

Bai X H, Liu K M, He X X, et al. Genetic diversity analysis of northern *Rhynchocypris lagowskii* populations based on genotyping by sequencing technology [J]. *Acta Hydrobiologica Sinica*, 2025, 49(11): 102503. [白晓慧, 刘克明, 何晓旭, 等. 基于GBS技术的北方地区拉氏大吻鲢群体遗传多样性分析 [J]. 水生生物学报, 2025, 49(11): 102503.]

基于GBS技术的北方地区拉氏大吻鲢群体遗传多样性分析

白晓慧¹ 刘克明¹ 何晓旭¹ 尤宏争¹ 骆小年² 汪笑宇¹ 薄其康¹ 张 韦¹

(1. 天津市水产研究所, 天津 300221; 2. 大连海洋大学, 辽宁 116023)

摘要: 为探究北方地区拉氏大吻鲢(*Rhynchocypris lagowskii*)群体的遗传多样性, 研究利用基因分型技术(Genotyping-by-sequencing, GBS), 对天津宁河、天津蓟州、天津沟河和辽宁太子河4个拉氏大吻鲢群体共109尾个体进行全基因组水平的单核苷酸多态性(Single nucleotide polymorphism, SNP)标记开发和遗传特征分析。通过测序共获得SNP标签3723530个, 经过滤后共得到高质量SNP位点133536个用于遗传多样性分析。4个群体的期望杂合度(H_e)为0.0397—0.0884, 观测杂合度(H_o)为0.0394—0.0818, 核苷酸多样性(P_i)为0.0406—0.0908, 多态信息含量(PIC)值为0.0327—0.0717, 群体间的遗传分化系数(F_{ST})为0.0039—0.8148, 群体内近交系数(F_{is})为0.0567—0.5480。遗传结构分析表明4个群体聚为3个明显的类群, 天津宁河群体和天津蓟州群体的个体聚为一支, 辽宁太子河群体和天津沟河群体的个体分别聚为一支。4个群体均发现了群体特异性SNP标记, 可以作为拉氏大吻鲢群体鉴定、群体遗传、性状关联分析和进化分析的候选分子标记。种群历史动态研究表明, 4个群体在第四纪冰川期均经历了有效群体大小(N_e)急剧下降过程。

关键词: GBS; 遗传多样性; SNP标记; 遗传结构; 种群动态; 拉氏大吻鲢

中图分类号: Q347 **文献标识码:** A **文章编号:** 1000-3207(2025)11-102503-10

doi: 10.3724/1000-3207.2025.2025.0097 **CSTR:** 32229.14.SSSWXB.2025.0097



拉氏大吻鲢(*Rhynchocypris lagowskii*), 也称拉氏鲢, 俗称柳根鱼, 隶属鲤形目, 鲤科, 雅罗鱼亚科, 鲢属, 是一种小型淡水经济鱼类, 在我国黑龙江、辽河、黄河、淮河、海河及长江中上游的支流等水域均有分布^[1]。拉氏大吻鲢肉质细嫩, 味道鲜美, 有很高的营养价值和经济价值, 被农业农村部遴选为2022年十大特色水产种质资源^[2, 3]。同时, 也是中国特有珍稀冷水鱼类——秦岭细鳞鲑(*Brachymystax tsinlingensis*)的主要饵料生物^[4]。近年来, 由于环境污染和过度捕捞, 其自然种群数量在逐渐减少, 导致市场供不应求^[3]。拉氏大吻鲢在东北地区人工繁殖和养殖已获得成功, 并取得了良好的经济效益^[5-7], 但养殖的苗种多为野生亲本驯化繁殖, 未经过系统的选育。目前, 对拉氏大吻鲢的研究主要集中在形态特征与分类^[8]、人工繁殖和养殖技术^[5, 6]、肌肉营养成分分析^[3]、染色体核型^[9]等。关于拉氏

大吻鲢群体遗传研究较少, 张帅^[9]基于线粒体D-loop基因分析了哈尔滨、吉林和本溪3个拉氏大吻鲢群体的遗传多样性, 杨培民等^[10]用线粒体D-loop和Cyt b基因序列研究了吉林延吉、敦化和辽宁法台拉氏大吻鲢群体的遗传多样性和遗传结构, 以上研究均表明东北地区的拉氏大吻鲢群体遗传多样性较高, 尚未见应用GBS技术对拉氏大吻鲢进行遗传多样性分析的报道。

传统群体遗传多样性分析多采用限制性片段长度多态性(RFLP)、扩增片段长度多态性(AFLP)、线粒体DNA(mtDNA)和微卫星(SSR)等遗传标记。单核苷酸多态性(SNP)是基于单核苷酸的突变, 在基因组中数量多、分布广, 且具有相对较高的遗传稳定性, 在动物遗传育种及遗传多样性评估中有着重要作用^[11]。随着分子生物学技术的发展, 涌现出多种可以获取SNP标记的高通量测序方法, 如

收稿日期: 2025-03-27; 修订日期: 2025-06-03

基金项目: 天津市科技计划(22ZYCGSN00820)资助 [Supported by the Tianjin Science and Technology Plan Project (22ZYCGSN00820)]

作者简介: 白晓慧(1982—), 女, 博士, 高级工程师; 研究方向为水产养殖与遗传育种。E-mail: bxh123bxh@163.com

通信作者: 张韦(1981—), 女, 推广研究员; 研究方向为水产养殖与遗传育种研究。E-mail: zhang-wei-316@163.com

GBS、2b-RAD、dd-RAD、SLAF等^[12]。GBS技术是简化基因组技术的一种, 可根据研究目的灵活调整所需捕获酶切位点的片段数(Tags), 从而获得大量高效的SNPs 标记。该方法成本低、效率高, 是近几年来水产动物种群进化、遗传结构分析和基因组选择的有效工具。已应用于半滑舌鳎(*Cynoglossus semilaevis*)^[13]、三角鲂(*Megalobrama terminalis*)、翘嘴鲌(*Culter alburnus*)、蒙古鲌(*Chanodichthys mongolicus*)^[14]、太平洋褶柔鱼(*Todarodes pacificus*)^[15]、托氏瑁螺(*Umbonium thomasi*)^[16]等水产动物SNP标记的开发和遗传多样性研究。本研究基于GBS 测序技术分析4个不同拉氏大吻鲢群体的遗传多样性和遗传结构, 旨在探明北方拉氏大吻鲢群体的遗传多样性水平和亲缘关系, 为拉氏大吻鲢种质资源保护、评价和群体选育提供依据。

1 材料与方法

1.1 实验材料

4个群体拉氏大吻鲢用于GBS分析, 2个养殖群体分别采自天津市宁河区养殖场(H群体)和天津市蓟州区养殖场(L群体), H群体2016年引自黑龙江又经4代繁育, L群体2021年引自辽宁辽阳, 又经1代繁育。2个野生群体分别采自天津市蓟州区沟河(Y群体)和辽宁省本溪市太子河(B群体; 表1)。实验鱼活体运回实验室, 用MS-222麻醉后, 剪取尾鳍迅速放于液氮中速冻, -80℃保存备用。

1.2 基因组DNA提取

利用天根生化科技(北京)有限公司的海洋动物组织基因组DNA提取试剂盒提取拉氏大吻鲢基因组DNA。DNA 样品经1%琼脂糖凝胶电泳和超微量紫外分光光度计(NanoDrop One)检测合格后, 将合格的DNA样品稀释至50 ng/μL, -20℃保存备用。

1.3 GBS文库构建

稀释好的DNA样品送至上海欧易生物医学科技有限公司进行建库测序^[17]。基因组DNA用限制

性内切酶*Pst*I-HF和*Msp*I进行双酶切, 设置酶切组和对照组, 37℃反应2h。酶切产物用T4 DNA连接酶加接头, 22℃连接2h, 使用0.7倍体积的Sera-Mag beads对连接产物进行纯化回收, 去除小于300 bp的小片段。对纯化后的连接产物进行PCR扩增, 程序为95℃ 30s, (95℃ 30s, 62℃ 20s, 68℃ 15s), 16个循环, 68℃ 5min, 4℃保存。使用 Qubit 测定 PCR 产物浓度, 对浓度大于5 ng/μL的合格样品, 取相同总量的DNA混合建库, 使用Illumina Novaseq PE150测序平台对混合好的文库进行测序。

1.4 数据质控及SNP鉴定

测序的原始数据(Raw data)去除接头, 过滤删除不含有限制性内切酶识别位点的序列, 含N量超过8%的碱基序列及低质量reads, 获得HQReads用于统计分析。使用Stacks软件^[18]对获得的GBS测序数据(HQReads)进行聚类及SNP分型, 并使用vcftools软件(version 0.1.14)^[19]对获得的SNP分型结果进行过滤, 剔除非二等位点, MAF 小于 0.01 的位点和SNP分型缺失率高于20%的位点, 得到样本的SNP分型结果。

1.5 群体遗传学分析

使用GCTA软件(version: 1.26.0)^[20]对获得的SNP标记进行主成分分析(Principal Component Analysis, PCA)。采用R包genepop^[21]计算每个群体的多态位点比例(P_o)、观测杂合度(H_o)、期望杂合度(H_e)、多态信息含量(PIC), 群体间遗传分化系数(F_{ST})和遗传距离(DR), 采用vcftools软件(Version 0.1.14)计算核苷酸多样性(P_i)。利用 ADMIXTURE软件^[22]进行群体结构分析, 根据 CV (Cross Validation Error) 确定的最优 K 值, 使用pong软件^[23]对所有个体进行种群结构聚类。通过treebest软件(Version: 1.9.2)^[24]计算距离矩阵, 采用邻接法(Neighbor-joining method)构建进化树, 进化树的可靠性通过bootstrap 法进行检验(重复1000次)。

1.6 特异性SNP位点筛选

对4个群体中缺失型的SNP位点进行过滤, 按照组内一致性100%进行筛选, 即组内所有样本均发生突变记为发生突变。对未筛选到特异性SNP位点的群体, 使用vcftools软件计算群体间每个SNP位点的 F_{ST} 值, 筛选 F_{ST} 值>0.25的位点, 即高度分化位点, 作为两个群体的差异SNP位点。

1.7 种群历史动态分析

使用 easySFS (<https://github.com/isaacovercast/easySFS>)软件和stairway plot (<https://github.com/xiaoming-liu/stairway-plot-v2/tree/master>)软件进行种群历史动态分析, 估算各群体的有效群体大小(N_e)。拉

表1 拉氏大吻鲢样本采集信息

Tab. 1 Sample information of *R. lagowskii*

群体 Population	采样地点 Sampling site	经纬度 Latitude and longitude	样本数量 Sample No.	采样部位 Sampling position
天津宁河群 体(H群体)	天津市宁河区 养殖场	39.40°N, 117.84°E	30	尾鳍
天津蓟州群 体(L群体)	天津市蓟州区 养殖场	40.03°N, 117.45°E	28	尾鳍
天津沟河群 体(Y群体)	天津市蓟州区 沟河	40.19°N, 117.45°E	26	尾鳍
辽宁太子河 群体(B群体)	辽宁省本溪市 太子河	41.13°N, 124.15°E	25	尾鳍

氏大吻鲢2年性成熟, 因此世代间隔时间(g)为2年, 每代核苷酸突变速率(μ)参考鲤(*Cyprinus carpio*)^[25], 为 3.51×10^{-9} 。

2 结果

2.1 数据质控及SNP鉴定

109个拉氏大吻鲢样本通过测序共得到数据量70.32 Gb, 过滤去除低质量数据后, 得到高质量数据56.44 Gb, 平均每个样本0.52 Gb, 其中Q20 $\geq$ 96.18%, Q30 $\geq$ 90.08%, 符合数据分析要求。测序共获得SNP标签3723530个, 平均测序深度28.01x, 经过滤后共得到SNP位点133536个。

2.2 遗传多样性分析

4个拉氏大吻鲢群体的遗传多样性指标见表2。4个群体 H_o 在0.0394—0.0818, H_e 在0.0397—0.0884, 天津蓟州群体 H_o 最高, 辽宁太子河群体 H_e 最高, 天津沟河群体杂合度水平最低。4个群体的PIC值在0.0327—0.0717, P_i 在0.0406—0.0908, 两个指标均以辽宁太子河群体最高, 天津沟河群体最低。天津沟河群体近交系数最大, 为0.5480, 天津宁河群体近交系数最小, 为0.0439。总体来看, 辽宁太子河群体和天津蓟州群体的遗传多样性水平要高于天津宁河群体和天津沟河群体。

2.3 群体间遗传分化和基因流

4个群体间的遗传分化系数(F_{ST})为0.0039—0.8148 (表3), 其中天津宁河群体和天津蓟州群体

表2 拉氏大吻鲢群体遗传多样性参数

群体 Population	观测杂合度 H_o	期望杂合度 H_e	多态信息含量 PIC	核苷酸多样性 P_i	近交系数 F_{is}
天津宁河群体(H群体)	0.0613	0.0643	0.0548	0.0655	0.0567
天津蓟州群体(L群体)	0.0818	0.0854	0.0717	0.0872	0.0636
天津沟河群体(Y群体)	0.0394	0.0397	0.0327	0.0406	0.5480
辽宁太子河群体(B群体)	0.0761	0.0884	0.0740	0.0908	0.1517

表3 拉氏大吻鲢群体的 F_{ST} 值(对角线下)与基因流 N_m (对角线上)

Tab. 3 Pairwise F_{ST} values (below diagonal) and N_m (above diagonal) among *R. lagowskii* populations

群体 Population	辽宁太子河 群体(B群体)	天津宁河 群体(H群体)	天津蓟州 群体(L群体)	天津沟河 群体(Y群体)
辽宁太子河 群体(B群体)	—	0.0883	0.0883	0.0662
天津宁河 群体(H群体)	0.7390	—	63.8526	0.0587
天津蓟州 群体(L群体)	0.7390	0.0039	—	0.0568
天津沟河 群体(Y群体)	0.7907	0.8098	0.8148	—

间的遗传分化最小, F_{ST} 值为0.0039。天津宁河群体和天津蓟州群体与其他2个群体的 F_{ST} 值在0.7390—0.8098, 遗传分化极大。进一步进行群体间基因流(N_m)分析显示, 天津宁河群体和天津蓟州群体间的基因流最大为63.8526, 二者与辽宁太子河群体和天津沟河群体间的基因流较小, 在0.0568—0.0883。

2.4 群体遗传结构分析

使用GCTA软件对获得的SNP标记进行主成分分析, 得到各群体样品的主成分聚类情况。如图1所示, 4个群体聚为3个明显的类群, 天津宁河群体和天津蓟州群体的个体聚为一群, 辽宁太子河群体和天津沟河群体的个体分别聚为2个不同的类群。

采用邻接法构建进化树, 4个群体的个体聚为3大支, 天津宁河群体和天津蓟州群体的个体聚为一支, 辽宁太子河群体和天津沟河群体的个体分别聚为一支(图2), 与主成分分析结果一致。

应用ADMIXTURE软件, 根据CV error (Cross validation error)最低点对应的K值来确定最佳分群数为3, 即所有样品划分为3个亚群, 遗传结构图如图3所示, 不同颜色代表不同的祖先血统, 在柱状图中的颜色比例代表样本包含的血统比例。由图3可知, 天津宁河群体和天津蓟州群体的个体来自相同的祖先血统, 辽宁太子河群体和天津沟河群体具有相对独立的祖先血统, 与系统进化树的分析结果相一致。

2.5 群体特异性SNP标记分析

对4个群体中缺失型的位点进行过滤, 按照组内一致性100%进行筛选, 即组内所有样本均发生突变记为发生突变。最终经过过滤, 辽宁太子河群体共筛选到42个contig57个特异性SNP位点, 其中

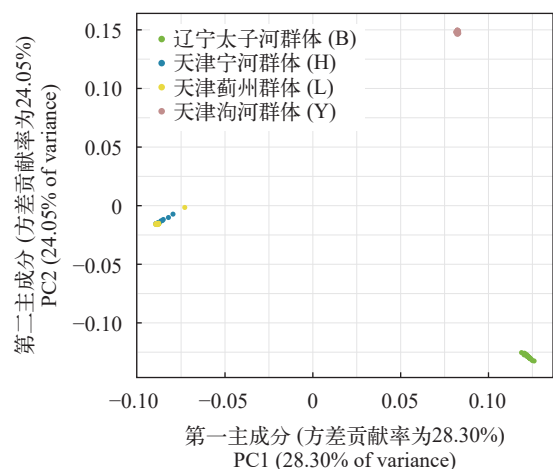


图1 拉氏大吻鲢4个群体的主成分分析

Fig. 1 Principal Components Analysis of four populations of *R. lagowskii*

contig124088突变位点最多,有6个位点发生突变。天津沟河群体共筛选到33个contig45个特异性SNP位点,其中contig56419突变位点最多,有5个位点发生突变,位点信息见表4。按照组内归一化方法,天津宁河群体和天津蓟州群体未筛选到特异性SNP位点,进一步筛选2个群体间高度分化(F_{ST} 值 >0.25)的位点,共得到9个contig12个SNP位点(表5)。根据贝叶斯信息准则(BIC)最佳建树模型SYM+ASC+G4构建系统进化树,可以较好地地区分天津宁河和天津蓟州2个群体。以上这些位点可以作为拉氏大吻鲃群体遗传分析和种群鉴定的候选分子标记。

2.6 种群历史动态

在4个拉氏大吻鲃群体中,辽宁太子河群体、

天津宁河群体和天津蓟州群体有效群体大小(N_e)变化趋势较一致,约在1000万年前经历了种群扩张,在约500万—600万年前 N_e 达到最大值,此后迅速下降,辽宁太子河群体 N_e 值下降最大,有效群体大小明显低于天津宁河群体和天津蓟州群体,直到4000年前左右,3个群体 N_e 值趋于平稳。天津沟河群体在约80万年前 N_e 达到最大值,此后缓慢下降,在1万年前 N_e 值趋于平稳。天津沟河群体在顶峰时的 N_e 值明显低于辽宁太子河群体、天津宁河群体和天津蓟州群体顶峰时的 N_e 值。在经历种群收缩后,约7万—8万年前至今,天津沟河群体 N_e 值高于其他3个群体,辽宁太子河群体 N_e 值最小(图4)。

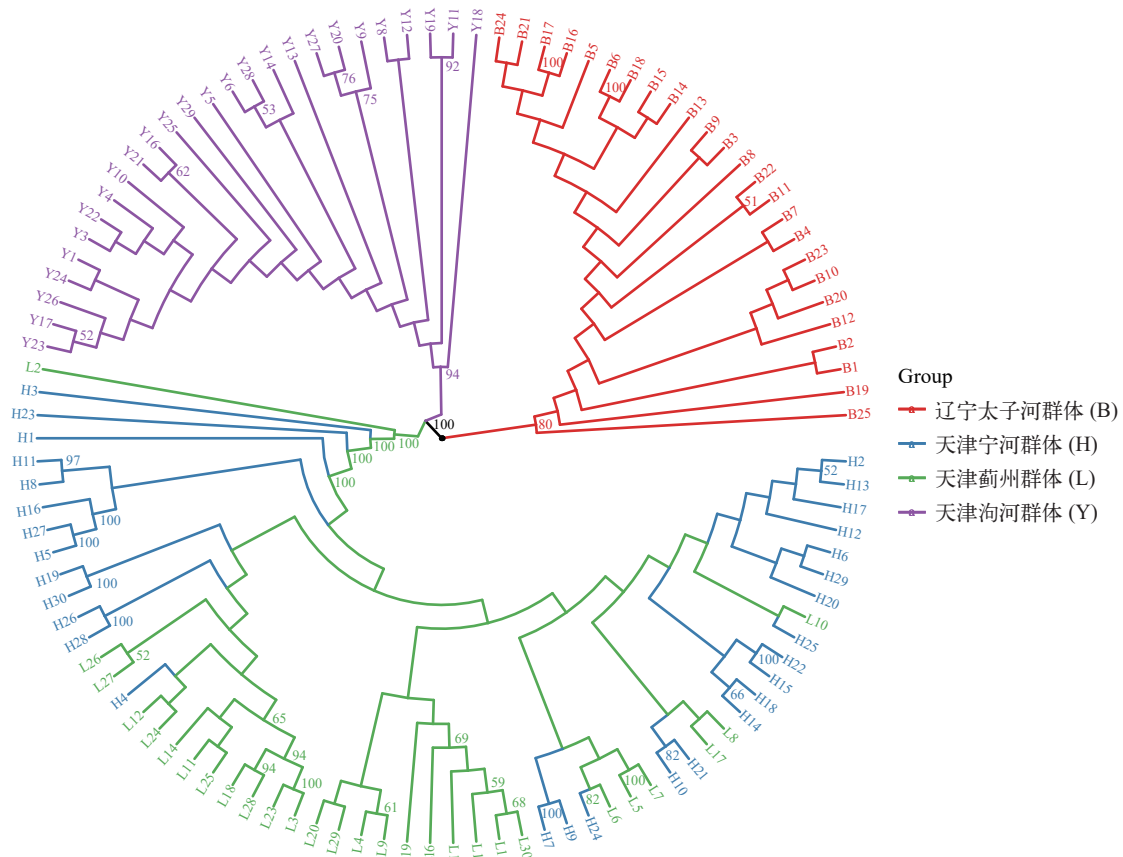


图2 4个拉氏大吻鲃群体的系统进化树

Fig. 2 Phylogenetic tree of four *R. lagowskii* populations

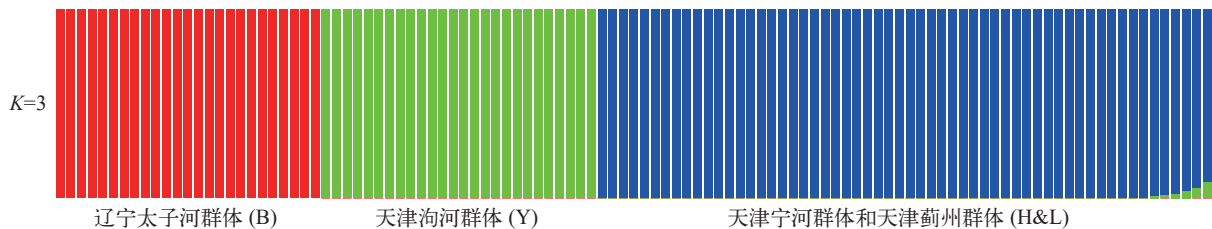


图3 K 值为3时的拉氏大吻鲃群体遗传结构图

Fig. 3 Genetic structure of *R. lagowskii* populations when K value is three

表 4 拉氏大吻鲢群体特异性SNP位点

Tab. 4 Specific SNP loci of *R. lagowskii* populations

辽宁太子河群体(B群体)				天津沟河群体(Y群体)			
Contig	POS	REF	ALT	Contig	POS	REF	ALT
contig8071	169	C	G	contig67	235	G	A
contig10725	119	G	T	contig7242	127	T	C
contig12947	272	A	C	contig14064	239	G	A
contig16578	177	A	G	contig16578	207	C	G
contig16923	148	A	G	contig16868	210	C	T
	205	A	C				
contig22163	186	A	T	contig16923	158	A	C
contig32454	207	C	T	contig20447	121	A	G
contig38778	267	A	C	contig21155	170	C	T
contig39454	105	T	C	contig40375	100	T	A
	239	A	G				
contig40375	204	C	T	contig41943	126	A	T
contig41020	215	G	C		194	G	A
contig41481	142	G	A	contig48378	195	A	C
contig43328	127	G	C		197	G	A
				contig49731	216	C	T
				contig56419	188	T	G
					200	C	T
					208	T	C
					237	T	C
					244	C	G
contig45138	151	T	C	contig58677	102	G	T
contig48126	229	A	G		214	C	G
contig49038	204	G	A	contig63059	164	A	C
contig49731	108	T	C		166	A	G
contig55825	187	A	T	contig71339	151	A	G
	190	G	C	contig81083	158	C	A
contig56419	192	G	A	contig87654	117	C	T
	206	T	C				
contig67697	123	G	T	contig89634	204	A	G
contig71027	218	C	G	contig90055	198	G	T
contig76406	161	C	T		247	A	T
contig81083	187	A	G	contig92214	137	C	G
contig90055	128	G	A	contig98978	176	C	T
contig90336	120	G	C	contig116251	199	T	C
	177	A	G	contig118188	217	T	C
contig97273	217	T	A	contig118643	257	T	C
contig97435	196	G	C				
contig105667	151	T	G	contig121453	101	G	T
contig109124	118	A	G		173	C	G
	119	G	A	contig124088	196	A	G
contig111190	223	A	G	contig129277	114	C	A
contig113654	195	G	A	contig130943	211	C	T
contig115007	142	C	G	contig131243	108	T	C
contig116322	153	G	A	contig137656	106	A	G
	158	A	C		193	G	T
	201	T	C	contig151944	138	G	C
contig119589	111	A	G		224	T	C
	237	T	A	contig158578	205	C	A
contig124088	200	T	G				
	232	T	G				
	244	C	T				
	253	T	A				

续表 4

辽宁太子河群体(B群体)				天津沟河群体(Y群体)			
Contig	POS	REF	ALT	Contig	POS	REF	ALT
contig129277	255	A	G				
	261	C	T				
	183	G	T				
contig133179	240	G	A				
contig136314	234	T	C				
contig137656	170	C	T				
contig141515	177	C	T				
	112	T	C				
contig150612	183	G	A				
contig151944	127	T	C				

注: Contig. 变异位点存在的重叠群; POS. 变异位点相对于参考基因组所在的位置; REF和ALT. 参考序列的碱基和相对应位置的突变碱基

Note: Contig. contig with mutation sites; POS. The position of the mutation site relative to the reference genome; REF& ALT. The base of the reference sequence and the mutated base at the corresponding position

表 5 天津宁河群体和天津蓟州群体间 F_{ST} 值>0.25的SNP位点

Tab. 5 SNP loci of F_{ST} >0.25 in Ninghe and Jizhou *R. lagowskii* populations

Contig	POS	REF	ALT	天津宁河群体(H群体)	天津蓟州群体(L群体)	F_{ST} 值
contig84061	137	G	A	AG GG AA	AA	0.39
	126	A	T	AT TT	AA	0.39
	71	A	C	AC CC AA	AA	0.38
	134	G	A	AG AA GG	GG	0.37
contig36093	78	T	A	AT TT AA	AT TT AA	0.37
contig4751	202	C	A	CC AC	CC AC AA	0.34
contig136154	213	C	A	AA AC	AA AC CC	0.31
contig54411	213	G	A	GG AG	GG AG AA	0.31
contig192971	139	G	T	TT GG GT	TT GG GT	0.28
contig54899	215	T	G	TT GG GT	TT GT	0.28
contig129435	140	C	T	CC CT	CC CT TT	0.26
contig118998	271	A	G	AA AG	AA AG GG	0.25

3 讨论

3.1 遗传多样性

遗传多样性是评估种群资源状况的重要参数,也是群体选育中基础群体选择的重要依据,遗传多样性高的种群其生存适应能力较强。GBS技术通过高通量测序可以获得覆盖整个基因组信息的大量 SNP 变异位点信息,特别适合大样本量,无参考基因且重复序列含量较高的物种的遗传多样性分析^[26, 27]。贾磊等^[13]通过GBS技术分析了半滑舌鲷野生群体和养殖群体的遗传多样性,结果表明野生群体的总体遗传多样性略高于养殖群体,个别养殖群体高于野生群体。罗慧等^[15]基于GBS技术筛选出的45266个高质量的SNP位点分析了青海湖水域6个地理群体青海湖裸鲤的遗传特征,结果表明6个群体属于一个集群,种群结构单一。本研究应用GBS技术获得了133536个高质量的SNP位点,分析

了北方地区4个拉氏大吻鲃群体的遗传多样性和遗传结构。在遗传多样性参数中,杂合度、多态信息含量(PIC)、核苷酸多样性(P_1)均是衡量群体遗传多样性的重要参数。当 $H_e > H_0$ 时,群体内可能发生近交或受到了人工选择的影响。本研究中4个群体的 H_e 均略高于 H_0 ,说明无论是野生群体还是养殖群体都有一定程度的近交。其中,天津沟河群体的近交系数(F_{is})最大,多态信息含量(PIC)最低,为中度多态性水平;其他3个群体的近交系数较小, PIC值均大于0.05,为高度多态性水平^[28]。在4个群体中,天津宁河群体原始群体2016年引自黑龙江,天津蓟州群体2021年从辽宁辽阳引进,其原始群体2013年引自黑龙江,两个群体即使经过多年的人工养殖,仍保持较高的遗传多样性水平。天津沟河群体和辽宁太子河群体虽同为野生群体,但天津沟河群体的杂合度、多态信息含量(PIC)、核苷酸多样性(P_1)等

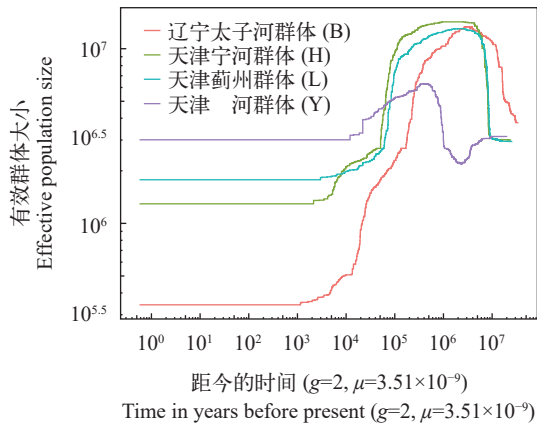


图4 四个拉氏大吻鲃群体有效群体大小历史变化图

Fig. 4 Historical changes in effective population size of four *R. lagowskii* populations

遗传多样性指标均明显低于辽宁太子河群体。张帅等^[9]和杨培民等^[10]基于线粒体D-loop和Cyt *b*基因研究了黑龙江、吉林、辽宁地区拉氏大吻鲃群体的遗传多样性,结果表明东北地区的拉氏大吻鲃群体遗传多样性较高,与本研究结果相一致。沟河为天津蓟运河的支流,在蓟运河的上游,2013年以来干旱少雨,河道周边基本为农田,由于畜禽养殖、水产养殖和农药化肥施用等活动,水质污染严重,2020年5个水质监测断面水质为V类或劣V类^[29]。有研究表明河流水质及河岸带等环境因素是限制拉氏大吻鲃幼鱼生存的重要因素^[30],由此推断,干旱少雨和水质污染可能给拉氏大吻鲃的栖息地造成了破坏。环境污染会引起种群规模的减少,由于随机的遗传漂变,降低种群的遗传多样性水平。即使种群最后实现了对新环境的适应,并恢复到原来的种群数量,也会由于奠基者效应(Founder effects),造成遗传来源单一,遗传变异性大大降低^[31]。因此,通过生态修复和水域生态恢复等手段,提高蓟运河水体的自净能力,实现蓟运河生态的全面复苏保护至关重要^[29]。

3.2 群体遗传结构

群体的遗传分化系数(F_{ST})是揭示种群间遗传分化程度的重要指标, F_{ST} 值在0—0.05表示群体间遗传分化很弱,在0.05—0.15表示遗传分化中等,0.15—0.15表示分化较大, F_{ST} 值大于0.25表示分化极大。天津宁河群体和天津蓟州群体间的 F_{ST} 值为0.0039,几乎没有遗传分化,两个群体间的基因流、系统进化树和遗传结构分析均显示两个群体的个体来自相同的祖先群体,两个群体的祖先群体均来自黑龙江,在不同的养殖环境下,祖先群体中的某些SNP位点仍然具有遗传稳定性,说明SNP标记是进行遗传多样性分析的准确且高效的遗传标记。

但经过多年不同环境下的人工选择和环境压力,群体内也发生了一定的遗传变异,从进化树可以看出,虽然两个群体聚为1大支,但在其内部每个群体内多数个体还是先聚为1小支,再和另一群体个体聚在一起。辽宁太子河群体和天津沟河群体间及二者与天津宁河群体和天津蓟州群体间的 F_{ST} 值均在0.7以上,遗传分化很大,表明辽宁、黑龙江和天津3个地理群体的拉氏大吻鲃遗传差异较大,可能是群体间栖息地环境的不同导致的。拉氏大吻鲃分布范围较广,北起黑龙江南至黄河、长江等水系的支流中均有分布,比较转录组研究表明,拉氏大吻鲃受温度、溶氧、pH等环境因子影响的基因得到了正向选择,在长期进化过程中产生对不同温度的适应性^[8],从而形成了相对独立的遗传结构。

3.3 SNP标记应用分析

本研究的4个群体均发现了群体特异性SNP标记,辽宁太子河群体和天津沟河群体依据单个的SNP位点就可以区分,天津宁河群体和天津蓟州群体虽同为一个祖先群体,但本研究筛选的高度分化的12个SNP位点组合,可以区分93%以上的两个群体个体,这些位点可以作为拉氏大吻鲃群体鉴定的分子标记。SNP标记是遗传变异中最常见的一种,在生物基因组中分布广泛,作为最新型的遗传标记技术,已在水产动物亲子鉴定、生物性状的关联分析、群体遗传分析等领域广泛应用。游淦元等^[32]基于192个SNP标记建立了长毛对虾亲子鉴定技术;研究人员已鉴定出多个与鱼类病毒病相关的候选SNP位点及抗逆性状相关的SNP位点^[33—34];谢玲莉等^[35]基于全基因组重测序获得的6548523个SNP位点,探讨了亚洲地区草鱼种群的遗传多样性和适应性进化。本研究获得的SNP标记也可作为拉氏大吻鲃群体遗传、性状关联分析和进化分析的候选分子标记,未来应加大采样范围,采用微卫星标记、基因组重测序等技术,更全面准确的揭示拉氏大吻鲃不同地理群体间可能存在的遗传差异及与环境因子相关的本地适应性现象。

3.4 种群历史动态分析

有效种群扩张和收缩的波动被认为是大多数鱼类的普遍特征。Li等^[36]利用高质量的基因组组装和测序数据,估算了约1000万年前至1万年前12个代表性硬骨鱼类的有效群体规模。结果表明,在新近纪和第四纪期间,大多数硬骨鱼类都经历了多轮种群收缩和扩张。本研究中除天津沟河群体外,其他3个群体均经历一轮扩张和收缩。多轮种群收缩和扩张可能与冰川期和间冰期的交替有关,且对生活在水温较高的水域上游的物种如尼罗罗

非鱼(*Oreochromis niloticus*)等鱼类的 N_e 值影响较大^[36]。拉氏大吻鲢温度耐受范围较广, 对温度变化有较强的适应性^[37], 冰川期和间冰期的温度变化可能对其有效群体大小影响有限。但自冰川期(1.2万—11万年前)开始以来, 日益剧烈的冰川作用, 使大多数鱼类的 N_e 值急剧下降, 本研究也有相似的结果。天津沟河群体有效群体经历两轮收缩, 可能与第四纪黄渤海地区发生的多次海侵海退事件有关, 海水的侵蚀会对淡水鱼的种群产生影响^[38]。本研究的种群历史动态分析是基于简化基因组数据, 与重测序相比数据量较小, 可能对结果产生一定的影响, 后续将基于深度基因组重测序, 对更多的群体开展分析, 更详细解析拉氏大吻鲢的种群分化和历史动态。(作者声明本文符合出版伦理要求)

参考文献:

- [1] Xu Z Y. Biological characteristics and breeding prospects of *Phoxinus lagowskii* [J]. *China Fisheries*, 2015(2): 62-64. [徐忠源. 拉氏鲢的生物学特性及养殖前景 [J]. 中国水产, 2015(2): 62-64.]
- [2] Zu X J, Liu Y H, Liu T G, et al. The analysis of the flesh rate and nutrient composition in the muscle of *Phoxinus lagowskii* Dybowski under artificial breeding [J]. *Journal of Inner Mongolia Agricultural University* (Natural Science Edition), 2016, 37(4): 6-13. [祖岫杰, 刘艳辉, 刘铁钢, 等. 池养拉氏鲢含肉率及营养成分分析 [J]. 内蒙古农业大学学报(自然科学版), 2016, 37(4): 6-13.]
- [3] Liu Y H, Zu X J, Liu T G, et al. Comparative analysis of nutrient composition in the muscle of *Phoxinus lagowskii* Dybowski with wild and artificial breeding [J]. *Journal of Jilin Agricultural University*, 2017, 39(1): 82-88. [刘艳辉, 祖岫杰, 刘铁钢, 等. 野生及养殖拉氏鲢营养成分比较分析 [J]. 吉林农业大学学报, 2017, 39(1): 82-88.]
- [4] Gou N N, Jin T Z, Zhang J L, et al. Population characteristics and seasonal variation of *Brachymystax lenok tsinlingensis*'s Main Bait Organism *Lagowskiella lagowskii* in national nature reserves of rare aquatic wildlife of Heihe River [J]. *Acta Agriculturae Boreali-occidentalis Sinica*, 2018, 27(9): 1258-1264. [苟妮娜, 靳铁治, 张建禄, 等. 黑河国家级自然保护区秦岭细鳞鲑主要饵料生物—拉氏鲢种群特征及其季节变化 [J]. 西北农业学报, 2018, 27(9): 1258-1264.]
- [5] Luo X N, Li J, Jin G H, et al. Artificial propagation of fat minnow *Phoxinus lagowskii* Dybowski [J]. *Fisheries Science*, 2013, 32(11): 673-675. [骆小年, 李军, 金广海, 等. 拉氏鲢人工繁殖试验 [J]. 水产科学, 2013, 32(11): 673-675.]
- [6] Luo X N, Li J, Yang P M, et al. Fry and juvenile rearing of fat minnow *Phoxinus lagowskii* Dybowski in an earthen pond [J]. *Fisheries Science*, 2014, 33(3): 186-189. [骆小年, 李军, 杨培民, 等. 拉氏鲢池塘苗种培育试验 [J]. 水产科学, 2014, 33(3): 186-189.]
- [7] Liu Y H, Chen W Q, Liu T G, et al. The current breeding status and development prospects of *Phoxinus lagowskii* [J]. *Fisheries Science & Technology Information*, 2015, 42(5): 255-257. [刘艳辉, 陈伟强, 刘铁钢, 等. 拉氏鲢的养殖现状及发展前景 [J]. 水产科技情报, 2015, 42(5): 255-257.]
- [8] Xie P. Morphological comparison and temperature adaptation of *Phoxinus* species in China [D]. Wuhan: Huazhong Agricultural University, 2019: 100. [谢鹏. 中国鲢属鱼类形态学比较和温度适应的研究 [D]. 武汉: 华中农业大学, 2019: 100.]
- [9] Zhang S, Jiang Y S, Wang F, et al. Karyotype of *Phoxinus lagowskii* [J]. *Journal of Guangdong Ocean University*, 2016, 36(3): 106-109. [张帅, 姜玉声, 王丰, 等. 洛氏鲢染色体核型分析 [J]. 广东海洋大学学报, 2016, 36(3): 106-109.]
- [10] Yang P M, Hu Z Y, Jin G H, et al. Genetic diversity and differentiation of 3 geographic populations of fat minnow *Phoxinus lagowskii* Dybowski revealed by D-loop and Cyt b gene analyses [J]. *Chinese Journal of Fisheries*, 2017, 30(4): 7-12. [杨培民, 胡宗云, 金广海, 等. 用线粒体D-loop和Cyt b基因序列分析拉氏鲢3个群体的遗传结构和遗传分化 [J]. 水产学杂志, 2017, 30(4): 7-12.]
- [11] 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. [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.]
- [12] Davey J W, Hohenlohe P A, Etter P D, et al. Genome wide genetic marker discovery and genotyping using next-generation sequencing [J]. *Nature Reviews Genetics*, 2011, 12(7): 499-510.
- [13] Jia L, Liu H, Zhang B, et al. Analysis of Genetic diversity of wild and cultivated populations of *Cynoglossus semilaevis* based on genotyping-by-sequencing technology [J]. *Transactions of Oceanology and Limnology*, 2020, 42(1): 107-114. [贾磊, 刘皓, 张博, 等. 基于GBS技术的半滑舌鳎野生与养殖群体遗传多样性分析 [J]. 海洋湖沼通报, 2020, 42(1): 107-114.]
- [14] Liu K, Feng X Y, Shen Y B, et al. Genetic structure analysis of *Megalobrama terminalis*, *Culter alburnus*, *Chanodichthys mongolicus* and their hybrids based on genotyping by sequencing [J]. *Journal of Fisheries of China*, 2021, 45(8): 1307-1316. [刘凯, 冯晓宇, 沈玉帮, 等. 基于基因分型测序(GBS)技术分析鲢鲈鱼类及其杂交子代的遗传结构 [J]. 水产学报, 2021, 45(8): 1307-1316.]
- [15] Liu Y, Hu X X, Lu H J, et al. Population genetics of *Todarodes pacificus* (Cephalopoda: Ommastrephidae) in the northwest Pacific Ocean via GBS sequencing [J]. *Open Life Sciences*, 2024, 19(1): 20220876.
- [16] Gao L, Bao X B, Yu S M, et al. Analysis of the genetic characteristics of *Umbonium thomasi* populations along

- the Yellow and Bohai Seas using GBS [J]. *Journal of Fishery Sciences of China*, 2020, **27**(2): 204-212. [高磊, 鲍相渤, 于思梦, 等. 基于GBS技术的黄渤海沿岸托氏瑁螺遗传特征分析 [J]. 中国水产科学, 2020, **27**(2): 204-212.]
- [17] Qi P, Gimode D, Saha D, *et al.* UGbS-Flex, a novel bioinformatics pipeline for imputation-free SNP discovery in polyploids without a reference genome: finger millet as a case study [J]. *BMC Plant Biology*, 2018, **18**(1): 117.
- [18] Rochette N C, Catchen J M. Deriving genotypes from RAD-seq short-read data using Stacks [J]. *Nature Protocols*, 2017, **12**(12): 2640-2659.
- [19] Danecek P, Auton A, Abecasis G, *et al.* The variant call format and VCFtools [J]. *Bioinformatics*, 2011, **27**(15): 2156-2158.
- [20] Yang J, Lee S H, Goddard M E, *et al.* GCTA: a tool for genome-wide complex trait analysis [J]. *The American Journal of Human Genetics*, 2011, **88**(1): 76-82.
- [21] Rousset F. Genepop'007: a complete re-implementation of the genepop software for Windows and Linux [J]. *Molecular Ecology Resources*, 2008, **8**(1): 103-106.
- [22] Alexander D H, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals [J]. *Genome Research*, 2009, **19**(9): 1655-1664.
- [23] Behr A A, Liu K Z, Liu-Fang G, *et al.* Pong: fast analysis and visualization of latent clusters in population genetic data [J]. *Bioinformatics*, 2016, **32**(18): 2817-2823.
- [24] Vilella A J, Severin J, Ureta-Vidal A, *et al.* EnsemblCompara GeneTrees: complete, duplication-aware phylogenetic trees in vertebrates [J]. *Genome Research*, 2009, **19**(2): 327-335.
- [25] Xu P, Zhang X F, Wang X M, *et al.* Genome sequence and genetic diversity of the common carp, *Cyprinus carpio* [J]. *Nature Genetics*, 2014, **46**(11): 1212-1219.
- [26] Nimmakayala P, Levi A, Abburi L, *et al.* Single nucleotide polymorphisms generated by genotyping by sequencing to characterize genome-wide diversity, linkage disequilibrium, and selective sweeps in cultivated watermelon [J]. *BMC Genomics*, 2014, **15**(1): 767.
- [27] You H Z, Meng M H, Ma L, *et al.* Analysis of genetic diversity and population structure of three wild *Urechis unicinctus* populations based on GBS sequencing [J]. *Journal of Sichuan Agricultural University*, 2024, **42**(5): 1145-1151. [尤宏争, 孟睦涵, 马林, 等. 基于GBS测序的3个野生单环刺螠群体遗传多样性和群体结构分析 [J]. 四川农业大学学报, 2024, **42**(5): 1145-1151.]
- [28] Botstein D, White R L, Skolnick M, *et al.* Construction of a genetic linkage map in man using restriction fragment length polymorphisms [J]. *The American Journal of Human Genetics*, 1980, **32**(3): 314-331.
- [29] Wang Z, Liu H T. Practical significance of river management to the water environment of rivers and lakes: a case study of ten years changes in the Jiyun River [J]. *Water Resources Development Research*, 2024, 1-13. [王卓, 刘海涛. 河务管理对河湖水环境的现实意义分析 以蓟运河十年变化为例 [J]. 水利发展研究, 2024, 1-13.]
- [30] Kang X, Zhang Y, Zhang N, *et al.* Assessment of habitat suitability of juvenile *Phoxinus lagowskii* in Taizi river [J]. *Asian Journal of Ecotoxicology*, 2011, **6**(3): 310-320. [康鑫, 张远, 张楠, 等. 太子河洛氏鲮幼鱼栖息地适宜度评估 [J]. 生态毒理学报, 2011, **6**(3): 310-320.]
- [31] Zheng S Y. Influence of environmental pollution on biodiversity [J]. *Bulletin of Biology*, 2002, **37**(5): 24. [郑世英. 环境污染对生物多样性的影响 [J]. 生物学通报, 2002, **37**(5): 24.]
- [32] You S Y, Yang L, Zeng Q M, *et al.* Development of SNP markers and parentage assignment techniques in *Penaeus penicillatus* [J]. *Oceanologia Et Limnologia Sinica*, 2023, **54**(4): 1165-1171. [游淞元, 杨乐, 曾庆民, 等. 长毛对虾 (*Penaeus penicillatus*) SNP标记开发及亲子鉴定技术研究 [J]. 海洋与湖沼, 2023, **54**(4): 1165-1171.]
- [33] Huang Y, Li Z, Li M, *et al.* Fish genomics and its application in disease-resistance breeding [J]. *Review in Aquaculture*, 2024(17): e12973.
- [34] Nie M M, Hu J W, Lu Y L, *et al.* Cold effect analysis and screening of SNPs associated with cold-tolerance in the olive flounder *Paralichthys olivaceus* [J]. *Journal of Applied Ichthyology*, 2019, **35**(4): 924-932.
- [35] Xie L L, Shen Y B, Gui L, *et al.* Genetic diversity and adaptation mechanism of *Ctenopharyngodon idella* in Asian by whole-genome resequencing [J]. *Journal of Fisheries of China*, 2025, **49**(1): 38-52. [谢玲莉, 沈玉帮, 桂朗, 等. 基于全基因组重测序技术探究亚洲草鱼遗传多样性及其适应机制 [J]. 水产学报, 2025, **49**(1): 38-52.]
- [36] Li J, Bian C, Yi Y, *et al.* Temporal dynamics of teleost populations during the Pleistocene: a report from publicly available genome data [J]. *BMC Genomics*, 2021, **22**(1): 490.
- [37] Bo Q K, Bai X H, Wang X Y, *et al.* Effect of temperature acclimation on growth and thermal tolerance of juvenile *Rhynchocypris lagowskii* [J]. *Fisheries Science*, 2025, **44**(1): 83-90. [薄其康, 白晓慧, 汪笑宇, 等. 温度驯化对拉氏大吻鲃幼鱼生长和温度耐受性的影响 [J]. 水产科学, 2025, **44**(1): 83-90.]
- [38] Xiong W, Huang L, Zhang Y, *et al.* Quaternary transgression process controlled by tectonic subsidence over the last 1.35 Ma: New insights from the eastern Bohai Sea [J]. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 2025(657): 112602.

GENETIC DIVERSITY ANALYSIS OF NORTHERN *RHYNCHOCYPRIS LAGOWSKII* POPULATIONS BASED ON GENOTYPING BY SEQUENCING TECHNOLOGY

BAI Xiao-Hui¹, LIU Ke-Ming¹, HE Xiao-Xu¹, YOU Hong-Zheng¹, LUO Xiao-Nian², WANG Xiao-Yu¹,
BO Qi-Kang¹ and ZHANG Wei¹

(1. Tianjin Fisheries Research Institute, Tianjin 300221, China; 2. Dalian Ocean University, Liaoning 116023, China)

Abstract: *Rhynchocypris lagowskii* is a small freshwater economic fish species selected by the Ministry of Agriculture and Rural Affairs as one of the top 10 characteristic aquatic germplasm resources in 2022. To investigate the genetic diversity of the *R. lagowskii* populations in northern China and provide references for its conservation and genetic breeding, Genotyping by sequencing (GBS) was used to develop single nucleotide polymorphism (SNP) markers and analyze genetic characteristics of 109 individuals from four *R. lagowskii* populations in Ninghe, Jizhou, Juhe (Tianjin), and Taizihe (Liaoning). Through sequencing, a total of 3723530 SNP markers were obtained, and 133536 high-quality SNP loci were screened for genetic analysis after filtering. The expected heterozygosity (H_e) of the four populations was 0.0397—0.0884, the observed heterozygosity (H_o) was 0.0394—0.0818, the nucleotide diversity (P_i) was 0.0406—0.098, the polymorphic information content (PIC) value was 0.0327—0.0717, the genetic differentiation coefficient (F_{ST}) was 0.0039—0.8148, and the inbreeding coefficient (F_{is}) within population was 0.0567—0.5480. Genetic structure analysis showed that four populations clustered into three distinct groups, with individuals from the Ninghe and Jizhou populations in Tianjin clustered into one branch, and individuals from the Taizihe and Juhe populations in Liaoning and Tianjin clustered into one branch, respectively. Population specific SNP markers were identified for all four populations, serving as candidate molecular markers for species identification, population genetics, trait association analysis, and evolutionary studies. Population history dynamics research indicates that the four populations of *R. lagowskii* all underwent a process characterized by a sharp decline in effective population size (N_e) during the Quaternary glacial period.

Key words: GBS; Genetic diversity; SNP marker; Genetic structure; Population dynamics; *Rhynchocypris lagowskii*