

- gosity of chromosome segments in finite populations[J]. Theor Point Biol 1971, 2: 125 - 141.
- [20] GODDARD M E. Mapping genes for quantitative traits using linkage disequilibrium [J]. Genet Sel 1991, 23 (Suppl 1): 131 - 134.
- [21] MEUWESSEN T H R, HAYES B, GODDARD M E. Prediction of total genetic value using genome wide dense marker maps[J]. Genetics 2001, 157: 1 819 - 1 829.
- [22] SHORT T H R, SOUTHWOOD M E, McLAREN D O, et al. Effect of the estrogen receptor locus on reproduction and production traits in four commercial pig lines[J]. J Anim Sci 1997, 75: 3 138 - 3 142.
- [23] ISRAEL C, WELLER J I. Estimation of candidate gene effects in dairy cattle populations[J]. J Dairy Sci 1998, 81: 1 653 - 1 662.
- [24] HAYES B, BOWMAN P J, GODDARD M E. Linkage disequilibrium and accuracy of predicting breeding values from marker haplotypes [J]. Proc Assoc Advmt Anim Breed Genet 2001, 14: 269 - 272.
- [25] HAYES B, GODDARD M E. Evaluation of marker assisted selection in pig enterprises[J]. Livest Prod Sci 2003, 81: 197 - 211.
- [26] HOSPITAL F A B L, CAUSSE L M, CHARCOSSET A. Use of markers in plant breeding lessons from genotype building experiments [A]. Proc 7th Wld Congr Genet Appl Livest Prod. Electronic Communication 22 - 05 [C]. 2002.
- [27] SOLLER M, MEDJUGOMA I. A successful marriage making the transition from quantitative trait locus mapping to marker assisted selection[A]. DEKKERS C M, LAMONT S J, ROTHSCILD M E. From Jay Lush to genomics: visions for animal breeding and genetics Dept Animal Science[C]. Ames: Iowa State University 1999.
- [28] JEON J, CARLBORG O, TORNSTEN A, et al. A paternally expressed QTL affecting skeletal and cardiac muscle mass in pigs maps to the GF2 locus[J]. Nat Genet 1999, 21: 157 - 158.
- [29] De KONING D, RATTINK A B, van AREN-DONK H B, et al. Genome wide scan for body composition in pigs reveals important role of imprinting[J]. Proc Natl Acad Sci 2000, 97: 7 947 - 7 950.
- [30] De KONING D J. Identification of (non-) Mendelian factors affecting pork production[D]. The Netherlands: Wageningen University, 2001.

(Edited by ZHOU Zhihong)

译文

分子遗传技术与数量遗传技术结合使遗传改良速度最大化

Jack C. M. DEKKERS

(Department of Animal Science and Center for Integrated Animal Genomics
239D Killee Hall Iowa State University Ames IA 50011 USA)

摘要: 随着分子遗传学的发展, 为家畜的遗传改良提供了新的手段, 可以直接选择影响性状的基因或包含基因的染色体区域, 即所谓数量性状基因座 (QTL)。QTL的检测和应用依赖于可以分型的遗传标记和 QTL之间是否存在连锁不平衡。本文对畜群中存在的或可以在试验畜群中产生的连锁不平衡的特性、及其在 QTL检测和标记辅助选择中的作用进行了综述, 并且概述了标记辅助选择的不同策略的优缺点。主要的观点是利用标记辅助选择加快遗传改良的速度是可行的, 特别是可以利用群体范围的连锁不平衡时更是如此, 但是这需要认真地分析并在现有育种方案中落实。

关键词: 分子遗传; 数量遗传; 数量性状基因座; 标记辅助选择

到目前为止, 大部分家畜数量性状的遗传进展, 是在对影响这些性状的基因数量和效应未知的情况下, 根据表型值计算出来的 EBV 进行选种而取得的。在利用数量遗传学方法进行遗传改良时, 性状的遗传结构实际上是当作“黑箱”处理的。尽管如此, 主要家畜品种仍取得了, 而且将继续取得可观的遗传进展, 很明显利用数量遗传方法进行选择是有效的。但是, 利用数量遗传学方法取得的巨大成功, 并不能否认在对数量性状黑箱了解后可以取得更大的进展。通过在 DNA 水平研究个体的遗传组成, 分子遗传学为我们提供了实现这一目标的工具。在遗传选择中, 分子水平数据对我们有用是因为基因型的遗传力等于 1 (假定没有基因型判型错误), 所有性别和个体的分子数据都可以测定, 可以在生命早期获得, 也可以在几乎没有表型信息时进行等。本文的目的是对在遗传改良中基因检测和利用的策略作一综述。

1 数量性状基因座 (QTL) 检测的原理

分子遗传学在遗传改良中的应用依赖于测定个体特定座位基因型的能力。如 Dekkers^[1]指出, 有 3 种检测基因座类型的方法:

- 1) 直接标记: 可以直接对功能基因进行分型的基因座;
- 2) 连锁不平衡 (LD) 标记: 功能突变造成的群体连锁不平衡的基因座;
- 3) 连锁平衡 (LE) 标记: 功能突变造成的群体连锁不平衡的基因座。

对这些基因座的检测和利用影响数量性状的基因座 (QTL), 依赖于是否能够发现标记座位的基因型和表型之间的关联。通过比较不同基因型个体表型平均值的差异, 可以发现是否存在这种关联。如果不同基因型个体间的表型值存在差异, 则说明该标记和一个 QTL 连锁, 或者说, 在理想的情况下, 该标记就是一个 QTL (直接标记)。不过, 这不等于说每个与 QTL 连锁的标记在不同基因型间的表型值都存在差异; 除了连锁以外, 第 2 种情况是, 在标记和 QTL 存在连锁不平衡 (LD) 时, 不同标记基因型个体间表型值也有差异。LD 的概念对于 QTL 检测和应用 QTL 进行选择 (标记辅助选择, MAS) 都很重要, 下面我们就要解释这个问题。

1.1 连锁不平衡

考虑一个标记座位, 其等位基因分别为 M 和 m , QTL 的等位基因分别为 Q 和 q 它们位于同一条染色体上, 即标记和 QTL 是连锁的。一个个体如果在这 2 个座位都是杂合的, 则其基因型为 $MmQq$ 。2 个座位的等位基因在 2 条同源染色体上是以单倍型的形式存在的。一个 $MmQq$ 基因型的个体可能有 2 种单倍型: MQ 和 mq 即 2 条同源染色体; 另一种可能是 2 个异源单倍型为 Mq 和 mQ 。连锁座位在同源染色体中的不同排列称为标记-QTL 的连锁相。单倍型中排列的重要性在于, 假如没有重组, 则后裔只能从一个亲本遗传一种单倍型。

连锁平衡 (LE) 和连锁不平衡 (LD) 取决于群体中不同单倍型的相对频率。在一个连锁平衡群体中, 2 个座位的等位基因随机地分配到不同的单倍型。换句话说, 标记等位基因 M 的染色体或单倍型携带 QTL 等位基因 Q 的概率不会比标记等位基因 m 所在染色体携带的概率高。用技术术语讲就是, MQ 单倍型的频率等于 M 的频率和 Q 的频率之积。因此, 如果标记和 QTL 处于连锁平衡状态, 知道个体标记基因型是没有意义的, 因为此时它不能提供 QTL 基因型方面的信息。但是, 如果标记和 QTL 为连锁不平衡状态, 则 M 和 m 标记等位基因所在的染色体携带 QTL 等位基因 Q 的概率不同, 因此, 这时不同标记基因型间的表型值可能会有不同。

标记和 QTL 之间的连锁不平衡是 QTL 检测和 MAS 的基础。因此, 对 LD 是否存在和不平衡程度的了解是十分重要的。群体中产生 LD 的主要因素是突变、选择、漂变 (近交)、迁移或杂交。使 LD 下降的主要因素为重组, 在每个世代, 重组都会使亲本的单倍型重排。图 1 显示的是重组使 LD 随世代数增加而降低的情况。LD 下降的速度与座位间的重组率有关, 即和它们在染色体上的距离有关。对于紧密连锁的座位, 任何 LD 都可以维持很长时间, 但对于非紧密连锁有座位 (如 $r=0.1$), LD 会随世代数增加而很快下降。

虽然在整个群体范围内标记和其连锁的 QTL 间可能为连锁平衡状态, 但即使对于非紧密连锁的座位之间, 家系内 LD 总是存在的。考虑一头双杂合子单倍型公畜 MQ 和 mq (图 2)。该公畜的基因型和近交系间杂交 F_1 的基因型相同, 该公畜会产生 4 种类型的配子: 非重组型 MQ 和 mq , 重组型 Mq 和 mQ 。因

为非重组型的频率高, 偏离的程度随重组率的不同而不同, 该公畜产生的配子处于 LD 状态, 即使距离远的座位也是如此(图 1), 因为这只是一个世代的重组。不过, 这种类型的 LD 仅存在于家系内; 另一头 Mq/mQ 公畜的后裔, 也会表现出 LD, 但这种 LD 的方向相反, 因为公畜的标记 -QTL 连锁相不同(图 2)。另一方面, MQ/mQ 和 Mq/mq 公畜家系不处于 LD 状态, 因为这些家系中 QTL 不分离。这种不同家系的 4 种类型 LD 数据混在一起时会相互抵消, 这样整个群体就会处于连锁平衡状态。不过, 家系内 LD 可以用来检测 QTL 和 MAS 但需要考虑连锁相的不同, 后面我们将阐述这个问题。

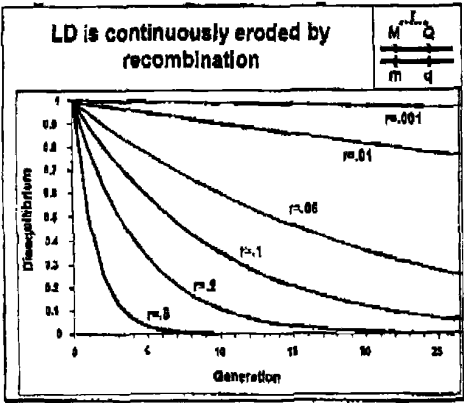


图 1 LD 随世代下降的图示

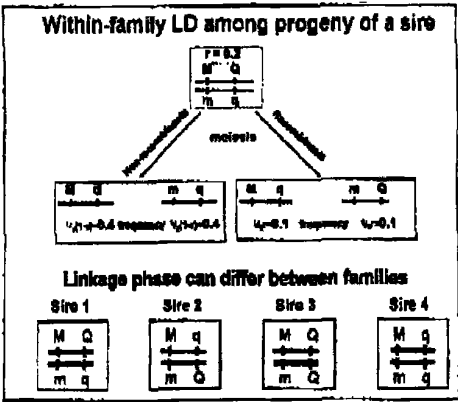


图 2 家系内标记-QTL 的 LD

1.2 利用 LD 来检测 QTL

利用遗传标记检测 QTL 的方法依赖于识别标记和表型之间的关联。相关情况只有在标记和 QTL 处于 LD 状态时才行, 因此依赖于所分析群体中 LD 的类型和大小。标记是 LD 还是 LE 与群体结构和群体中 LD 的大小有关。对于品种杂交来说, LD 广泛存在, 即使是与 QTL 距离为 10~20 cM 的标记, 在群体范围内它们之间也存在着 LD。但是, 对于闭

锁的远交群来说, LD 的大小有限, 与 QTL 存在 LD 的标记必须与功能突变位点相近(根据群体历史不同, 在 1~5 cM 间)。在闭锁的繁育群内, 找到 LD 标记比较困难, 需要高密度的标记图谱, 利用候选基因法或精细定位方法^[2]。检测出功能突变位点是最难的, 只有很少的成功例子^[2]。下面阐述各种情形下 QTL 检测的原理。

1.2.1 杂交群中存在群体范围的 LD 基因不同的 2 个品种杂交, 由于单倍型频率的差异, F_1 群体在距离较大的座位存在的 LD, 在 F_2 经过 1 个世代的重组的 F_2 群体中仍广泛存在 LD(图 1)。这样, 基因不同的 2 个品种杂交后, 可以应用分散在基因组中有限的标记(大约每 15~20 cM), 利用基因组扫描进行 QTL 检测, 这是利用 F_2 或回交品种或品系进行 QTL 定位的基础(如 Malek 等^[3,4])。这些广泛存在的 LD 可供检测与标记有一定距离的 QTL 但同时也限制我们不能精确定位 QTL 的位置。

品系杂交初期形成的合成系, 在群体范围内存在较多的 LD。随着杂交后代代数的增加, LD 的大小会有所减少, 因此所覆盖的距离比 F_2 群体中的要短(图 1)。这时进行 QTL 定位需要更密的标记图谱, 这样可以更精细地确定 QTL 的位置。

1.2.2 远交群中家系内 LD 因为不同家系间标记和 QTL 间的连锁相不同, 利用家系内 LD 进行 QTL 检测时, 需要对家系内, 而不是整个群体的标记效应进行拟合。与 F_2 或回交群相似, 家系内 LD 也是广泛存在的, 因此, 可以用覆盖整个基因组有限的与 QTL 有一定距离的显著标记进行 QTL 定位。因此, 利用品种杂交或者半同胞家系分析, LE 标记可以在基因组范围内检测出来, 这时仅需要粗略的标记图谱(标记间间隔 20 cM)。文献报道了很多利用这一方法检测出 QTL 的成功例子。

1.2.3 远交群中群体范围的 LD 群体中有足够多的 LD 可用于遗传改良, 这纯粹是由于创造和降低 LD 的各种因素共同作用的结果, 因此这是每个群体选育过程中随机抽样的结果。在这一基础上, 除了紧密连锁的基因座以外, 经过多个世代闭锁的群体一般处于平衡状态。因此, 在这些群体中, 只有与 QTL 紧密连锁的标记才有可能与表型相关联(图 1), 即使如此, 也由于存在随机抽样效应, 不能保证一定会存在关联。

有 2 种策略可以检测出在群体水平与 QTL 连锁的 LD 标记:

1) 评估与性状关联的基因(候选基因)内部或附近的标记;

2) 利用标记间距 0.5~2.0 M 的高密度连锁图谱进行基因组扫描。

很明显,这 2 种方法是否成功依赖于群体中 LD 的大小。通常人类群体中的研究很少发现 LD 在 1 M 之内。因此,在人类群体中需要许多标记才能依据群体范围内的 LD 进行 QTL 定位。因为家畜群体中存在选择和近交,因此,利用群体范围的 LD 进行 QTL 定位成功的机会要大一些。实际上,Famir 等^[5]在荷兰黑白花奶牛群中发现了许多 LD,其区间超过 5 M。其他家畜中也有相似的结果(如家禽,Heifetz 等^[6])。家畜中广泛存在的 LD 对于 QTL 检测是有利的,但对于检测出这些 QTL 座出现的自发突变则是不利的;由于广泛存在 LD 与自发突变有一定距离的标记也可以与表型产生关联。

候选基因法利用的是物种丰富的基因组信息(如人类、小鼠),其他物种中的突变效应,以前检测出的 QTL 区域,以及根据性状的生理学基础,可以鉴别出对性状生理过程起重要作用的基因。利用猪种中基因图谱和多态性,在闭锁繁育群中可以评估候选基因的基因型和表型的关联(Rothschild 等^[7])。

利用基因组技术的最新进展,可以对整个基因组测序,已经有多个畜种正在进行这项工作,鸡和牛的测序工作已经完成,猪基因组的测定正在进行。此外,测序还可用于鉴定基因组中大量的位点,包括单核苷酸多态(SNPs),即 DNA 碱基的变异。例如,通过 3 种家鸡与红色原鸡的比较,已经找出了 280 万的 SNPs(国际鸡多态性图谱联盟 ICPMG, 2004)。这方面的知识连同基因型测定方面成本的下降,使我们可以利用高密度标记 LD 图谱进行 QTL 检测。

2 标记辅助选择

Dekkers 和 Hospital^[8]对标记辅助选择在家畜和植物育种中应用的原则和局限进行了综述,现将 MAS 在家畜应用中的几个问题阐述如下:

2.1 应用标记辅助选择的性状

影响单基因性状和数量性状的座位或染色体区域可以通过分子标记进行鉴别。单基因性状包括遗传缺陷、疾病和外貌特征。为了 QTL 的检测和应用,可将数量性状划分为下列几类:a) 常规记录性状,可进一步分为非限性性状、限性性状和一些宰后性状;b) 难以记录的性状(采食量,产品质量);c) 无法记

录的性状(抗病力)。上面的 3 类数量性状检测 QTL 的可能性大小随着获得表型数据的难易程度,由 a)、b)、c) 的顺序依次下降。基于同样的原因,检测影响 a 类性状的 QTL 通常采用需较多表型数据的基因组扫描法,而检测影响其余 2 类性状的 QTL 一般选择候选基因法。就利用标记辅助选择所获得的额外遗传改进而言,c 类性状最高,a 类性状最低,尤其对于选择 2 种性别都具有常规记录的性状。借助标记辅助选择获得额外遗传改进的能力与通过常规方法获得遗传进展的能力成反比^[9]

2.2 应用分子遗传信息进行选择的一般策略

一旦发现与标记连锁的 QTL,就可利用表型与基因型间的关联来估计标记的效应。估计结果可作为每个候选个体的“分子评分”,分子评分可用于预测个体育种值和个体选择。分子评分的组成和计算依赖于所应用的 LD 类型和标记在选择中如何应用。

前面讲述的 3 种基因座类型不仅检测方法不同,而且在选择计划中的应用方法也不同。直接分子标记、以及程度较低的 LD 分子标记,可以在整个群体中通过基因型进行选择,而使用 LE 分子标记就必须考虑不同家系间标记和 QTL 之间不同的连锁相。因此,应用标记进行选择的难易程度和能力大小与它们检测的难度相反,从功能突变、ID 标记到 LE 标记依次提高了标记选择的能力。依据这 3 类标记选择的方法,分别被称为基因辅助选择(GAS)、LD 标记辅助选择和 LE 标记辅助选择

除了分子评分,个体也可以获得与性状有关的其他基因的(多基因)综合效应育种值。表 1 列出了遗传改良中应用分子遗传信息的各种方法和策略,详细内容在本文后面介绍。

总的来说,在选择的时候,要合理应用分子信息和表型信息,主要有下面 3 种选择策略:

- 1) 仅依据分子评分值的选择;
- 2) 顺序选择,在根据分子评分值的选择基础上,再依据表型值估计的 EBV 进行选择;
- 3) 根据分子评分值和常规 EBV 组合成一个指数进行选择。

只根据分子评分值的选择忽略了其他基因(多基因)对性状的影响信息,其选择效率最低,除非所有影响该性状的基因作用都包含在分子评分值中。但是,该策略不需要另外的表型值,仅需要估计标记评分值,当表型值很难记录或记录成本过高时,这种方法很有意义(例如疾病性状、肉质等)。

如果候选个体的表型和分子信息均可以利用，选择反应低相似；2步选择时分子评分低但表型 EBV 指数选择的效果应该比顺序选择的要好，其原因与 2 高的个体不会被选中。这些策略(及其他的)还依赖于其他因素，比如市场和成本。

表 1 遗传改良计划中应用分子信息的策略

类型	选择计划	标记要求	计算分子评分值需要的信息	分子评分内容	选择标准
品种间 ¹⁾	基因导入	距离	目的座位标记等位基因的品	有 无目的等位基因	分子评分
	· 导入选择	QTL<10 cM	系来源		
		基因组标记间距		占受体品种等位	分子评分(受体性
	· 恢复选择	~20 cM	标记等位基因的品系来源	基因的% ²⁾	状的表型值)
	· 横交	距离 QTL <5 ~10 cM	目的座位标记等位基因的品 系来源	目的等位基因数 量	分子评分
品种内	合成系开发	距离 QTL<10 cM	QTL 标记等位基因的品系来源 估计 QTL 标记 效应	QTL 标记估计结果的和	分子评分 表型 EBV
	LD -MAS /GAS	距离	QTL 标记的基因型		
	直接标记	QTL<1 ~2 cM	QTL 标记效应的估计	QTL 标记估计结果的和	分子评分
	候选基因				表型 EBV
	ID -标记				
	LE -MAS	距离 QTL<10 cM	标记等位基因的亲本来源 家系内 QTL 标记效应的评估值	QTL 标记估计结果的和	分子评分 表型 EBV

1)所有都是 LD -MAS 利用杂交种广泛存在的 ID; 2)要十分强调标记与目标座位的连锁, 以减少连锁副作用 (linkage drag)

2 3 利用品种差异的标记辅助改良

分子信息用于整合不同品种的优良特性有很多途径。如前所述, 所有有赖于杂交产生的广泛 LD, 这些 LD 可用于 MAS。如果造成性状品种间变异的主要是由少数基因引起的, 则可以应用标记辅助导入策略; 如果是由许多基因引起的, 则在合成系内利用标记辅助选择是首选的改良方法。下面将进一步阐述这些策略。

2 3 1 标记辅助导入 在肉质性状的改良中, 导入计划的目的是将一个或多个肉质基因 (目的基因) 从具有优良肉质、但性能较差的一个品种 (供体品种) 引入到另一个具有高生产性能、但缺乏目的基因的品种 (受体品种)。可以通过最初的 F₁ 代杂交, 然后与受体品种多次回交, 最后进行一代或几代的横交。回交的目的是获得一批携带一个拷贝的原供体 QTL 等位基因、其他大部分基因组与受体品种相似的个体。整个导入过程是通过连续多个世代与受体品种回交完成的, 经过对供体基因组的“稀释”; 同时选择仅携带原供体 QTL 基因的个体作为下一代的亲本以维持原供体的特定基因。横交的目的是固定供体的 QTL 等位基因。最终的结果是产生一个类似于受体品种, 同时携带 2 个拷贝供体 QTL 等位基因的群体。

标记辅助导入的有效性受到 2 种情况的限制, 一是鉴定回交或横交个体是否具有携带目的基因的

能力, 二是鉴定回交个体是否大部分基因组为受体基因组的能力, 特别是在目的基因区域附近。后者影响恢复大部分基因组为受体基因组所需要的回交代数。分子遗传学的发展大大提高了导入过程这 2 个阶段的效率。提高回交阶段的有效性有 2 种方法: i) 鉴定目的基因的携带者 (导入选择); ii) 提高恢复供体的基因背景能力 (恢复选择)。横交阶段的效率可以通过目的基因的导入选择来提高。需要注意的是无论导入选择还是恢复选择都不需要估计标记或者 QTL 的效应, 其选择效率仅仅取决于品种间差异, 以及标记与由于广泛存在的 LD 造成的品种差异间的关联。

对于导入选择和横交选择, 所依据都是根据目的基因存在与否的分子评分 (只选择带有目的等位基因的个体) (表 1)。如果不能直接确定目的基因的基因型, 则利用杂交群中丰富的 LD 用与该 QTL 连锁图距小于 10 cM 的侧翼标记来鉴别。所用标记的等位基因必须是品种特异的, 以便确认其来源。导入选择的效果取决于目标基因的数量和这些基因座的置信区间。后者决定了需要导入基因组的区间大小。这 2 个因素都对需要的群体数量有很大影响, 只有达到一定群体数量时才能在回交阶段鉴别出携带目的基因的个体, 以及在横交阶段鉴别出纯合子。对于多目的基因的标记辅助导入, 在回交阶

段,可以利用基因金字塔策略来减少所需的群体数量^[10-11]。

对于恢复选择,选用的标记要覆盖整个基因组,密度小于 20 cM,因此大部分影响性状的基因与标记间的图距在 10 cM 以内。此外,标记必须有品种特异的等位基因,这样才能追踪其品种来源。然后用标记基因型来估计某个体含有受体基因组的比例,结果可以用来分子评分表示(表 1),选留比例最高的个体。将导入选择和恢复选择结合起来,可以在目标座位附近选择受体品种,在基因组的其他片段选择供体品种。但是,导入选择不仅选择了目标座位,还选择了供体品种中与目标座位连锁的其他座位,这些座位可能对生产性能是不利的。为了减少这种所谓的目标座位附近的连锁副作用(linkage drag),在恢复选择时应突出对目标座位相邻标记的分子评分(除了用于导入选择的一些侧翼标记)。

大多数研究都只考虑单 QTL 的标记辅助导入(MAI)^[10],但实际上通常有多个 QTL 需要同时导入。有些研究^[11]表明,必须有大群体才能在回交时获得足够数量在所有的 QTL 座位都为杂合子的个体,这导致 MAI 在动物育种计划中难以应用。但是在很多情况下,没有必要立即固定 QTL 等位基因。回交阶段的目标可锁定在提高受体品种中含有供体有利 QTL 等位基因的基因频率,直到达到足够大以便在其后的回交中能很好地进行选择。对这种情形,在回交阶段可以根据分子评分值来选留个体,分子评分值的计算是依据在 n 个导入 QTL 中供体期望带有的等位基因数,它由标记基因型确定: $MS = \sum_{i=1}^n P(Q_i)$ (Q_i),这里, $P(Q_i)$ 是指携带供体 QTL i 等位基因的个体的概率。不考虑双重组,如果某个个体在 QTL 两侧的 2 个标记上携带 2、1、0 个供体等位基因,记概率 $P(Q_i)$ 为 1、0.5、0。在横交阶段,可以根据相似的分子评分值进行选择。Chaiwong 等^[12]评估了这些策略的有效性,研究表明尽管在群体大小有限的情况下,回交阶段可能无法保持 50% 的频率,但是 MAI 仍然能使多 QTL 等位基因频率达到在下一步的回交中可以进行选择的程度。

2.3.2 标记辅助合成系培育 Lande 等^[13]提出了在 2 个近交系杂交所产生的杂交群中的辅助标记选择策略。该策略利用系间或品种间杂交时最初存在的群体范围 LD 进行标记辅助选择。因此,在 F_2 代中检测到的 QTL-标记关联可以提供进行几代选择的信息,直到 QTL 或者标记被固定或者不平衡消失。

Zhang 等^[14]评估了在依据 BIUP EBV 选择的情形下标记辅助选择的应用价值,他们比较了如下 3 种选择策略。

MAS 选择依据为来自标记效应的分子评分值;
BIUP: 选择依据为由表型值估计的 BLUP EBV;
COMB: 选择依据为有标记和表型信息估计的 EBV 组合成的指数。

假定一个由 2 个近交系杂交的群体,有 100 个 QTL,100 个标记、基因组长度为 2 000 cM。在 F_2 代分 2 步估计标记的效应。第一步,在另外一个同样的 F_2 代分离群体检测最大效应的标记;然后,为获得无偏估计,这些标记的效应在选择后的 F_2 代群体重新估计,后者用来估计以标记为基础的 EBV,将它作为在以后的选择过程中的分子评分值。Zhang 等^[14]发现,指数选择 (COMB) 的选择反应最大,其次是 BLUP EBV 选择和仅用标记进行的选择。随着世代进展,所有策略的选择进展都下降,这是因为模拟用的是有限位点,随着选择的进行它们都被固定了。因为重组使标记和 QTL 之间的连锁不平衡消失,因此, MAS 策略的选择反应随世代数下降的速度较快。但是,只用标记进行的选择却取得了稳定的选择反应。与 BIUP 和 COMB 策略比较,仅用标记进行的选择有可能将来应用于肉质性状的选择,因为肉质性状可以不要求连续的表型值评定。Zhang 等考虑的是近交系杂交的理想情况。虽然 2 个近交系中感兴趣的性状没有差异,但它们在所有座位的等位基因均为不同的纯合子。利用品种间杂交来改良肉质一般说来有不同的方法,杂交群可以大量增加连锁不平衡。然而,随着 2 个品种大多数 QTL 的分离,会降低连锁不平衡。不过,即使用商品种猪的杂交,基因频率有足够大差异的品种杂交也可以检测到相当数量的 QTL^[3-4,15]。另外,发现有些有利的 QTL 来自于许多 QTL 的平均效应较低的品种^[4]。

2.4 利用品种内变异的遗传改良

大部分猪的选择方案都是将重点放在品种内或品系内的遗传改进,然后进行品系间的杂交。品种内选择需要品种内个体间存在差异,而不是像前面所讨论过的品种间的差异。本部分的目的是从常规选择入手,以品种内选择为基础,讨论肉质性状的遗传改良策略,分连锁平衡标记和连锁不平衡标记 2 个选择方法。

2.4.1 基于家系内 LD 的 LE-MAS 要应用家系内 QTL 和连锁标记间的连锁不平衡,需要获得标记

效应,或者至少要搞清每个家系的标记-QTL连锁相。这就需要知道家系内成员的标记基因型和表型资料。如果标记和QTL之间的连锁不紧密,表型记录就必须从候选个体的近亲获得,因为这种不紧密的基因型和表型关联在基因重组中会被削弱。借助后裔资料的标记-QTL效应或连锁相可以通过简单的统计测验计算出来,这种方法是对遗传了双亲等位基因后裔的平均表型值进行比较。或者对标记辅助选择的动物模型进行改造,将标记信息用于复杂家系的遗传评估^[19]。这种模型最终可获得结合了多基因EBV和QTL效应信息的BLUP育种值。由于选择是针对包含肉质性状在内的表型性状,必须将这些估计值,结合表型性状育种值组成经济选择指数。

连锁平衡-标记辅助选择方法需要进行大量的表型和基因型测定,这使该方法的经济可行性受到质疑。而且,在将标记辅助选择应用于准确估计QTL效应前,必须获得几个世代的数据资料。家系内连锁不平衡法估计QTL效应的另一个障碍是它需要了解在群体内分离的QTL区间。由于大部分关于猪QTL图谱的研究都是基于品种间杂交的模型,所以关于品种内QTL分离的信息是有限的。因此,在实施标记辅助选择之前必须开展品种内QTL图谱的研究。尽管这种研究可能集中在以前在品种间杂交中所鉴定出来的QTL区间上,仍需要足够大规模的群体进行检测和验证群体内分离现象。Evans等^[17]最近开展了这方面的研究,他们利用高度分化品种杂交组成的杂交群所定位的QTL区域确实证实商品群体中也出现分离。Spehn等^[18]围绕在奶牛育种方案中应用QTL信息进行了讨论。

2.4.2 根据群体范围内LD进行的ID-MAS 尽管在闭锁群体内功能基因以外的标记可能不会与一个QTL间存在大量的连锁不平衡,但因为遗传漂变、选择、突变和混群等因素,那些与QTL紧密连锁的标记极有可能与这个QTL存在相当部分的群体范围内的连锁不平衡^[19~21]。在家畜中,如Famir^[9]对奶牛研究所阐述的一样,这种可能性在一个有效群体含量较小的选择群体中比较高。连锁不平衡的程度能够通过使用连锁紧密的标记的单倍型来提高。高密度的标记图谱如每隔1~2 cM一个标记的图谱也将有可能包含与QTL紧密连锁的群体范围内连锁不平衡标记,正如Meuwissen等^[21]最近在模拟中所阐述的一样,他们发现在一个有效群体含量为100的群体中,标记间隔为1或2 cM、覆盖整个基因组足够的

连锁不平衡条件下,根据标记单倍型和至少500个个体表型值间存在的关联,对几个世代的育种值估计可能具有很高的准确性。伴随着新的高通量技术的应用,这种方法的成本将逐步降到一个可以接受的水平。另外,在一个有限的区域中,利用QTL区域内已鉴定的饱和标记密度,这种方法也可以使用。

对于与QTL在群体范围内存在连锁不平衡的标记来说,如果有多个连锁标记可以追踪QTL的话,选择可以直接采用标记的基因型或标记的单倍型信息。然而,有必要在选择条件下估计群体中标记的效应,了解群体中存在的连锁不平衡程度和连锁相,以及避免可能的QTL与背景基因组间的互作。由于同样的原因,利用常规方法重新估计QTL效应是非常必要的。估计QTL效应需要群体中随机样本的标记基因型和肉质性状的表型数据,并且在动物模型中应该以标记基因型和单倍型作为固定效应^[22~23]。

2.4.3 ID-MAS与LE-MAS的比较 分子遗传学在育种应用上的一个非常重要的问题就是能否实现连锁平衡-标记辅助选择、连锁不平衡-标记辅助选择或GAS的应用。最终要求是检测最少的连锁平衡标记和最大限度发现功能突变座位。然而,一旦发现了一个功能突变座位,就要求在其他群体中估计和验证该突变的效应是否比连锁平衡-标记低,因为后者需要测定世代群体的表型和基因型而不是一个随机样本。在连锁平衡-标记辅助选择中,对常规遗传评估过程的基因型数据的完整性的要求也比连锁不平衡-标记辅助选择和GAS高很多,而且对测定表型和基因型的个体以及分析方法要求都很高。连锁不平衡-标记辅助选择中对遗传评估的要求比GAS略高,因为连锁不平衡-标记辅助选择要求鉴别和分析标记的单倍型和验证标记-QTL连锁相。

然而对前面提到的一个给定的QTL,连锁平衡-标记辅助选择包括利用有限的标记对整个基因组进行分析和对QTL效应进行估计,这对采用高密度标记基因型分析的连锁不平衡-标记辅助选择也是可行的。Meuwissen等^[21]研究了利用贝叶斯混合模型分析,通过高密度基因分型技术构建的标记单倍型,他们获得了高度精确的MV值。

通过连锁平衡-标记辅助选择的方法在一个特定的QTL提高遗传改进速度的机率最小,因为可用于估计家系内QTL效应的信息是有限的,而连锁不

平衡-标记辅助选择和 GAS 的效应估计信息则是来自整个家系数据资料。连锁不平衡-标记辅助选择的准确性比 GAS 略低, 因为标记-QTL 间的不完全连锁不平衡和大量其他效应的存在(标记单倍型与 QTL 基因型)^[24], 因此对 GAS 进行知识产权保护和产品多样化的可能性是最大的, 而连锁平衡-标记辅助选择则是有限的。

3 育种计划中标记辅助选择的整合

很明显成功应用一个标记辅助选择方案需要一系列完善的策略和方法, 这些策略必须符合商业目标和市场要求。这种方法包括 DNA 的采集、表型记录、系谱记录、基因型分型、基因型和表型数据库的建立, 以及分析工具和质量控制系统。标记辅助选择的应用还要求慎重考虑经济和商业风险。标记辅助选择的经济分析需要一个广泛全面的方法, 它旨在评价标记辅助选择的经济可行性和应用的最优方案。一个成功的例子是 Hayes 等^[25]对一个综合性猪场核心育种群中的连锁平衡-标记辅助选择进行了深入的经济分析。在经济评价中考虑的因素有 QTL 检测、为多性状育种目标对 QTL 区域鉴定的标记辅助选择、基因分型的成本和猪场生产过程的其他收益。结果他们发现对于假定的成本和价格参数, 应用连锁平衡-标记辅助选择是可行的。他们还发现, 特别是当 QTL 检测建立在小样本基础上时, 在 QTL 检测过程中的阈值应严格。这样在应用过程中基因分型的成本就会降低, 同时选择假阳性个体的机率也会最小。

然而 Hayes 等^[25]是在生产水平上对借助标记辅助选择所获得的经济回报进行评估, 它只是额外遗传改进的一部分, 大多数商业育种措施获益都是通过增加育种群的市场份额或优良遗传资源。通常标记辅助选择的应用对市场份额的影响比所获得的遗传改进更大。

标记辅助选择成功与否还依赖于 QTL 效应在不同群体和不同环境中的一致性。Hospital 等^[26] 2002 年对植物中标记辅助导入研究发现, 简单性状主基因的效应是固定的, 但复杂性状(如产量)的 QTL 效应则不是固定的。这种不一致的效应目前在一些研

究深入的家畜基因中都存在。例如, ESR 基因在一些猪的产仔数研究中并没有发现有效应, 并且已证实存在基因与品系或环境间的互作^[7]。出现这些不一致的研究结果的原因有群体的统计误差, 如假阳性或不可靠结果(小样本量), 过高估计 QTL 的真实效应, 群体间 LD 标记-QTL 连锁相的不稳定性, 以及基因与环境的互作和上位效应等。这就说明需要对目标群体在一定的环境下, 不断进行基因或 QTL 效应估计和监测, 要求对表型资料的记录方面给予持续的高度重视。另外, 还应该研究在实施核心群育种计划条件下, 在商品生产群中如何估计基因效应的策略, 尤其是包括杂交利用的情况。这也为利用标记辅助选择获得非加性效应、以及设计交配模式提供了机遇。

假定标记效应是不确定的, 在不妨碍通过常规选择方法实现整体育种目标的条件下, 使用分子遗传信息似乎是明智的。这里有一个在选择群体中应用标记辅助选择的重要概念, 它在常规选择方法中不使用或较少使用^[27]。一个极好的例子是基于标记的早期选择, 即在获得个体本身或后代记录之前, 利用要进行测定的全同胞家系个体的标记信息进行早期选择。在这种情况下, 常规选择方法无法获得选择所需的资料, 因为 EBV 的计算需要系谱信息, 而该信息对全同胞家系中的所有个体都是相同的。然而, 家系内成员的遗传标记是不同的, 这就为选择提供了基础, 而不是采用随机选择的方法。

最近对基因和 QTL 图谱的相关研究已经发现 QTL 的表达方式可能不是孟德尔方式。特别在几项研究中已经发现在猪中一些基因和 QTL 属于根据不同配子来源而表现的印记遗传方式^[28-29]。今后的研究毫无疑问将鉴别出其他更多的影响 QTL 遗传和表达的表现遗传现象。在设计选择方案时就需要考虑这些效应。尽管它们一方面使选择方案变得复杂, 但它们也为选择提供了利用更多信息的机会。例如, De Koning^[30]就提出采用印记 QTL 和性别连锁 QTL 合并的方法, 有可能通过品种间一套有计划的杂交策略来达到为市场提供多元化产品的目的。

参考文献(见原文)

【译者 张 豪, 朱良瑞, 梅盈杰; 责任编辑 周志红】