



•综述•

物种形成过程中的分化基因组岛及其形成机制

孙琼¹, 王嵘^{1,2}, 陈小勇^{1,2*}

1. 华东师范大学生态与环境科学学院, 上海 200241; 2. 华东师范大学浙江天童森林生态系统国家野外科学观测研究站, 上海 200241

摘要: 理解物种形成机制是生态和进化领域的重要任务。得益于测序技术的快速发展, 越来越多研究发现分化种群(亚种、物种)间的基因组常呈现异质性分化景观, 存在分化基因组岛, 这被认为是基因流存在下的歧化选择引起的, 支持基因流存在下的成种假说。然而, 基因渐渗、祖先多态性的差异分选、连锁选择等其他进化过程也可导致分化基因组岛的形成。现有实证研究在解析分化基因组岛的形成机制时, 往往忽略了上述其他进化过程的作用。为此, 本文在辨析分化基因组岛相关概念的基础上, 总结了利用种群基因组数据鉴定分化基因组岛的方法, 对比了不同进化过程形成分化基因组岛的特征, 指出在区分不同机制时联用基因渐渗程度、绝对分化指数(d_{XY})、相对节点深度(RND)、重组率等多个指标的必要性, 归纳了物种形成过程中分化基因组岛形成机制解析的研究思路, 并对未来在生殖隔离机制上的深入探索以及实证研究的整合分析等方面进行了展望。

关键词: 成种; 基因流; F_{ST} ; 基因组分化景观; 种群基因组

孙琼, 王嵘, 陈小勇 (2022) 物种形成过程中的分化基因组岛及其形成机制. 生物多样性, 30, 21383. doi: 10.17520/biods.2021383.

Sun Q, Wang R, Chen XY (2022) Genomic island of divergence during speciation and its underlying mechanisms. Biodiversity Science, 30, 21383. doi: 10.17520/biods.2021383.

Genomic island of divergence during speciation and its underlying mechanisms

Qiong Sun¹, Rong Wang^{1,2}, Xiaoyong Chen^{1,2*}

1 School of Ecological and Environmental Sciences, East China Normal University, Shanghai 200241

2 Zhejiang Tiantong Forest Ecosystem National Observation and Research Station, East China Normal University, Shanghai 200241

ABSTRACT

Background & Aims: Understanding the mechanisms underlying the formation of species is a major task of ecological and evolutionary studies. Aided by the rapid development of sequencing technologies during the last two decades, a growing body of research has revealed a heterogeneous genomic landscape of divergence with the existence of genomic islands of divergence between closely related taxa. This pattern was interpreted as evidence for the speciation-with-gene-flow model, and genomic islands in divergence landscape were hypothesized to be caused by divergent selection in the face of gene flow. However, genomic islands can also arise from other evolutionary processes, such as introgression, divergent sorting of ancient polymorphisms, and linked selection. Previous empirical studies often focused on part of evolutionary processes when discerning the mechanism governing the formation of genomic islands. Therefore, this paper aims to clarify genomic island of divergence and related concepts, summarize the pros and cons of different methods of identifying genomic islands, compare the predicted properties of genomic islands of divergence under different evolutionary processes, and propose a research route for discerning the mechanism contributing to genomic islands during speciation. This paper may provide a guide for future studies on genomic islands and their underlying mechanisms.

Progress: In recent years, various empirical and model-based approaches have been proposed to identify genomic islands. After comparing the pros and cons of these approaches, we think the significance test on observed F_{ST} using the null distribution of F_{ST} under different recombination rate according to the inferred best-fitting demographic model is the most reasonable. Based on similarities and differences of the predicted properties of genomic islands of divergence

收稿日期: 2021-09-22; 接受日期: 2022-01-18

基金项目: 国家自然科学基金(31870356; 31630008)

* 通讯作者 Author for correspondence. E-mail: xychen@des.ecnu.edu.cn

under different evolutionary processes, we point out the necessity of exploiting multiple indicators, such as introgression level, absolute divergence (d_{XY}), relative node depth (RND), and recombination rate, when discerning the mechanism contributing to genomic islands during speciation.

Prospects: We suggest more attention should be paid to the ecological and genetic basis of reproductive isolation after discerning the mechanism underlying the formation of genomic islands. Meanwhile, future researches should be standardized to facilitate integrative analysis in a comparative framework, thus improving our understandings of the frequency of speciation with gene flow in the natural world.

Key words: speciation; gene flow; F_{ST} ; genomic landscape of divergence; population genome

物种多样性是地球最显著的特征之一, 据估计, 地球上有数以百万计的物种(Mora et al, 2011; Costello et al, 2013), 而物种形成(简称成种)是形成地球上物种多样性的基本过程(Nosil & Feder, 2012)。揭示成种机制是生态与进化研究的重要目标, 其中一个争论已久的热点问题是: 基因流存在下成种能否发生(吕昊敏等, 2015; Wolf & Ellegren, 2017)。传统的成种理论认为自然选择在物种形成中发挥着重要作用。由于不同区域间环境相差较大, 种群间产生遗传上的分化, 进而形成不同的物种。1942年, Ernst Mayr提出“生物学种”概念, 认为基因流的同质化作用阻碍种群间分化及生殖隔离的形成, 而成种需要地理障碍来阻断基因流的同质化作用(Mayr, 1942)。因此, 异域成种被认为是物种形成的基本模式, 而邻域成种和同域成种相对少见(Rundle & Nosil, 2005; Mallet et al, 2009; Smadja & Butlin, 2011; Nosil, 2012), 这种观点占据主导地位长达半个多世纪。

直到20年前, 成种观念发生了重大变化, 吴仲义(Wu, 2001)提出基因流存在下的成种(speciation with gene flow)过程, 从基因水平为成种模式的检验提供了理论依据。他认为: (1)在较强的基因流下, 种群间分化在基因组大部分区域较低; (2)与适应有关的位点在歧化选择(divergent selection)的作用下, 可以抵御基因流的同质化效应(图1A), 从而维持超出中性期望的分化程度; (3)与受歧化选择位点紧密连锁的区域, 因分化搭车(divergence hitchhiking)效应也表现出较高的分化程度(图1B)。因此, 基因组不同区域的分化程度呈镶嵌格局, 亦被称为异质性基因组分化景观(heterogeneous genomic landscape of divergence/differentiation) (Nosil et al, 2008; Ellegren et al, 2012)。其中, 分化基因组岛(genomic island of divergence/differentiation)这一隐喻概念, 被用于指代基因组中分化程度超出中性期望的区

域(图1C) (Harr, 2006; Nosil et al, 2009)。如果受歧化选择位点的连锁区域出现新的有利突变, 那么分化程度超出中性期望的区域(分化基因组岛)将进一步增大(Wu, 2001; Nosil et al, 2009; Feder et al, 2012b; Yeaman et al, 2016)。随时间推移, 受歧化选择位点逐渐增加, 导致整个基因组水平的基因流降低、生殖隔离程度逐渐升高而成种; 由于基因流引起的同质化效应减弱, 中性分化的背景水平逐渐升高, 从而导致分化基因组岛变小, 最终将消失(Wu, 2001; Feder & Nosil, 2010; Feder et al, 2012a)。

21世纪初, 高通量测序技术发展迅猛, 越来越多利用种群基因组数据的成种机制研究均观测到近缘物种及亚种、变种、生态型等种下类群之间的基因组分化景观具有异质性, 并通过某些方法鉴定出“分化基因组岛”, 如富贵椰子(*Howea belmoreana*)和平叶棕(*H. forsteriana*) (Papadopulos et al, 2019)、基于分子证据划分的冈比亚按蚊(*Anopheles gambiae*) M型和S型(Turner et al, 2005)、草原向日葵(*Helianthus petiolaris*)的沙丘型和非沙丘型(Andrew & Rieseberg, 2013; Todesco et al, 2020)、妊丽鱼一种(*Astatotilapia* sp.)的沿岸型和底栖型(Malinsky et al, 2015)等, 均被认为符合基因流存在下的成种假说的预测。在这些研究中, 分化基因组岛被认为确实与基因流存在下的成种假说的描述一致, 包含了与适应相关的成种基因, 抵御了基因流的同质化作用, 且与歧化选择和生殖隔离有关(Wu & Ting, 2004; Vijay et al, 2016)。因此, 分化基因组岛的鉴定被认为是验证成种机制的重要环节, 尤其是对于基因流存在下的成种这一假说的验证, 这也引起了越来越多研究者的关注。

然而, 分化基因组岛并非一定是基因流存在下的歧化选择(divergent selection in the face of gene flow)所导致的。不存在基因流时, 近期的选择性清除(recent selective sweep)、长期的连锁选择(long-

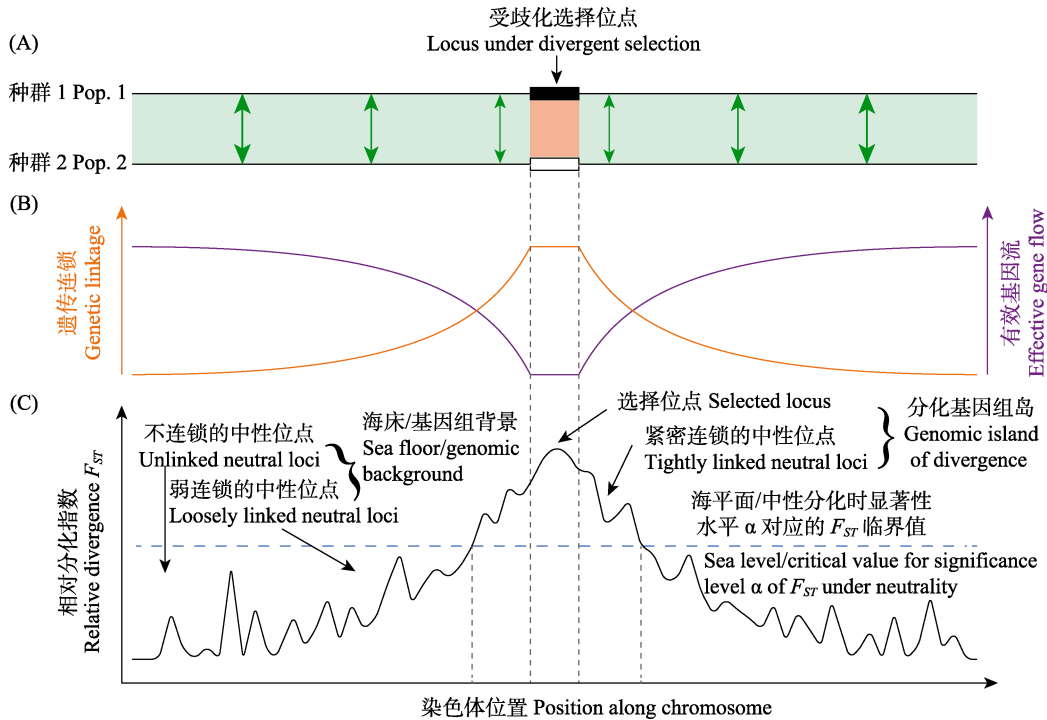


图1 基因流存在下的成种与分化基因组岛。(A)基因流存在下的成种过程初期在基因水平的图示。刚开始分化的两个种群染色体上仅少量位点(此处仅展示1个)受歧化选择,其等位基因在种群间的交换受限。双向箭头粗细表示有效基因流强弱。(B)中性位点与受歧化选择位点的遗传连锁及中性位点的有效基因流在染色体上的变化。(C)异质性基因组分化景观与分化基因组岛。“海平面”将“分化基因组岛”和“海床”分开来。改编自Wu (2001)、Nosil和Feder (2012)和Aeschbacher等(2017)。

Fig. 1 Speciation with gene flow and genomic island of divergence. (A) Conceptual diagram showing the early stage of the process of speciation with gene flow from the genic view. When Pop.1 and Pop.2 start to diverge, only a few loci on the chromosome (here only one is shown) are under divergent selection, and the exchange of the alleles at these loci between populations is restricted. The thickness of the double-headed arrows represents the strength of effective gene flow. (B) Variation of the genetic linkage between neutral loci and divergently selected locus and the effective gene flow at neutral sites along the chromosome. (C) Heterogeneous landscape of genomic divergence and genomic island of divergence. Sea level separates the genomic island of divergence and the sea floor. Adapted from Wu (2001), Nosil & Feder (2012), and Aeschbacher et al (2017).

term linked selection)、祖先多态性的差异分选 (divergent sorting of ancient polymorphisms)、基因渐渗等进化过程,以及影响其中某些过程的基因组内在结构(intrinsic genomic architecture)特征,例如重组率景观、潜在受选择位点密度(density of targets for selection)等,也可能导致或促进分化基因组岛形成(Noor & Bennett, 2009; Han et al, 2017; Ravinet et al, 2017; Semenov et al, 2019)。

因此,阐明分化基因组岛的形成原因是揭示物种形成机制的关键,对分化基因组岛形成机制的解析是近年进化领域的核心议题之一。本文将结合已有研究进展,阐述如何利用种群基因组数据鉴定分化基因组岛,如何辨别影响分化基因组岛形成的进化过程,以及如何利用生物信息学方法区分分化基因组岛形成的各种机制。本文将为分化基因组岛形成机制研究提供思路,加深对自然界中基因流存在

下的成种发生频率等重要问题的理解。

1 分化基因组岛的概念与鉴定

1.1 分化基因组岛及相关概念

Harr (2006)首次使用“分化基因组岛”(genomic island of differentiation)一词来指代小家鼠(*Mus musculus*)两个亚种 *M. musculus musculus* 和 *M. musculus domesticus* 基因组中分化程度较高的区域。后来,“分化基因组岛”被定义为基因组中分化程度(显著)超出中性背景期望的与大小无关的区域;而基因组其他区域(基因组背景, genomic background)被比喻成“海床”(sea floor),用于判断分化程度是否显著超出中性期望的阈值被比喻成“海平面”(sea level) (图1C) (Nosil et al, 2009; Nosil & Feder, 2012)。因为分化程度较高的区域被认为包含生殖隔离基因,所以分化基因组岛起初也被称为成种基因

组岛(genomic island of speciation) (Turner et al, 2005; Feder et al, 2012a; Cruickshank & Hahn, 2014; Burri et al, 2015)。随着认知的深入, 研究者们发现分化基因组岛可能与成种无关, 其他机制, 例如长期的连锁选择导致低重组区核苷酸多态性下降和分化程度升高, 也能导致分化基因组岛的形成(Noor & Bennett, 2009)。因此, 成种基因组岛逐渐被分化基因组岛这一更中性的术语取代(Harrison, 2012; Cruickshank & Hahn, 2014)。

Michel等(2010)认为当基因组中仅有少数几个分散位点受歧化选择时, 这些位点可以被称为分化基因组岛; 当基因组中有较多位点因受到歧化选择(即使是在成种初期)而超出中性期望分化水平时, 这些位点可以形成分化基因组群岛(archipelago)甚至大陆(continent)。与此同时, 他们也强调, 分化基因组岛和分化基因组大陆(genomic continent of divergence)并不互斥, 分化基因组大陆可以被视为较大的分化基因组岛。随着时间的推移, 上述超出中性期望阈值的基因组区域可能存在扩张或收缩的行为(Wu, 2001; Nosil et al, 2009; Feder et al, 2012b; Yeaman et al, 2016), 其大小也受鉴定方法及所采用的中性期望阈值的影响(Via, 2009; Feder & Nosil, 2010)。因此, 为行文方便, 文中将Michel等(2010)提出的岛屿-大陆连续体用“分化基因组岛”指代, 这并不与分化基因组岛的定义相悖(Nosil et al, 2009; Nosil & Feder, 2012)。

1.2 分化基因组岛的鉴定方法

现有鉴定“分化基因组岛”的方法主要包括两大类: 根据经验的方法和基于模型的方法(表1)。根据经验的方法有固定/近固定差异(fixed/nearly fixed difference)的局部聚集程度、固定指数(fixation index, F_{ST})的百分位数、 ZF_{ST} 、使用实际数据重抽样构造 F_{ST} 零分布进行检验等。基于模型的方法有隐马尔科夫模型(hidden Markov model, HMM)、BayeScan、根据假定的种群动态场景或推断的最优种群统计模型构建 F_{ST} 的零分布进行检验等。

早期研究使用固定/近固定差异的局部聚集程度来鉴定分化基因组岛。Turner等(2005)在冈比亚按蚊的M型和S型中利用杂交探针作为分子标记, 以一定数量的探针为滑窗进行基因组扫描, 使用卡方检验判断每个滑窗内的含有固定/近固定差异的探

针数量是否显著高于随机情况下的期望值, 将相邻的显著滑窗合并为分化基因组岛, 最后发现有3个区域显著偏高(Bonferroni校正后 $P < 0.05$)。Harr(2006)使用类似的方法, 只是分子标记类型(SNP)和显著性水平($P \leq 0.015$)有所差异。这种方法忽略了基因组中其他很多非固定/近固定差异的信息, 可能并不适用于分化时间短而固定/近固定差异较少的研究体系中。

基于受歧化选择位点的遗传分化高于中性位点这一原理, 之后的很多研究借用衡量种群间分化的 F_{ST} (图2)或标准化后的 F_{ST} (ZF_{ST})等参数, 以实际数据计算得到的 F_{ST} 分布的经验阈值来鉴定分化基因组岛。有些研究将滑窗 F_{ST} 是否高于全基因组 F_{ST} 的95%或99%等更高的百分位数作为判定条件(之一) (Renaut et al, 2013; Poelstra et al, 2014; Ma et al, 2018; Battey, 2020)。与之类似, Campagna等(2017)计算了基因组滑窗 F_{ST} 的均值(0.008)和标准差(0.015)之后, 采用超出 F_{ST} 均值约12个标准差的阈值(0.2)作为分化基因组岛的判定条件之一(另一条件为滑窗内是否含有 $F_{ST} \geq 0.85$ 的SNP)。根据全基因组 F_{ST} 均值和标准差鉴定等同于使用Z标准化后的 F_{ST} (ZF_{ST})鉴定, 即以滑窗是否满足 $ZF_{ST} \geq 2$ 或3等阈值为依据(Han et al, 2017; Wang et al, 2019; Choi et al, 2020; Behrens et al, 2021)。

一些研究通过实际数据重抽样手段构建 F_{ST} 零分布, 将显著性阈值作为近似的中性期望分化水平, 进而确定分化基因组岛。例如, 使用置换检验(permutation test)判断某位点/滑窗的 F_{ST} 是否显著偏离种群间单倍型随机置换或全基因组标记随机置换时的 F_{ST} (Michel et al, 2005; Berg et al, 2015; Feulner et al, 2015; Mořkovský et al, 2018)。类似地, Michel等(2010)和Ackiss等(2020)通过精确检验(exact test)判断实际的 F_{ST} 是否显著。其他重抽样手段(如自展法bootstrap)有时也用于构建 F_{ST} 零分布(Nadeau et al, 2012; Larson et al, 2017)。以上根据 F_{ST} 的观测分布或者重抽样得到的 F_{ST} 零分布来鉴定“分化基因组岛”的方法无本质区别, 采用的阈值近似地视为中性背景分化水平上界(即“海平面”), 但可能与实际情况相差甚远, 例如当有效种群规模较小时, 随机漂变亦可导致 F_{ST} 显著增高。因此, 此类方法的鉴定结果可能不符合分化基因组岛定义中的

表1 不同分化基因组岛鉴定方法的比较

Table 1 Comparison of different methods for identifying genomic island of divergence

方法 Methods	优点 Pros	缺点 Cons	示例文献 Example references
基于经验的方法 Empirical approach			
固定/近固定差异的局部聚集程度 Extent of regional clustering of fixed/nearly fixed differences	非常简单便捷 Very simple and convenient	未充分考虑种群统计的影响; 仅使用固定/近固定差异的分子标记进行分析, 忽略了基因组中大量的其他变异信息 The effect of population demography is not fully considered; Only fixed or nearly fixed differences are used in the analysis, with numerous differences in the genome ignored	Turner et al, 2005; Harr, 2006
F_{ST} 百分位数 Percentile of F_{ST}	非常简单便捷 Very simple and convenient	未充分考虑种群统计的影响 The effect of population demography is not fully considered	Renaut et al, 2013; Poelstra et al, 2014; Ma et al, 2018; Battey, 2020
Z标准化后的 F_{ST} (ZF_{ST})大小 Value of Z-transformed F_{ST} (ZF_{ST})	非常简单便捷 Very simple and convenient	未充分考虑种群统计的影响 The effect of population demography is not fully considered	Han et al, 2017; Wang et al, 2019; Choi et al, 2020; Behrens et al, 2021
根据实际数据重抽样构造 F_{ST} 零分布检验观测 F_{ST} 的显著性 Significance test on observed F_{ST} using the null distribution of F_{ST} built by resampling real data	简单便捷 Simple and convenient	未充分考虑种群统计的影响 The effect of population demography is not fully considered	Michel et al, 2005; Nadeau et al, 2012; Berg et al, 2015; Feulner et al, 2015; Larson et al, 2017; Mořkovský et al, 2018; Ackiss et al, 2020
基于模型的方法 Model-based approach			
隐马尔科夫模型 Hidden Markov model (HMM)	简单便捷; 无需设置滑窗大小, 考虑了相邻分子标记之间的非独立性 Simple and convenient; No need to set the sliding window size, and the non-independence among neighboring markers is modeled	未充分考虑种群统计的影响; 每种隐含状态对应观测值的分布、高/低分化水平两种隐含状态互相的转移概率是否为0等方面的假设对鉴定结果有一定影响 The effect of population demography is not fully considered; The identification result is affected by the assumptions such as the distribution of observations corresponding to each hidden state and whether the transition probability between the two hidden states of high/low divergence is 0	Turner et al, 2005; Hofer et al, 2012; Soria-Carrasco et al, 2014; Marques et al, 2016; Morales et al, 2018
BayeScan	简单便捷 Simple and convenient	当研究对象中有经历过严重瓶颈效应的种群时, 鉴定结果存在较高的假阳性率 High false positive rate when studied population underwent a strong bottleneck	Michel et al, 2010; Andrew & Rieseberg, 2013; Larson et al, 2019; McCulloch et al, 2021
根据假定的种群动态场景构建不同杂合度下 F_{ST} 的零分布检验观测 F_{ST} 的显著性 Significance test on observed F_{ST} using the null distribution of F_{ST} under different heterozygosity according to the assumed population demographic scenario	简单便捷; 相关软件内置了不同种群动态场景 Simple and convenient; Different population demographic models are built in related softwares	假定的种群动态场景可能与现实场景相差甚远, 导致假阳性率较高 The assumed population demographic model may be very different from the real scenario, resulting in a high false positive rate	Michel et al, 2010; Tsumura et al, 2012; Bradbury et al, 2013; Hudson et al, 2013
根据推断的最优种群统计模型构建 F_{ST} 的零分布检验观测 F_{ST} 的显著性 Significance test on observed F_{ST} using the null distribution of F_{ST} according to the inferred best-fitting demographic model	构造 F_{ST} 零分布时使用的种群统计模型更符合现实场景 The demographic model used to build the null distribution of F_{ST} is more realistic	步骤有些繁琐; 未考虑重组率对 F_{ST} 分布的影响 Somewhat complex; The effect of recombination rate on the distribution of F_{ST} is ignored	Malinsky et al, 2015; Wang et al, 2016
根据推断的最优种群统计模型构建不同重组率下 F_{ST} 的零分布检验观测 F_{ST} 的显著性 Significance test on observed F_{ST} using the null distribution of F_{ST} under different recombination rate according to the inferred best-fitting demographic model	构造 F_{ST} 零分布时使用的种群统计模型更符合现实场景; 考虑了重组率对 F_{ST} 分布的影响 The demographic model used to build the null distribution of F_{ST} is more realistic, and the effect of recombination rate on distribution of F_{ST} is considered	步骤繁琐 Complicated	目前尚无示例文献 No example reference is available yet

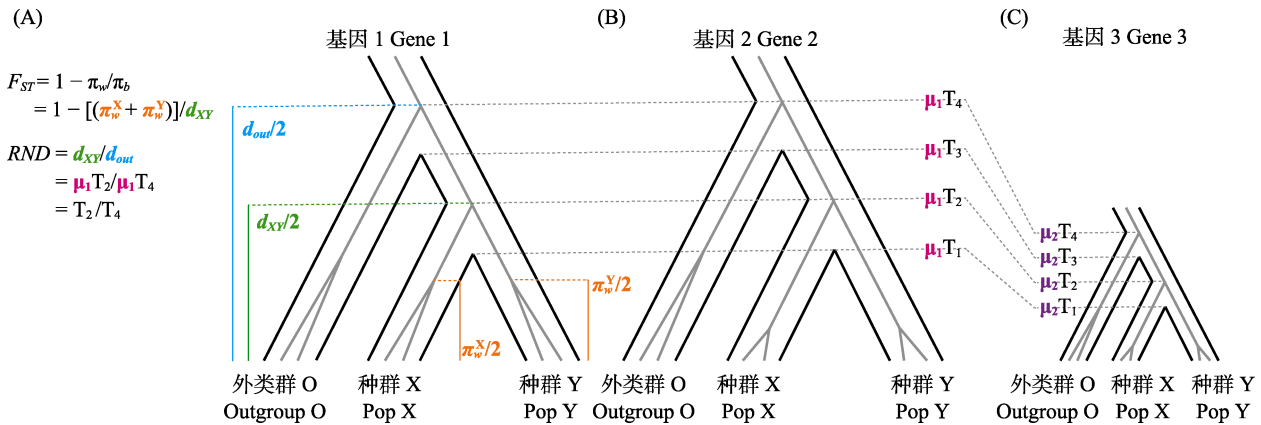


图2 种群内核苷酸多样性(π_w)、种群间核苷酸序列分化/绝对分化指标(π_b 或 d_{XY})、固定指数(F_{ST})和相对节点深度(RND)图示。X、Y为两个正在分化的种群/物种，O为外类群。物种树(黑色)内是示例基因的基因树(灰色)。此处物种树和基因树的大小均与示例基因突变率 μ 成正比， $\mu_1 > \mu_2$ 。T₁、T₃为X与Y、O与XY祖先的种群分化时间/成种时间，T₂、T₄为X与Y、O与XY祖先的示例基因分化时间。 π_w^X 、 π_w^Y 、 d_{XY} (即 π_b)为X种群内、Y种群内、X与Y种群间两两配对序列的平均核苷酸差异。 F_{ST} 可表示为 $1 - \pi_w / \pi_b$ 或类似形式。 RND 校正了 μ 对序列分化的影响。 d_{out} 是O与X、Y的 d_{OX} 和 d_{OY} 的均值。基因1和基因2的 d_{XY} 相等，但基因2的 π_w 更小，因此其 F_{ST} 更高(A和B)。基因2和基因3的 F_{ST} 、T₂、T₄和 RND 均相等，但 μ 不同，基因3的 d_{XY} 更小(B和C)。改自 Cruickshank和Hahn (2014)和Rosenzweig等(2016)。

Fig. 2 Illustration of within-population nucleotide diversity (π_w), between-population sequence divergence/absolute divergence (π_b or d_{XY}), fixation index (F_{ST}) and relative node depth (RND). X and Y are two diverging populations or species. The gene tree (grey) of an example gene is contained within the species tree (black). The sizes of species tree and gene tree displayed are both directly proportional to the mutation rate (μ) of the example gene, $\mu_1 > \mu_2$. T₁ and T₃ are the divergence time or speciation time between X and Y, and between O and the ancestor of X and Y, respectively. T₂ and T₄ are the divergence time of example gene between X and Y, and between O and the ancestor of X and Y, respectively. π_w^X , π_w^Y , and absolute divergence (d_{XY}) represent the average number of nucleotide differences of pairwise sequences within X, within Y, and between X and Y, respectively. F_{ST} is calculated as $1 - \pi_w / \pi_b$ or similar formulas. RND is an estimate of sequence divergence with the effect of μ corrected. d_{out} is the average of absolute divergence between O and X and between O and Y. d_{XY} of gene 1 is the same as that of gene 2, but F_{ST} of gene 2 is higher due to lower π_w (A and B). There are no differences in F_{ST} , T₂, T₄, and RND between gene 2 and gene 3, but d_{XY} of gene 3 is lower due to lower μ (B and C). Adapted from Cruickshank & Hahn (2014) and Rosenzweig et al (2016).

“显著超出中性背景期望”这一要点。

广泛应用于序列分析的HMM也可用于鉴定分化基因组岛。它不依赖于对滑窗大小的设置，可以将相邻分子标记之间的非独立性考虑进来(Hofer et al, 2012)。HMM由隐含状态(hidden states)、符号/观测值(symbols/observations)、不同隐含状态间的转移概率(transition probabilities)以及各隐含状态到各观测值的输出概率(emission probabilities)构成。鉴定分化基因组岛时，隐含状态通常设置为高(分化基因组岛)、中(基因组背景)和低三种分化水平(Hofer et al, 2012; Soria-Carrasco et al, 2014; Marques et al, 2016; Morales et al, 2018)，也有研究仅设置为高和低两种分化水平(Turner et al, 2005)。观测值输入的数据可以是离散变量(如固定/近固定变异与否)(Turner et al, 2005)，也可以是连续变量，如logit(F_{ST})(Soria-Carrasco et al, 2014)、 F_{ST} 对应的中性检验P值及其变型(Hofer et al, 2012; Marques et

al, 2016; Morales et al, 2018)。每种隐含状态对应的观测值都满足一定的分布，但对于特定的观测值，每种隐含状态下都有可能发生，只是概率不同，这往往导致分化程度显著超出中性期望的位点未被判定为分化基因组岛，并且鉴定出来的分化基因组岛可能包含分化程度未显著超出中性期望的长片段(Hofer et al, 2012; Marques et al, 2016)。另外，HMM中每种隐含状态所对应观测值的分布、高/低分化水平两种隐含状态互相的转移概率是否为0等方面的假设，对分化基因组岛鉴定结果也有一定的影响。

BayeScan基于衡量种群间遗传分化的 F_{ST} 来鉴定潜在受选择位点(Foll & Gaggiotti, 2008)。一些研究将BayeScan鉴定出的潜在受歧化选择位点视为分化基因组岛(Michel et al, 2010; Andrew & Rieseberg, 2013; Larson et al, 2019; McCulloch et al, 2021)。BayeScan将每个位点的 F_{ST} 拆分成所有位点

共享的组分(β 组分)和位点特异的组分(α 组分), 构造包含 α 组分或不含 α 组分(作为中性模型)两种模型, 设置后一种模型的先验优势(prior odds), 接下来对于每个位点使用reversible-jump MCMC算法估计两种模型的后验优势(posterior odds), 计算得到用于判断 α 组分是否显著的贝叶斯因子(Bayes factor), 当某位点的 α 组分显著大于0时, 认为该位点受到歧化选择。然而, 当所研究的种群经历过严重的瓶颈事件时, BayeScan的结果存在较多的假阳性(Foll & Gaggiotti, 2008)。

根据种群动态进行溯祖模拟的方法与以上方法有本质区别, 一定程度上为分化基因组岛的鉴定提供了更可靠的手段。其中, 通过比较不同杂合度下 F_{ST} 的零分布与观测分布以检测非中性位点的方法(Beaumont & Nichols, 1996), 被较多的研究用于鉴定分化基因组岛(Michel et al, 2010; Tsumura et al, 2012; Bradbury et al, 2013; Hudson et al, 2013)。这一方法在假定的诸如无限岛屿模型或脚踏石模型等种群动态场景之下进行溯祖模拟, 然后计算得到不同杂合度下 F_{ST} 的零分布, 与观测分布进行比较检验, 从而得到不同杂合度下的显著非中性位点。内置这一方法的软件有LOSITAN、Fdist2、Arlequin等。但是, 假定的种群动态场景及设置的参数会影响所产生的零分布, 可能会导致较高的假阳性率(Beaumont & Nichols, 1996; Excoffier et al, 2009; Hofer et al, 2012)。而 $\delta a \delta i$ (Gutenkunst et al, 2009)和Fastsimcoal2 (Excoffier et al, 2013)等软件则可以通过构建不同的种群统计模型(demographic model) (如设置基因流有无或有效种群规模是否改变等), 并利用种群基因组数据筛选出最优模型来推断种群历史动态(包含有效种群规模、基因流有无及强度等重要参数), 比上述方法更符合实际的种群动态场景。根据最优的种群统计模型, 使用溯祖模拟软件可以模拟中性条件下滑窗 F_{ST} 的零分布, 据此得到显著性阈值, 再将观测的 F_{ST} 与之比较则可以鉴定出显著高于中性期望分化水平的位点, 即分化基因组岛(Malinsky et al, 2015; Wang et al, 2016)。然而, Booker等(2020)通过计算机模拟发现即使在中性条件下, 重组率变异也会影响 F_{ST} 的变异程度, 低重组率导致 F_{ST} 呈右偏分布, 造成在全基因组水平使用单一 F_{ST} 阈值鉴定离群滑窗时, 离群滑窗会偏向于

富集在低重组区。Booker等(2020)、Stevison和McGaugh(2020)都建议将所有滑窗根据重组率高低划分成不同区间, 使用可考虑重组率的溯祖模拟软件模拟不同重组率下 F_{ST} 的零分布, 再比较各重组率区间下 F_{ST} 的零分布与观测分布, 从而鉴定出(离群滑窗以及)分化基因组岛。其中, 可考虑重组率的溯祖模拟软件有MSMS等, 更多介绍详见高峰和李海鹏(2016)的综述。

综上所述, 很多研究倾向于使用更简单便捷的方法鉴定分化基因组岛(表1), 但结果可能存在较高的假阳性率, 并非严格意义上的分化基因组岛。而通过比较不同重组率下基于最优种群统计模型溯祖模拟产生的 F_{ST} 的零分布与观测到的 F_{ST} 分布的差异, 来鉴定分化基因组岛是更恰当的方法, 也符合成种研究中人们更关注非中性过程的事实。

2 分化基因组岛的形成机制

由上文可知, 已有研究大都使用 F_{ST} 衡量种群间的分化程度, 并以此作为分化基因组岛鉴定的依据。 F_{ST} 表征的是种群间变异和总变异的相对大小, 受到种内遗传多样性的影响, 属于相对分化指标(图2) (Charlesworth, 1998; Cruickshank & Hahn, 2014)。在种群基因组研究中通常使用 $\pi_b/(\pi_b + \pi_w)$ 或 $(\pi_b - \pi_w)/\pi_b$ 等公式计算 F_{ST} , 其中 π_b 、 π_w 分别表示种群间核苷酸序列分化(即种群间两两配对序列的核苷酸差异均值)、种群内核苷酸多样性(即种群内两两配对序列的核苷酸差异均值), π_b 亦可用 d_{XY} 来表示(图2) (Hudson et al, 1992; Cruickshank & Hahn, 2014; Irwin et al, 2018; Tavares et al, 2018)。亲缘关系紧密但形态、生境、行为等方面分化明显的姊妹种(或种下类群, 如亚种、变种、生态型、基于分子证据划分的分子型(molecular form)等)是成种研究的主要系统, 为方便起见, 下文以姊妹种为例展开论述。基因流存在下的成种假说(更具体地说, 基因流存在下的歧化选择这一进化过程)常用于解释姊妹种基因组中分化基因组岛的形成(图1) (Turner et al, 2005; Malinsky et al, 2015; Marques et al, 2016; Choi et al, 2020)。然而, 即使在姊妹种间不存在基因流的情况下, 来自其他物种的基因渐渗、祖先多态性的差异分选、近期的选择性清除、长期的连锁选择等进化过程也可能导致某些位点的分化程度(F_{ST})显

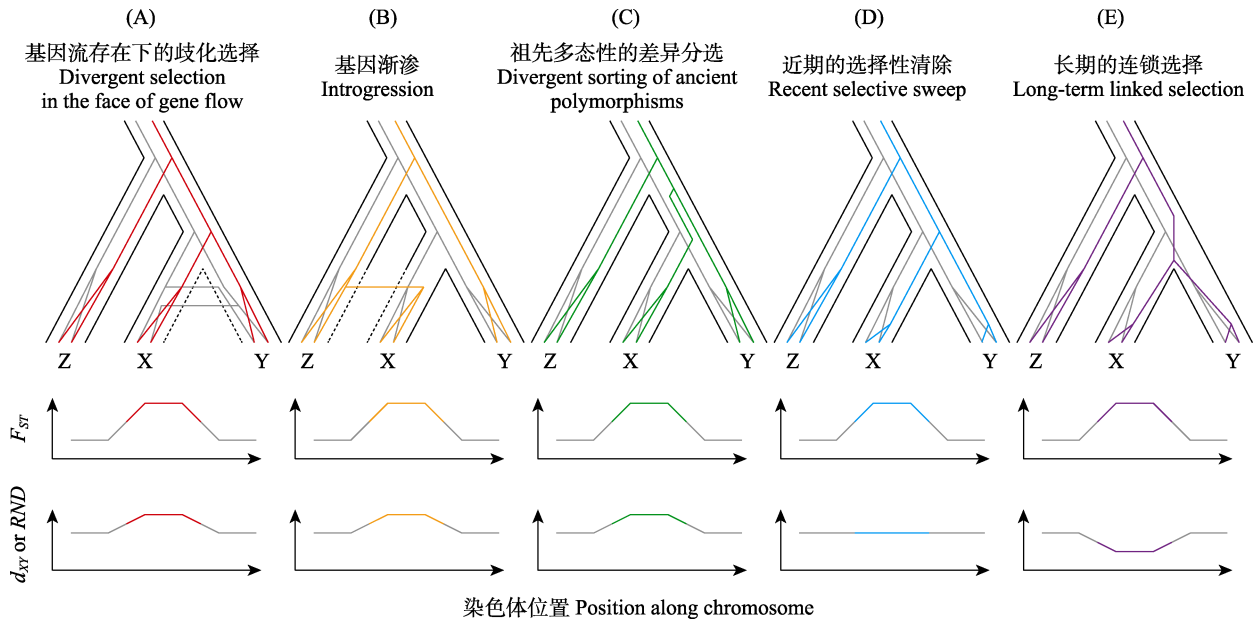


图3 可导致分化基因组岛形成的不同进化过程示意图。X、Y为两个正在分化的种群/物种。第二种机制中(B), 近缘物种Z向X有基因渐渗, 其他机制中Z向X或Y均无基因渐渗(A、C、D、E)。每个物种树内包含两个示例基因的基因树, 彩色基因树表示示例基因是经历了对应进化过程的位点或紧密连锁位点, 灰色基因树表示弱连锁或不连锁的中性位点。物种树下的图表示每种机制两类位点 F_{ST} 、 d_{XY} (假设基因组中 μ 无异质性)和 RND (无需对 μ 的异质性做出假设)的差异。改自Cruickshank和Hahn (2014)和Han等(2017)。

Fig. 3 Schematic diagram illustrating different evolutionary processes that could lead to the formation of genomic island of divergence. X and Y are two diverging populations or species. Z is a related species, from which some genetic materials are introgressed into X in scenario (B) in contrast to the other four scenarios (A, C, D, E). The gene trees of two example genes are contained within the species tree (black). The colored gene tree represents the locus underlying the corresponding evolutionary process or its tightly linked loci, and the grey gene tree represents the loosely linked or unlinked loci. The graphs below species trees show the difference of F_{ST} , d_{XY} (on the assumption of no variation of μ), and RND (without the assumption of no variation of μ) between two kinds of loci under various evolutionary processes. Adapted from Cruickshank & Hahn (2014) and Han et al (2017).

著高于中性期望, 即呈现分化基因组岛(图3) (Noor & Bennett, 2009; Han et al, 2017; Ravinet et al, 2017; Semenov et al, 2019)。下文中, 我们将按照图3中的顺序依次介绍与分化基因组岛形成有关的各种机制。前三种机制主要通过使 d_{XY} 偏高形成分化基因组岛, 而后两种机制则主要通过降低 π_w 。

2.1 基因流存在下的歧化选择

基因流存在下的歧化选择指基因流存在的情况下, 歧化选择的作用可降低基因组中某些位点的有效基因流, 导致基因组不同区域分化程度的变化。基因流存在下的成种假说认为, 在成种过程初期, 基因组大部分区域的基因流不受限制, 基因流强度较大, 等位基因的交换使得这些区域被同质化, 从而维持较低的分化程度(图1, 图3)。但是, 基因组中存在少量位点与环境适应有关而受到不同方向的选择(即歧化选择), 这些位点的等位基因交换受限, 有效基因流强度降低, 基因流的同质化作用较

弱, 从而维持超出中性期望的分化程度。与受歧化选择位点紧密连锁的区域受搭车效应(hitchhiking)的影响, 其分化程度也偏高, 从而和歧化选择位点共同构成分化基因组岛(图1, 图3) (Wu, 2001; Feder et al, 2012a)。在上述过程中, 基因流与歧化选择均发挥着重要的作用, 基因组的背景基因流水平相似, 歧化选择导致目标位点的基因流降低, 使得基因组中呈现有效基因流强度的变化, 进而影响基因组的分化景观。基因流存在下的歧化选择主要是通过影响种间核苷酸序列分化(d_{XY})来影响位点的分化程度(图3)。对于与歧化选择位点不连锁或弱连锁的位点而言, 等位基因的频繁交换使种间的 d_{XY} 降低, 而歧化选择位点及紧密连锁区域的 d_{XY} 因基因流受限则保持较高水平 (Charlesworth et al, 1997; Cruickshank & Hahn, 2014)。因此, 受歧化选择位点及紧密连锁位点的 F_{ST} 显著高于中性期望。

目前已有不少实证研究, 报道基因流存在下的

歧化选择可以导致分化基因组岛的形成。欧洲的小嘴乌鸦(*Corvus corone*)和冠小嘴乌鸦(*C. cornix*)处于早期分化阶段, 存在杂交区。Poelstra等(2014)通过分析种群基因组数据发现, 这两个物种的基因流并不仅仅局限于由形态界定的杂交区(morphological hybrid zone), 杂交区两侧不同物种的种群在基因组水平的分化程度很低, 但基因组中含有少量分化基因组岛。这些分化基因组岛中包含与色素沉积和视觉感知有关的受歧化选择基因, 抵御了基因流的同质化作用, 维持了两种乌鸦在表型上的分化。Malinsky等(2015)根据基因流存在、分化基因组岛 d_{XY} 偏高和高分化区域聚集等结果, 认为在坦桑尼亚马索科湖中同域分布的妊丽鱼一种(*Astatotilapia* sp.)的两个生态型(即沿岸型和底栖型)符合基因流下的成种模型, 而分化基因组岛中存在与暗视觉、形态建成等相关的适应性基因, 其基因型影响着两个生态型的交配偏好。夏威夷岛上多型铁心木(*Metrosideros polymorpha*)的两个变种 *M. polymorpha* var. *newellii* 和 *M. polymorpha* var. *glaberrima* 在生境和形态方面有所分化, 前者叶片较窄, 生长于河岸地带, 而后者叶片较宽, 生长于邻近的森林中。Choi等(2020)证明了这两个变种开始分化后存在持续的基因流, 并发现全基因组的分化程度较低(F_{ST} 中位数为0.043), 但分化基因组岛的 d_{XY} 偏高且富集了正选择信号, 认为歧化选择与基因流共同塑造了两个变种的基因组分化景观。

2.2 基因渐渗

基因渐渗是指两个亲本物种的杂交后代与亲本之一反复回交, 使另一亲本的某些遗传物质渗入回交亲本基因库中的现象或过程(Harrison & Larson, 2014; 程祥等, 2020)。作为适应性遗传变异的重要来源之一, 基因渐渗在适应性进化和物种形成中起到了不可忽视的作用(Hedrick, 2013)。例如, 从尼安德特人(*Homo neanderthalensis*)渗入到现代人(*H. sapiens*)中的一些有利等位基因, 有助于现代人适应非洲之外的环境(Sankararaman et al, 2014; Racimo et al, 2015); 外来物种加拿大底鳉(*Fundulus heteroclitus*)所携带的适应污染环境的遗传变异, 近期渐渗到墨西哥湾当地种大底鳉(*F. grandis*), 使得大底鳉也能够严重污染的休斯顿航道存活下来, 而不至于局部灭绝(Oziolor et al, 2019)。检测基因渐

渗可使用ABBA-BABA检验和TreeMix等方法, 更多方法的详细介绍可参考毛建丰等(2017)和程祥等(2020)的综述。

当来自其他近缘物种的某些等位基因渗入姊妹种的其中一个并逐渐固定时, 渐渗位点在姊妹种间分化程度偏高, 可成为分化基因组岛。基因渐渗带来的远缘序列, 主要通过增加姊妹种间核苷酸序列分化(d_{XY}), 从而使得 F_{ST} 超出中性期望(图3)(Guerrero & Hahn, 2017; Choi et al, 2020)。不少研究发现来自其他物种向姊妹种一方的基因渐渗可以导致基因组中出现高分化区域(Richards & Martin, 2017; Teng et al, 2017)。巴哈马群岛中的新普罗维登斯岛和圣萨尔瓦多岛相距约300 km。Richards和Martin (2017)的调查显示, 来自新普罗维登斯岛上鳞属物种*Cyprinodon laciniatus*的基因渐渗促进了圣萨尔瓦多岛鳞属的适应性辐射, 例如*C. laciniatus*将颅面发育相关基因(*ski*等)的遗传变异渐渗至捕食软体动物的*C. brontotheroides*中, 导致*C. brontotheroides*和同域分布的食鳞的*C. desquamator*在颌骨形态和营养生态位上发生分化; 这两个物种基因组中某些区域因来自*C. laciniatus*的基因渐渗而具有较高的 d_{XY} , 致使 F_{ST} 偏高, 但*C. brontotheroides*和*C. laciniatus*渐渗位点的 d_{XY} 偏低。

2.3 祖先多态性的差异分选

姊妹种的祖先种基因组中某些位点可能由于经历了长期的平衡选择(long-term balancing selection)而存在序列差异程度较高的不同单倍型, 即存在较高的祖先多态性(ancient/ancestral polymorphisms)。如果这些不同的单倍型在成种过程中通过谱系分选(lineage sorting)被不均等地分配到后代谱系(即姊妹种), 则祖先多态性发生了差异分选(divergent sorting)(图3)(Guerrero & Hahn, 2017; Han et al, 2017)。Guerrero和Hahn (2017)将成种过程形象地比喻成一个“筛子”(sieve), 祖先多态性被成种过程所筛选。祖先多态性的差异分选这一机制也能导致姊妹种间基因组中分化基因组岛的形成(图3)(Guerrero & Hahn, 2017; Han et al, 2017)。某些位点上较高的祖先多态性造成姊妹种间核苷酸序列分化(d_{XY})偏高, 从而呈现 F_{ST} 显著高于中性期望的现象(Cruickshank & Hahn, 2014; Guerrero & Hahn, 2017; Han et al, 2017)。

近年来,不少案例研究证实了祖先多态性的差异分选在分化基因组岛形成中的作用。[Han等\(2017\)](#)利用加拉帕戈斯群岛中6个岛屿和遥远的科科斯岛上多种达尔文地雀的种群基因组数据,分析了12个物种对或种群对,发现在达尔文地雀辐射进化早期(大约一百万年前),与喙性状(喙形和大小)有关的 *ALX1* 和 *HMGA2* 位点可能因受到频率依赖选择(frequency-dependent selection)而在祖先种内维持了不同的单倍型,随后不同单倍型被差异分选至近期分化的不同后代物种,从而使得具有不同喙性状的物种对在 *ALX1* 和(或) *HMGA2* 位点的 d_{XY} 和 F_{ST} 显著高于基因组背景水平。[Wang等\(2019\)](#)发现十字花科的加拿大岩芥(*Boechea stricta*)的两个亚种分化基因组岛 d_{XY} 偏高,且大多数分化基因组岛的序列分化时间早于两个亚种的分化时间,认为分化基因组岛是祖先多态性差异分选的结果。

2.4 近期的选择性清除

非直接受选择的位点可因相邻位点受选择而呈现核苷酸多态性降低的现象称为连锁选择,包括选择性清除和背景选择两种形式([Cutter & Payseur, 2013](#); [Cruickshank & Hahn, 2014](#); [Rettelbach et al, 2019](#))。选择性清除是指正选择使有利变异在种群中的频率快速上升,与之连锁的等位基因频率因搭车效应也快速上升,导致连锁位点核苷酸多样性下降([Smith & Haigh, 1974](#));而背景选择是指有害突变受到负选择作用从种群中被清除,从而导致连锁位点核苷酸多样性下降([Charlesworth et al, 1993](#))。

近期的选择性清除指选择性清除仅发生于现存种群(extant populations/current-day populations)中,而在这些种群的祖先种群(ancestral population)中未发生,即祖先种群的核苷酸多样性没有受连锁选择的影响([Cruickshank & Hahn, 2014](#)),与后文将介绍的长期的连锁选择机制相区分(图3)。即使不存在基因流,姊妹种对局域生境的生态适应引起的近期选择性清除也能导致分化基因组岛的形成([Han et al, 2017](#))。近期的选择性清除并未影响种间核苷酸序列分化(d_{XY}),而是通过降低种内的核苷酸多样性使得某些区域的 F_{ST} 高于中性期望(图3) ([Charlesworth, 1998](#); [Cruickshank & Hahn, 2014](#); [Han et al, 2017](#))。[Cruickshank和Hahn \(2014\)](#)重新分析了以往报道的多个物种对,均发现分化基因组岛的 d_{XY} 和基因组

背景的 d_{XY} 并没有显著差异,但种内的核苷酸多样性显著偏低,使用IMa2软件分析发现物种对间并不存在基因流,因此,认为成种事件之后发生的局域适应引发的近期选择性清除是导致分化基因组岛形成的原因。

2.5 长期的连锁选择

长期的连锁选择是指连锁选择在祖先种内就已发挥作用的过程,包括持续的背景选择(ongoing background selection)与频繁的选择性清除(recurrent selective sweep) ([Burri, 2017](#); [Han et al, 2017](#))。与近期的选择性清除机制不同,在长期的连锁选择中,祖先种内和现存种内均持续或频繁发生的连锁选择(recurrent bouts of linked selection)不仅降低了现存姊妹种内的核苷酸多样性 π_w , 还降低了姊妹种的祖先种内的核苷酸多样性,导致姊妹种间核苷酸序列分化(d_{XY})偏低(图3) ([Nachman & Payseur, 2012](#); [Cruickshank & Hahn, 2014](#); [Han et al, 2017](#))。当长期的连锁选择导致某位点种内核苷酸多样性降低程度高于 d_{XY} 降低程度时,该位点将表现出显著偏高的 F_{ST} , 成为分化基因组岛。

长期的连锁选择可以导致分化基因组岛形成的观点也得到了不少案例支持。[Burri等\(2015\)](#)调查了姬鹀属(*Ficedula*)多个独立的物种对或同一物种的种群对,发现几乎所有物种对和种群对的分化基因组岛出现在相同的位置,分化基因组岛具有显著偏低的 d_{XY} 、 π_w 和重组率,认为低重组率导致较强的长期连锁选择效应是分化基因组岛形成的主要原因。类似地,[Wang等\(2016\)](#)分析了欧洲山杨(*Populus tremula*)和美洲山杨(*P. tremuloides*)的种群基因组数据,发现二者基因组中的分化基因组岛也表现出 d_{XY} 、 π_w 和重组率偏低的特征,另外 F_{ST} 和重组率显著负相关,支持长期的连锁选择在塑造两种山杨的异质性基因组分化景观中发挥重要作用。

3 分化基因组岛形成机制的推断方法

基因组的遗传分化及遗传多样性格局经常用于推断哪种过程主导着分化基因组岛的形成。然而,从上文中不难发现,一些机制下形成的分化基因组岛在遗传分化及遗传多样性格局上具有相似的表现,因此,需要纳入更多信息才可能将不同机制区分开来。下文将详细介绍如何利用种群基因组数据,

判断在分化基因组岛形成中发挥主要作用的机制。

在不同机制下, 分化基因组岛的 d_{XY} 表现不完全相同, 一些研究尝试根据 d_{XY} 将主要影响种内核苷酸多样性的两种机制(近期的选择性清除、长期的连锁选择)与其他三种机制区分开(Malinsky et al, 2015; Ma et al, 2018; Choi et al, 2020)。 d_{XY} 表示种间两两配对序列的平均核苷酸差异, 不受种内核苷酸多样性的影响, 属于绝对分化指标, 亦可用 π_b 表示(Nei, 1987)。在严格异域模型下, d_{XY} 的期望值 $E(d_{XY}) = 2\mu t + \theta_{Anc}$, 与种群分化时间 t 、突变率 μ 、祖先种内核苷酸多样性 θ_{Anc} 正相关($\theta_{Anc} = 4N_e^{Anc}\mu$, 此处 N_e^{Anc} 表示祖先种有效种群规模); 而存在基因流的情况下, d_{XY} 与有效基因流负相关(Cruickshank & Hahn, 2014)。基因渐渗带来远缘DNA, 基因流存在下的歧化选择阻碍等位基因的交流, 祖先多态性的差异分选位点具有较高的 θ_{Anc} , 因此, 这三种机制导致分化基因组岛与基因组背景相比均具有更高的 d_{XY} , 长期的连锁选择由于降低了 θ_{Anc} 而使分化基因组岛的 d_{XY} 偏低, 而近期的选择性清除并不会导致分化基因组岛和基因组背景的 d_{XY} 存在显著差异(图3)(Guerrero & Hahn, 2017; Han et al, 2017)。

值得注意的是, d_{XY} 受突变率的影响(图2)。基于 d_{XY} 准确区分部分机制的前提假设是基因组的突变率不存在异质性, 然而这一假设并不符合实际。因此, 仅仅依据 d_{XY} 而不考虑基因组位点间突变率的差异, 可能会得出错误推断。鉴于此, 一些学者采用相对节点深度(relative node depth, RND)来排除突变率异质性对机制推断带来的影响(图2)(Burri et al, 2015; Wang et al, 2016; Wang et al, 2019)。RND相当于使用突变率校正后的 d_{XY} , 根据 $RND = d_{XY}/d_{out}$ 计算, 其中 $d_{out} = (d_{XO} + d_{YO})/2$, d_{XO} 和 d_{YO} 与 d_{XY} 类似, 分别指姊妹种X、Y和外类群O的种间两两配对序列的平均核苷酸差异(Feder et al, 2005)。

当分化基因组岛的RND显著高于基因组背景时, 可以排除近期的选择性清除和长期的连锁选择这两种机制(图3, 图4)。接下来, 可以通过比较分化基因组岛和基因组背景的渐渗程度将基因渐渗这一机制与基因流存在下的歧化选择、祖先多态性的差异分选进行区分。如果分化基因组岛的渐渗程度显著高于基因组背景, 一般认为分化基因组岛主要由基因渐渗这一机制导致; 反之, 则基本可以确定

基因渐渗不是分化基因组岛形成的主要原因(图4)(Malinsky et al, 2015; Guerrero & Hahn, 2017; Choi et al, 2020)。Malinsky等(2015)通过ABBA-BABA检验和fineSTRUCTURE分析发现, 美妊丽鱼(*Astatotilapia calliptera*)存在其姆巴卡河种群向马索科湖中沿岸型*Astatotilapia* sp.的基因渐渗, 但沿岸型和同域分布的底栖型分化基因组岛的基因渐渗程度(以Patterson's D和 f_{dM} 两个指标衡量)并未显著偏高, 因此, 基因渐渗不是这两个生态型分化基因组岛形成的主要原因。然而, Martin等(2015)指出Patterson's D在使用全基因组数据计算时可靠性才比较高, 不适用于表征基因组滑窗的渐渗程度, 并给出了改进后的 f_d 指数。其他一些指标(如 d_f , RND_{min}等)也可用于衡量基因组滑窗的基因渐渗程度(Rosenzweig et al, 2016; Pfeifer & Kapan, 2019; 程祥等, 2020; Malinsky et al, 2021)。

基因流存在下的歧化选择与祖先多态性的差异分选引起的分化基因组岛在 d_{XY} (或RND)、基因渐渗程度等指标上的表现相似。不同的是, 基因流存在下的歧化选择在分化基因组岛形成中起主导作用时, 姊妹种的同域(或近距离)种群对由于基因流发生机会更大, 相对分化 F_{ST} 会低于异域(或远距离)种群对, 而祖先多态性的差异分选并不会导致这样的结果(图4)(Burri et al, 2015; Wang et al, 2019)。除此之外, 同域种群对间共享过量的衍生等位基因(derived alleles)、同域种群对间分化基因组岛 d_{XY} 相对于基因组背景值的增加幅度比异域种群对间更大等证据也支持基因流存在下的歧化选择这一机制(Burri et al, 2015; Han et al, 2017; Ma et al, 2018; Wang et al, 2019)。若祖先多态性的差异分选在分化基因组岛形成中起主导作用, 姊妹种所携带的不同单倍型在姊妹种分化之前就已经长期存在于祖先种内, 表现为分化基因组岛的分化时间要显著早于姊妹种分化时间(Guerrero & Hahn, 2017; Wang et al, 2019)。这一特征可作为辅助证据支持祖先多态性的差异分选在分化基因组岛形成中的作用。姊妹种的分化时间可使用世代周期 G 、突变率 μ 等参数基于 $\delta a \delta i$ 、fastsimcoal2等软件筛选出来的最优种群统计模型进行计算。分化基因组岛的分化时间可利用公式 $T = d_a/2\mu$ 估计(Wang et al, 2019), 其中种间的净核苷酸差异 $d_a = d_{XY} - (\pi_w^X + \pi_w^Y)/2$ (π_w^X 和 π_w^Y 分别指姊

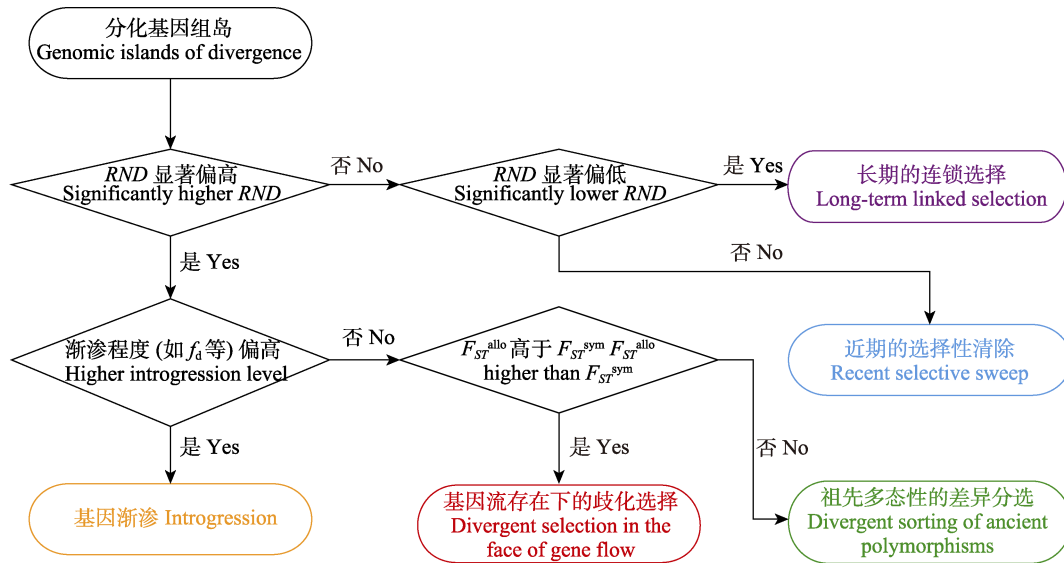


图4 分化基因组岛形成机制解析逻辑流程图。 F_{ST}^{sym} 和 F_{ST}^{allo} 分别表示姊妹种同域种群对、异域种群对的 F_{ST} 。此图仅展示了关键证据(辅助性证据见正文)。
 Fig. 4 Flow chart illustrating the logic of discerning the mechanism contributing to genomic islands of divergence. F_{ST}^{sym} and F_{ST}^{allo} are abbreviations for F_{ST} of sympatric and allopatric populations of focal sister species, respectively. Here, only the key pieces of evidence are listed (see text for complementary evidence).

妹种X、Y的核苷酸多样性) (Cruickshank & Hahn, 2014)。

当分化基因组岛的RND未显著高于基因组背景时, 分化基因组岛的形成机制可能是近期的选择性清除或长期的连锁选择(图3, 图4)。虽然这两种机制均可降低姊妹种内的核苷酸多样性 π_w , 但二者不同的是, 近期的选择性清除并没有降低祖先种内的核苷酸多样性 θ_{Anc} , 所以不会导致 d_{XY} (或RND)偏低 (Cruickshank & Hahn, 2014; Han et al, 2017)。因此, 当分化基因组岛的RND处于中等水平时, 可推断出近期的选择性清除是主导分化基因组岛形成的进化过程; 而当分化基因组岛的RND显著低于基因组背景时, 暗示着长期的连锁选择起主导作用(图4)。

除 d_{XY} 和RND偏低之外, 长期的连锁选择导致分化基因组岛与基因组背景相比, 还可能具有重组率偏低或潜在受选择位点密度(常以编码序列密度来衡量)偏高等特征。由于连锁选择的强度与重组率负相关且与潜在受选择位点的密度正相关, 在重组率低、潜在受选择位点密度高的基因组区域中, 核苷酸多样性受长期的连锁选择的负面影响更强 (Cutter & Payseur, 2013; Burri, 2017)。因此, 这些区域与重组率高、潜在受选择位点密度低的区域相比, 其分化程度更容易超出中性期望分化水平上限, 进

而形成分化基因组岛(Burri et al, 2015; Vijay et al, 2016; Wang et al, 2016)。当分化基因组岛具有重组率低、潜在受选择位点密度高两个表现或其中之一时, 可以作为辅助证据支持长期的连锁选择主导分化基因组岛形成这一结论。其中, 重组率估计主要有3种手段, 分别基于家系数据、配子分型数据和种群基因组数据(Smith & Fearnhead, 2005; Peñalba & Wolf, 2020), 三者所需样本量和分辨率等方面有所不同。在分化基因组岛形成机制的研究中, 由于缺乏家系数据, 大都直接使用种群基因组数据估计种群水平的重组率 ρ ($\rho = 4N_e r$, 其中 N_e 表示有效种群规模, r 为重组率)。另外, 基因组滑窗的潜在受选择位点密度可由滑窗内基因密度或编码序列密度等指标表示(Burri et al, 2015; Rettelbach et al, 2019; Wang et al, 2019)。

4 总结与展望

借助测序技术巨大变革的契机, 从基因组的变异格局破译物种形成机制成为进化生态学研究热点。基因流存在下的成种假说中所强调的基因流存在下的歧化选择这一进化过程常用于解释异质性基因组分化景观中的分化基因组岛, 然而其他一些机制也能导致分化基因组岛的形成。大部分现有


的实证研究并未提供充分的证据来区分不同机制, 以至于结论的可信度略显不足(Cruickshank & Hahn, 2014; Guerrero & Hahn, 2017; Semenov et al, 2019)。例如, Malinsky等(2015)在研究*Astatotilapia* sp.同域分布的两个生态型分化基因组岛时, 由于未考虑导致 d_{XY} 更高的其他进化过程, 其所认为的形成机制是基因流存在下的歧化选择的证据并不充分。后来Guerrero和Hahn (2017)的再分析认为这两个生态型分化基因组岛更有可能是祖先多态性的差异分选导致的。本文梳理了分化基因组岛形成机制及其推断方法, 为类似研究的开展提供了借鉴。


然而, 解析种群基因组数据以了解姊妹种分化基因组岛的形成机制并不是成种研究的终极目标。维持生殖隔离的生态和遗传基础应是成种研究的重要因素, 但是很多实证研究在推断分化基因组岛的形成机制之后, 并未深入挖掘姊妹种间建立完全(或不完全)生殖隔离的生态与遗传基础(Turner et al, 2005; Choi et al, 2020; Hirase et al, 2021)。姊妹种间的遗传分化和先验的生态/表型差异的关联, 并不能直接证明分化基因组岛与生殖隔离的因果关系。如前所述, 鉴定得到的分化基因组岛并非一定与成种直接有关; 其次, 即使分化基因组岛是由基因流存在下的歧化选择这一机制所导致的, 其包含的多个候选基因也不一定都与生殖隔离有关, 有可能是与生殖隔离基因紧密连锁的位点; 另外, 分化基因组岛还可能存在不少未知功能的序列(Wolf & Ellegren, 2017)。这给准确理解生殖隔离的生态与遗传基础带来了困难, 而功能验证则为解决这些困难提供了路径(Poelstra et al, 2014; Todesco et al, 2020; Wang et al, 2021)。未来需要将种群基因组测序与受控选择实验、转录组分析以及日益发展的基因编辑技术等多方面手段整合, 以验证分化基因组岛中的候选基因在生殖隔离中发挥的功能, 确定姊妹种间建立生殖隔离的生态与遗传基础。


另外, 基因流存在下的成种假说提出后, 其在自然界中发生的频率尚不清楚(Nosil, 2008; Smadja & Butlin, 2011; Wolf & Ellegren, 2017)。目前分化基因组岛形成机制研究主要集中于少数几个类群, 未来需要分析更广泛的类群和不同类群中更多的姊妹种, 判断基因流存在下的歧化选择在姊妹种间分化基因组岛形成中发挥的作用。与此同时, 现有研

究使用的遗传变异数据类型和分析方法等存在一定程度的差异, 难以将不同的研究放在可比较的框架下进行整合分析。因此, 在未来研究中, 需制定标准分析流程、统一遗传变异数据类型、基于更完善的理论模型来规范结果的解读, 解答基因流存在下的成种事件在自然界中的常见程度等重要问题。

ORCID

孙琼  <https://orcid.org/0000-0003-3673-5677>

王嵘  <https://orcid.org/0000-0003-4652-2149>

陈小勇  <https://orcid.org/0000-0002-4795-8940>

参考文献

- Ackiss AS, Larson WA, Stott W (2020) Genotyping-by-sequencing illuminates high levels of divergence among sympatric forms of coregonines in the Laurentian Great Lakes. *Evolutionary Applications*, 13, 1037–1054.
- Aeschbacher S, Selby JP, Willis JH, Coop G (2017) Population-genomic inference of the strength and timing of selection against gene flow. *Proceedings of the National Academy of Sciences, USA*, 114, 7061–7066.
- Andrew RL, Rieseberg LH (2013) Divergence is focused on few genomic regions early in speciation: Incipient speciation of sunflower ecotypes. *Evolution*, 67, 2468–2482.
- Batley CJ (2020) Evidence of linked selection on the Z chromosome of hybridizing hummingbirds. *Evolution*, 74, 725–739.
- Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society of London Series B: Biological Sciences*, 263, 1619–1626.
- Behrens KA, Girasek QL, Sickler A, Hyde J, Buonaccorsi VP (2021) Regions of genetic divergence in depth-separated *Sebastes* rockfish species pairs: Depth as a potential driver of speciation. *Molecular Ecology*, 30, 4259–4275.
- Berg PR, Jentoft S, Star B, Ring KH, Knutsen H, Lien S, Jakobsen KS, André C (2015) Adaptation to low salinity promotes genomic divergence in Atlantic cod (*Gadus morhua* L.). *Genome Biology and Evolution*, 7, 1644–1663.
- Booker TR, Yeaman S, Whitlock MC (2020) Variation in recombination rate affects detection of outliers in genome scans under neutrality. *Molecular Ecology*, 29, 4274–4279.
- Bradbury IR, Hubert S, Higgins B, Bowman S, Borza T, Paterson IG, Snelgrove PVR, Morris CJ, Gregory RS, Hardie D, Hutchings JA, Ruzzante DE, Taggart CT, Bentzen P (2013) Genomic islands of divergence and their consequences for the resolution of spatial structure in an exploited marine fish. *Evolutionary Applications*, 6, 450–461.
- Burri R (2017) Interpreting differentiation landscapes in the light of long-term linked selection. *Evolution Letters*, 1,

- 118–131.
- Burri R, Nater A, Kawakami T, Mugal CF, Olason PI, Smeds L, Suh A, Dutoit L, Bureš S, Garamszegi LZ, Hogner S, Moreno J, Qvarnström A, Ružić M, Sæther SA, Sætre GP, Török J, Ellegren H (2015) Linked selection and recombination rate variation drive the evolution of the genomic landscape of differentiation across the speciation continuum of *Ficedula* flycatchers. *Genome Research*, 25, 1656–1665.
- Campagna L, Repenning M, Silveira LF, Fontana CS, Tubaro PL, Lovette IJ (2017) Repeated divergent selection on pigmentation genes in a rapid finch radiation. *Science Advances*, 3, e1602404.
- Charlesworth B (1998) Measures of divergence between populations and the effect of forces that reduce variability. *Molecular Biology and Evolution*, 15, 538–543.
- Charlesworth B, Morgan MT, Charlesworth D (1993) The effect of deleterious mutations on neutral molecular variation. *Genetics*, 134, 1289–1303.
- Charlesworth B, Nordborg M, Charlesworth D (1997) The effects of local selection, balanced polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided populations. *Genetical Research*, 70, 155–174.
- Cheng X, Li LL, Xiao Y, Chen XY, Hu XS (2020) Advances in the methods of detecting interspecific gene introgression and their applications. *Scientia Sinica Vita*, 50, 1388–1404. (in Chinese with English abstract) [程祥, 李玲玲, 肖钰, 陈晓阳, 胡新生 (2020) 种间基因渐渗检测方法及其应用研究进展. *中国科学: 生命科学*, 50, 1388–1404.]
- Choi JY, Purugganan M, Stacy EA (2020) Divergent selection and primary gene flow shape incipient speciation of a riparian tree on Hawaii Island. *Molecular Biology and Evolution*, 37, 695–710.
- Costello MJ, May RM, Stork NE (2013) Can we name earth's species before they go extinct? *Science*, 339, 413–416.
- Cruickshank TE, Hahn MW (2014) Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular Ecology*, 23, 3133–3157.
- Cutter AD, Payseur BA (2013) Genomic signatures of selection at linked sites: Unifying the disparity among species. *Nature Reviews Genetics*, 14, 262–274.
- Ellegren H, Smeds L, Burri R, Olason PI, Backström N, Kawakami T, Künstner A, Mäkinen H, Nadachowska-Brzyska K, Qvarnström A, Uebbing S, Wolf JBW (2012) The genomic landscape of species divergence in *Ficedula* flycatchers. *Nature*, 491, 756–760.
- Excoffier L, Dupanloup I, Huerta-Sánchez E, Sousa VC, Foll M (2013) Robust demographic inference from genomic and SNP data. *PLoS Genetics*, 9, e1003905.
- Excoffier L, Hofer T, Foll M (2009) Detecting loci under selection in a hierarchically structured population. *Heredity*, 103, 285–298.
- Feder JL, Egan SP, Nosil P (2012a) The genomics of speciation-with-gene-flow. *Trends in Genetics*, 28, 342–350.
- Feder JL, Gejji R, Yeaman S, Nosil P (2012b) Establishment of new mutations under divergence and genome hitchhiking. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences*, 367, 461–474.
- Feder JL, Nosil P (2010) The efficacy of divergence hitchhiking in generating genomic islands during ecological speciation. *Evolution*, 64, 1729–1747.
- Feder JL, Xie XF, Rull J, Velez S, Forbes A, Leung B, Dambroski H, Filchak KE, Aluja M (2005) Mayr, Dobzhansky, and Bush and the complexities of sympatric speciation in *Rhagoletis*. *Proceedings of the National Academy of Sciences, USA*, 102, 6573–6580.
- Feulner PGD, Chain FJJ, Panchal M, Huang Y, Eizaguirre C, Kalbe M, Lenz TL, Samonte IE, Stoll M, Bornberg-Bauer E, Reusch TBH, Milinski M (2015) Genomics of divergence along a continuum of parapatric population differentiation. *PLoS Genetics*, 11, e1004966.
- Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, 180, 977–993.
- Gao F, Li HP (2016) Application of computer simulators in population genetics. *Hereditas (Beijing)*, 38, 707–717. (in Chinese with English abstract) [高峰, 李海鹏 (2016) 群体遗传学模拟软件应用现状. *遗传*, 38, 707–717.]
- Guerrero RF, Hahn MW (2017) Speciation as a sieve for ancestral polymorphism. *Molecular Ecology*, 26, 5362–5368.
- Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD (2009) Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genetics*, 5, e1000695.
- Han F, Lamichhaney S, Grant BR, Grant PR, Andersson L, Webster MT (2017) Gene flow, ancient polymorphism, and ecological adaptation shape the genomic landscape of divergence among Darwin's finches. *Genome Research*, 27, 1004–1015.
- Harr B (2006) Genomic islands of differentiation between house mouse subspecies. *Genome Research*, 16, 730–737.
- Harrison RG (2012) The language of speciation. *Evolution*, 66, 3643–3657.
- Harrison RG, Larson EL (2014) Hybridization, introgression, and the nature of species boundaries. *Journal of Heredity*, 105, 795–809.
- Hedrick PW (2013) Adaptive introgression in animals: Examples and comparison to new mutation and standing variation as sources of adaptive variation. *Molecular Ecology*, 22, 4606–4618.
- Hirase S, Yamasaki YY, Sekino M, Nishisako M, Ikeda M, Hara M, Merilä J, Kikuchi K (2021) Genomic evidence for speciation with gene flow in broadcast spawning marine invertebrates. *Molecular Biology and Evolution*, 38,

- 4683–4699.
- Hofer T, Foll M, Excoffier L (2012) Evolutionary forces shaping genomic islands of population differentiation in humans. *BMC Genomics*, 13, 107.
- Hudson AG, Vonlanthen P, Bezault E, Seehausen O (2013) Genomic signatures of relaxed disruptive selection associated with speciation reversal in whitefish. *BMC Evolutionary Biology*, 13, 108.
- Hudson RR, Slatkin M, Maddison WP (1992) Estimation of levels of gene flow from DNA sequence data. *Genetics*, 132, 583–589.
- Irwin DE, Milá B, Toews DPL, Brelsford A, Kenyon HL, Porter AN, Grossen C, Delmore KE, Alcaide M, Irwin JH (2018) A comparison of genomic islands of differentiation across three young avian species pairs. *Molecular Ecology*, 27, 4839–4855.
- Larson WA, Dann TH, Limborg MT, McKinney GJ, Seeb JE, Seeb LW (2019) Parallel signatures of selection at genomic islands of divergence and the major histocompatibility complex in ecotypes of sockeye salmon across Alaska. *Molecular Ecology*, 28, 2254–2271.
- Larson WA, Limborg MT, McKinney GJ, Schindler DE, Seeb JE, Seeb LW (2017) Genomic islands of divergence linked to ecotypic variation in sockeye salmon. *Molecular Ecology*, 26, 554–570.
- Lyu HM, Zhou RC, Shi SH (2015) Recent advances in the study of ecological speciation. *Biodiversity Science*, 23, 398–407. (in Chinese with English abstract) [吕昊敏, 周仁超, 施苏华 (2015) 生态物种形成及其研究进展. *生物多样性*, 23, 398–407.]
- Ma T, Wang K, Hu QJ, Xi ZX, Wan DS, Wang Q, Feng JJ, Jiang DC, Ahani H, Abbott RJ, Lascoux M, Nevo E, Liu JQ (2018) Ancient polymorphisms and divergence hitchhiking contribute to genomic islands of divergence within a poplar species complex. *Proceedings of the National Academy of Sciences, USA*, 115, E236–E243.
- Malinsky M, Challis RJ, Tyers AM, Schiffels S, Terai Y, Ngatunga BP, Miska EA, Durbin R, Genner MJ, Turner GF (2015) Genomic islands of speciation separate cichlid ecomorphs in an East African crater lake. *Science*, 350, 1493–1498.
- Malinsky M, Matschiner M, Svoldal H (2021) Dsuite - Fast D-statistics and related admixture evidence from VCF files. *Molecular Ecology Resources*, 21, 584–595.
- Mallet J, Meyer A, Nosil P, Feder JL (2009) Space, sympatry and speciation. *Journal of Evolutionary Biology*, 22, 2332–2341.
- Mao JF, Ma YP, Zhou RC (2017) Approaches used to detect and test hybridization: Combining phylogenetic and population genetic analyses. *Biodiversity Science*, 25, 577–599. (in Chinese with English abstract) [毛建丰, 马永鹏, 周仁超 (2017) 结合系统发育与群体遗传学分析检验杂交是否存在的技术策略. *生物多样性*, 25, 577–599.]
- Marques DA, Lucek K, Meier JI, Mwaiko S, Wagner CE, Excoffier L, Seehausen O (2016) Genomics of rapid incipient speciation in sympatric threespine stickleback. *PLoS Genetics*, 12, e1005887.
- Martin SH, Davey JW, Jiggins CD (2015) Evaluating the use of ABBA-BABA statistics to locate introgressed loci. *Molecular Biology and Evolution*, 32, 244–257.
- Mayr E (1942) *Systematics and the Origin of Species*. Columbia University Press, New York.
- McCulloch GA, Foster BJ, Dutoit L, Harrop TWR, Guhlin J, Dearden PK, Waters JM (2021) Genomics reveals widespread ecological speciation in flightless insects. *Systematic Biology*, 70, 863–876.
- Michel AP, Guelbeogo WM, Grushko O, Schemerhorn BJ, Kern M, Willard MB, Sagnon NF, Costantini C, Besansky NJ (2005) Molecular differentiation between chromosomally defined incipient species of *Anopheles funestus*. *Insect Molecular Biology*, 14, 375–387.
- Michel AP, Sim S, Powell THQ, Taylor MS, Nosil P, Feder JL (2010) Widespread genomic divergence during sympatric speciation. *Proceedings of the National Academy of Sciences, USA*, 107, 9724–9729.
- Mora C, Tittensor DP, Adl S, Simpson AGB, Worm B (2011) How many species are there on Earth and in the ocean? *PLoS Biology*, 9, e1001127.
- Morales HE, Pavlova A, Amos N, Major R, Kilian A, Greening C, Sunnucks P (2018) Concordant divergence of mitogenomes and a mitonuclear gene cluster in bird lineages inhabiting different climates. *Nature Ecology & Evolution*, 2, 1258–1267.
- Mořkovský L, Janoušek V, Reif J, Řídl J, Pačes J, Choleva L, Janko K, Nachman MW, Reifová R (2018) Genomic islands of differentiation in two songbird species reveal candidate genes for hybrid female sterility. *Molecular Ecology*, 27, 949–958.
- Nachman MW, Payseur BA (2012) Recombination rate variation and speciation: Theoretical predictions and empirical results from rabbits and mice. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences*, 367, 409–421.
- Nadeau NJ, Whibley A, Jones RT, Davey JW, Dasmahapatra KK, Baxter SW, Quail MA, Joron M, French-Constant RH, Blaxter ML, Mallet J, Jiggins CD (2012) Genomic islands of divergence in hybridizing *Heliconius* butterflies identified by large-scale targeted sequencing. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences*, 367, 343–353.
- Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.
- Noor MAF, Bennett SM (2009) Islands of speciation or mirages in the desert? Examining the role of restricted recombination in maintaining species. *Heredity*, 103, 439–444.

- Nosil P (2008) Speciation with gene flow could be common. *Molecular Ecology*, 17, 2103–2106.
- Nosil P (2012) *Ecological Speciation*. Oxford University Press, New York.
- Nosil P, Egan SP, Funk DJ (2008) Heterogeneous genomic differentiation between walking-stick ecotypes: “Isolation by adaptation” and multiple roles for divergent selection. *Evolution*, 62, 316–336.
- Nosil P, Feder JL (2012) Genomic divergence during speciation: Causes and consequences. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences*, 367, 332–342.
- Nosil P, Funk DJ, Ortiz-Barrimentos D (2009) Divergent selection and heterogeneous genomic divergence. *Molecular Ecology*, 18, 375–402.
- Oziolor EM, Reid NM, Yair S, Lee KM, Guberman VerPloeg S, Bruns PC, Shaw JR, Whitehead A, Matson CW (2019) Adaptive introgression enables evolutionary rescue from extreme environmental pollution. *Science*, 364, 455–457.
- Papadopulos AST, Igea J, Dunning LT, Osborne OG, Quan XP, Pellicer J, Turnbull C, Hutton I, Baker WJ, Butlin RK, Savolainen V (2019) Ecological speciation in sympatric palms: 3. Genetic map reveals genomic islands underlying species divergence in *Howea*. *Evolution*, 73, 1986–1995.
- Peñalba JV, Wolf JBW (2020) From molecules to populations: Appreciating and estimating recombination rate variation. *Nature Reviews Genetics*, 21, 476–492.
- Pfeifer B, Kapan DD (2019) Estimates of introgression as a function of pairwise distances. *BMC Bioinformatics*, 20, 207.
- Poelstra JW, Vijay N, Bossu CM, Lantz H, Ryll B, Müller I, Baglione V, Unneberg P, Wikelski M, Grabherr MG, Wolf JBW (2014) The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. *Science*, 344, 1410–1414.
- Racimo F, Sankararaman S, Nielsen R, Huerta-Sánchez E (2015) Evidence for archaic adaptive introgression in humans. *Nature Reviews Genetics*, 16, 359–371.
- Ravinet M, Faria R, Butlin RK, Galindo J, Bierne N, Rafajlović M, Noor MAF, Mehlig B, Westram AM (2017) Interpreting the genomic landscape of speciation: A road map for finding barriers to gene flow. *Journal of Evolutionary Biology*, 30, 1450–1477.
- Renaut S, Grassa CJ, Yeaman S, Moyers BT, Lai Z, Kane NC, Bowers JE, Burke JM, Rieseberg LH (2013) Genomic islands of divergence are not affected by geography of speciation in sunflowers. *Nature Communications*, 4, 1827.
- Rettelbach A, Nater A, Ellegren H (2019) How linked selection shapes the diversity landscape in *Ficedula* flycatchers. *Genetics*, 212, 277–285.
- Richards EJ, Martin CH (2017) Adaptive introgression from distant Caribbean islands contributed to the diversification of a microendemic adaptive radiation of trophic specialist pupfishes. *PLoS Genetics*, 13, e1006919.
- Rosenzweig BK, Pease JB, Besansky NJ, Hahn MW (2016) Powerful methods for detecting introgressed regions from population genomic data. *Molecular Ecology*, 25, 2387–2397.
- Rundle HD, Nosil P (2005) Ecological speciation. *Ecology Letters*, 8, 336–352.
- Sankararaman S, Mallick S, Dannemann M, Prüfer K, Kelso J, Pääbo S, Patterson N, Reich D (2014) The genomic landscape of Neanderthal ancestry in present-day humans. *Nature*, 507, 354–357.
- Semenov GA, Safran RJ, Smith CCR, Turbek SP, Mullen SP, Flaxman SM (2019) Unifying theoretical and empirical perspectives on genomic differentiation. *Trends in Ecology & Evolution*, 34, 987–995.
- Smadja CM, Butlin RK (2011) A framework for comparing processes of speciation in the presence of gene flow. *Molecular Ecology*, 20, 5123–5140.
- Smith JM, Haigh J (1974) The hitch-hiking effect of a favourable gene. *Genetical Research*, 23, 23–35.
- Smith NGC, Fearnhead P (2005) A comparison of three estimators of the population-scaled recombination rate: Accuracy and robustness. *Genetics*, 171, 2051–2062.
- Soria-Carrasco V, Gompert Z, Comeault AA, Farkas TE, Parchman TL, Johnston JS, Buerkle CA, Feder JL, Bast J, Schwander T, Egan SP, Crespi BJ, Nosil P (2014) Stick insect genomes reveal natural selection’s role in parallel speciation. *Science*, 344, 738–742.
- Stevison LS, McGaugh SE (2020) It’s time to stop sweeping recombination rate under the genome scan rug. *Molecular Ecology*, 29, 4249–4253.
- Tavares H, Whibley A, Field DL, Bradley D, Couchman M, Copsey L, Elleouet J, Burrus M, Andalo C, Li MM, Li Q, Xue YB, Rebocho AB, Barton NH, Coen E (2018) Selection and gene flow shape genomic islands that control floral guides. *Proceedings of the National Academy of Sciences, USA*, 115, 11006–11011.
- Teng HJ, Zhang YH, Shi CM, Mao FB, Cai WS, Lu L, Zhao FQ, Sun ZS, Zhang JX (2017) Population genomics reveals speciation and introgression between brown Norway rats and their sibling species. *Molecular Biology and Evolution*, 34, 2214–2228.
- Todesco M, Owens GL, Bercovich N, Légaré JS, Soudi S, Burge DO, Huang KC, Ostevik KL, Drummond EBM, Imerovski I, Lande K, Pascual-Robles MA, Nanavati M, Jahani M, Cheung W, Staton SE, Muños S, Nielsen R, Donovan LA, Burke JM, Yeaman S, Rieseberg LH (2020) Massive haplotypes underlie ecotypic differentiation in sunflowers. *Nature*, 584, 602–607.
- Tsumura Y, Uchiyama K, Moriguchi Y, Ueno S, Ihara-Ujjino T (2012) Genome scanning for detecting adaptive genes along environmental gradients in the Japanese conifer, *Cryptomeria japonica*. *Heredity*, 109, 349–360.

孙琼等: 物种形成过程中的分化基因组岛及其形成机制

- Turner TL, Hahn MW, Nuzhdin SV (2005) Genomic islands of speciation in *Anopheles gambiae*. *PLoS Biology*, 3, e285.
- Via S (2009) Natural selection in action during speciation. *Proceedings of the National Academy of Sciences, USA*, 106, 9939–9946.
- Vijay N, Bossu CM, Poelstra JW, Weissensteiner MH, Suh A, Kryukov AP, Wolf JBW (2016) Evolution of heterogeneous genome differentiation across multiple contact zones in a crow species complex. *Nature Communications*, 7, 13195.
- Wang BS, Mojica JP, Perera N, Lee CR, Lovell JT, Sharma A, Adam C, Lipzen A, Barry K, Rokhsar DS, Schmutz J, Mitchell-Olds T (2019) Ancient polymorphisms contribute to genome-wide variation by long-term balancing selection and divergent sorting in *Boechera stricta*. *Genome Biology*, 20, 126.
- Wang J, Street NR, Scofield DG, Ingvarsson PK (2016) Variation in linked selection and recombination drive genomic divergence during allopatric speciation of European and American aspens. *Molecular Biology and Evolution*, 33, 1754–1767.
- Wang ZF, Jiang YZ, Bi H, Lu ZQ, Ma YZ, Yang XY, Chen NN, Tian B, Liu BB, Mao XX, Ma T, DiFazio SP, Hu QJ, Abbott RJ, Liu JQ (2021) Hybrid speciation via inheritance of alternate alleles of parental isolating genes. *Molecular Plant*, 14, 208–222.
- Wolf JBW, Ellegren H (2017) Making sense of genomic islands of differentiation in light of speciation. *Nature Reviews Genetics*, 18, 87–100.
- Wu CI (2001) The genic view of the process of speciation. *Journal of Evolutionary Biology*, 14, 851–865.
- Wu CI, Ting CT (2004) Genes and speciation. *Nature Reviews Genetics*, 5, 114–122.
- Yeaman S, Aeschbacher S, Bürger R (2016) The evolution of genomic islands by increased establishment probability of linked alleles. *Molecular Ecology*, 25, 2542–2558.

(责任编辑: 郭亚龙 责任编辑: 李会丽)