

分类号：
学号：20222013015

密级：
单位代码：10759

石河子大学

硕士学位论文



绵羊适应性相关转座元件的鉴定及组织表达分析

学位申请人	梁千千
指导教师	韩吉龙副教授
申请学位门类级别	农学硕士
学科、专业名称	畜牧学
研究方向	动物遗传育种与繁殖
所在学院	动物科技学院

中国·新疆·石河子
2026年6月

分类号：
学号：20222013015

密级：
单位代码：10759

石河子大学

硕士学位论文



绵羊适应性相关转座元件的鉴定及组织表达分析

学位申请人	梁千千
指导教师	韩吉龙副教授
申请学位门类级别	农学硕士
学科、专业名称	畜牧学
研究方向	动物遗传育种与繁殖
所在学院	动物科技学院

中国·新疆·石河子
2026年6月

**Identification and Tissue Expression Analysis of Transposable
Elements Associated with Adaptation in Sheep**

A Dissertation Submitted to

Shihezi University

In Partial Fulfillment of the Requirements

for the Degree of

Master of Agriculture

By

Liang qian-qian

(Animal Genetics, Breeding and Reproduction)

Dissertation Supervisor: Assoc. Prof. Han Ji-long

August, 2026

课题来源

石河子大学国际科技合作项目

课题名称：中亚肥臀型绵羊转座元件微进化分析及环境适应性机制

课题编号：GJHZ202307

石河子大学学位论文独创性声明及使用授权声明

学位论文独创性声明

本人所提交的学位论文是在我导师的指导下进行的研究工作及取得的研究成果。据我所知，除文中已经注明引用的内容外，本论文不包含其他个人已经发表或撰写过的研究成果。对本文的研究做出重要贡献的个人和集体，均已在文中作了明确的说明并表示谢意。

研究生签名：梁千仟

时间：2026年5月21日

使用授权声明

本人完全了解石河子大学有关保留、使用学位论文的规定，学校有权保留学位论文并向国家主管部门或指定机构送交论文的电子版和纸质版。有权将学位论文在学校图书馆保存并允许被查阅。有权自行或许可他人将学位论文编入有关数据库提供检索服务。有权将学位论文的标题和摘要汇编出版。保密的学位论文在解密后适用本规定。

研究生签名：梁千仟

时间：2026年5月21日

导师签名：韩吉龙

时间：2026年5月21日

摘要

目的：绵羊是最早被人类驯化的家畜之一，在全球畜牧业中占有重要地位。中国幅员辽阔、环境多样，在不同自然条件下形成了众多具有独特特征和良好适应能力的本土绵羊品种。解析其适应性背后的遗传机制，对于保护我国绵羊的遗传多样性、科学利用绵羊遗传资源具有重要意义。转座元件是推动基因组适应进化的潜在动力。本研究基于泛基因组策略，对绵羊转座元件进行全基因组鉴定与组织表达分析，旨在挖掘与本土适应性相关的转座元件衍生结构变异（TEV），为解析绵羊环境适应的分子机制提供新线索。

方法：本研究首先基于 NCBI 数据库中所有已组装的 16 个绵羊基因组及 1 个参考基因组（Oar_v4.0），采用“map-to-pan”迭代组装策略构建绵羊泛基因组，通过序列比对鉴定非参考基因组来源的新序列，并利用结构变异（SV）检测流程识别插入、缺失等变异类型，分析其组成特征及转座元件来源。在此基础上，对哈萨克斯坦主要绵羊品种和多浪羊进行全基因组重测序，并结合收集的 324 只代表性绵羊样本（包括 12 只亚洲摩弗伦羊和 7 只盘羊）进行结构变异（SV）分型，构建 SV 数据集，进而开展群体遗传结构分析，解析绵羊群体的遗传分化与迁徙历史。进一步采用冗余分析（RDA）和群体分化指数（FST）的方法，筛选与环境适应性（寒冷、干旱、海拔）及重要性状（繁殖力、季节性发情）相关的候选位点。为进一步解析转座元件的生物学功能，本研究采集绵羊 11 种组织（包括尾脂、皮肤、胸腺、大脑皮层、肝脏等）及尾部发育关键时期（胚胎期 E40-60，对应尾部延伸阶段；E70-75，对应脂肪沉积阶段）的样本进行转录组测序。通过系统分析转座元件的时空表达谱，揭示其组织特异性与发育阶段动态表达特征；进一步采用共表达分析，探究转座元件与邻近基因的表达相关性，并结合其基因组定位特征，为解析其潜在的顺式调控机制提供线索。

结果：（1）泛基因组分析成功构建了绵羊泛基因组，共鉴定出 378,900 个结构变异，其中 63.42% 来源于转座元件衍生，同时鉴定出 254 Mb 非参考序列，从中预测获得 1,421 个新基因。（2）群体遗传结构分析揭示了家羊自西亚起源后，向西形成欧洲古老支系，向东经中亚扩散至东亚，向南进入非洲的迁徙历史。（3）通过 RDA 分析，共鉴定出与寒冷、干旱、高海拔适应相关的 SVs 分别有 2,417 个、2,856 个和 2,595 个；筛选出 *IKZF2*、*ADCY1*、*C2H9orf43*、*CMSS1*、*KIAA1715*、*ITGB5* 等基因附近的候选 SV 位点。（4）选择信号分析筛选出与季节性发情相关的变异（位于 *VIPR2*、*MC2R* 基因附近），以及与高繁殖力相关的变异（位于 *FSHR*、*BMP2* 基因附近）。（5）转录组分析显示，转座元件呈现显著的时空特异性表达模式，并筛选获得三个与尾型分化相关的候选转座元件：*DCT* 基因 UTR3 区的 DNA/TcMar-Tigger 元件及 *ALOX15* 基因 UTR3 区的 LINE/RTE-BovB 和 LINE/L1 元件。（6）尾脂发育动态分析表明，尾部伸长阶段（E40-45 至 E55-60）转座元件以上调为主；尾部脂肪形成阶段（E55-60 至 E70-75）转座元件表达呈现与基因相反的模式，仅 371 个上调、5,183 个显著下调。功能富集结果显示，尾部伸长阶段转座元件相关基因富集于形态发生过程，尾部脂肪形成阶段转变为生长与分化调控通路。基于此，进一步筛选获得 10 个候选转座元件。

结论：本研究整合泛基因组学与转录组学技术，系统揭示了转座元件在绵羊适应性进化与尾脂沉积中的双重功能。在基因组层面，转座元件贡献了绵羊 63.42% 的结构变异，是群体遗传多样性重要来源，基于此构建的群体遗传结构清晰地勾勒出家羊的迁徙历史。同时，部分与环境适应性相关

的结构变异定位于关键功能基因调控区，提示转座元件可能通过插入突变参与寒冷、干旱、高海拔等环境胁迫的适应性应答。在表达调控层面，转座元件呈现时空特异性和发育阶段特异性的时空表达模式，主要分布于非编码区域，通过与邻近基因存在广泛的共表达相关性（以正相关为主）。结合基因组定位分析，推测其可通过作为替代启动子、增强子或边界元件发挥顺式调控作用，从而影响尾部椎骨发育及脂肪沉积过程。上述发现为解析转座元件介导的环境适应与尾部脂肪沉积协同调控机制提供了新的线索。

关键词：绵羊；泛基因组；转座元件；环境适应性；尾脂沉积

Abstract

Objective: Sheep were among the first domesticated livestock and play a significant role in global animal husbandry. China has a vast territory and diverse environments, which have given rise to numerous indigenous sheep breeds with unique characteristics and strong adaptability under different natural conditions. Elucidating the genetic mechanisms underlying their adaptability is of great importance for conserving the genetic diversity of Chinese sheep and for the scientific utilization of sheep genetic resources. Transposable elements (TE) are potential drivers of genomic adaptive evolution. This study employs a pan-genome strategy to perform genome-wide identification and tissue expression analysis of sheep TE, aiming to uncover transposable element-derived structural variants (TEV) associated with local adaptability and to provide new clues for deciphering the molecular mechanisms of environmental adaptation in sheep.

Method: In this study, we first constructed a sheep pan-genome using a “map-to-pan” iterative assembly strategy based on all 16 assembled sheep genomes available in the NCBI database, along with one reference genome (Oar_v4.0). Through sequence alignment, we identified novel sequences not present in the reference genome, and employed a structural variant (SV) detection pipeline to characterize variation types such as insertions and deletions, analyzing their compositional features and transposable element origins. On this basis, whole-genome resequencing was performed for major Kazakh sheep breeds and Duolang sheep. Combined with collected genomic data from 324 representative sheep samples (including 12 Asian mouflon and 7 argali), we performed SV genotyping to construct an SV dataset, followed by population genetic structure analysis to elucidate the genetic differentiation and migratory history of sheep populations. Furthermore, redundancy analysis (RDA) and population differentiation index (FST) approaches were applied to identify candidate loci associated with environmental adaptability (cold, drought, altitude) and important traits (fecundity, seasonal estrus). To further investigate the biological functions of transposable elements, we collected samples from 11 tissues (including tail fat, skin, thymus, cerebral cortex, liver, etc.) and from key developmental stages of the tail (embryonic stages E40–60, corresponding to tail elongation; E70–75, corresponding to fat deposition) for transcriptome sequencing. Through systematic analysis of the spatiotemporal expression profiles of transposable elements, we revealed their tissue-specific and developmental stage-specific dynamic expression patterns. Co-expression analysis was further employed to explore the expression correlation between transposable elements and neighboring genes, and combined with their genomic localization features, to provide insights into their potential cis-regulatory mechanisms.

Result: (1) Pan-genome analysis successfully constructed a sheep pan-genome, identifying a total of 378,900 structural variants (SV), of which 63.42% were derived from transposable elements. Additionally, 254 Mb of non-reference sequences were identified, from which 1,421 novel genes were predicted. (2) Population genetic structure analysis revealed the migratory history of domestic sheep, originating in West Asia, subsequently giving rise to ancient European lineages to the west, dispersing eastward into East Asia

via Central Asia, and migrating southward into Africa. (3) Through RDA analysis, a total of 2,417 SVs associated with cold adaptation, 2,856 SVs associated with drought adaptation, and 2,595 SVs associated with high-altitude adaptation were identified. Candidate SV loci near genes such as *IKZF2*, *ADCY1*, *C2H9orf43*, *CMSS1*, *KIAA1715*, and *ITGB5* were screened out. (4) Selection signature analysis identified variants associated with seasonal estrus (located near the *VIPR2* and *MC2R* genes), as well as variants associated with high fecundity (located near the *FSHR* and *BMP2* genes). (5) Transcriptome analysis revealed that transposable elements exhibit significant spatiotemporally specific expression patterns. Three candidate transposable elements associated with tail type differentiation were identified: a DNA/TcMar-Tigger element in the 3' UTR region of the *DCT* gene, as well as an LINE/RTE-BovB element and an LINE/L1 element in the 3' UTR region of the *ALOX15* gene. (6) Dynamic analysis of tail fat development showed that during the tail elongation stage (E40-45 to E55-60), transposable elements were predominantly upregulated. During the tail fat formation stage (E55-60 to E70-75), transposable elements exhibited an expression pattern opposite to that of genes, with only 371 upregulated and 5,183 significantly downregulated. Functional enrichment results indicated that genes associated with transposable elements during the tail elongation stage were enriched in morphogenetic processes, whereas during the tail fat formation stage, they shifted to pathways involved in growth and differentiation regulation. Based on these findings, 10 candidate transposable elements were further identified.

Conclusion: This study integrates pan-genomic and transcriptomic technologies to systematically reveal the dual role of transposable elements in the adaptive evolution and tail fat deposition of sheep. At the genomic level, transposable elements contribute to 63.42% of structural variants in sheep, serving as a significant source of population genetic diversity. The population genetic structure constructed based on these variants clearly delineates the migratory history of domestic sheep. Furthermore, some structural variants associated with environmental adaptability are located in regulatory regions of key functional genes, suggesting that transposable elements may participate in adaptive responses to environmental stresses such as cold, drought, and high altitude through insertional mutagenesis. At the expression regulatory level, transposable elements exhibit spatiotemporally specific and developmental stage-specific expression patterns, are predominantly distributed in non-coding regions, and show extensive co-expression correlations (predominantly positive) with neighboring genes. Combined with genomic localization analysis, it is hypothesized that transposable elements may exert cis-regulatory roles by acting as alternative promoters, enhancers, or boundary elements, thereby influencing caudal vertebral development and fat deposition processes. These findings provide new insights into the coordinated regulatory mechanisms of environmental adaptation and tail fat deposition mediated by transposable elements.

Key words: Sheep; Pan-genome; Transposable elements; Environmental adaptability; Tail Fat Deposition

目录

摘要	I
Abstract	IV
缩略词表	VIII
第一章 绪论	1
1.1 文献综述	1
1.1.1 绵羊适应性的研究进展	1
1.1.2 转座元件的研究进展	4
1.1.3 泛基因组的研究进展	10
1.2 本研究的目的是和意义	14
1.3 研究内容	14
1.4 技术路线	14
第二章 绵羊泛基因组的构建及转座元件鉴定	16
2.1 材料与方法	16
2.1.1 试验样品	16
2.1.2 构建绵羊图泛基因组	20
2.1.3 非参考序列的提取和注释	20
2.1.4 基因组转座元件库构建	20
2.1.5 基因组转座元件特征分析	22
2.1.6 绵羊重测序样本的 SV 检测	23
2.1.7 基于泛基因组变异的群体遗传结构分析	24
2.1.8 RDA 分析环境适应性	25
2.1.9 选择信号分析	25
2.1.10 与适应性相关候选位点验证	26
2.2 结果	27
2.2.1 绵羊基因组中转座元件的特征分析	29
2.2.2 泛基因组中的结构变异统计	37
2.2.3 基于泛基因组的结构变异群体分型	40
2.2.4 群体遗传结构分析	42
2.2.5 筛选与环境适应性相关的 SV	44
2.2.6 环境适应性相关的 SV 验证	53

2.3 讨论	57
2.4 结论	58
第三章 绵羊早期胚胎尾脂形成中转座元件的动态表达	60
3.1 材料与方法	60
3.1.1 公共数据下载	60
3.1.2 实验动物样品采集	61
3.1.3 RNA 提取、文库构建与测序	62
3.1.4 测序数据质控、比对、组装与定量	62
3.1.5 转座元件-基因共表达相关性分析	62
3.1.6 差异表达基因及转座元件邻近基因的功能注释与通路富集分析	63
3.1.7 候选转座元件的验证	63
3.2 结果	64
3.2.1 不同组织转座元件表达谱	64
3.2.2 不同尾型差异转座元件分析	65
3.2.3 不同发育时期尾脂差异转座元件分析	66
3.2.4 候选转座元件表达模式的 RT-qPCR 验证	74
3.3 讨论	75
3.4 结论	77
第四章 全文结论	78
第五章 论文创新点	79
参考文献	80
附录	94

缩略词表

英文缩略词	英文名称	中文全称
CNV	Copy Number Variation	拷贝数变异
DNA	DeoxyriboNucleic Acid	脱氧核糖核酸
ERV	Endogenous Retrovirus	内源性逆转录病毒
FST	Fixation Index	群体间分化指数
GO	Gene ontology	基因本体
ICE	Integrating conjugative elements	整合结合元件
KEGG	Kyoto Encyclopedia of Genes and Genomes	京都基因和基因组百科全书
LINE	Long interspersed nuclear elements	长散在核元件
lncRNA	Long Noncoding RNA	长非编码 RNA
LTR	Long terminal repeat	长末端重复序列
Mb	megabase	兆碱基
mRNA	Messenger RNA	信使 RNA
PAV	Presence/Absence Variation	存在/缺失变异
PCA	Principal components analysis	主成分分析
QTL	Quantitative Trait Locus	数量性状位点
RDA	Redundancy Analysis	冗余分析
RNA	Ribonucleic Acid	核糖核酸
RNA-seq	RNA sequencing	转录组测序
RT-qPCR	Reverse Transcription Quantitative PCR	逆转录实时荧光定量 PCR
SINE	Short interspersed nuclear elements	短散在核元件
siRNA	Small interfering RNA	小干扰 RNA
SNP	Single Nucleotide Polymorphism	单核苷酸多态性
SV	Structural Variation	结构变异
TE	Transposable element	转座元件
TEV	Transposable Element-Related Structural Variants	转座元件相关结构变异

第一章 绪论

1.1 文献综述

1.1.1 绵羊适应性的研究进展

绵羊作为最早被人类驯化的家养动物之一,其驯化历史可追溯至约 12,000 年前的近东和中东地区^[1],随着人类的迁徙扩散至全球的各个生态区域。不同地区的绵羊为适应各自的生存压力,逐渐演化出独特的形态特征。我国具有丰富的地形地貌和多样的生态环境,绵羊经过长期对我国不同生态环境(寒冷、干旱、高海拔等)的适应,使它们在外形特征方面出现了明显的适应性改变^[2]。

全球气候变化深刻影响着家畜在极端环境下的生存,高温与强降水等天气现象会显著削弱家畜的生产性能和繁殖能力,给畜牧业带来严重冲击^[3]。同时,热应激与高湿环境会导致家畜死亡率上升、生产性能下降,进而造成畜牧业经济损失。因此,解析家畜的关键基因与适应机制,对于科学评估其生产性能、维护遗传多样性具有关键作用^[4]。

1.1.1.1 绵羊寒冷适应性研究进展

低温环境对家畜的生存与繁殖构成重要挑战,其中新生羔羊受影响最为明显,在寒冷条件下易出现初生重偏低、死亡率增高等问题^[5]。就绵羊而言,冷适应的形态特征主要表现为被毛密度、厚度、长度以及尾型的变化。被毛密度高、长度长可有效减少热量散失,有助于适应寒冷环境^[6]。绵羊的脂尾具有脂肪沉积功能,而脂肪代谢是动物抵御寒冷的重要途径之一^[7]。脂尾形态是绵羊适应寒冷的关键性状之一。在绵羊驯化过程中,寒冷气候作为主要选择压力,促使代谢相关基因中与冷适应性相关的功能变异得以保留。Ahmad 等对比了高寒地区 Changthangi 绵羊与平原地区绵羊品种,在受选择信号区域内识别出 UCP 基因家族,该家族参与寒冷诱导的非颤抖性产热,提示其与 Changthangi 的冷适应性相关^[8]。Jiao 等基于阿勒泰羊与湖羊的转录组分析,鉴定出 *UCP1*、*ADRB3*、*ADORA2A*、*ATP2A1*、*RYR1* 和 *IP6K1* 等产热相关候选基因,这些基因涉及脂尾脂肪组织褐变、脂肪分解及下丘脑体温调节中枢产热等功能^[9],进一步佐证了绵羊尾型与冷适应性之间的潜在关联。此外,Fan 等通过比较藏羊夏冬两季的瘤胃微生物群落结构发现,冬季藏羊瘤胃微生物群落更倾向于分解营养价值较低的饲草,以维持宿主的能量代谢^[10]。

1.1.1.2 绵羊干旱适应性研究进展

干旱环境引发的水源短缺抑制植被生长,进而降低饲草营养价值^[11]。基于此,能量代谢与营养调控或成为家畜适应干旱环境的关键生理机制。Mwacharo 等通过对沙漠肥尾绵羊、东非肥尾羊、西亚肥尾羊及欧洲绵羊的全基因组比较,识别出 172 个与干旱适应相关的候选基因,揭示了肥尾绵羊通过多种途径实现干旱适应的复杂遗传机制^[12]。Yang 等对 77 只绵羊进行基因组分析,筛选出 15 个可能参与沙漠气候适应的功能基因^[13]。其中 *ANXA6*、*GPX3*、*GPX7*、*PTGS2*、*CPA3*、*CPVL*、*ECE1*、*CALM2*、*CACNA2D1*、*KCNJ5* 和 *COX2* 参与调控肾细胞及肾脏血管的水分重吸收与滞留,*RAP1A*、*SLC4A4*、*CPA3*、*CPB1* 则与胰腺分泌相关^[13]。我国塔克拉玛干沙漠与塔里木盆地地区气候干旱少雨,却孕育出独具特色的地方绵羊品种。Zhang 等基于五个农业生态区绵羊群体的选择信号分析与转录组测序,揭示了沙漠绵羊品种在极端干旱环境下维持常年发情的遗传调控机制^[3]。Wang 等结合塔里木盆地三个本土绵羊品种与三个引进品种的芯片数据,通过选择信号分析鉴定出多个与繁殖性状、生长性状及干旱适应相关的基因,揭示了该区域绵羊的遗传特征^[14]。干旱环境常伴随强日照、强紫外线辐射与高温等多重因素,这些环境条件会直接影响家畜的生理机能和繁殖表现。因此,家畜在干旱环境下的适应能力,通常与其对炎热、高海拔等极端环境的适应性紧密相关。目前,关于绵羊适应干旱环境所涉及的功能基因及其遗传基础,仍有待深入探究。

1.1.1.3 绵羊高海拔适应性研究进展

青藏高原平均海拔超过 4000 米,其高海拔环境带来低氧等特殊生态压力,该地区每次呼吸的平均含氧量仅为海平面的 60%左右^[15]。低氧条件易干扰动物新陈代谢,增加高原疾病的发生风险。我国重要牧区主要分布于高海拔地区,因此,阐明绵羊高海拔适应性相关基因及其作用机制,对于推动绵羊产业发展具有重要意义。

藏羊作为我国青藏高原特有的畜禽遗传资源,以长被毛为典型表型特征,具备耐受高海拔寒冷环境的能力^[16]。卢晓丽等研究显示,藏羊通过增加血液中白细胞数量与血红蛋白含量来实现对高原极端环境的适应^[17]。Wei 等对中国七个地方绵羊群体开展全基因组扫描,采用 *Fst* 与 *XP-EHH* 方法识别出 236 个基因所在的染色体区域呈现正向选择特征,其中 *CRYAA*、*EPAS1* 和 *LONP1* 参与缺氧应答过程^[18],Edea 等对五个分布于不同海拔的埃塞俄比亚绵羊群体进行选择信号分析,发现 *MITF*、*FGF5*、*MTOR*、*TRHDE* 和 *TUBB3* 等候选基因通过参与呼吸系统发育及平滑信号通路 (smoothened signaling pathway),赋予其对极端海拔条件的适应能力^[19]。Hu 等基于全基因组序列、单核苷酸多态性、染色体 DNA 及 Y 染色体变异信息,对藏绵羊的基因组变异进行分析,在涉及缺氧和紫外线信号通路以及角形态特征的相关基因中检测到强烈的选择信号,并进一步

揭示了适应性渗入在藏绵羊环境适应进化中的关键作用^[20]。此外, *FGF7* 可通过增强肺功能以应对高原环境引发的肺损伤, 提示绵羊在高海拔地区存在新的适应机制^[21]。

1.1.1.4 绵羊尾部脂肪沉积的适应性选择

考古证据表明, 脂尾羊的历史可追溯为 5000 年前。肥尾羊的形成被认为是瘦尾羊在面对环境胁迫时形成的适应性品种^[22]。我国地方绵羊品种以脂尾型为主, 源于蒙古系和哈萨克系的地方脂尾(臀)型绵羊广泛分布于全国各地。地方脂尾型绵羊保留较好的适应性更适合传统放牧饲养方式: 如远距离转场, 熬过漫长严寒的冬季。绵羊尾臀部脂肪性状的性状具有复杂的遗传和分子调控机制, 是进化用于动态储存和释放能量的组织器官^[23]。脂肪组织具有较强的可塑性, 能响应生理与环境信号, 在代谢、结构和表型上进行重塑的特性^[24]。为维持机体内稳态, 脂肪组织可响应环境温度变化从表型和代谢特征方面发生适应性改变, 也可根据机体营养状态(过剩/不足)进行扩增或缩小, 从而保障机体代谢健康与环境适应性^[25]。地方绵羊尾臀脂肪在季节变换过程中呈现出显著的动态重塑特征: 夏秋季能量供应充足时, 尾部脂肪细胞沉积脂质, 体积增大; 冬季则动用储存脂质供能, 尾臀脂肪组织缩小, 以应对饲草短缺并抵御寒冷^[26]。这种尾型性状的异质性及可塑性, 作为绵羊适应复杂环境的关键生物学基础。

1.1.1.5 适应性相关转座元件的研究

转座元件作为基因组中丰度最高的可移动遗传单元, 近年来在生物适应性进化研究中的作用日益受到关注。研究表明, 转座元件在植物、鸟类、鱼类及哺乳动物中均参与了宿主对环境胁迫的适应过程, 其调控机制呈现出多样性和物种特异性。

在植物适应干旱环境的研究中, 玉米 *DRESH8* 转座元件插入变异被证实与耐旱性密切相关^[27]。Sun 等^[28]研究发现, *Gypsy* 转座元件插入在玉米驯化过程中受到选择, 携带 *DRESH8* 的现代玉米品系表现出约 4% 的籽粒长度增加, 且其核苷酸多样性仅为野生近缘种大刍草的三分之一, 提示该位点在驯化过程中发生了选择性清除。值得注意的是, 尽管插入序列仅占玉米基因组的 1.2%, 但高达 41.7% 的耐旱相关性状关联位点映射至这些插入区域, 表明转座元件插入在调控耐旱性与产量性状的平衡中发挥普遍作用。类似地, 葡萄泛基因组分析发现, 野生群体与驯化群体间转座元件变异频率显著分离的区域中, 驯化群体基因富集于有机物合成通路(如酚类和糖的合成), 而野生群体则富集于抗病和抗逆相关通路, 揭示了转座元件在葡萄驯化过程中经历了不同的选择压力^[29]。

在动物适应性进化研究中, 鸟类 *LTR* 转座元件的研究显示, 尽管鸟类基因组中以 *LINE* 元件为主, 但具有澳大利亚进化历史的物种中 *LTR* 元件占主导地位, 提示转座元件组成特征与物种进化历史密切相关^[30]。鱼类研究为转座元件“爆发”与物种分化的关联提供了直观证据: 鳗鲡属、胡鲠属、石斑鱼属等多个鱼类类群的转座元件“爆发”时

间与相应共同祖先分化时间重叠，支持转座元件活化可能驱动物种分化的假说。在低氧适应方面，棕色田鼠（地下鼠）与布氏田鼠（地面鼠）的对比研究表明，两类田鼠在急性低氧条件下反转录转座元件的差异表达模式不同：棕色田鼠的转座元件通过促进细胞分裂分化、调节细胞周期应对低氧胁迫，而布氏田鼠则通过促进细胞凋亡、血管生成并抑制 DNA 复制修复来适应低氧^[31]。这揭示转座元件在不同生态型物种的低氧适应中发挥了差异化的调控作用。

转座元件不仅在环境适应中发挥作用，也参与生殖发育调控。西部食蚊鱼卵巢发育转录组分析显示，共鉴定到 74,390 个表达转座元件，且在卵黄形成时期表达量显著高于其他时期。这些活化转座元件主要分布在内含子及基因邻近区域，附近基因显著富集于排卵相关通路^[32]。通过 RT-qPCR 验证发现，*ADAMTS3*、*WNT4a* 等卵巢发育相关基因与转座元件呈现一致表达趋势^[32]，提示转座元件可能作为调控元件参与生殖过程。在谷子驯化研究中，转座元件插入数目在栽培种中显著增多，且更多插入基因内部；基于转座子标记的全基因组关联分析成功鉴定到与刚毛颜色、分蘖数、开花时间等驯化性状相关的显著位点，进一步证实转座元件在驯化过程中的功能重要性^[33]。

1.1.2 转座元件的研究进展

1.1.2.1 转座元件的定义

转座元件（Transposable elements, TE）是一类可在基因组中移动的散布重复序列，具备在染色体上发生位置迁移并进行自我复制的能力。20 世纪 40 年代，美国科学家 Barbara McClintock 通过对玉米种子颜色变化的研究^[34]，这一发现首次证明，存在一类可在基因组中“跳跃”并发挥基因调控作用的遗传元件。该成果深刻影响了基因组结构与动力学的相关研究，颠覆了 DNA 作为静态分子的传统观念，McClintock 因此被授予诺贝尔生理学或医学奖。

1.1.2.2 转座元件的分类

转座元件是定位于染色体上的一类基本遗传单位，具备在基因组内移动并自主复制的能力，本质为由结构与转座机制多样化的可移动序列构成。其从源位点迁移至靶位点的过程，即为其最具特征的转座行为。这类元件从源位点转移至靶位点，即其标志性的转座行为。依据转座机制与结构特征的不同，研究者已提出多种转座元件分类方式。1989 年，Finnegan 首次提出了真核生物中转座元件的分类系统^[35]，根据转座机制将其分为两类：以 RNA 为中间体的 I 类（逆转座元件）和以 DNA 为中间体的 II 类（DNA 转座元件）。其中，I 类转座元件执行“复制-粘贴”式的转座过程，II 类转座元件则遵循“剪切-粘贴”机制。受当时研究条件所限，Finnegan 提出的分类体系存在一定的局限性。