

第12章

纯系品种选育与杂种优势利用

王建康

中国农业科学院作物科学研究所

wangjiankang@caas.cn

<http://www.isbreeding.net>

本章的主要内容

- § 12.1 自交过程中的选择与纯系品种选育
- § 12.2 近交衰退与杂种优势
- § 12.3 配合力与双列杂交设计
- § 12.4 轮回选择与群体改良

§ 12.1 自交过程中的选择与 纯系品种选育

- § 12.1.1 纯系品种选育的一般过程
- § 12.1.2 重组近交家系的群体均值与亲本选择
- § 12.1.3 自交过程中的选择

纯系品种的选育

- 纯系品种的育种过程（包括杂交品种育种过程中的自交系选育），其实就是优良的纯合重组基因型的选育过程。自交是一种极端形式的近交，本节内容对其他形式的近交也是适用的，只不过纯系的产生速度慢一些而已。因此，这一章提到的纯系，有的地方也称为自交系或近交系。
- 对于目标是选育杂交种的育种过程，优良重组自交系选育只是育种的第一步。在得到优良的自交系后，还要测试这些自交系与其它自交系杂交 F_1 的表现，以选育出表现优异的杂交品种。

纯系品种选育的一般过程

- 纯系品种选育一般从两个或多个优良纯系（称为亲本，parents or parental lines）间的杂交开始，但有时也从轮回选择群体中产生纯系。可采用的杂交类型有：
 - 单交（single cross），如亲本 $P_1 \times$ 亲本 P_2 ；
 - 有限回交（limited backcross），如亲本 P_1 和亲本 P_2 产生的 F_1 与其中的一个亲本回交1或2次；
 - 顶交又称三交（top cross or three-way cross），如亲本 P_1 和亲本 P_2 产生的 F_1 与第三个亲本 P_3 杂交，即 $(P_1 \times P_2) \times P_3$ ；
 - 双交（double cross），即两个 F_1 再杂交，如 $(P_1 \times P_2) \times (P_3 \times P_4)$ 等。
- 以单交和有限回交在植物育种中的应用最为普遍。

CIMMYT 面包小麦育种流程图

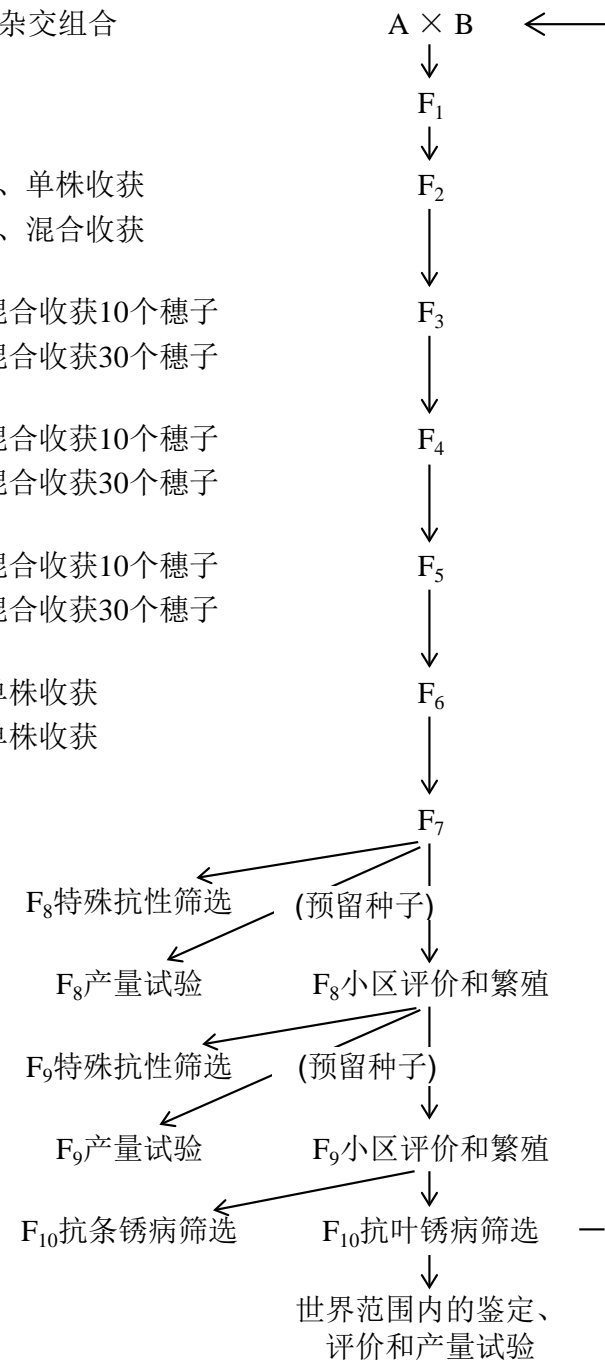
育种试验地点

育种工作的简单描述

世代名称和流程

包含修饰系谱 (MODPED) 和选择混合 (SELBLK) 两种选择方法

育种试验地点	育种工作的简单描述
Toluca	从大约200个亲本材料中配置大约1000个杂交组合
Cd. Obregon	每个中选的F ₁ 混合收获
Toluca	修饰系谱: 每个F ₂ 群体选择30~80个单株、单株收获 选择混合: 每个F ₂ 群体选择30~80个单株、混合收获
Cd. Obregon	修饰系谱: 每个F ₃ 家系选择10个单株、混合收获10个穗子 选择混合: 每个F ₃ 家系选择30个单株、混合收获30个穗子
Toluca	修饰系谱: 每个F ₄ 家系选择10个单株、混合收获10个穗子 选择混合: 每个F ₄ 家系选择30个单株、混合收获30个穗子
Cd. Obregon	修饰系谱: 每个F ₅ 家系选择10个单株、混合收获10个穗子 选择混合: 每个F ₅ 家系选择30个单株、混合收获30个穗子
Toluca	修饰系谱: 每个F ₆ 家系选择10个单株、单株收获 选择混合: 每个F ₆ 家系选择40个单株、单株收获
Cd. Obregon	小区混合收获
Toluca and El Batan	小区混合收获
Cd. Obregon	小区混合收获
Toluca and El Batan	小区混合收获
Cd. Obregon	小区混合收获
Toluca and El Batan	小区混合收获



选育纯系的常用方法和穿梭育种

- 常用的选育纯系方法包括系谱法（pedigree method）、混和法（bulk population selection）、一粒传（single seed descend）以及修饰系谱和混和方法（modified pedigree and bulk method）等。
- 为加速分离世代的纯合过程，在分离世代早期，如 F_1 、 F_2 和 F_3 等，也可采用加倍单倍体技术。
- 除了常规的表型选择外，对于遗传力高且难以鉴定的性状，在不同的育种阶段还可考虑利用分子标记辅助选择（marker assisted selection）的方法。
- CIMMYT面包小麦选育的流程图中，育种群体在墨西哥的Toluca和Cd. Obregon两个地点交替种植和选择，CIMMYT称这一选择过程为穿梭育种（Shuttle Breeding）。

纯系选育过程中的决策问题

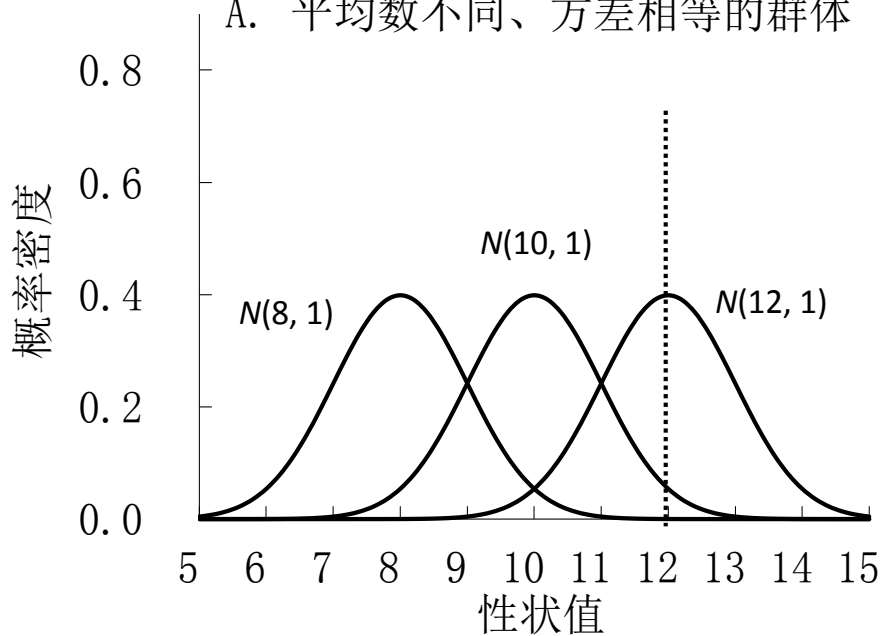
- 纯系品种的选育过程一般都包含多个自交世代。每个世代中，育种家都要做出这样或那样的抉择。例如，在一个育种计划刚开始时，育种家需要根据所占有的资源决定育种规模，需要确定利用什么样的亲本进行杂交、以及配置多少个杂交组合；对于杂种 F_1 世代，需要决定是立即自交，还是先回交然后再自交，或者是与第三个亲本顶交后再自交；对于每个分离世代，需要考虑根据哪些性状对后代群体进行选择，采用多大的选择强度；随着材料的纯合，需要决定在什么样的环境、多少个地点开展产量试验等等。
- 此外，育种家还需考虑如何把不同的优良性状结合在一起，如何评价育种材料的广适应性以及如何采用新的育种技术等等，其最终目的都是为了尽快得到满足一定育种目标的基因型。

理想的育种群体

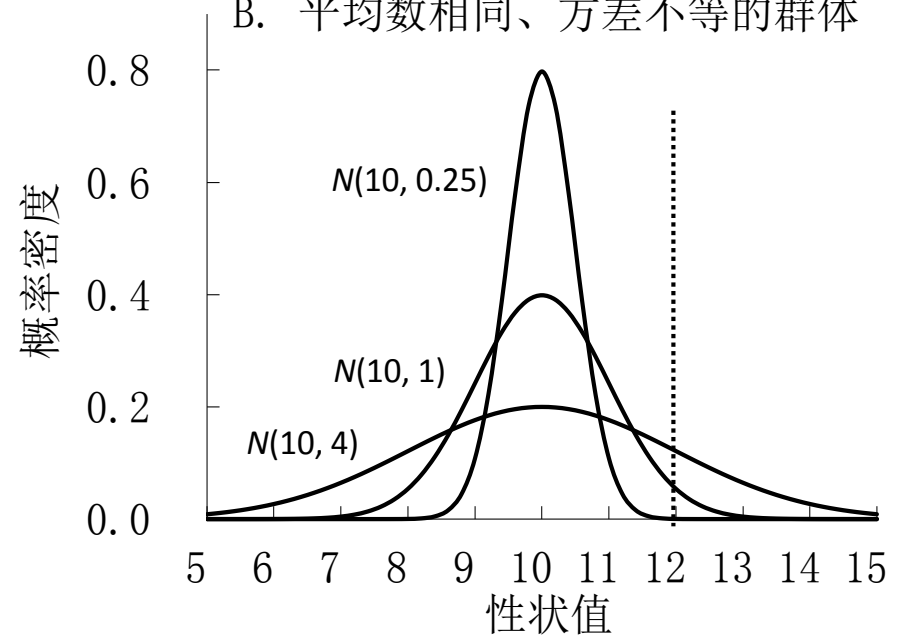
- 从流程图12.1可以看出，纯系品种的选育就是通过杂交创造遗传分离群体，通过重复自交使基因型不断纯合，纯合过程中通过选择不断地提高有利基因以及基因组合的频率，使育种群体逐渐接近或达到目标基因型。
- 创造适宜的育种群体是品种选育的基础。那么，什么样的群体才是理想的、易于获得目标基因型的育种群体呢？简单地讲，理想的育种群体需要满足以下两个条件：（1）具有较高的平均数（假定高值是育种目标）和较大的遗传方差；（2）选择可以在育种群体中鉴定出遗传上优良的个体或后代，即选择是有效的。

不同均值和方差的育种群体

A. 平均数不同、方差相等的群体



B. 平均数相同、方差不等的群体



- 大多数纯系品种的选育过程，都与流程图12.1类似，涉及亲本选配、后代选择和多环境产量比较试验等过程，有的育种家也形象地称为“配、选、比”。实际工作中，如何创制理想的育种群体呢？

重组近交家系的群体均值

- 考虑一个座位上的两个等位基因 A_1 和 A_2 ，频率分别为 p 和 q 。在近交系数为 F 的群体中，三种基因型 A_1A_1 、 A_1A_2 和 A_2A_2 的频率分别为 p^2+pqF 、 $2pq(1-F)$ 和 q^2+pqF ，基因型值分别用 $m+a$ 、 $m+d$ 和 $m-a$ 表示，群体平均数如下：

$$\mu_F = m + (p - q)a + 2pq(1 - F)d$$

$$\mu_{\text{RIL}} = m + (p - q)a$$

F₁衍生近交家系群体均值与亲本的关系

- 假定亲本A的基因型为A₁A₁，亲本B的基因型为A₂A₂。在亲本A中，等位基因A₁和A₂的频率分别用p_A=1和q_A=0表示。在亲本B中，两个等位基因的频率分别用p_B=0和q_B=1表示。因此，

$$\mu_A = m + (p_A - q_A)a \quad \mu_B = m + (p_B - q_B)a$$

$$\mu_{\text{RIL}(A \times B)} = m + \left[\frac{1}{2}(p_A + p_B) - \frac{1}{2}(q_A + q_B) \right] a = \frac{1}{2} \mu_A + \frac{1}{2} \mu_B$$

BC₁衍生近交家系群体均值与亲本的关系

- 假定A × B的F₁与亲本A回交，在不存在选择的情况下，由回交世代衍生的重组近交家系与回交群体有相同的基因频率，因此重组近交家系的平均数为：

$$\mu_{\text{RIL}(A \times B) \times A} = m + \left[\left(\frac{3}{4} p_A + \frac{1}{4} p_B \right) - \left(\frac{3}{4} q_A + \frac{1}{4} q_B \right) \right] a = \frac{3}{4} \mu_A + \frac{1}{4} \mu_B$$

- 推广到独立遗传的多基因，公式12.3和12.4也都是成立的。这表明亲本的基因型值越高，其后代的平均数也越高。育种家经常选择两个表现优良的亲本进行杂交（即优 × 优组合），其目的就是为了得到高平均数的育种群体。同时，如果杂交组合中的一个亲本表现较差，从公式12.4可以看出，与优良亲本的回交可以提高优良亲本在后代群体均值中的权重，进而提高后代均值。

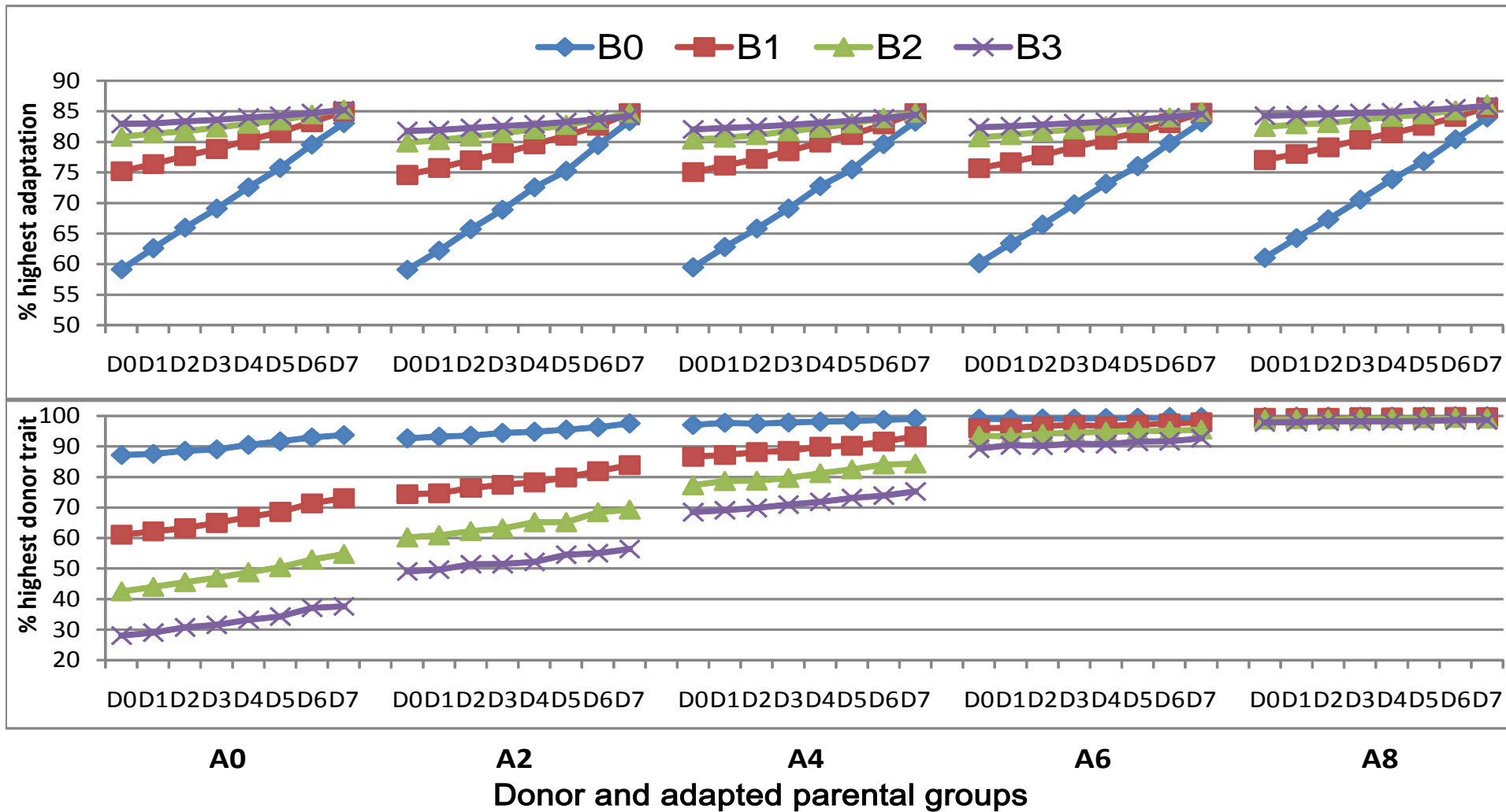
高遗传方差的重要性

- 理想育种群体不仅均值要高、而且其中所包含的加性遗传方差还要大。因此，在选择优良亲本进行杂交、保证后代群体具有高均值的同时，还要考虑亲本之间在遗传上的差异和互补性。
- 只有这样，后代中才有可能出现同时优于两个亲本的基因型，即出现超亲分离；然后再通过选择的手段，将具有超亲分离的基因型从众多后代基因型中鉴定出来。
- 从这个意义上讲，育种也就是不断地进行基因间的重组、鉴定并选择综合性状优于双亲的超亲分离后代基因型的过程。

适宜的回交次数

- 与优良亲本的回交可以提高重组近交家系的群体平均数，直到优良亲本达到完全恢复时为止，但那时候的后代中已不再具有任何遗传变异，选择也不再有效。因此，有必要确定究竟要回交多少代才算合适。
- 回交育种是转育基因的有效方法。随着育种工作的开展，供体亲本的适应性也在不断提高。除轮回亲本中表现欠佳性状的有利基因外，供体亲本有时也会携带一些影响产量和适应性的有利基因。这时，回交代数越多，供体亲本中有利基因丢失的可能性就越大。

回交育种的模拟比较研究



自交过程中的选择

- 纯系选育过程中，不同自交世代的遗传方差可分解为自交家系间方差和家系内方差两部分。家系间选择的有效性依赖于家系间方差中的加性成分，家系内选择的有效性依赖于家系内方差中的加性成分。

$$F_n = 1 - \left(\frac{1}{2}\right)^{n-2} \quad V_{F_n} = (1 + F_n)V_A + (1 + F_n)(1 - F_n)V_D$$

$$V_{bF_{n-1:n}} = (1 + F_{n-1}) \left[V_A + \frac{(1 - F_n)^2}{1 - F_{n-1}} V_D \right]$$

$$V_{wF_{n-1:n}} = (1 - F_n)(V_A + V_D)$$

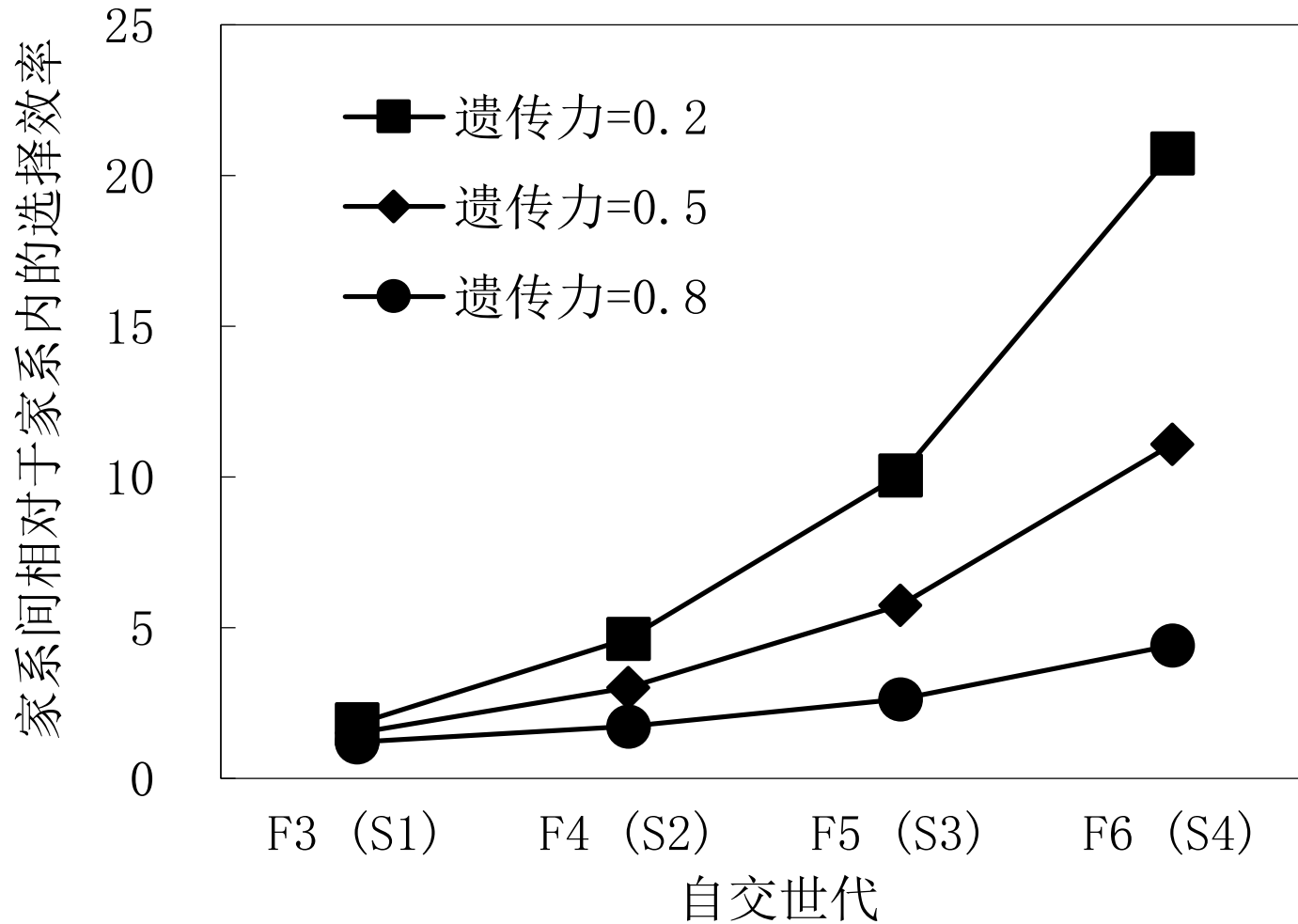
家系间、家系内的遗传力和相对选择效率

- 假定不同世代的家系间和家系内，除加性方差外均具有相同的非加性和环境方差，用 V_E 表示。

$$h_{F_2}^2 = \frac{V_A}{V_A + V_E} \quad h_{bF_{n-1:n}}^2 = \frac{2F_n V_A}{2F_n V_A + V_E}$$

$$h_{wF_{n-1:n}}^2 = \frac{(1 - F_n)V_A}{(1 - F_n)V_A + V_E} \quad RE = \frac{h_{bF_{n-1:n}}^2}{h_{wF_{n-1:n}}^2}$$

不同自交世代和遗传力下家系间选择相对于家系内选择的效率



§ 12.2 近交衰退与杂种优势

- § 12.2.1 杂种优势与作物育种
- § 12.2.2 近交对群体均值和遗传方差的影响
- § 12.2.3 杂交群体的均值和遗传方差
- § 12.2.4 杂种优势的遗传基础

近交衰退与杂种优势

- 近交会不断降低亚群体内的遗传多样性，形成亚群体之间的分化。对以随机交配为主的物种来说，近交往往会引起个体繁殖力和生活力的下降，这一生物学现象称为近交衰退（inbreeding depression）。
- 生物学家（包括遗传学家和育种家）在观察到近交衰退这一现象的同时，还观察到两个近交系或两个种之间的杂交往往表现出一定的优势，近交引起的繁殖力和生活力下降，往往会在近交系之间的杂种中得到恢复，这就是杂种优势（hybrid vigor or heterosis）。

近交衰退与杂种优势的关系

- 近交衰退和杂种优势其实是一个问题的两个方面，近交衰退程度高的物种，其杂种优势一般也比较高；杂种优势程度高的物种，其近交衰退的程度一般也比较大。
- 由于长期选择的作用，在以自交繁殖方式为主的物种中，近交衰退和杂种优势的程度要远低于随机交配物种。
- 基因型接近纯合的近交系，即使发生进一步的近交，群体结构也不会发生大的变化，因此也就不再能够观察到近交衰退。但两个纯系之间的杂交仍会出现杂种优势。

杂种优势与杂交种选育

- 由于杂种优势的普遍存在，通过自交或近交选育近交系，然后通过配制近交系之间的杂交组合来利用杂种优势，就成为一种重要的育种方法，在玉米、高粱、水稻、油菜、棉花等重要作物的育种中已经取得巨大成就，这些作物的杂交品种已经在农业生产中发挥了巨大作用。
- 杂种优势是杂种超越亲本的一种遗传现象，一个杂种可以是来自两个群体间的杂交，但更普遍地是指两个或多个自交系间的杂交，一个杂交种可以是单交（single cross）、三交或顶交（three-way cross）、以及双交（double cross）等。

杂种优势的不同含义

- 目前对杂种优势的含义有三种不同的看法。（1）凡是 F_1 基因型值高于亲本中亲值，均应视为杂种优势，这种杂种优势概念与显性效应有类似的含义。（2）杂种优势和显性是同一生理和遗传现象，只是表现的程度有所不同，只有杂种表现比高值亲本更高，或比低值亲本更低，才是杂种优势。因此杂种优势和超显性有相同的含义。（3）认为应对 F_1 进行连续自交， F_1 的表现比最高的重组近交系还大，或比最低的重组近交系还小时，才有杂种优势。
- 对任何一种作物来说，如果杂种优势充分大，生产杂交种经济可行，那么就可以考虑选育杂交种作为品种。

不同作物的产量性状超中亲杂种优势

$$H = (F_1 - MP) / MP \times 100\%$$

作物	超中亲杂种优势H		
	平均	最小	最大
异花授粉			
玉米（美国）	121	92	240
玉米（欧洲）	129	112	143
黑麦	192	86	329
常异花授粉			
Faba豆	60	22	95
春油菜	30	20	50
冬油菜	50	20	80
自花授粉			
水稻	46	3	106
大豆	13	7	19
小麦	9	-14	106

近交衰退或杂种优势的参照群体

- 近交衰退或杂种优势都是一个相对的概念，需要一个参照物才能对它们进行度量。下一节将看到，杂种优势可以用两个亲本的中亲值作参照。近交衰退的参照群体是一个无限大的随机交配群体。
- 现假定该参照群体通过任意一种近交交配系统（包括有限大小群体的随机交配），最终分化成众多亚群体。单个亚群体经历长期的近交后，近交系数趋近于1，每个座位都会被固定在一个等位基因上。如不考虑选择，则等位基因在所有亚群体中的频率仍然等于原始参照群体的基因频率。这时，近交衰退可以从亚群体相对于参照群体在均值上的差异来表示。

近交群体的均值和遗传方差计算

基因型	所有亚群体的频率	基因型值	频率×基因型值
A_1A_1	$p^2 + pqF$	a	$p^2a + pqaF$
A_1A_2	$2pq(1 - F)$	d	$2pqd - 2pqdF$
A_2A_2	$q^2 + pqF$	$-a$	$-q^2a - pqaF$
群体均值	$\mu_F = (p - q)a + 2pqd - 2pqdF$		
遗传方差	$V_F = 2pqa^2(1 + F) + 2pq(1 - 2pq)(1 - F)d^2 + 4pq(q - p)ad(1 - F)$		

单座位加显性遗传模型下，随机交配群体、近交群体的均值和遗传方差

$$\mu_0 = (p - q)a + 2pqd$$

$$V_0 = 2pqa^2 + 2pq(1 - 2pq)d^2 + 4pq(q - p)ad$$

$$\mu_F = (p - q)a + 2pqd - 2pqdF$$

$$V_F = 2pqa^2(1 + F) + 2pq(1 - 2pq)(1 - F)d^2 + 4pq(q - p)ad(1 - F)$$

近交群体均值和遗传方差的变化量

$$\Delta\mu = \mu_F - \mu_0 = -2pqdF \quad (\text{单基因模型})$$

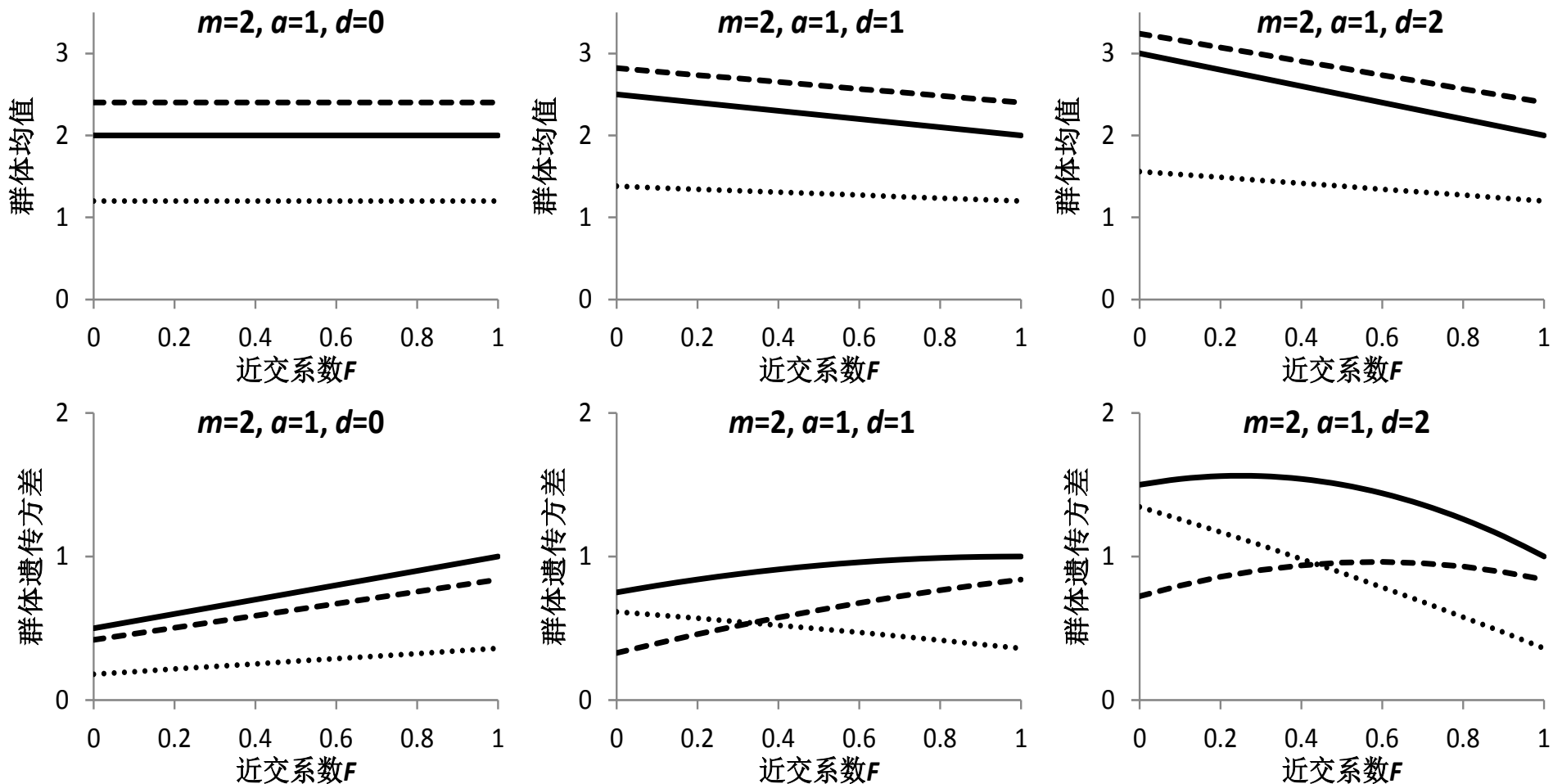
$$= -2(\Sigma pqd)F \quad (\text{多基因模型})$$

$$\Delta V = V_F - V_0 = 2pqa^2F - 2pq(1 - 2pq)d^2F - 4pq(q - p)adF \quad (\text{单基因模型})$$

近交群体均值和遗传方差的变化量

- 可以看出，如不考虑基因频率的变化，近交衰退只与显性效应有关，显性效应越大，近交衰退越严重。也就是说，只要存在显性效应，即杂合基因型的表现不等于两种纯合基因型的平均数，就会出现近交衰退，衰退的方向为隐性纯合基因型的表现。
- 同时还可以看出，近交衰退是近交系数的线性函数。这一线性关系在玉米株高和产量等性状上都可以观察到，说明了多基因加显性模型在数量遗传研究中的有效性。
- 与均值的变化相比，方差的变化要复杂得多，既与加显性效应有关，同时又与它们的乘积项有关。近交后的方差不一定总是变小。

单基因模型下群体均值（上）和群体方差（下）随近交系数的变化



家系内的加性方差遗传力

- 为简单起见，下面只考虑加性效应对遗传方差的贡献。如用 p_0 和 q_0 分别表示参照群体中等位基因 A_1 和 A_2 的频率，在不考虑显性效应时，参照群体的遗传方差全部来自加性效应。

$$V_G = V_A = 2p_0q_0a^2$$

$$V_{wF} = 2E(pq)a^2 = 2p_0q_0(1-F)a^2 = (1-F)V_A$$

$$h_{wF}^2 = \frac{(1-F)V_A}{(1-F)V_A + V_E} = \frac{(1-F)V_A}{V_A + V_E - FV_A} = \frac{h^2(1-F)}{1-h^2F}$$

近交家系间的加性遗传方差和总方差

$$V_{bF} = V[(p - q)a] = a^2V(1 - 2q) = 4a^2V(q) = 4a^2 \times p_0q_0F = 2FV_A$$

$$V_F = (1 + F)V_A$$

杂种优势的度量

- 参照群体经过一段时间的近交分化后，如果打破亚群体之间的隔离，一代随机交配就能使群体重新回到HW平衡。因此，在无选择和无连锁不平衡的情况下，亚群体分化形成的近交家系之间的广泛互交，得到的群体与分化前的参照群体是等价的，群体均值也立即恢复到分化前的均值，杂交群体在均值上的增加量，正好弥补了近交群体在均值上的减少量。从遗传学理论上讲，近交衰退和杂种优势是一个问题的正反两个方面。
- 在把近交衰退和杂种优势看作是同一个遗传学问题时，杂种优势就可以通过近交家系之间、大量杂交形成群体的均值与近交家系群体均值之间的差异进行度量。在实际应用中，育种家关心的往往是两个特定群体甚至是两个特定近交系之间的杂种优势，这种理论上的度量方法在实际育种应用中具有其局限性。

亲本群体的均值和中亲值

- 下面介绍另外一种基于中亲值的杂种优势度量方法。仍从单座位加显性遗传模型开始，用 p 和 q 分别表示亲本群体 P_1 中等位基因 A_1 和 A_2 的频率，亲本群体 P_2 的等位基因频率用 $p+x$ 和 $q-x$ 表示。两个亲本群体均值和中亲值分别为：

$$\mu_{P_1} = (p - q)a + 2pqd$$

$$\mu_{P_2} = [(p + x) - (q - x)]a + 2(p + x)(q - x)d$$

$$= \mu_{P_1} + 2xa - 2(q - p)xd + x^2d$$

$$\mu_{\bar{P}} = \frac{1}{2} \mu_{P_1} + \frac{1}{2} \mu_{P_2} = \mu_{P_1} + xa - (p - q)xd + x^2d$$

两个亲本群体的配子基因型频率及杂交群体的均值和遗传方差计算

亲本群体 P ₁ 的配子		亲本群体 P ₂ 的配子		杂交 F ₁ 后代群体	
配子型	频率	配子型	频率	后代基因型	频率
A ₁	p	A ₁	$p + x$	A ₁ A ₁	$p^2 + px$
A ₂	q	A ₂	$q - x$	A ₁ A ₂	$2pq - (p - q)x$
				A ₂ A ₂	$q^2 - qx$
杂交 F ₁ 后代均值				$\mu_{F_1} = [(p - q) + x]a + [2pq - (p - q)x]d$	

杂交F₁群体的均值和中亲优势

$$\mu_{F_1} = \mu_{P_1} + xa - (p - q)xd$$

$$H_{F_1} = \mu_{F_1} - \mu_{\bar{P}} = x^2d \quad (\text{单基因模型})$$

$$= \Sigma(x^2d) \quad (\text{多基因模型})$$

杂交 F_1 群体的均值和中亲优势

- 近交衰退一样，中亲优势也只依赖于每个座位上基因的显性效应。加性基因不会引起近交衰退，也不会引起杂种优势。中亲优势依赖于两个亲本群体基因频率差异的平方。基因频率无差异的两个群体，杂交也就不会产生杂种优势，基因频率在一个群体中为0、另一个群体中为1时，这时的两个亲本都是自交系，中亲杂种优势达到最高。
- 推广到多基因时，如果不同座位的基因显性效应有不同的方向，这时的中亲优势会相互抵消。因此，看不到中亲优势的杂交 F_1 中，不一定代表无显性效应存在。当所有等位基因频率在两个亲本中只有0和1两种取值时，中亲优势等于显性效应的代数。

杂交F₂群体的均值和中亲优势

$$\begin{aligned}\mu_{F_2} &= [(p + \frac{1}{2}x) - (q - \frac{1}{2}x)]a + 2(p + \frac{1}{2}x)(q - \frac{1}{2}x)d \\ &= (p - q)a + 2pqd + xa - (px - qx - \frac{1}{2}x^2)d\end{aligned}$$

$$H_{F_2} = \mu_{F_2} - \mu_{\bar{P}} = \frac{1}{2}x^2d = \frac{1}{2}H_{F_1}$$

- 从上面的公式看到，杂种F₂的中亲优势只有杂种F₁的一半。在一个遗传研究群体中，如果没有观察到这种关系，意味着其他遗传因素的存在，如连锁和上位性互作等。

杂种优势的其他度量方法

- 为方便起见，在杂种优势的一些理论研究中，多采用中亲优势。
- 而在实际应用中，人们还采用超中亲的百分数来表示杂种优势，如表12.2。
- 育种家有时关心的可能还有超高亲优势，或相对于某对照品种的优势。因此，杂种优势有时也常常用超高亲或超对照品种的百分数表示。

杂种优势与亲本遗传差异的关系

- 公式12.28和公式12.30均表明，两个亲本之间基因频率的差异越大、差异的座位越多，亲本间的遗传差异越大，杂种优势也就越高。
- 但是，很多研究都表明，这一结论仅在一定范围内成立。超过一定的范围，随着遗传差异的增加，杂种优势还会下降，甚至出现杂种不育的现象。

玉米杂交F₁的平均单株产量 (lb) 与亲本自交系遗传差异的关系

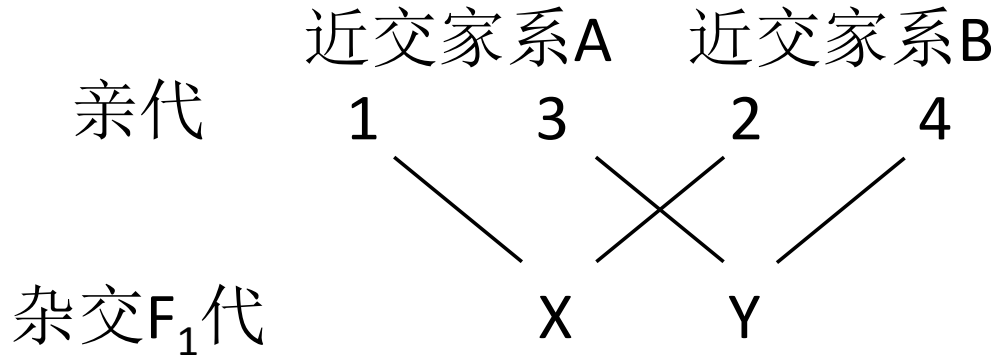
杂交亲本	F ₁ 产量	F ₂ 产量	F ₁ 中亲优势	F ₁ 高亲优势
群体内随机交配	0.352			
同一个地区的两个自交系	0.392	0.359	0.040	0.023
美国东南×中西群体	0.488	0.456	0.076	0.063
美国东南×波多黎各群体	0.520	0.454	0.134	0.102
美国中西×波多黎各群体	0.520	0.456	0.140	0.114
美国东南×墨西哥群体	0.405	0.372	0.081	-0.013
美国中西×墨西哥群体	0.380	0.355	0.062	-0.026
波多黎各×墨西哥群体	0.323	0.313	0.031	-0.031

遗传方差在杂交F₁间和杂交F₁内的分解

- 如不考虑选择，近交家系产生雌雄配子的频率与参照群体完全一样，把大量近交家系间的杂交F₁联合起来看的话，其实等同于参照群体。这也是§ 10.3.4中，把双亲DH或RIL家系之间的互交F₁看作是双亲F₂群体的原因。
- 因此，杂交F₁群体的总遗传方差与参照群体完全相等，考虑上位性互作的总遗传方差为：

$$V_G = V_A + V_D + V_{AA} + V_{AD} + V_{DD} + \dots$$

杂交F₁间的协方差



$$f_{12} = f_{14} = f_{23} = f_{34} = 0 \quad f_{13} = f_{24} = F$$

$$r = 2f_{XY} = \frac{1}{2}(f_{13} + f_{14} + f_{23} + f_{24}) = F$$

$$u = f_{13}f_{24} + f_{14}f_{23} = F^2$$

杂交F₁内的方差

$$V_{bc} = \text{Cov}(X, Y) = FV_A + F^2V_D + F^2V_{AA} + F^3V_{AD} + F^4V_{DD} + \dots$$

$$V_{wC} = (1 - F)V_A + (1 - F^2)V_D + \\ (1 - F^2)V_{AA} + (1 - F^3)V_{AD} + (1 - F^4)V_{DD} + \dots$$

- 杂交F₁间遗传方差的构成中，只有加性方差是参照群体的F倍，其它方差成分的系数都是近交系数的平方或高次方。因此，当近交系数较低时，杂交F₁间的差异主要由加性效应所决定，其它遗传效应的贡献较小。
- 当近交系数接近1时，杂交F₁间的差异由所有遗传效应决定，杂交F₁内几乎没有遗传变异。同时还可以看到，当近交系数为0.5时，杂交F₁间的遗传方差等同于参照群体的全同胞家系。

杂种优势的遗传基础

- 关于杂种优势的遗传基础，遗传学上有两种截然不同的假说，即显性假说（dominance hypothesis）和超显性假说（overdominance hypothesis）。
- 但两者有一个共同点，均认为杂种优势是不同遗传结构的亲本杂交后，彼此的遗传物质（基因）相互作用的结果。

杂种优势遗传基础的显性假说

- 显性假说认为显性基因的聚合和等位基因间的互作是造成杂种优势的原因，杂种优势是由于杂合子中隐性不利基因被掩盖而引起的，杂种优势是多对基因共同作用的结果。
- 假定三种基因型 A_1A_1 、 A_1A_2 和 A_2A_2 的遗传值分别为 a 、 d 和 $-a$ ，由单个座位引起的杂种优势等于杂合基因型的值减去2个纯合基因型的平均数，即中亲杂种优势等于显性效应 d 。因此，对单个座位来说，只有在显性效应存在时才会出现杂种优势。显性假说认为，单个座位上杂合型的表现不会超过最好的亲本。但如果考虑多个座位，由于有利基因在亲本中经常呈分散分布， F_1 的表现就会超过最好的亲本。

两个基因座位下显性产生的杂种优势

- 如假定两个自交系的基因型分别为 $A_1A_1B_2B_2$ 和 $A_2A_2B_1B_1$ ，两个座位上有相同的加性效应 a 和显性效应 d ，，不考虑中亲值 m 的话，两个自交系的基因型值为0，而杂种 F_1 的基因型值为 $2d$ ，超过两个亲本的基因型值。

$$\begin{array}{ccc} \text{自交系1: } A_1A_1B_2B_2 & \times & \text{自交系2: } A_2A_2B_1B_1 \\ \text{基因型值} = a + (-a) = 0 & & \text{基因型值} = (-a) + a = 0 \\ & \downarrow & \\ \text{F}_1\text{杂种: } A_1A_2B_1B_2 & & \\ \text{基因型值} = 2d & & \end{array}$$

显性假说的两种反对意见

- 第一，如果显性假说成立，实践中应该能够选育到与杂交种表现同样优良的自交系。在图12.6中，如果是完全显性（即 $d=a$ ），基因型 $A_1A_1B_1B_1$ 应该和杂种 $F_1: A_1A_2B_1B_2$ 有同样的表现。但截至目前，至少在玉米上还没有选育到与优良杂交种表现相当的自交系。
- 第二，在显性效应存在时， F_2 群体的分布应该是偏态的。而实际中，对于具有杂种优势的许多性状来说，它在 F_2 群体中表现为对称的正态分布。但是，如果考虑到控制性状的基因很多，它们之间还存在连锁这一事实，则对上面的质疑会做出较好的解释。由于存在多基因间的连锁，很难选择到所有座位上都是优良基因的重组基因型；多个基因存在时， F_2 群体的分布偏度是很微小的。

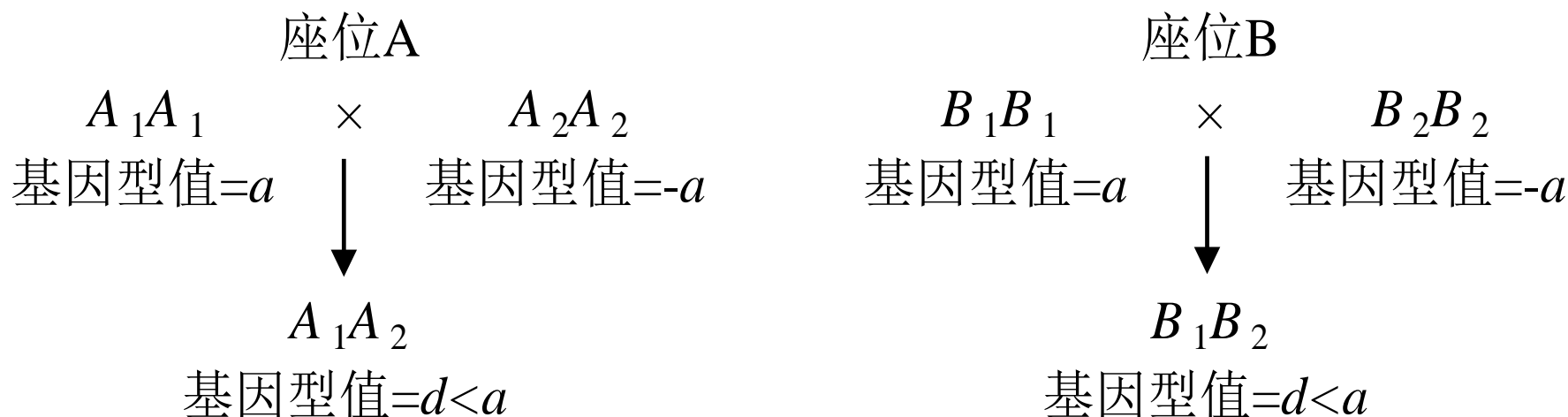
杂种优势遗传基础的超显性假说

- 超显性假说认为等位基因内的相互作用是造成杂种优势的原因，杂种优势是由于杂合基因型优于两个纯合基因型而引起的。假定三种基因型 A_1A_1 、 A_1A_2 和 A_2A_2 的遗传值为 a 、 d 和 $-a$ ，杂种优势是由于 $d > a$ 引起。对超显性来说，选育到与杂交种一样优良的自交系是不可能的。但基因间的紧密连锁会引起假超显性（pseudo-overdominance）。

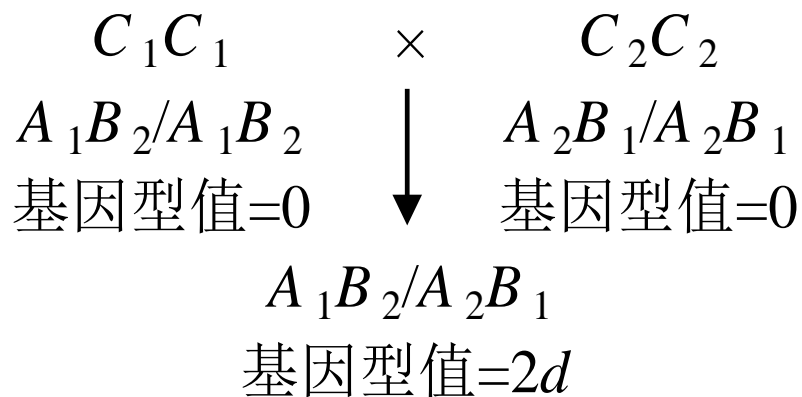
紧密连锁引起的假超显性

- 假定两个座位A和B完全连锁，等位基因 A_1 和 B_2 连锁在一起，用 C_1 表示配子基因型 A_1B_2 ；等位基因 A_2 和 B_1 连锁在一起，用 C_2 表示配子基因型 A_2B_1 。当两个座位有相同的加性效应 a 和显性效应 d 时， C_1C_1 和 C_2C_2 的基因型值均为0，但杂种 F_1 (C_1C_2) 基因值为 $2d$ ，超过两个亲本的基因型值。因此，如果我们不知道两个基因连锁在一起影响同一个性状的话，就会错误地得出超显性的结论。

互斥连锁引起的假超显性



假超显性



无显性的上位性互作

均值 $m=4$ 、两个座位的加性效应分别等于1和2、它们之间的加性 \times 加性上位性互作效应等于2、其它效应均等于0时

座位A	座位B			行平均
	B_1B_1	B_1B_2	B_2B_2	
A_1A_1	9	5	1	5
A_1A_2	6	4	2	4
A_2A_2	3	3	3	3
列平均	6	4	2	4

上位性效应对杂种优势的作用

- 不同座位之间的上位性效应也会产生杂种优势。考虑包含加性×加性上位性互作效应的9种基因型值（表12.6），对座位A来说， A_1A_2 的基因型值等于 A_1A_1 和 A_2A_2 基因型值的平均，说明在座位A上不存在显性；对位点B一样， B_1B_2 的基因型值等于 B_1B_1 和 B_2B_2 基因型值的平均，说明在座位B上也不存在显性。
- 现在假定有自交系 $A_1A_1B_2B_2$ 和 $A_2A_2B_1B_1$ ，它们的基因型值分别为1和3，但杂种 F_1 ($A_1A_2B_1B_2$)的基因型值为4，不难得到杂种 F_2 的平均基因型值也等于2。因此，在不存在显性只存在上位性效应时，也能引起杂种优势。
- 但在这一例子中，我们可以找到比杂种表现更好的纯合自交系 $A_1A_1B_1B_1$ ，在这种情况下，杂交种也许并不是必须的。

杂种优势的分子遗传学基础

- 过去的经验数据大多支持杂种优势是由部分显性或完全显性造成的，上位性和超显性的作用较小。但最近借助分子数据的研究又似乎表明，各种遗传效应都在杂种优势中起作用。
- Tang et al. (1995) 利用一个水稻回交重组近交家系群体，对株高、穗长、抽穗期、成熟期及一些产量构成性状的QTL做图研究表明，显性效应是杂种优势的主要遗传基础。
- Luo et al. (2001) 利用另一个水稻重组近交家系群体，对产量构成性状QTL做图的研究却表明，超显性效应是近交衰退和杂种优势的主要成因。

杂种优势的分子遗传学基础

- Zhou et al. (2012) 利用水稻优良杂交种Shanyou 63的永久F₂群体，结合超高密度连锁图谱和基因型数据，对杂种优势的遗传基础进行了探索。结果表明，杂种优势的程度、不同遗传成分对杂种优势的贡献，在性状间有很大差异。产量的杂种优势最强，接下去依次是每穗粒数、粒重和单株分蘖数。对于产量、每穗粒数和粒重的杂种优势，超显性的贡献最大。显性与显性上位对于单株分蘖数和粒重较为重要，对其它两个性状也有作用。显性效应对所有性状的贡献都是相对较小的。

杂种优势的分子遗传学基础

- 对玉米的研究同样表明，不存在一种遗传模型能够解释不同群体、不同性状上表现出的杂种优势。
- Tang et al. (2010) 利用玉米优良杂交种‘豫玉22’的永久F₂群体，结合覆盖全基因组261个SSR标记的基因型数据，研究了玉米产量及产量构成性状杂种优势的遗传基础。在产量、穗长、穗行数和百粒4四个性状中，共定位到13个与杂种优势有关的座位，上位性检测得到的加加上位互作QTL最多。结果表明，显性效应和加加互作是决定‘豫玉22’杂种优势的主要因素。
- Larièpe et al. (2012) 利用多个重组近交家系群体，开展NCIII遗传交配设计，对玉米重要农艺性状进行QTL定位研究，检测出的产量QTL大多具有明显的超显性。

杂种优势的利用

- 育种和农业生产中，人们最关心的当然还是产量性状的杂种优势。产量也许是终极的、复杂程度最高的一个数量性状，它的形成涉及到众多生理过程和发育阶段，无疑受大量基因的影响。因此，期望用一种遗传模型来解释所有与产量和杂种优势有关的基因座位，是不切合实际的。
- 与任何遗传研究一样，研究结果依赖于采用的遗传材料和群体类型。利用不同物种、不同材料、不同群体，得到的结论不尽一致，也是正常的。要想完全说明杂种优势的遗传基础，需要把单个基因在染色体上的位置和功能、以及不同基因之间的相互关系了解清楚。
- 但是，从育种实践来看，对杂种优势的遗传基础缺乏完整的了解，并没有妨碍杂交种的选育和在生产上的推广利用。

§ 12.3 配合力与双列杂交设计

- § 12.3.1 一般配合力和特殊配合力
- § 12.3.2 双列杂交遗传交配设计
- § 12.3.3 不完全双列杂交设计
- § 12.3.4 包含正反交（有或无自交）的完全双列杂交
- § 12.3.5 包含正交（有或无自交）的完全双列杂交

近交系的一般配合力

- 利用参照群体衍生的大量近交系为亲本，两两随机杂交构成一个新的群体。育种家在很多时候关心一个特定的近交系与其它近交系杂交，是否能够产生大量的优良杂交组合？两个特定的近交系结合在一起是否能够产生出一个优良的杂交组合？
- 对一个特定的近交系，如果利用其它近交系与其杂交，这些杂交组合的平均表现与所有组合形成群体的均值之间的差异，称为该近交系的一般配合力（general combining ability, GCA）。

两个近交系的特殊配合力

- 如用 X 表示近交系 $P_1 \times$ 近交系 P_2 杂交 F_1 的表现， X 表示所有组合的群体均值，离差就可以按照公式12.34进行分解。其中，前两项为两个亲本的一般配合力，最后一项称为两个亲本 P_1 和 P_2 的特殊配合力（specific combining ability, SCA）。

$$X - \bar{X} = GCA_{P_1} + GCA_{P_2} + SCA$$

遗传方差在配合力上的分解

- 从效应分解公式12.34就能得到方差分解公式12.35。如果区分杂交组合的亲本，组合之间的方差就能被分解为父本一般配合力方差、母本一般配合力方差和特殊配合力方差三部分。如不区分杂交组合的亲本，组合之间的方差被分解为一般配合力方差和特殊配合力方差两部分。

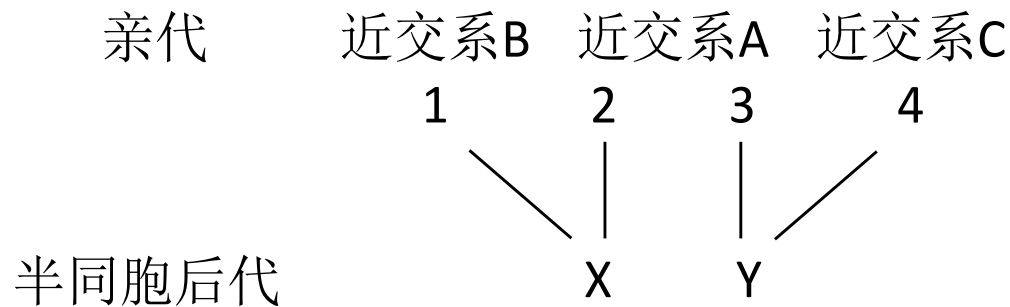
$$V_X = V_{GCA_M} + V_{GCA_F} + V_{SCA} = V_{GCA} + V_{SCA} = V_{bc}$$

一般配合力的方差

- 从定义可以看出，一般配合力方差相当于以一个近交系作为父本、其它近交系作为母本，杂交产生半同胞家系间的协方差。

一般配合力的方差

- 右图表示近交系A与另外两个近交系B和C的半同胞后代X和Y的系谱。亲代个体之间的共祖先系数、以及后代X和Y之间的共祖先系数计算如下。



$$f_{23} = F$$

$$f_{12} = f_{13} = f_{14} = f_{24} = f_{34} = 0$$

$$r = 2f_{XY} = \frac{1}{2}(f_{13} + f_{14} + f_{23} + f_{24}) = \frac{1}{2}F$$

$$u = f_{13}f_{24} + f_{14}f_{23} = 0$$

一般配合力的方差

- 根据公式8.38，得到父本一般配合力方差的表达式12.36。在不存在母本效应和性连锁效应的假定下，雌雄亲本具有相同的一般配合力方差，即公式12.37。从公式12.32给出的组合间方差减去父母本的一般配合力方差，就得到特殊配合力方差公式12.38。

$$V_{GCA_F} = V_{GCA_M} = \frac{1}{2} FV_A + \frac{1}{4} F^2V_{AA} + \dots$$

$$V_{SCA} = F^2V_D + \frac{1}{2} F^2V_{AA} + F^3V_{AD} + F^4V_{DD} + \dots$$

- 从中可以看出，一般配合力之间的差异，是由加性效应和加加互作效应引起的；特殊配合力之间的差异，是由非加性效应引起的。

双列杂交遗传交配设计

- 开展配合力分析的常用遗传交配设计是双列杂交（diallel cross design or diallel mating design）。
- 根据交配设计的亲本是否相同、是否包含反交、是否包含自交等因素，双列杂交又可以分成很多类型。根据具体的遗传研究问题，配合力效应有时可以被看作是固定模型，有时又需要被看作随机模型。不同效应模型的分析方法也不完全相同。因此，双列杂交设计及其分析方法包含着丰富的内容。

完全双列杂交及其类型

- 在完全双列杂交（complete diallel design）中，配制杂交组合的亲本是同一组近交系。设有 p 个亲本近交系，在它们之间进行所有的杂交和自交，共得到 p^2 个组合，其中包含 $p(p-1)/2$ 个正交组合， $p(p-1)/2$ 个反交组合和 p 个自交组合。
- 如果可以忽略细胞质效应，或不考虑正反交效应，这时可以只配制 $p(p-1)/2$ 个正交组合，加上所有的自交组合，共有 $p(p+1)/2$ 个。
- 如不关心近交系自身的表现，同时细胞质效应又可以忽略不计，这时可以只配制 $p(p-1)/2$ 个正交组合。
- 如不关心近交系自身的表现，但还想研究正反交之间的差异，这时可以只配制 $p(p-1)/2$ 个正交组合和 $p(p-1)/2$ 个反交组合，共 $p(p-1)$ 个。

不完全双列杂交

- 在有些物种，如玉米的杂交育种中，存在明显的杂种优势群，优异杂交品种往往是不同优势群两个自交系的杂交组合。这时的配合力分析，一般基于两个优势群之间的杂交，而不是同一个优势群内自交系之间的杂交。在开展双列杂交设计时，通常选择一个优势群的 m 个自交系为父本，另一个优势群的 n 个自交系为母本，总的亲本数 $p=m+n$ ，在两组亲本间配置 $m \times n$ 个组合，这样的设计称为不完全双列杂交设计（incomplete diallel design）。

不完全双列杂交与NCII设计的异同

- 不完全双列杂交看上去与第10章的NCII设计十分类似，都可利用双因素带互作的方差分析模型进行分析。其实，它们之间还是有一定区别的，这里的区别主要考虑的是参照群体和研究目的上的差异。
- NCII设计的参照群体通常是一个随机交配群体，父本是参照群体的一组随机样本，母本是参照群体的另一组随机样本，配制杂交的亲本基因型一般都是高度杂合的。NCII设计的研究目的是通过全同胞和半同胞家系的方差分解，对参照群体的遗传方差进行分解，估计参照群体的遗传力。

不完全双列杂交与NCII设计的异同

- 不完全双列杂交的参照群体是产生亲本自交系的祖先群体，父本自交系和母本自交系可以具有相同的祖先群体，也可以具有不同的祖先群体，配制杂交组合的亲本基因型一般都是高度纯合的。
- 有时也把亲本自交系的大量随机杂交组合形成的群体，视为参照群体。当亲本来自不同的参照群体时，它们杂交 F_1 的参照群体其实是两个亲本参照群体互交之后的群体。
- 除对参照群体的遗传方差进行分解、估计参照群体的遗传力外，双列杂交往往还关心一般配合力和特殊配合力的估计、以及杂交种表现的预测等实际育种问题。

包含9个亲本的双列杂交设计

父 本	母 本								
	1	2	3	4	5	6	7	8	9
1	11	12	13	14	15	16	17	18	19
2	21	22	23	24	25	26	27	28	29
3	31	32	33	34	35	36	37	38	39
4	41	42	43	44	45	46	47	48	49
5	51	52	53	54	55	56	57	58	59
6	61	62	63	64	65	66	67	68	69
7	71	72	73	74	75	76	77	78	79
8	81	82	83	84	85	86	87	88	89
9	91	92	93	94	95	96	97	98	99

包含6个亲本、36个组合、两次重复 的完全双列杂交

父本	重复I						重复II					
	A	B	C	D	E	F	A	B	C	D	E	F
A	24.7	32.6	40.4	24.7	21.8	26.8	30.2	27.1	42.9	33.2	26.3	18.6
B	33.7	45.8	29.3	42.0	30.5	23.0	38.2	41.7	34.7	34.8	31.3	21.6
C	41.0	37.8	44.6	35.9	40.5	24.2	42.5	40.9	50.0	36.8	36.8	24.7
D	31.5	37.7	32.2	30.4	31.0	18.4	34.0	33.5	44.6	33.9	33.4	23.4
E	32.7	26.1	26.3	23.4	18.0	10.8	28.9	29.4	21.0	20.5	18.1	10.7
F	20.4	22.1	21.8	22.3	15.0	9.5	17.3	25.6	17.5	13.2	16.4	14.5

不完全双列杂交设计的统计分析

- 设有 m 个父本、 n 个母本，配制 $m \times n$ 个所有可能的杂交组合，每个组合有 r 个观测值。用 $i=1, 2, \dots, m$ 表示父本， $j=1, 2, \dots, n$ 表示母本， $k=1, 2, \dots, r$ 表示重复，下面的公式给出每个观测值 X_{ijk} 的线性分解。

$$X_{ijk} = \bar{X} + \text{GCA}_{M_i} + \text{GCA}_{F_j} + \text{SCA}_{ij} + \varepsilon_{ijk}$$

- 显然，与第9章介绍的基因型和环境互作模型、第10章介绍的NCII方差分析线性模型，除效应参数的表示符号有差异外，完全相同的。

利用一般配合力预测杂交组合

- 在得到亲本一般配合力的估计之后，就可以利用下面的公式预测每个杂交组合的表现。

$$\hat{X}_{ij} = \bar{X} + \text{GCA}_{M_i} + \text{GCA}_{F_j}$$

$i=1, 2, \dots, m$ 表示父本， $j=1, 2, \dots, n$ 表示母本

不完全双列杂交方差分析表

变异来源	自由度	均方	固定模型期望均方	随机模型期望均方
父本间	$m-1$	MS_{GCA_M}	$V_\varepsilon + mrV_{GCA_M}$	$V_\varepsilon + rV_{SCA} + nrV_{GCA_M}$
母本间	$n-1$	MS_{GCA_F}	$V_\varepsilon + mrV_{GCA_F}$	$V_\varepsilon + rV_{SCA} + mrV_{GCA_F}$
父本×母本	$(m-1)(n-1)$	MS_{SCA}	$V_\varepsilon + rV_{SCA}$	$V_\varepsilon + rV_{SCA}$
随机误差	$mn(r-1)$	MS_ε	V_ε	V_ε
总和	$mnr-1$			

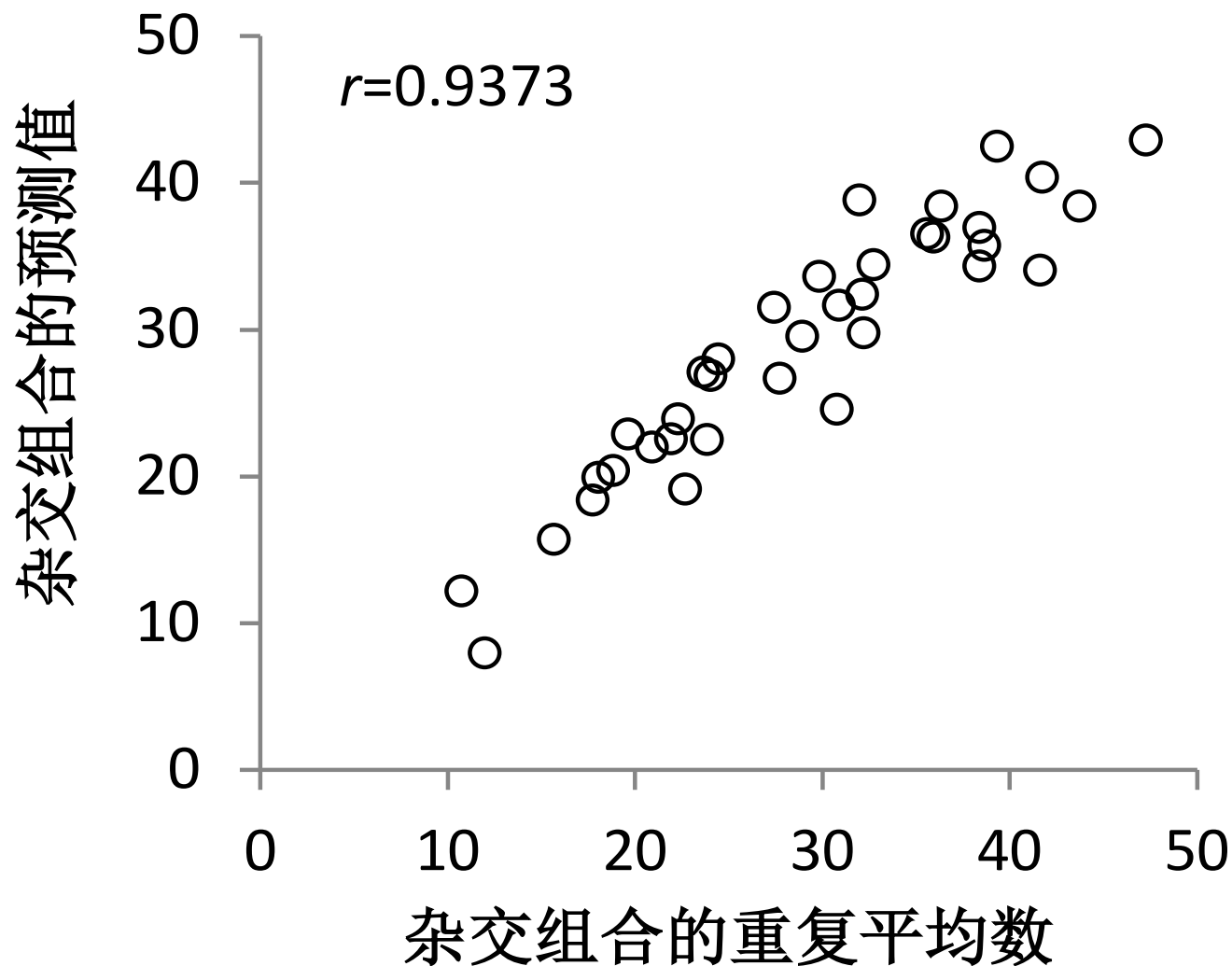
利用表12.9重复平均数得到的 亲本GCA和SCA估计值

SCA	1	2	3	4	5	6	父本GCA
A	-4.07	-3.77	7.62	-0.57	-2.80	3.59	0.26
B	-0.34	5.36	-6.81	4.11	-0.73	-1.58	5.03
C	1.37	-3.13	4.40	-2.03	2.93	-3.53	9.13
D	-1.66	-0.91	1.47	-0.26	2.46	-1.10	3.15
E	6.23	1.08	-3.43	-0.62	-1.85	-1.41	-6.69
F	-1.53	1.37	-3.24	-0.63	-0.01	4.03	-10.88
母本GCA	2.41	4.51	4.93	0.41	-2.26	-10.00	

利用表12.8数据的不完全双列杂交方差分析

变异来源	自由度	平方和	均方	F值 (固定模型)	P值	固定模型方差估计	随机模型方差估计
父本配合力	5	3381.78	676.36	58.28	<0.0001	55.40	53.94
母本配合力	5	1867.77	373.55	32.19	<0.0001	30.16	28.71
特殊配合力	25	726.32	29.05	2.50	0.0058	8.72	8.72
随机误差	36	417.78	11.60			11.60	11.60
总和	71	6393.64	90.05				

利用一般配合力的杂交组合预测值



包含正反交（有或无自交）的完全双列杂交

- 在包含正反交的完全双列杂交中，如果要区分自交系作为父本和母本的一般配合力，这时的分析方法与不完全双列杂交完全一样。
- 如果不区分父本和母本的一般配合力，也就是说，每个自交系不管作为父本还是母本，只有一个一般配合力。这时的分析方法就不能套用前面的不完全双列杂交。
- 设有 p 个亲本，在它们之间配制所有可能的杂交（即有自交），共 p^2 个组合，每个组合有 r 个观察值。用 $i=1, 2, \dots, p$ 表示父本， $j=1, 2, \dots, p$ 表示母本， $k=1, 2, \dots, r$ 表示重复，每个观测值仍用 X_{ijk} 表示。如无自交，相当于要求 $i \neq j$ ，每个组合仍有 r 个观察值。

包含正反交（有或无自交）的完全双列杂交

- 有、无自交的两种设计中，亲本 i 的一般配合力用 GCA_i 表示，满足总和为0的约束条件。亲本 i 和 j 的特殊配合力用 SCA_{ij} 表示（无自交时要求 $i \neq j$ ），对亲本 i ，满足第 i 行和 i 列总和均为0的约束条件。
- 随机效应用 ε_{ijk} 表示，如有自交，则要求对任意 i 和 j ， r 个误差效应之和等于0，即（ $i, j=1, 2, \dots, p$ ）；如无自交，则要求对任意 $i \neq j$ ， r 个误差效应之和等于0。
- 这里特别提醒一下，特殊配合力满足的约束条件与公式12.39是有本质区别的。

包含正反交完全双列杂交的线性模型

$$X_{ijk} = \bar{X} + \text{GCA}_i + \text{GCA}_j + \text{SCA}_{ij} + \varepsilon_{ijk} \quad (\text{有自交})$$

$i, j=1, 2, \dots, p$ 表示亲本, $k=1, 2, \dots, r$ 表示重复;

$$X_{ijk} = \bar{X} + \text{GCA}_i + \text{GCA}_j + \text{SCA}_{ij} + \varepsilon_{ijk} \quad (\text{无自交})$$

$i \neq j$

- 为便于看出有、无自交的差异, 公式前半部分给出有自交设计的观测值线性分解, 后半部分给出无自交设计的观测值线性分解。
- 下文的效应计算公式与此类似, 先给出有自交的情况, 然后给出无自交的情况。如两者的 i 或 j 取值范围完全相同, 无自交计算公式则略去 i 或 j 的取值范围; 如有差异, 无自交计算公式后面只给出这种差异。

重复平均和总平均的计算

$$\bar{X}_{ij\cdot} = \frac{1}{r} \sum_k X_{ijk} \quad (\text{有自交}), \quad i, j=1, 2, \dots, p;$$

$$\bar{X}_{ij\cdot} = \frac{1}{r} \sum_k X_{ijk} \quad (\text{无自交}), \quad i \neq j$$

$$\bar{X} = \frac{1}{p^2} \sum_{i,j} \bar{X}_{ij\cdot} = \frac{1}{p^2 r} \sum_{i,j,r} X_{ijk} \quad (\text{有自交})$$

$$\bar{X} = \frac{1}{p(p-1)} \sum_{i \neq j} \bar{X}_{ij\cdot} = \frac{1}{p(p-1)r} \sum_{i \neq j, k} X_{ijk} \quad (\text{无自交})$$

行平均和列平均的计算

$$\bar{X}_{i..} = \frac{1}{p} \sum_j \bar{X}_{ij}. \quad (\text{有自交}), \quad i=1, 2, \dots, p;$$

$$\bar{X}_{i..} = \frac{1}{p-1} \sum_{j \neq i} \bar{X}_{ij}. \quad (\text{无自交})$$

$$\bar{X}_{.j.} = \frac{1}{p} \sum_i \bar{X}_{ij}. \quad (\text{有自交}), \quad i=1, 2, \dots, p;$$

$$\bar{X}_{.j.} = \frac{1}{p-1} \sum_{i \neq j} \bar{X}_{ij}. \quad (\text{无自交})$$

样本均值与配合力的关系

$$\bar{X}_{ij.} = \bar{X} + \text{GCA}_i + \text{GCA}_j + \text{SCA}_{ij}$$

(有自交), i 固定, $j=1, 2, \dots, p$;

$$\bar{X}_{ij.} = \bar{X} + \text{GCA}_i + \text{GCA}_j + \text{SCA}_{ij}$$

(无自交), $j \neq i$;

$$\bar{X}_{.ji} = \bar{X} + \text{GCA}_j + \text{GCA}_i + \text{SCA}_{ji}$$

(有自交), i 固定, $j=1, 2, \dots, p$;

$$\bar{X}_{.ji} = \bar{X} + \text{GCA}_j + \text{GCA}_i + \text{SCA}_{ji}$$

(无自交), $i \neq j$;

样本均值与配合力的关系

$$p(\bar{X}_{i..} + \bar{X}_{.i.}) = 2p\bar{X} + 2p\text{GCA}_i$$

(有自交) , $i=1, 2, \dots, p$;

$$(p-1)(\bar{X}_{i..} + \bar{X}_{.i.}) = 2(p-1)\bar{X} + 2(p-2)\text{GCA}_i$$

(无自交)

样本均值与GCA关系及GCA计算

$$p(\bar{X}_{i..} + \bar{X}_{.i.}) = 2p\bar{X} + 2p\text{GCA}_i$$

(有自交), $i=1, 2, \dots, p$;

$$(p-1)(\bar{X}_{i..} + \bar{X}_{.i.}) = 2(p-1)\bar{X} + 2(p-2)\text{GCA}_i$$

(无自交)

$$\text{GCA}_i = \frac{1}{2}(\bar{X}_{i..} + \bar{X}_{.i.}) - \bar{X} \quad (\text{有自交}), \quad i=1, 2, \dots, p;$$

$$\text{GCA}_i = \frac{p-1}{p-2} \left[\frac{1}{2}(\bar{X}_{i..} + \bar{X}_{.i.}) - \bar{X} \right] \quad (\text{无自交})$$

SCA和剩余效应计算

$$SCA_{ij} = X_{ij\cdot} - \bar{X} - GCA_i - GCA_j$$

(有自交), $i, j=1, 2, \dots, p$;

$$SCA_{ij} = X_{ij\cdot} - \bar{X} - GCA_i - GCA_j$$

(无自交), $j \neq i$;

$$\varepsilon_{ijk} = X_{ijk} - \bar{X}_{ij\cdot}$$

(有自交), $i, j=1, 2, \dots, p, k=1, 2, \dots, r$;

$$\varepsilon_{ijk} = X_{ijk} - \bar{X}_{ij\cdot}$$

(无自交), $j \neq i$;

有自交的方差分析自由度与平方和

- 一般配合力共有 p 个，自由度为 $p-1$ 。共配置 p^2 个杂交组合，组合的自由度为 p^2-1 ，从中减去一般配合力自由度，就是特殊配合力自由度 $p(p-1)$ 。总自由度为 p^2r-1 ，从中减去组合的自由度 p^2-1 ，就是误差自由度，即 $p^2(r-1)$ 。
- 亲本一般配合力的平方和乘以 pr 、再乘以2之后，就是一般配合力平方和。之所以要再乘以2，是因为每个杂交组合中都包含两个一般配合力。所有 p^2 个特殊配合力的平方和乘以 r 之后，就是特殊配合力平方和。根据剩余效应计算误差平方和，最后就得到方差分析表。

包含正交、反交和自交的完全双列杂交方差分析表

亲本数为 p ，配制包括正交、反交和自交所有可能的 p^2 个组合，每个组合有 r 次重复观察值

变异来源	自由度	均方	固定模型期望均方	随机模型期望均方
GCA	$p-1$	MS_{GCA}	$V_{\varepsilon} + 2prV_{GCA}$	$V_{\varepsilon} + rV_{SCA} + 2prV_{GCA}$
SCA	$p(p-1)$	MS_{SCA}	$V_{\varepsilon} + rV_{SCA}$	$V_{\varepsilon} + rV_{SCA}$
随机误差	$p^2(r-1)$	MS_{ε}	V_{ε}	V_{ε}
总和	p^2r-1			

无自交的方差分析自由度与平方和

- 一般配合力共有 p 个，自由度为 $p-1$ 。共配置 $p(p-1)$ 个杂交组合，组合的自由度为 $p(p-1)-1$ ，从中减去一般配合力自由度，就是特殊配合力自由度 $p(p-2)$ 。总自由度为 $p(p-1)r-1$ ，从中减去组合的自由度 $p(p-1)-1$ ，就是误差自由度，即 $p(p-1)(r-1)$ 。
- 亲本一般配合力的平方和乘以 $(p-1)r$ 、再乘以2之后，就是一般配合力平方和。所有 $p(p-1)$ 个特殊配合力的平方和乘以 r 之后，就是特殊配合力平方和。根据剩余效应计算误差平方和，最后就得到方差分析表。

包含正、反交的完全双列杂交方差分析表

亲本数为 p ，配制包括正、反交的所有 $p(p-1)$ 个组合，
每个组合有 r 次重复观察值

变异来源	自由度	均方	固定模型期望均方	随机模型期望均方
GCA	$p-1$	MS_{GCA_M}	$V_\varepsilon + 2(p-1)rV_{GCA_M}$	$V_\varepsilon + rV_{SCA} + 2(p-1)rV_{GCA_M}$
SCA	$p(p-2)$	MS_{SCA}	$V_\varepsilon + rV_{SCA}$	$V_\varepsilon + rV_{SCA}$
随机误差	$p(p-1)(r-1)$	MS_ε	V_ε	V_ε
总和	$p(p-1)r-1$			

利用表12.9中36个组合重复平均数得到的GCA和SCA估计值

亲本	SCA						GCA
	A	B	C	D	E	F	
A	-4.07	-5.11	4.44	-3.01	-1.66	2.96	1.33
B	0.99	5.36	-8.65	3.00	1.75	-0.88	4.77
C	4.54	-1.30	4.40	-1.31	7.25	-0.98	7.03
D	0.79	0.20	0.74	-0.26	6.04	0.71	1.78
E	5.09	-1.40	-7.75	-4.21	-1.85	-3.18	-4.47
F	-0.89	0.67	-5.78	-2.44	1.77	4.03	-10.44

利用表12.9中36个组合、两次重复观测值的方差分析

变异来源	自由度	平方和	均方	F值	P值	固定模型方差估计	随机模型方差估计
一般配合力	5	4946.61	989.32	85.25	<0.0001	40.74	39.79
特殊配合力	30	1029.26	34.31	2.96	0.0011	11.35	11.35
随机误差	36	417.78	11.60			11.60	11.60
总和	71	6393.64	90.05				

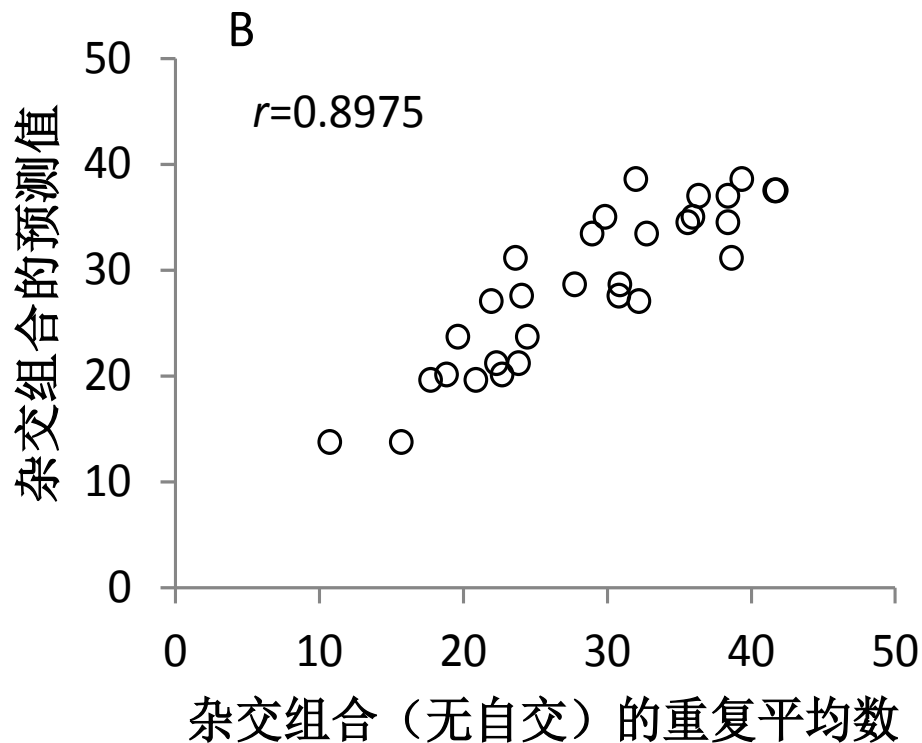
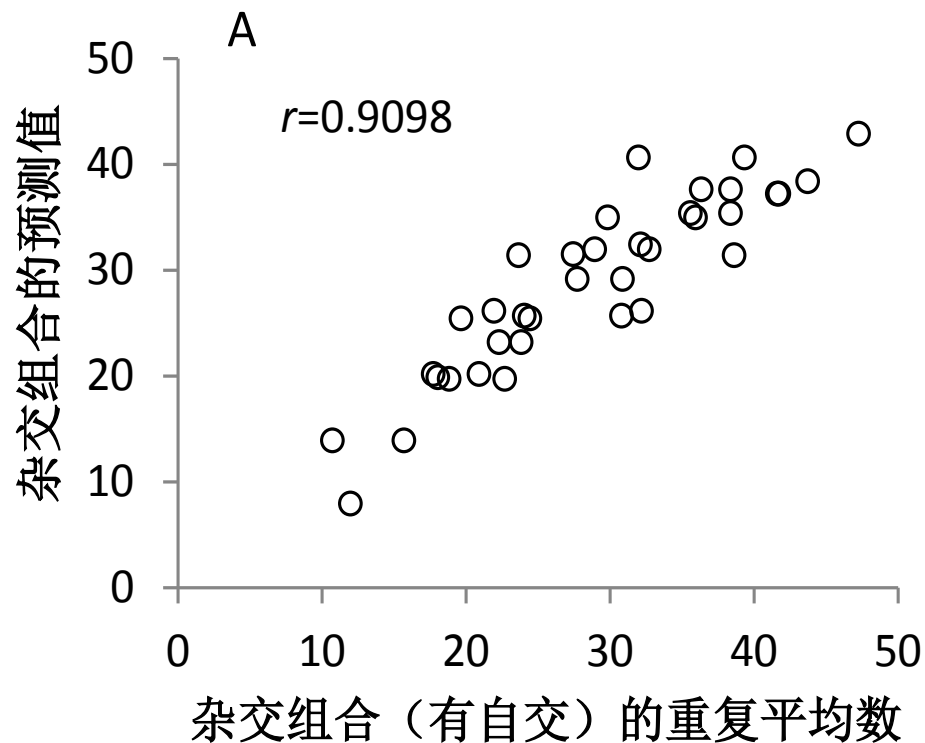
利用表12.9中15个正交、15个反交重复平均数得到的GCA和SCA估计值

亲本	SCA						GCA
	A	B	C	D	E	F	
A		-5.16	4.14	-4.48	-3.52	2.57	2.67
B	0.94		-6.59	3.89	2.25	1.09	3.75
C	4.24	0.76		-0.65	7.51	0.74	6.24
D	-0.68	1.09	1.40		5.14	1.27	2.16
E	3.23	-0.90	-7.50	-5.11		-3.02	-3.69
F	-1.28	2.64	-4.06	-1.88	1.93		-11.13

利用表12.9中15个正交、15个反交两次重复平均数得到的方差分析结果

变异来源	自由度	平方和	均方	F值	P值	固定模型方差估计	随机模型方差估计
一般配合力	5	4047.88	809.58	67.27	<0.0001	39.88	38.85
特殊配合力	24	782.01	32.58	2.71	0.0034	10.27	10.27
随机误差	30	361.04	12.03			12.03	12.03
总和	59	5190.92					

利用一般配合力的杂交组合预测值



包含正交（有或无自交）的完全双列杂交

- 设有 p 个亲本，在它们之间配制 $p(p-1)/2$ 个正交和 p 个自交组合，共 $p(p+1)/2$ 个，每个组合有 r 个观察值。用 $i=1, 2, \dots, p$ 表示父本、 $j=i, i+1, \dots, p$ 表示母本、 $k=1, 2, \dots, r$ 表示重复。亲本 i 的一般配合力用 GCA_i 表示， $i=1, 2, \dots, p$ ，满足总和为0的约束条件。亲本 i 和 j 的特殊配合力用 SCA_{ij} 表示， $j=i, i+1, \dots, p$ 。第 i 行和第 i 列的特殊配合力满足和为0的约束条件，需要注意的是特殊配合力 SCA_{ii} 在行列求和时占了两份。
- 如无自交，相当于 $j=i+1, \dots, p$ ，每个组合仍有 r 个观察值。一般配合力满足总和为0的约束条件，第 i 行和第 i 列的特殊配合力满足和为0的约束条件。
- 随机效应用 ε_{ijk} 表示，对于任意有意义的 i 和 j ， r 个误差效应之和均等于0。

包含正交完全双列杂交的线性模型

$$X_{ijk} = \bar{X} + \text{GCA}_i + \text{GCA}_j + \text{SCA}_{ij} + \varepsilon_{ijk} \quad (\text{有自交})$$

$i=1, 2, \dots, p$ 和 $j=i, i+1, \dots, p$ 表示亲本,
 $k=1, 2, \dots, r$ 表示重复;

$$X_{ijk} = \bar{X} + \text{GCA}_i + \text{GCA}_j + \text{SCA}_{ij} + \varepsilon_{ijk} \quad (\text{无自交})$$

$j=i+1, \dots, p$

重复平均和亲本平均的计算

$$\bar{X}_{ij\cdot} = \frac{1}{r} \sum_k X_{ijk} \quad (\text{有自交}), \quad i=1, 2, \dots, p, \quad j=i, i+1, \dots, p$$

$$\bar{X}_{ij\cdot} = \frac{1}{r} \sum_k X_{ijk} \quad (\text{无自交}), \quad j=i+1, \dots, p;$$

$$\bar{X}_i = \frac{1}{p+1} \left(\sum_{j<i} \bar{X}_{ji\cdot} + 2\bar{X}_{ii\cdot} + \sum_{i<j} \bar{X}_{ij\cdot} \right)$$

(有自交), $i=1, 2, \dots, p$)

$$\bar{X}_i = \frac{1}{p-1} \left(\sum_{j<i} \bar{X}_{ji\cdot} + \sum_{i<j} \bar{X}_{ij\cdot} \right) \quad (\text{无自交})$$

总平均、各种平均与配合力的关系

$$\bar{X} = \frac{1}{p} \sum_i \bar{X}_i = \frac{1}{\frac{1}{2} p(p+1)r} \sum_{i \leq j, k} X_{ijk} \quad (\text{有自交})$$

$$\bar{X} = \frac{1}{p} \sum_i \bar{X}_i = \frac{1}{\frac{1}{2} p(p-1)r} \sum_{i < j, k} X_{ijk} \quad (\text{无自交})$$

$$(p+1)\bar{X}_i = (p+1)\bar{X} + (p+2)\text{GCA}_i \\ (\text{有自交}), \quad i=1, 2, \dots, p;$$

$$(p-1)\bar{X}_i = (p-1)\bar{X} + (p-2)\text{GCA}_i \\ (\text{无自交})$$

GCA的计算

$$\text{GCA}_i = \frac{p+1}{p+2} (\bar{X}_i - \bar{X}) \quad (\text{有自交}), \quad i=1, 2, \dots, p;$$

$$\text{GCA}_i = \frac{p-1}{p-2} (\bar{X}_i - \bar{X}) \quad (\text{无自交})$$

SCA和剩余效应计算

$$SCA_{ij} = \bar{X}_{ij\cdot} - \bar{X} - GCA_i - GCA_j$$

(有自交) , $i=1, 2, \dots, p, j=i, i+1, \dots, p$

$$SCA_{ij} = \bar{X}_{ij\cdot} - \bar{X} - GCA_i - GCA_j$$

(无自交) , $j=i+1, \dots, p$

$$\varepsilon_{ijk} = X_{ijk} - \bar{X}_{ij\cdot} \text{ (有自交)}$$

$i=1, 2, \dots, p, j=i, i+1, \dots, p, k=1, 2, \dots, r;$

$$\varepsilon_{ijk} = X_{ijk} - \bar{X}_{ij\cdot} \text{ (无自交) , } j=i+1, \dots, p$$

有自交的方差分析自由度与平方和

- 一般配合力有 p 个，自由度为 $p-1$ 。包含自交，共有 $p(p+1)/2$ 个组合，组合的自由度为 $p(p+1)/2-1$ ，从中减去一般配合力自由度，就是特殊配合力自由度，即 $p(p-1)/2$ ，正好等于正交组合的个数。总自由度为 $p(p+1)r/2-1$ ，从中减去组合的自由度，就是误差自由度，即 $p(p+1)(r-1)/2$ 。
- 亲本一般配合力的平方和乘以、再乘以2之后，就是一般配合力平方和。特殊配合力的平方和乘以重复数 r 之后，就是特殊配合力平方和。根据剩余效应估计值计算误差平方和。

包含正交和自交的完全双列杂交方差分析表

亲本数为 p ，配制 $p(p-1)/2$ 个正交和 p 个自交，共
 $p(p+1)/2$ 个，每个组合有 r 次重复观察值

变异来源	自由度	均方	固定模型期望均方	随机模型期望均方
GCA	$p-1$	MS_{GCA}	$V_{\varepsilon} + (p+1)rV_{GCA}$	$V_{\varepsilon} + rV_{SCA} + (p+1)rV_{GCA}$
SCA	$\frac{1}{2}p(p-1)$	MS_{SCA}	$V_{\varepsilon} + rV_{SCA}$	$V_{\varepsilon} + rV_{SCA}$
随机误差	$\frac{1}{2}p(p+1)(r-1)$	MS_{ε}	V_{ε}	V_{ε}
总和	$\frac{1}{2}p(p+1)r-1$			

无自交的方差分析自由度与平方和

- 一般配合力共有 p 个，自由度为 $p-1$ 。共配置 $p(p-1)/2$ 个杂交组合，组合的自由度为 $p(p-1)/2-1$ ，从中减去一般配合力自由度，就是特殊配合力自由度 $p(p-3)/2$ 。总自由度为 $p(p-1)r/2-1$ ，从中减去组合的自由度，就是误差自由度，即 $p(p-1)(r-1)/2$ 。
- 亲本一般配合力的平方和乘以 $(p-1)r/2$ 、再乘以2之后，就是一般配合力平方和。特殊配合力的平方和乘以 r 之后，就是特殊配合力平方和。根据剩余效应计算误差平方和，最后就得到方差分析表。

只包含正交的完全双列杂交方差分析表

亲本数为 p ，配制包括 $p(p-1)/2$ 个正交组合，每个组合有 r 次重复观察值

变异来源	自由度	均方	固定模型期望均方	随机模型期望均方
GCA	$p-1$	MS_{GCA}	$V_{\varepsilon} + (p-1)rV_{GCA}$	$V_{\varepsilon} + rV_{SCA} + (p-1)rV_{GCA}$
SCA	$\frac{1}{2}p(p-3)$	MS_{SCA}	$V_{\varepsilon} + rV_{SCA}$	$V_{\varepsilon} + rV_{SCA}$
随机误差	$\frac{1}{2}p(p-1)(r-1)$	MS_{ε}	V_{ε}	V_{ε}
总和	$\frac{1}{2}p(p-1)r-1$			

重复平均数、一般配合力和特殊配合力估计

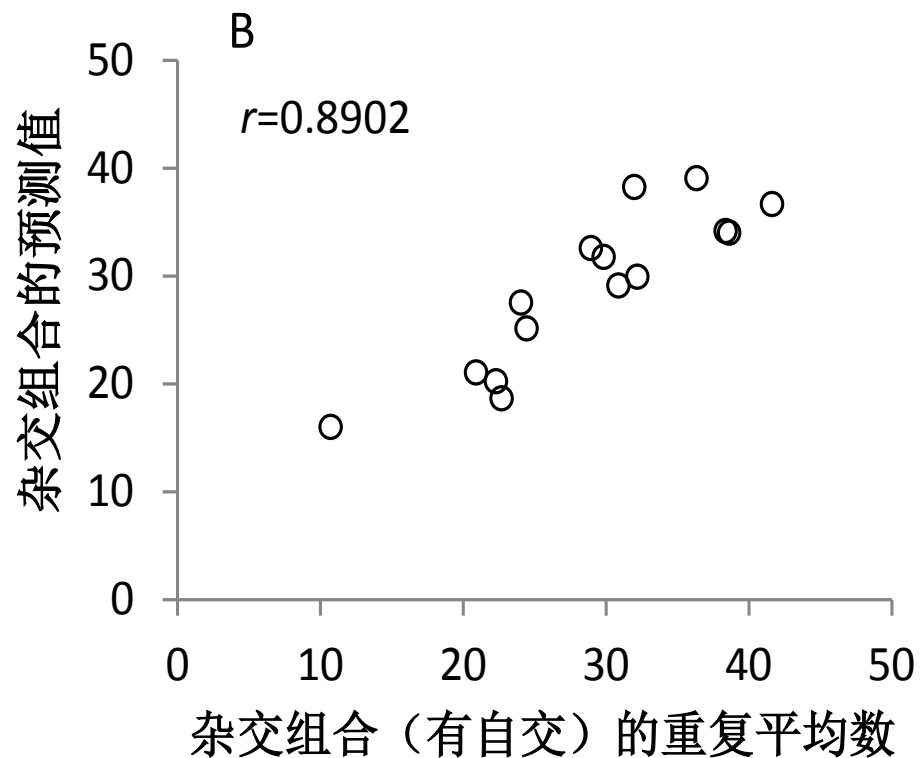
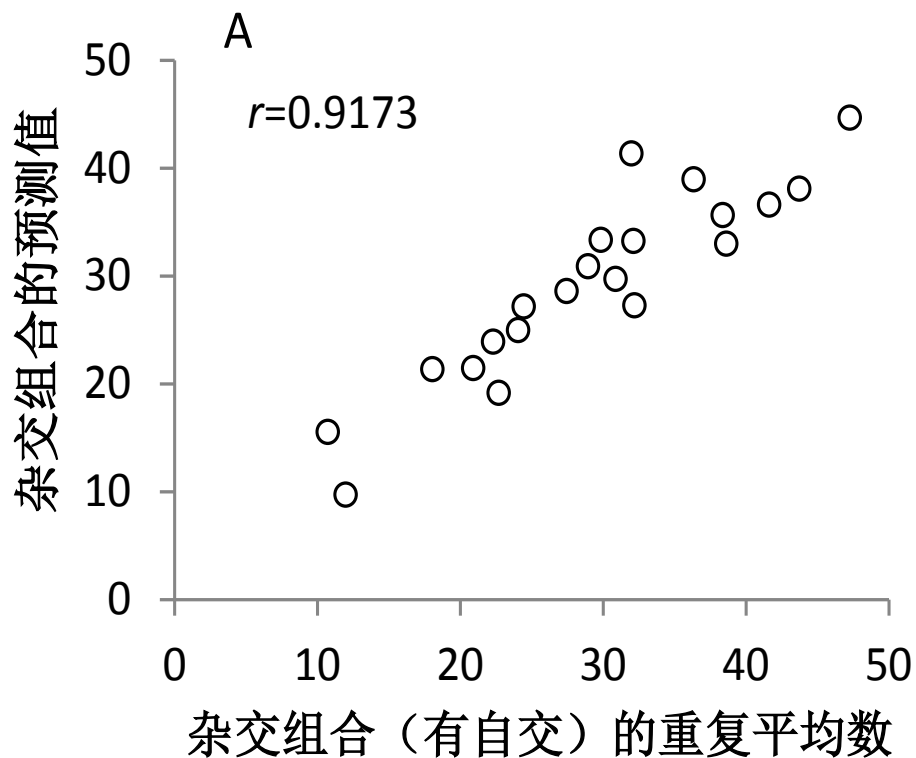
6个亲本的15个正交和6个自交完全双列杂交设计

亲本	两次重复的平均数						行列平均
	A	B	C	D	E	F	
A	27.45	29.85	41.65	28.95	24.05	22.70	28.87
B		43.75	32.00	38.40	30.90	22.30	34.42
C			47.30	36.35	38.65	24.45	38.24
D				32.15	32.20	20.90	31.59
E					18.05	10.75	24.66
F						12.00	17.87
总平均	29.28						29.28
亲本	SCA						GCA
A	-1.12	-3.57	4.88	-1.99	-0.84	3.76	-0.35
B		5.47	-9.62	2.60	1.16	-1.50	4.50
C			2.33	-2.79	5.56	-2.69	7.85
D				-1.17	4.94	-0.42	2.02
E					-3.16	-4.51	-4.04
F						2.68	-9.98

利用表12.8中15个正交及有自交和无自交两次重复观测数据的方差分析

设计类型	变异来源	自由度	平方和	均方	F值	P值	方差估计
有自交	GCA	5	2826.65	565.33	54.11	<0.0001	39.63
	SCA	15	608.82	40.59	3.88	0.0023	15.07
	随机误差	21	219.42	10.45			10.45
	总和	41	3654.89	98.99			
无自交	GCA	5	1876.98	375.40	34.61	<0.0001	36.46
	SCA	9	393.43	43.71	4.03	0.0086	16.43
	随机误差	15	162.68	10.85			10.85
	总和	29	2433.09	70.95			

利用一般配合力的杂交组合预测值



§ 12.4 轮回选择与群体改良

- § 12.4.1 轮回选择育种方法与群体改良
- § 12.4.2 群体内轮回选择的遗传进度
- § 12.4.3 群体间轮回选择的遗传进度
- § 12.4.4 杂种优势预测与配合力选择

轮回选择育种方法与群体改良

- 作为动植物育种的一个重要方法，轮回选择（recurrent selection）是对育种群体的周期性改良。轮回选择通过表型评价、基因型鉴定等手段，选择群体中的优良个体，在这些优良个体间互交形成下一代的育种群体，并不断重复这个过程。
- 换句话说，轮回选择是不断重复“鉴定、选择、重组”过程的一种群体改良方法。

轮回选择的优点

- 广义地说，图12.1表示的自交系选择过程也可被看作轮回选择的一种。但一般意义上，轮回选择指的是一个群体内的改良或两个群体间的交互改良，改良后的群体可以作为开放授粉品种或作为选育新近交系的来源。
- 对图12.1的系谱选择来说，由于自交使得基因很快得以固定，不同基因之间没有充分的时间发生重组。轮回选择从两个方面弥补这一不足：一是通过周期性的选择，不断提高群体中有利基因的频率；二是通过不断重组打破有利基因与不利基因之间的紧密连锁，保持群体的遗传变异，以保证下一轮选择的遗传进度。

玉米蛋白质和油份含量的长期选择试验

- 国际上最著名的轮回选择，是美国Illinois大学100多年来关于玉米蛋白质含量和油份含量的选择试验。这项玉米群体改良最早是由C. G. Hopkins于1896年开始的，原始群体是“Burr's White”，Hopkins在田间挑选163个穗子，于实验室内测定其油分和蛋白质含量，原始群体的油分含量为4.7%，蛋白质含量为10.9%。
 - 按高油目标，选择24个穗子形成“Illinois高油品系”；
 - 按低油目标，选择12个穗子形成“Illinois低油品系”；
 - 按高蛋白目标，选择24个穗子形成“Illinois高蛋白品系”；
 - 按低蛋白目标，选择12个穗子形成“Illinois低蛋白品系”。
- 以这4个群体作为第一个轮回的杂交亲本，随后在4个方向上分别进行选择。

玉米蛋白质和油份含量的长期选择试验

- 经70代的轮回选择后，高油群体的油份含量达16.6%，低油群体的油份含量达0.4%，分别相当于原始群体的354.8%和8.5%；高蛋白群体的蛋白质含量达26.6%，低蛋白群体的蛋白质含量达4.4%，分别相当于原始群体的224.0%和40.4%。
- 第48代时，在上述基础上又增添了逆向选择，至70代时高油转低油方向下降了8.85%，低油转高油方向上升了2.38%；高蛋白转低蛋白方向下降了8.5%，低蛋白转高蛋白方向上升了9.6%；高油转低油再转高油方向则由约9%回到14.02%。

玉米蛋白质和油份含量的长期选择试验

- 这项试验在进行到第76代时，Dudley对试验结果作了系统分析，结果表明76代后仍看不出选择进度有停滞的迹象，选择仍可不断地获得遗传进度。第100代的结果验证了这一观点。
- 从已实现的遗传进度推算加性遗传方差，再估计涉及的基因数目和增效基因频率，结果表明高油含量与低油含量材料间的有效因子在50个左右，高蛋白含量与低蛋白含量间的有效因子在120个左右；高油含量与低油含量间存在相引连锁，在这个群体中，仍是以加性方差为主，显性方差约仅为加性方差的38~46%。

群体间和群体内轮回选择

- 由于杂交和自交都比较容易实现，轮回选择在玉米育种中得到广泛的研究和应用。玉米轮回选择分群体内和群体间两种类型。
- 群体内轮回选择的目的是，通过不断提高群体内有利等位基因频率来提高群体的平均表现，一般基于个体自身表现或其自交后代的平均表现开展选择；改良后的群体可以直接作为品种（称群体品种）或作为自交系选育的材料来源。
- 群体间轮回选择以选育杂交种为目的，交互轮回选择的两个群体一般是两个互补的杂种优势群体，一般基于个体和自交后代的测交表现开展选择，两个群体互为测验种。然后从改良后的群体中选育自交系，在不同来源的自交系间配置杂交组合、选育优良杂交种。

常用的玉米轮回选择方法

轮回选择方法	测试后代类型	重组后代类型
群体内改良		
混合选择, 花粉控制	单个植株	单个植株
混合选择, 无花粉控制	单个植株	开放授粉种子
半同胞家系选择	半同胞家系	半同胞家系
修饰穗行选择	半同胞家系	开放授粉种子
半同胞家系/ S_1 选择	半同胞家系	S_1 家系
全同胞家系选择	全同胞家系	全同胞家系
自交家系选择	S_1 或 S_2 家系	S_1 或 S_2 家系
群体间改良		
交互轮回选择	半同胞家系	S_1 家系
交互全同胞选择	全同胞家系	S_1 家系
测交选择	测交组合	S_1 家系

常用的轮回选择方法的预期遗传进度

轮回选择方法 (V_A 的系数)	单周 期年 数	产量		穗长	
		每年	每个育 种周期	每年	每个育 种周期
混合选择, 无花粉控制 (1/2)	1	0.054	0.054	8.9	8.9
混合选择, 花粉控制 (1)	1			17.8	17.8
修饰穗行选择 (1/8)	1	0.175	0.175	5.6	5.6
半同胞家系选择 (1/4)	2	0.175	0.350	5.6	11.3
半同胞家系/ S_1 选择 (1/2)	2	0.350	0.701	11.3	22.5
全同胞家系选择 (1/2)	2	0.225	0.451	7.9	15.8
自交 S_1 家系选择 (1)	2	0.361	0.722	11.3	22.7
自交 S_2 家系选择 (3/2)	3	0.328	0.985	5.2	15.7

杂种优势预测与配合力选择

- 在杂交种选育过程中，一旦确定了一个作物的杂种优势有利用价值之后，育种家关心的其实已不再是杂种优势本身，而是杂交组合的具体表现，即如何从众多的组合中发现优良杂交组合。
- 优良杂交组合可能是表现较好的两个自交系间的杂交，也可能是表现不太好的两个自交系间的杂交，也就是说难以从自交系的自身表现推测杂交种的产量和杂种优势。
- 实践中，由于自交系的数量很大，要在田间测试所有可能的杂交组合是不太可能的，因此杂交种表现的预测一直是杂交育种中的一个重要课题。

杂种优势预测与配合力选择

- 利用自交系表现是最早研究的一个预测方法。以玉米为例，杂交种的亲本常常来自不同的杂种优势群，假定杂种优势群P1中有100个自交系，杂种优势群P2中有150个自交系，如果可用两个自交系的平均产量预测杂交种的产量，那么我们只需调查群体P1中的100个自交系和P2中的150个自交系的产量，就能得到P1和P2间的15000个杂交种的产量表现。
- 实际上，杂交种的产量和自交系产量间的相关很低甚至没有相关性，其主要原因是非加性效应（如显性效应和上位性效应等）的存在，当然性状遗传力较低也是一个原因。近些年，人们还探索了利用分子标记的基因型鉴定数据，来预测杂交种的表现。

杂种优势预测与配合力选择

- 植物育种中，每个双亲杂交群体都很容易产生数百，甚至数千个基因型互不相同的重组近交家系。例如，双亲的差异座位数为 n ，纯合基因型的个数为 2^n 。当 $n=10$ 时， $2^n=1024$ ； $n=15$ 时， $2^n=32768$ 。
- 就目前来看，准确预测两个自交系的杂交种表现还很困难。从公式12.36~公式12.38我们看到，加性方差存在于一般配合力中，显性方差或非加性方差存在于特殊配合力中。同时，近交程度越高，加性方差前面的系数越大，说明一般配合力的选择效果越好。

杂种优势预测与配合力选择

- 因此，可以利用一般配合力方差、特殊配合力方差及组合间遗传方差的构成，对于加性方差明显的育种群体，可以通过早期世代的个体表现和测交表现，对一般配合力进行选择，进而及早淘汰一般配合力较低的自交系，提高杂交组合群体的平均表现。
- 特殊配合力是由非加性方差决定的，只有在配制了杂交组合之后才能进行估计。因此，在自交系的纯合过程中，难以对特殊配合力开展早期世代选择。特殊配合力的有效预测与选择，仍然是遗传育种中有待深入研究的一个重要科学问题。