富川瑶族自治县葛坡镇
E51024	海子水牛	江苏省盐城市射阳县射阳海子	S06548093	富钟水牛	
F61634	海子水牛	水牛种牛场	CW59412823	富钟水牛	
S9	盐津水牛	云南省昭通市大关县天星镇	CW59412651	富钟水牛	
S11	盐津水牛			CW59412883	西林水牛
S15	德宏水牛	云南省德宏傣族景颇族自治州	CW59412867	西林水牛	广西壮族自治区百色市西林县那佐乡
S16	德宏水牛	芒市风平镇	S06640648	西林水牛	
S7	槟榔江水牛	国家槟榔江水牛保种场	S06546387	西林水牛	
S8	槟榔江水牛			CW59161554	

续表1

编号	品种名称	采集地点	编号	品种名称	采集地点
S13	滇东南水牛	云南省红河哈尼族彝族自治州	ML1494	摩拉水牛	
S14	滇东南水牛	蒙自市芷村镇	ML1426	摩拉水牛	
S026	鄱阳湖水牛	江西省庐山市蛟塘镇	ML1702	摩拉水牛	广西壮族自治区水牛研究所种畜场
S028	鄱阳湖水牛		ML1652	摩拉水牛	
S023	信丰山地水牛	江西省赣州市信丰县万隆乡	ML1029	摩拉水牛	
S038	信丰山地水牛		N1257	尼里/拉菲水牛	
X14	峡江水牛	江西省吉安市峡江县金江一乡	N1147	尼里/拉菲水牛	广西壮族自治区水牛研究所种畜场
X21	峡江水牛		N1177	尼里/拉菲水牛	
Y08451340	福安水牛		NL1344	尼里/拉菲水牛	
Y08449637	福安水牛	福建省福安市赛岐镇	N1187	尼里/拉菲水牛	

1.2 基因组 DNA 提取

将血液样本充分颠倒混匀，用移液器吹打 3~4 次后取 200 μL 样品于深孔板，向深孔板中分别加入 300 μL 裂解液和 5 μL 蛋白酶 K，然后置于 65 $^{\circ}\text{C}$ 烘箱裂解 1 h。裂解完成后，向每个孔中加入 380 μL 异丙醇和 100 μL 磁珠，然后每孔使用 500 μL 的 75% 乙醇洗脱 3 次，使用紫外分光光度计检测 DNA 浓度，1% 琼脂糖凝胶电泳检测 DNA 质量。

1.3 引物合成

根据 SNP 位点序列信息，结合水牛基因组信息，获取 SNP 位点前后各 100 bp 特异性 DNA 序列，使用引物设计工具 SNPway 设计 KASP 反应引物及桑格 (Sanger) 反应引物，引物均由武汉天一华煜基因科技有限公司合成。

1.4 PCR 扩增与分型

将 DNA 样品加到 384 孔 PCR 反应板上，每块 PCR 反应板上均需加 2 个及以上的阴性对照。配置 KASP 基因分型混合液：Primer Mix 1.5 μL ，2 \times KASP Master Mix 2.5 μL ，ddH₂O 1.35 μL 。将 KASP 基因分型混合液加到已含 DNA 模板的 PCR 反应板上，将反应板密封后 1 500 r/min 离心 10 s，更换可荧光透视膜密封 PCR 反应板。PCR 反应程序：94 $^{\circ}\text{C}$ 15 min；94 $^{\circ}\text{C}$ 20 s，65 $^{\circ}\text{C}$ 60s，10 个循环；94 $^{\circ}\text{C}$ 20 s，57 $^{\circ}\text{C}$ 60 s，40 个循环；荧光读取 30 $^{\circ}\text{C}$ 60 s。

1.5 Sanger 测序验证

根据 KASP 扩增结果，选取分型成功的位点，将 KASP 分型结果差异较大的 4 个水牛样本，进行 PCR 扩增测序验证，PCR 扩增体系 30 μL ：2 \times Taq Mix 15 μL 、上下游引物各 1 μL (浓度为 20 ng/ μL)、模板 DNA 1 μL 、ddH₂O 12 μL 。PCR 反应程序：95 $^{\circ}\text{C}$ 30 s；95 $^{\circ}\text{C}$ 10 s，60 $^{\circ}\text{C}$ 30 s，72 $^{\circ}\text{C}$ 30 s，35 个循环；72 $^{\circ}\text{C}$ 5 min。反应产物送武汉天一华煜基因科技有限公司进行 Sanger 测序，并使用 SNP analysis 对测序原

始数据进行分析，并与 KASP 分型结果进行比对。

1.6 SNP 位点遗传多样性分析

根据 KASP 测序分型结果数据计算遗传多样性参数，包括多态性信息含量值 (PIC)、观察杂合度 (Ho)、期望杂合度 (He)、香农信息指数 (SHA) 和有效等位基因 (Ne)。

$$PIC = 1 - \sum_{i=1}^n P_i^2 - \sum_{i=1}^{n-1} \sum_{j=i+1}^n 2P_i P_j^2,$$

式中， n 为等位基因数量， P_i 代表第 i 个等位基因的频率； P_j 代表第 j 个等位基因的频率。

1.7 群体遗传结构分析

Nei 氏标准遗传距离是基于进化模型提出的遗传距离测度，个体或群体之间的距离越大，亲缘关系越远；反之，距离越小，亲缘关系越近^[13-14]。计算公式如下： $D = -\ln(I)$ ， $I = \sum(x_i \cdot y_i)$ ，式中， D 代表两个群体间的遗传距离， I 为两个群体间的遗传一致度， x_i 和 y_i 分别代表两个群体中第 i 个等位基因的频率。

为进一步分析群体结构，并对 14 个水牛品种确定最佳的类群分组，选择使用 Structure 2.3.4 进行贝叶斯聚类^[15]。具体参数如下：混合模型设定迭代次数 (MCMC) 的重复次数为 20 000；设定等位变异频率特征数 (遗传群体数) $K = 2 \sim 20$ ；Burn-in 周期为 10 000，每个 K 值重复 20 次，并利用 Structure Harvester 算出最佳 ΔK 值，即为最佳群体分层情况。

1.8 水牛分子身份证的构建

使用 1.3~1.7 中筛选获得的 34 个 SNPs 位点进行数字编码，作为构建分子身份证基本信息。具体方法为每个 SNP 标记对应 3 种等位基因 (XY、XX 和 YY)，分别用 11、10、01 代表等位基因多态性，为二进制编码分子身份证，然后在进行十进制编码，即获得每个品种的分子身份证。

2 结果

2.1 KASP 标记的开发和筛选

基于前期研究 10 头不同品种的水牛基因组重测序数据结果, 分析比对出不同品种间有差异且品种内一致的 SNP 位点, 按照突变注释文件 (MAF) >0.4、缺失率 <0.1、杂合度 <0.2、上下游 200 bp 无序列变异、两等位基因位点、PIC >0.25、染色体分布情况

清晰、在基因组上均匀分布这 8 个标准筛选, 共计获得 284 个候选 SNPs 位点。对 284 个 SNPs 位点进行 KASP 引物合成, 并在 10 个已测序水牛中进行基因分型效率评估, 最终保留 95 个分群清晰、基因分型准确性高、重复性好的位点 (图 1)。在 KASP 分型成功的位点中挑选 4 个结果差异较大对应样本, 使用 Sanger 测序验证, 最终筛选得到 60 个符合质控标准的 KASP 标记, 分布于 10 个染色体上 (表 2)。

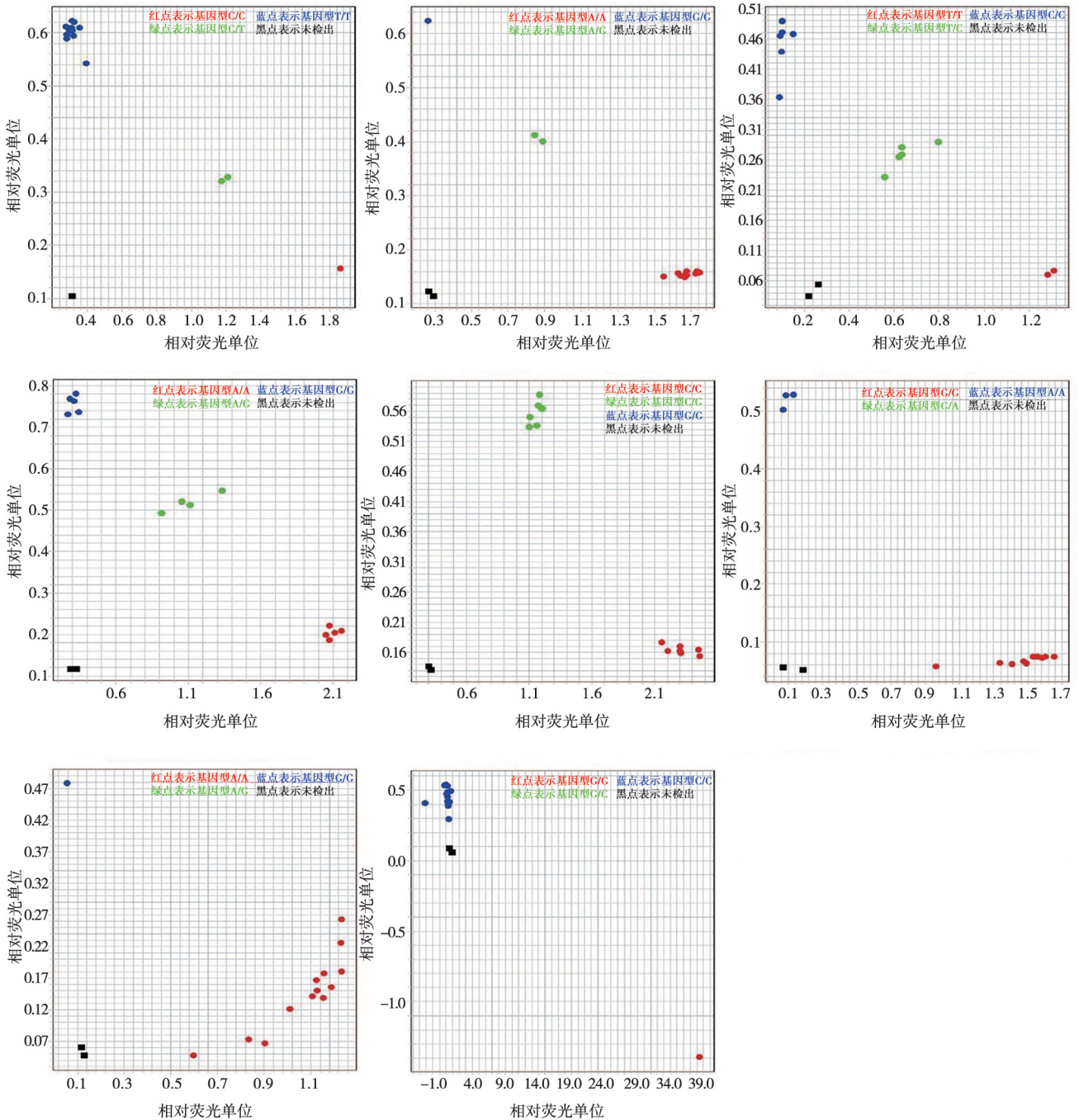


图 1 部分 SNP 位点的 KASP 标记初筛基因分型

表2 60个 SNPs 标记位点信息

位点名称	染色体	位置	参考位点	突变位点	突变方式	分型情况	位点情况
K026-001	NC_059157.1	106 235 084	C	T	c.156-72488G>A	√	-
K026-004	NC_059157.1	121 401 270	A	G	c.49+11505T>C	√	-
K026-007	NC_059157.1	193 259 853	T	C	n.141-223A>G	√	-
K026-008	NC_059157.1	198 084 426	C	T	c.174+317C>T	×	+
K026-010	NC_059157.1	198 099 993	C	T	c.281+1051C>T	√	-
K026-011	NC_059158.1	3 090 037	C	T	c.1050+7704C>T	×	+
K026-013	NC_059158.1	3 097 328	C	T	c.1050+14995C>T	√	+
K026-016	NC_059158.1	7 467 226	C	T	c.-179+29293G>A	×	+
K026-020	NC_059158.1	8 051 325	G	A	c.40+22340G>A	×	+
K026-023	NC_059158.1	114 688 540	T	A	n.114688540T>A	√	-
K026-028	NC_059158.1	148 990 370	G	A	n.148990370G>A	√	-
K026-029	NC_059158.1	151 205 183	A	G	n.151205183A>G	√	-
K026-030	NC_059158.1	152 258 833	A	G	c.234+33857T>C	√	+
K026-032	NC_059158.1	152 402 450	C	G	c.83-109609G>C	√	+
K026-033	NC_059158.1	152 417 411	C	G	c.83-124570G>C	√	+
K026-038	NC_059159.1	27 933 986	A	G	c.*2927A>G	√	-
K026-039	NC_059159.1	45 090 885	T	G	c.2442-929T>G	×	+
K026-040	NC_059159.1	45 094 235	C	T	c.2785-76C>T	√	-
K026-041	NC_059159.1	45 101 545	A	T	c.3746-529A>T	√	-
K026-042	NC_059159.1	45 114 294	T	G	c.-209T>G	√	-
K026-043	NC_059159.1	45 739 049	A	T	c.305-251T>A	√	+
K026-047	NC_059159.1	70 135 556	A	G	n.70135556A>G	√	+
K026-048	NC_059159.1	144 359 402	G	C	n.144359402G>C	√	-
K026-051	NC_059159.1	156 244 679	C	G	c.935-1845C>G	×	+
K026-053	NC_059159.1	158 113 918	C	G	n.158113918C>G	√	+
K026-054	NC_059159.1	160 845 640	A	C	c.-3428T>G	√	-
K026-056	NC_059159.1	167 874 602	G	T	c.1446-32650C>A	√	-
K026-059	NC_059160.1	71 540 533	C	T	n.71540533C>T	√	-
K026-062	NC_059160.1	87 945 350	G	A	n.463-1504C>T	√	-
K026-063	NC_059160.1	88 098 942	C	T	c.-4929C>T	√	+
K026-067	NC_059160.1	151 359 575	C	G	n.*4327G>C	√	+
K026-069	NC_059160.1	157 443 868	A	G	n.201+17015T>C	√	-
K026-070	NC_059160.1	162 188 046	T	C	c.235-80T>C	√	+
K026-074	NC_059161.1	6 000 791	T	C	n.6000791T>C	√	+
K026-078	NC_059161.1	17 229 105	A	G	n.408-3650A>G	√	+
K026-079	NC_059161.1	17 247 053	C	T	c.-5166C>T	√	+
K026-080	NC_059161.1	17 504 975	T	C	n.17504975T>C	√	+
K026-085	NC_059161.1	31 133 817	A	G	c.37+44280A>G	√	+
K026-087	NC_059161.1	32 347 867	T	C	n.32347867T>C	√	+
K026-092	NC_059163.1	6 226 574	T	A	n.6226574T>A	√	+

续表2

位点名称	染色体	位置	参考位点	突变位点	突变方式	分型情况	位点情况
K026-094	NC_059163.1	6 235 556	C	G	n.6235556C>G	√	+
K026-096	NC_059163.1	19 955 552	C	T	c.248-49328C>T	√	+
K026-098	NC_059163.1	24 515 935	C	T	c.981+192G>A	√	+
K026-105	NC_059163.1	47 865 007	G	A	n.47865007G>A	√	+
K026-108	NC_059163.1	95 360 936	G	T	n.95360936G>T	√	+
K026-110	NC_059164.1	34 661 515	A	G	n.34661515A>G	√	+
K026-116	NC_059164.1	45 703 990	T	A	n.816-7962T>A	√	+
K026-117	NC_059164.1	45 705 057	C	T	n.816-6895C>T	√	+
K026-118	NC_059164.1	45 720 584	C	T	n.864-8285C>T	√	+
K026-119	NC_059164.1	45 722 363	G	C	n.864-6506G>C	√	+
K026-142	NC_059167.1	25 517 679	C	T	c.-2961C>T	√	+
K026-145	NC_059167.1	25 546 777	G	A	c.1082+224C>T	√	-
K026-155	NC_059167.1	88 432 978	A	G	c.252+3274T>C	√	+
K026-156	NC_059167.1	92 947 860	A	G	c.1205-18757A>G	√	+
K026-158	NC_059167.1	93 000 155	T	C	n.93000155T>C	√	+
K026-162	NC_059167.1	99 079 342	C	A	n.99079342C>A	√	+
K026-165	NC_059167.1	99 095 238	G	A	n.99095238G>A	√	+
K026-167	NC_059168.1	97 080 042	G	A	c.* 4139G>A	√	-
K026-171	NC_059168.1	100 966 207	G	A	c.-4419G>A	√	+
K026-256	NC_059175.1	61 145 356	A	G	c.38-73608A>G	√	-

注: √代表分型成功, ×代表分型未成功, -代表位点缺失, +代表位点没有缺失。

2.2 SNP 位点多样性分析

40 个样本中, 对 60 个 SNPs 标记进行分型, 结果有 6 个 SNPs 分型失败 (表 2), 成功率为 90%。对 54 个 SNPs 位点进行遗传多样性分析, 结果显示 N_e 数值 1.5 以上的占比 46.3% (图 2A), 平均值为 1.522 (图 2F), 表明等位基因在群体中分布较为均匀; SHA 为 0.503 (0.000~0.692) (图 2F), 集中在 0.4~0.8 之间 (85.2%) (图 2B); H_o 和 H_e 平均值分别为 0.182 (0.000~0.475)、0.328 (0.100~0.499) (图 2C、图 2D、图 2F); PIC 在 0.200~0.400 之间的 SNP 比例为 85.1%, 平均值为 0.269。

2.3 群体遗传结构分析

KASP 标记的分型显示其中有 20 个 SNPs 位点在 40 个样本中有部分缺失 (表 2), 剩余 34 个 SNPs 位点进行遗传结构分析, 通过计算, 14 个水牛品种 D 在 0.133~1.000 之间, 平均为 0.508 (表 3)。其中盐津水牛和信丰山地水牛之间的 D 最小为 0.133, 表明这 2 个水牛品种之间的亲缘关系最近; 尼里/拉菲水牛和多数水牛地方品种之间的 D 均为 1, 表明这 2

个类型水牛品种之间的亲缘关系最远。进化树结果显示同一品种不同样本都在聚类在一起, 14 个水牛品种分为 4 个类群 (图 3A), 其中槟榔江水牛和尼里/拉菲水牛与其他地区水牛在不同的分支上 (图 3A), 摩拉水牛虽然与另外 11 种水牛在一个分支上, 但其距离较远, 说明其亲缘关系较远。Structure 分析结果显示, 当 $K=2$ 时, ΔK 取得最大值 (图 3B), 可推断 40 份水牛样本分为 2 个类群 (图 3C), 其中槟榔江水牛、尼里/拉菲水牛和摩拉水牛为 1 个类群, 其他 11 个地区的水牛品种为 1 个类群, 与进化树聚类分析结果一致。

2.4 水牛分子身份证的构建

利用 34 个 SNPs 位点, 按照固有顺序统计所有 SNP 位点在 14 个水牛品种标准样本的基因型, 并进行 11、10、01 二进制编码, 继而对这些组合进行十进制赋值, 获得每个水牛品种的分子身份证 (表 4)。结果显示在二进制和十进制条件下, 同一个品种不同样本之间具有相同的分子身份证, 但不同品种之间的分子身份证差异较大。

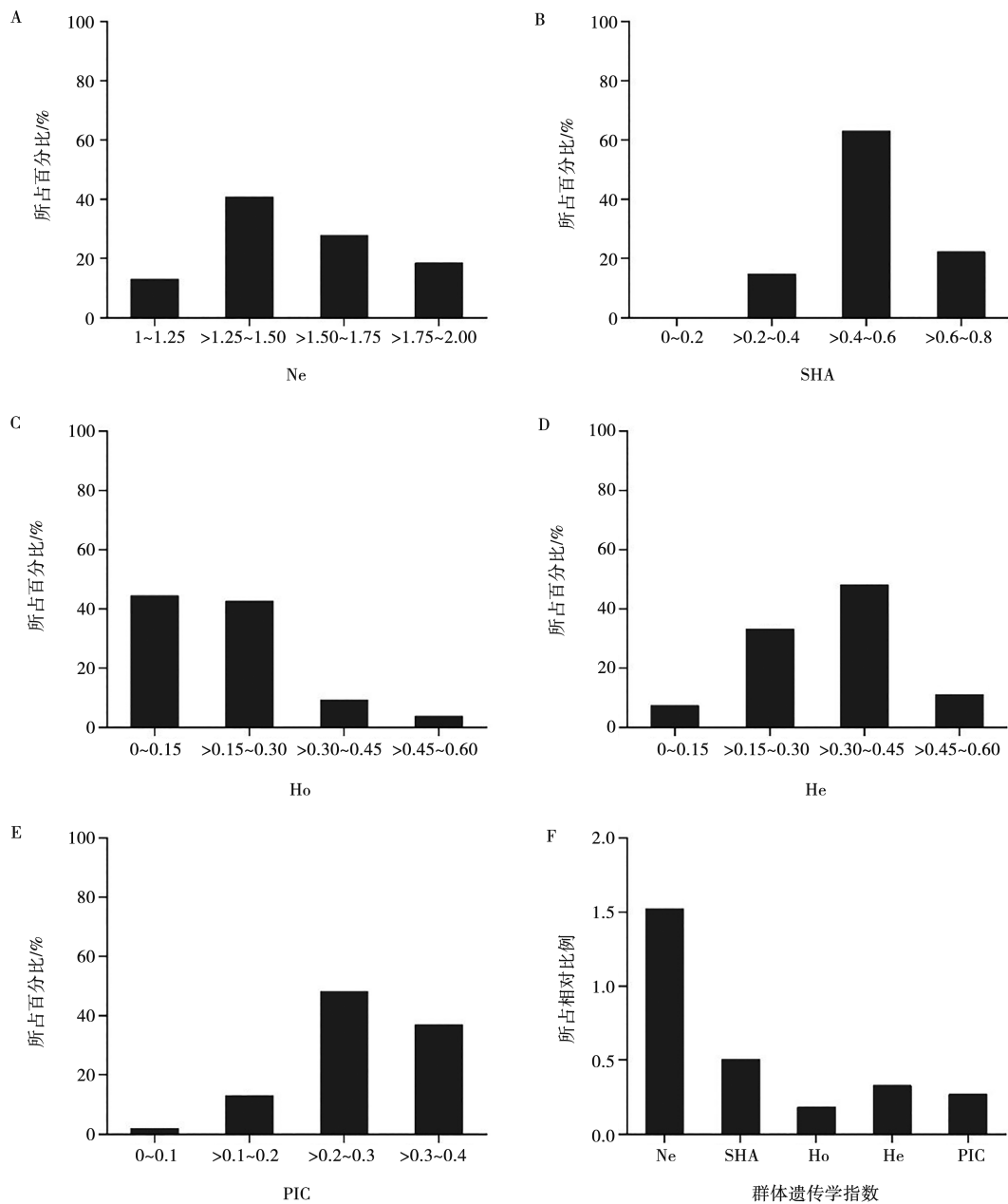


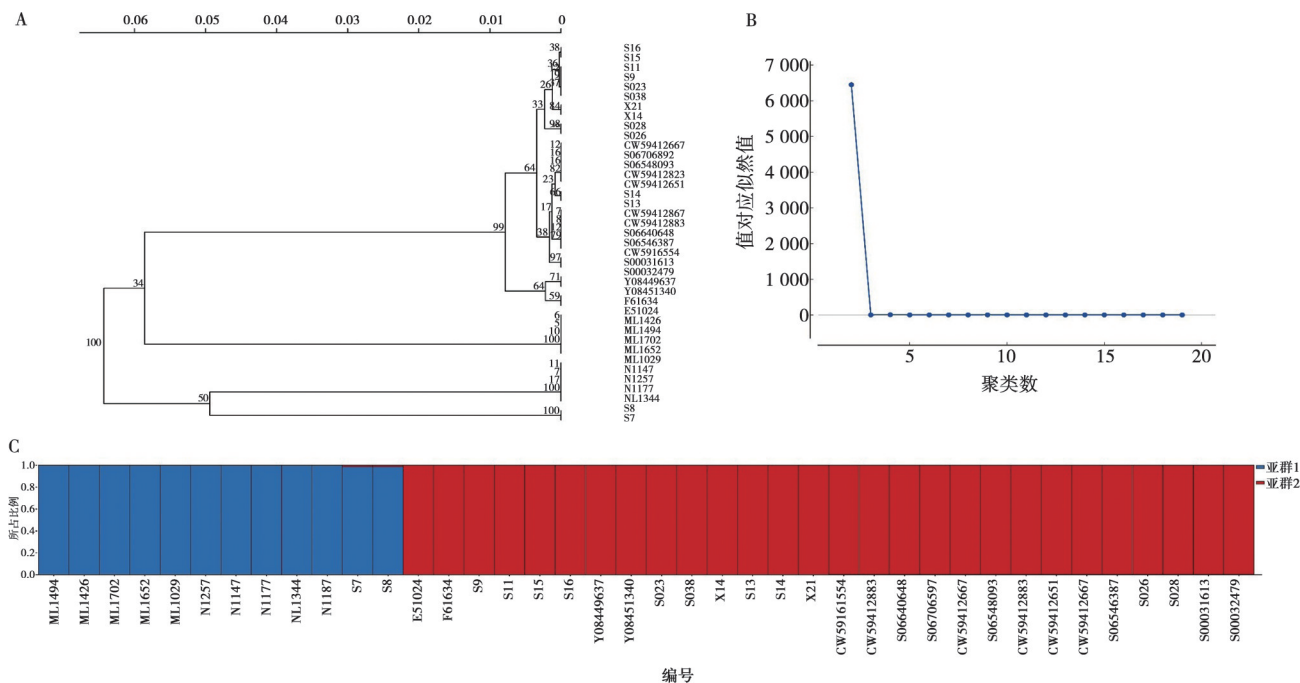
图2 54个KASP标记基于40份水牛样本的遗传多样性分析

表3 14个水牛品种之间的遗传距离

品种	盱眙山区水牛	海子水牛	盐津水牛	德宏水牛	槟榔江水牛	滇东南水牛	鄱阳湖水牛	信丰山地水牛	峡江水牛	福安水牛	富钟水牛	西林水牛	摩拉水牛	尼里/拉菲水牛
盱眙山区水牛	0.000	0.431	0.391	0.429	0.800	0.447	0.721	0.400	0.479	0.375	0.333	0.364	0.975	1.000
海子水牛	0.431	0.000	0.160	0.173	0.862	0.216	0.569	0.200	0.250	0.154	0.240	0.163	0.965	1.000
盐津水牛	0.391	0.160	0.000	0.184	0.868	0.191	0.543	0.133	0.191	0.125	0.178	0.174	0.940	1.000
德宏水牛	0.429	0.173	0.184	0.000	0.855	0.204	0.571	0.224	0.240	0.140	0.265	0.224	0.943	1.000
槟榔江水牛	0.800	0.862	0.868	0.855	0.000	0.865	0.794	0.885	0.843	0.857	0.857	0.840	0.875	0.810
滇东南水牛	0.447	0.216	0.191	0.204	0.865	0.000	0.533	0.234	0.213	0.146	0.200	0.234	0.939	1.000
鄱阳湖水牛	0.721	0.569	0.543	0.571	0.794	0.533	0.000	0.556	0.500	0.551	0.630	0.556	0.903	0.968
信丰山地水牛	0.400	0.200	0.133	0.224	0.885	0.234	0.556	0.000	0.234	0.204	0.222	0.217	0.959	1.000
峡江水牛	0.479	0.250	0.191	0.240	0.843	0.213	0.500	0.234	0.000	0.184	0.313	0.306	0.939	1.000

续表3

品种	盱眙山区水牛	海子水牛	盐津水牛	德宏水牛	槟榔江水牛	滇东南水牛	鄱阳湖水牛	信丰山地水牛	峡江水牛	福安水牛	富钟水牛	西林水牛	摩拉水牛	尼里/拉菲水牛
福安水牛	0.375	0.154	0.125	0.140	0.857	0.146	0.551	0.204	0.184	0.000	0.170	0.167	0.944	1.000
富钟水牛	0.333	0.240	0.178	0.265	0.857	0.200	0.630	0.222	0.313	0.170	0.000	0.182	0.935	1.000
西林水牛	0.364	0.163	0.174	0.224	0.840	0.234	0.556	0.217	0.306	0.167	0.182	0.000	0.959	1.000
摩拉水牛	0.975	0.965	0.940	0.943	0.875	0.939	0.903	0.959	0.939	0.944	0.935	0.959	0.000	0.800
尼里/拉菲水牛	1.000	1.000	1.000	1.000	0.810	1.000	0.968	1.000	1.000	1.000	1.000	1.000	0.800	0.000



A. 聚类分析; B. K 值与 ΔK 趋势变化情况; C. K=2 时群体遗传结构分析。

图 3 40 份水牛样本的聚类图和群体结构

表 4 14 个水牛品种分子身份证编码

品种名称	二进制	十进制
盱眙山区水牛	1111101101001111101011011010010101110010110011010101011111110111	2.924×10^{20}
海子水牛	101010101010010110101001101010010101100101100110101010100110100101	1.967×10^{20}
盐津水牛	1010101110100111101010111101001010110100101010101001100110100110	1.979×10^{20}
德宏水牛	10101011101001011010101101101001010110100101100110101010100110100110	1.979×10^{20}
槟榔江水牛	11010101101001100101111101111111111010011111111111101111110011001	2.463×10^{20}
滇东南水牛	101010111010010110101001111010010101100101100110101011100110110111	1.979×10^{20}
鄱阳湖水牛	101010011010010111110111101011010110100101100110101011100110100111	1.956×10^{20}
信丰山地水牛	10101001101001111010101101101001010111100101100110101010100110100111	1.956×10^{20}
峡江水牛	10011001101001011011100111101001010111100101110110101011100110100110	1.771×10^{20}
福安水牛	10101010101001011010101111101001010111100101110110101010100110100111	1.967×10^{20}
富钟水牛	10101101101001111010101111010010101101001010110101010100111100111	2.002×10^{20}
西林水牛	10101010101001011010100101101001010110100101100110101011100110100101	1.967×10^{20}
摩拉水牛	0101010101101111101110110011101111011011101110111011101110110011010	9.850×10^{19}
尼里/拉菲水牛	1101010111010111010101011001011011111101010111001010111011001011011	2.465×10^{20}

3 讨论

3.1 水牛品种 SNP 标记位点 KASP 转化

目前在动物品种鉴定的研究中,多应用简单序列重复间区 (ISSR)、随机扩增多态性 DNA (RAPD)、序列相关扩增多态性 (SRAP) 等分子标记方法,而这些分子标记方法大都存在着操作繁琐、稳定性差、成本高、准确率不足等缺点。SNP 标记是当前遗传标记研究中最多也是最有前景的分子标记。该标记主要是指群体中变异频率大于 1% 的单个核苷酸改变而导致的核酸序列多态性,具体是指在 DNA 序列中的单个碱基的变异,形式包括单碱基的缺失、插入、转换及颠换。这是动植物基因组中最常见的遗传变异类型之一,在目前的研究中多用于构建高密度遗传图谱和标记辅助选择以及对目的性状相关的候选基因 SNP 的精确定位^[16]。早在 2010 年就提出,随着测序技术的发展,SNP 检测技术必将成为新一代动物遗传育种方法和标记辅助方法,并在牛的遗传育种中起至关重要的作用^[17]。KASP 是基于已知 SNP 的位点荧光基因分型技术,可用于对基因组 SNP 位点以及插入缺失多态性进行双等位基因分型^[18]。它无需对靶点即特异性引物/探针进行标记,不需要根据每个 SNP 位点合成特异的荧光引物,其独特的扩增阻滞突变系统 (PCR) PCR 原理,让所有的位点检测最终都使用通用荧光引物扩增,大大降低了成本。本研究应用 10 个水牛品种基因组重测序数据,筛选出不同品种间有差异且品种内一致的 SNP 位点,通过 KASP 初步筛选获得 95 个 SNPs 标记位点,结合 Sanger 测序方法,获得多态性标记为 60 个,占比为 62.5%;通过对 40 头水牛样本中进行标记分型,转化成功率为 90%,相比于其他来说具有较高的重复性和准确性^[19-20]。

3.2 SNP 位点多样性分析

利用筛选获得的 54 个 SNPs 标记位点对 40 头水牛样本进行遗传多态性分析,结果显示 SNP 位点的 H_o 和 H_e 平均值分别为 0.182、0.328, N_e 平均值为 1.522, SHA 平均值为 0.503, PIC 平均值为 0.269,由此表明,本研究筛选的位点在水牛品种间具有中度多态性。对所有位点进一步分析发现,有 20 个 SNPs 在部分品种中存在位点缺失,需要舍去。研究分析发现,摩拉水牛和尼里/拉菲水牛与其他水牛品种之间比较, D 均大于 0.5,这 2 种水牛属于河流型水牛品种,主要为引进品种^[2],与其他水牛品种在总胆固醇、脂蛋白等生理方面生化指标方面具有一定的差异^[21],这与沼泽型水牛 30 个微卫星 DNA 的群体遗传多样性分析结果相似^[22]。槟榔江水牛与其他 D 也均在 0.8 以上,是我国目前唯一发现的河流型水牛,

其遗传距离与引进的河流型水牛相近,但与国内沼泽型水牛有较大的遗传差异^[23]。除鄱阳湖水牛外,国内沼泽型的水牛地方品种之间的 D 均在 0.5 以下,说明本研究筛选的 34 个 SNPs 位点具有较高的准确性,准确率高达 92.9%,这些结果都说明了本试验所选用的 34 个 SNP 位点能够很好地揭示了水牛品种的遗传多样性。

3.3 水牛品种分子身份证的价值与应用

构建品种分子身份证通常是基于 DNA 指纹图谱信息,采用一定的编码规则对信息进行数值化和图像化处理,得到身份代码或二维码信息^[24]。在植物分子身份证构建研究方面,简单重复序列 (SSR) 标记技术被广泛应用于农作物^[25-26]、果树^[27-28]、蔬菜^[29] 等植物分子身份证构建研究方面,但在动物中应用较少。KASP 作为 SNP 分型检测方法,具有高效、经济、准确等特点,也在逐步替代 SSR,如在葡萄中应用 KASP 方法基于 22 个 SNPs 位点构建了 70 个品种的分子身份证^[30],在桂花中使用 KASP 技术基于 14 个 SNPs 位点构建了 89 个品种的指纹码^[31]。本研究首次将 KASP 技术结合 Sanger 测序方法筛选获得 34 个 SNPs 标记位点,并基于此分别采用二进制和十进制的方法构建了 14 个水牛品种的分子身份证,每个品种均具有唯一分子身份证。其中 10 个 SNPs 位点: K026-074、K026-080、K026-085、K026-092、K026-094、K026-108、K026-116、K026-117、K026-118、K026-119,在河流型水牛和沼泽型水牛中完全不一致,根据位点的分子身份证信息可以有效区分河流型水牛和沼泽型水牛,表明本研究筛选获得的 34 个 SNPs 位点使用 KASP 标记技术能够作为一种有效的技术手段,用于水牛品种资源的区分和分子身份证构建。

4 结论

应用 KASP-PCR 技术并结合遗传多样性和群体结构分析的研究,可成功开发出能精准区分全部供试 14 个水牛品种的 34 个 SNPs 位点,并构建了 14 个水牛品种具有唯一性的二进制和十进制分子身份证。研究结果为水牛品种精准鉴定和品种资源的保护提供技术支撑。

参考文献:

- [1] BORGHESE A. Buffalo production and research [J]. REU Technical series, 2005 (67): 1-2.
- [2] 黄加祥,黄锋,诸葛莹,等. 中国奶水牛产业发展综述: 发展历程及趋势建议 [J]. 中国奶牛, 2019 (12): 1-8.
- [3] 苗永旺,李大林,霍金龙,等. 中国水牛的遗传多样性与起源

- 分化 [J]. 中国牛业科学, 2008 (4): 16-20.
- [4] ABOUSOLIMAN I, REYER H, OSTER M, et al. Analysis of candidate genes for growth and milk performance traits in the Egyptian Barki sheep [J]. *Animals (Basel)*, 2020, 10 (2): 197.
- [5] 武泽文, 顾亚玲, 张娟, 等. A2 纯合基因型荷斯坦奶牛的筛选 [J]. 黑龙江畜牧兽医, 2019 (8): 44-47.
- [6] 赵睿骁. 金川多肋牦牛的分子标记筛选及分型 [D]. 成都: 西南民族大学, 2021.
- [7] 梁若冰, 谢岩, 范学华, 等. 中国荷斯坦公牛脊椎畸形综合征和尿酸合酶缺乏症遗传缺陷检测 [J]. 中国奶牛, 2014 (7): 22-26.
- [8] 吕小青, 劳兰兰, 赵凤, 等. 北京地区荷斯坦母牛群体 HH1 遗传缺陷基因抽样调研 [J]. 中国奶牛, 2019 (1): 12-14.
- [9] FONTANESI L, CALÒ D G, GALIMBERTI G, et al. A candidate gene association study for nine economically important traits in Italian Holstein cattle [J]. *Anim Genet*, 2014, 45 (4): 576-580.
- [10] COSENZA G, IANNACCONI M, AUZINO B, et al. Remarkable genetic diversity detected at river buffalo prolactin receptor (PRLR) gene and association studies with milk fatty acid composition [J]. *Anim Genet*, 2018, 49 (3): 159-168.
- [11] ZHANG Y, LIANG D, HUANG H, et al. Technical note: development and application of KASP assays for rapid screening of 8 genetic defects in Holstein cattle [J]. *J Dairy Sci*, 2020, 103 (1): 619-624.
- [12] SIECK R L, FULLER A M, BEDWELL P S, et al. Mandibulofacial dysostosis attributed to a recessive mutation of CYP26C1 in Hereford cattle [J]. *Genes (Basel)*, 2020, 11 (11): 1246.
- [13] NEI M. Analysis of gene diversity in subdivided populations [J]. *Proc Natl Acad Sci U S A*, 1973, 70 (12): 3321-3323.
- [14] NEI M, CHESSER R K. Estimation of fixation indices and gene diversities [J]. *Ann Hum Genet*, 1983, 47 (3): 253-259.
- [15] FALUSH D, STEPHENS M, PRITCHARD J K. Inference of population structure using multilocus genotype data; linked loci and correlated allele frequencies [J]. *Genetics*, 2003, 164 (4): 1567-1587.
- [16] 曲娟娟, 于敏, 董蕾, 等. 在动植物遗传育种中的应用 [J]. 东北农业大学学报, 2011, 42 (12): 9-13.
- [17] 谷继承, 刘丑生, 刘刚, 等. 牛遗传资源分子遗传多样性的研究进展 [J]. 中国畜牧兽医, 2010, 37 (9): 127-131.
- [18] HE C, HOLME J, ANTHONY J. SNP genotyping: the KASP assay [J]. *Methods Mol Biol*, 2014, 1145: 75-86.
- [19] 魏中艳, 李慧慧, 李骏, 等. 应用 SNP 精准鉴定大豆种质及构建可扫描身份证 [J]. 作物学报, 2018, 44 (3): 315-323.
- [20] 樊晓静, 于文涛, 蔡春平, 等. 利用 SNP 标记构建茶树品种资源分子身份证 [J]. 中国农业科学, 2021, 54 (8): 1751-1760.
- [21] 钟华配, 刘新勇, 李辉, 等. 不同品种河流型水牛血清生化指标的研究 [J]. 中国畜牧杂志, 2021, 57 (9): 70-73.
- [22] 邓廷贤, 庞春英, 朱鹏, 等. 沼泽型水牛微卫星 DNA 的遗传多样性分析 [J]. 中国畜牧兽医, 2016, 43 (11): 2989-2996.
- [23] 赵志军, 杨茂生, 袁跃云, 等. 槟榔江水牛产业化发展现状与对策 [J]. 养殖与饲料, 2018 (4): 90-91.
- [24] 陆育生, 彭程, 常晓晓, 等. 基于 SSR 标记的广东黄皮种质资源遗传多样性分析及分子身份证构建 [J]. 生物技术通报, 2023, 39 (12): 187-199.
- [25] 秦瑞英, 许学, 张立平, 等. 小麦 SSR 指纹图谱及品种身份证的构建: 基于毛细管电泳分析 [J]. 中国农学通报, 2017, 33 (34): 46-55.
- [26] 管俊娇, 杨晓洪, 王江民, 等. 云梗系列水稻品种分子身份证的构建 [J]. 江苏农业学报, 2018, 34 (6): 1201-1206.
- [27] 李益, 马先锋, 唐浩, 等. 柑橘品种鉴定的 SSR 标记开发和指纹图谱库构建 [J]. 中国农业科学, 2018, 51 (15): 149-159.
- [28] 李慧峰, 王涛, 冉昆. 利用 SSR 荧光标记构建 41 份山东省苹果资源分子身份证 [J]. 沈阳农业大学学报, 2020, 51 (1): 70-77.
- [29] 吴燕, 俞治州, 程春燕, 等. 黄瓜品种 (系) 遗传多样性分析及分子身份证构建 [J]. 分子植物育种, 2023 (4): 1-9.
- [30] 王富强, 张建, 温常龙, 等. 基于 KASP 标记的葡萄品种鉴定 [J]. 中国农业科学, 2021, 54 (13): 2830-2846.
- [31] 李书情. 基于全基因组 SNP 的桂花品种精准鉴定体系的构建及应用 [D]. 郑州: 河南农业大学, 2023.