

现 代 遗 传 学 丛 书

统计遗传学

Statistical Genetics

顾万春 著



科学出版社

www.sciencep.com

现代遗传学丛书

统计遗传学

顾万春 著

科学出版社

北京

内 容 简 介

本书是现代遗传学丛书之一。它运用遗传学与数学的理论和方法,归纳整合群体遗传学、数量遗传学、生态遗传学和分子遗传学等分支学科内容,系统深入地论述自然选择与人工选择、基因频率与概率、群体遗传结构与分化、繁殖机制与瓶颈效应、遗传漂变与样本策略、梯度变异与生态遗传、遗传参数与数学模型、选择指数与选择效率、分子标记与MAS以及遗传多样性测定评价等遗传学量化命题,阐明遗传变异规律及方法论,指导遗传与育种研究实践。全书系统清晰,内容翔实,图文、表配置有序。

本书可供综合性大学和林业、农学、医学、植物学类大专院校的遗传学、育种学、种质资源学、生态学、生物工程学等方面的教师、研究生、大学生以及有关科研人员参考。

图书在版编目(CIP)数据

统计遗传学/顾万春 著. —北京: 科学出版社, 2004.6

(现代遗传学丛书)

ISBN 7-03-012476-6

I. 统… II. 顾… III. 遗传学-应用统计学 IV. Q3-32

中国版本图书馆 CIP 数据核字(2004)第 001656 号

责任编辑: 马学海 吴寅泰 宛 楠/责任校对: 朱光光

排版制作: 科学出版社编务公司/责任印制: 安春生

封面设计: 王 浩

科 学 出 版 社 出 版

北京东黄城根北街16号

邮政编码:100717

<http://www.sciencep.com>

源海印刷有限责任公司印刷

科学出版社发行 各地新华书店经销

*

2004年7月第 一 版 开本: 787×1092 1/16

2004年7月第一次印刷 印张: 27 1/4

印数: 1—3 000

字数: 625 000

定价: 58.00 元

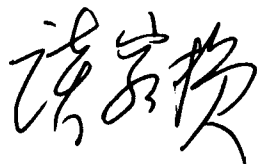
(如有印装质量问题, 我社负责调换〈新欣〉)

《现代遗传学丛书》序

《现代遗传学丛书》诞生于我国“科学的春天”。1978年，在中国遗传学会成立大会上，经科学出版社罗见龙、蒋伯宁先生提议，大会代表一致同意由我的老师——中国遗传学的先驱和奠基人之一李汝祺教授和本人主编这一套遗传学基础理论系列书，以供生命科学领域的科研人员、教师、研究生和高年级大学生阅读。根据当时的约定，本丛书的组稿原则是按学科发展的需要与可能，成熟一本列选一本。

本丛书的第一个分册是1981年出版的盛祖嘉教授的《微生物遗传学》，此书曾修订再版，印数已超过27000册。紧接着，发育遗传学创始人之一李汝祺教授亲自完成的《发生遗传学》(上、下册)于1985年问世，被我国遗传学界誉为“中国遗传学的经典著作”。我和李汝祺先生曾请李竞雄院士撰写《植物细胞遗传学》，此书在1993年由李竞雄与宋同明二位教授合作完成。本丛书于世纪之交又出版了由分别童克中、刘良式、盛志廉、陈瑶生和孟金陵等教授撰写的《基因及其表达》、《植物分子遗传学》、《数量遗传学》、《植物生殖遗传学》4个分册，受到读者的欢迎。今年本丛书还将出版张玉静教授主编的《分子遗传学》、童克中教授所著的《基因及其表达》第二版、吴常信院士主编的《动物遗传学》、盖钧镒教授主编的《植物数量性状遗传体系》、顾万春教授所著的《统计遗传学》，以及为推动遗传学出版物中符号的使用与国际接轨，由王金发教授主译的《TIG遗传命名指南》等。接下来将出版的还有吴旻院士主编的《肿瘤遗传学》、印木泉教授等编著的《毒理遗传学》、刘良式教授主编的《植物分子遗传学》第二版、杜若甫教授主编的《中国人群体遗传学》、陈竺和强伯勤院士主编的《基因组学》，以及盛祖嘉、陈永青和毛裕民教授编著的《微生物遗传学》第三版等。藉本丛书扩大开本之际，特作此序，并感谢童克中教授为丛书封面的设计提出反映学科内涵的创意。

希望上述遗传学家的这些著作能引出我国年轻一代遗传学者层出不穷的佳作，为推动我国生命科学基础学科更加健康和迅速地发展，为我国的科技现代化做出应有的贡献。



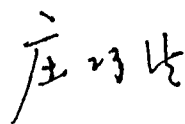
2000年4月19日

序

这是一本以多年生木本植物为研究对象，融汇群体遗传学、生态遗传学、数量遗传学、分子遗传学等分支学科的基础知识，将群体遗传结构的变化规律及其影响因素、样本概率、遗传多样性分析、群体与数量遗传参数估算以及选择反应与效率等理论与实践的重要问题，运用统计学、数学的方法予以贯通并量化，可用来指导遗传、育种研究的统计遗传学专著。

多年生木本植物，因野生性强、杂合度高、个体大和生育周期长等特点，其遗传背景、性状遗传与变异比较复杂，采取的遗传改良技术路线与农作物、畜禽有较大不同，需要从群体出发，讲究抽样策略，并借助于统计学原理和方法论。这方面的研究在国内都比较薄弱，这本《统计遗传学》的出版，对我国来说，确是一个良好的开端。作者得益于多年来从事遗传学和林木遗传育种学科研究与教学的积累，在消化吸收、移植和发展应用诸多遗传学原理与方法的基础上编著了这本书。书中对多年生木本植物的自然选择与人工选择、群体分布和模型、群体遗传结构与变异、交配机制与繁殖、遗传漂变与样本策略、梯度变异和生态遗传、遗传参数和数学模型、选择育种与优化模型、分子标记及量化评价等进行了较系统的阐述，拉近了遗传学理论，尤其是群体遗传学理论与多年生木本植物遗传改良的关系。一些新内容的撰写与编排也反映出作者在遗传学领域的较深造诣，譬如，选择作用与适合度评价，遗传漂变和样本技术，实用化的遗传多样性量化评价技术，梯度变异的原理与数学表述，隔代遗传的相似性，分子标记辅助选择的(量化)评价分析等，都是其中比较新颖和可资借鉴的内容。

书中所阐述的与遗传学有关的统计学方法，对于粗具群体或数量遗传学知识的读者来说，比较简便易懂，可以说是一本兼容基础理论与方法论的书籍。相信本书的出版将有助于提高我国遗传学、育种学研究的理论水平，将对木本植物遗传改良事业以及献身于动植物遗传育种的科研和教学人员作出深层贡献；也希望以此书为契机，触发我国更多的年轻一代学者研究，继续撰写出更新、更好的科学著作。



2002年10月于北京

前 言

统计遗传学是运用统计学与遗传学的频率与概率的基础理论,归纳遗传规律与遗传参数及其关系,并与遗传改良相结合的遗传学分支学科,是归纳逻辑指导下的整合遗传学。

总结近年来为研究生讲课和专业培训班讲课的讲稿,结合以往研究与教学的积累,经过滚动修改和补充,最终写成本书。全书依托遗传学理论与新成就,用统计学、数学与遗传学的概率与频率的原理,将群体与个体的遗传规律联系在一起,解析遗传学规律与量化特征的逻辑关系,归纳了群体遗传、数量遗传、生态遗传和分子遗传等多个遗传学分支的实用统计遗传学内容,是多学科分支交叉的新体系。

人们认识到,生命对环境反应的单元是个体,适应与进化单元是群体和物种。但长期以来了解生物遗传对环境适应的微观机制甚少,遗传学的群体研究与个体研究、微观研究与宏观研究是脱节的。

多年生木本植物的遗传改良迫切需要将群体与个体相结合的遗传学理论和方法论的指导,但长期以来因种种原因,这种需求或被分隔,或被忽视。为此,本书强调“物种/群体/家系/个体”在遗传学与育种学研究上的重要意义,贯通遗传单元组合的统计遗传学研究系统,定量化阐明遗传变异规律、遗传评价与遗传改良的多种参数。第一章至第八章从基因频率与统计概率入手,阐述选择作用、基因频率估算、适合度、群体分布与分割,迁移与基因流、近交与性状方差、交配系统、瓶颈效应和遗传漂变等内容;第九章综合介绍常用的遗传多样性参数;第十、十一章讲述基因梯度与梯度变异的统计遗传;第十二章至第十七章在概率与方差的基础上,分析性状遗传参数组合、选择效应与效率、指数选择等;第十八章简要介绍分子标记辅助选择(MAS)实验的统计遗传。为了便于对统计遗传学数学方法的掌握,书中备有主要参数符号及说明,附录给出了数学模型期望值、回归相关与组内相关以及梯度研究数学方法。

遗传科学研究与生物技术发展互为-体,已成为 21 世纪高新技术产业的重要支柱。基因组学、生物信息学和整合生物学将微观与宏观的遗传学连接统一起来。毋庸置疑,微观与宏观的遗传学的结合,是新世纪遗传学发展的必然趋势。书中第九章和第十八章的内容是适应这一发展趋势的归纳与例举。

本书在遵循遗传学理论与统计学原理的前提下,力求体现遗传系统的完整性,结合物种以下的遗传单元的遗传物质、遗传结构及遗传分化的组合系统,阐述遗传学的量化规律关系,力求构筑统计遗传学指导实践的完整性、系统性和实用性。书中对不少实用理论进行了拓广和深化。如,选择的三效应公式、选择适合度的分析评价、繁殖群体瓶颈效应模型、群体有效大小的模式、遗传多样性参数组合、梯度与梯度变异的耦合、概率与频率方差及性状遗传参数、调查遗传参数、遗传相似性模糊集、选择效率与预测、选择指数效率以及 MAS 实验的统计遗传等。全书在对遗传学理论与方法论进行较为系

统地深入论述的基础上，对统计遗传学理论与应用的重要进展有较充分的反映。

书中使用的遗传学词汇、统计公式及遗传参数符号较多，遗传参数均已附注英文，遗传学名词与词汇尽力标注英文；不同遗传学分支的习惯统计符号除用统计学惯用符号统一外，还保留一部分原有习惯符号，亦已简要说明。全书引用了较多的例证，除了模式实验例外，采用多年生木本植物较新的研究成果，将理论模型延伸到统计技术，兼顾遗传参数之间的量化关系，体现了理论与实践的结合。既具有针对性，又不失一般指导意义。

值本书出版之际，向对本书出版给予热情指导与帮助修改的中国科学院院士庄巧生先生，王明庥院士、范云六院士、董玉琛院士，表示诚挚的感谢。对在不同阶段为书稿做校对工作的胡新生博士、李志君博士、王军辉博士、李斌博士、李文英博士、兰彦平博士、明军博士、罗建勋博士生、程诗明博士生、林富荣博士生等，一并致以谢意。

顾万春

2004年6月于北京

统计遗传学主要参考符号

符号	说明	符号	说明
A_1, A_2, \dots, A_i	基因座上等位基因; 复等位基因	n	统计量样本; 等位基因数
A, A_e	等位基因数, 有效等位基因数	Δ	递增量; 收获量标符, 变化率
$p(p_i), q$	基因频率	GC	群体有效大小的效率指数
$u(P), \Delta P$	概率, p 与 $(p+dp)$ 之间的概率密度	G_{ST}, G_{IT}	基因分化系数
$f(x), \phi(u)$	正态分布密度函数, 密度函数	F_{ST}	(基因)固定指数
\hat{P}_T	样本平均频率的估算值	P_{IT}, P_{ST}	随机交配指数
$M\delta_p, M_i$	基因频率变化量的均值, 分子得分	$\alpha(\alpha_i)$	基因平均效应(值)
Q, d	基因频率离差的估算值; 离差	$a(A)$	加性效应; (a)平均剩余
Z	逻辑变换的基因频率, 位置概率函数	$d(D)$	显性效应; 显性离差
G_m	遗传复杂度	I, I_i	上位效应, 选择指数
δ, δ_T	遗传(平均)分化系数	\bar{x}	统计量平均值
$S(\alpha)$	基因频率的选择强度	d, d^2	误差, 显性的相关标记; 距离
P, Q, H, p_{ij}	基因型频率	σ^2	参数的方差
$\phi(x), E(x)$	频率函数	σ	参数的标准差
P, P^*, \bar{P}	概率, 表现型值; 多态位点百分率	S^2, V	统计量方差
u_k	基因(型)多样性	S, \sqrt{V}	统计量标准差
$G; y$	基因型值	$S_{\bar{d}}$	标准误差
$f(y_{ii})$	对应基因型值 y_{ii} 的频率分布	R	有效生殖率; 统计量极差
B, b, \tilde{b}	育种值; 回归系数	$r(r^2), r_1, r_{II}$	相关系数(确定系数); 组内相关系数; 选择正确度
μ	突变率; 参数平均值	i	世代数; 某界定的时间值
Δp	突变的变化率, 基因频率变化率	H^2, h^2	广义遗传力(率), 狭义遗传力, 多种 表达: $h_B^2, h_N^2, h_x^2, h_y^2$ 以及多元分析 遗传力 MH
ω	淘汰值	R	选择响应; 重复率, 对应于收缩系 数(C)
S	选择系数, 受精率, 隔离的某参数	GCA 与 SCA	一般配合力与特殊配合力
I, SD	选择差	r_G	遗传相关, 对应的表型相关(r_P) 和环境相关(r_E)
i	选择强度	P_{xy}, r_{ii}	通径系数(相关的分割系数)
W, W_{ii}, a	适合度, 马尔萨斯的适合度参数	GCV	遗传变异系数
m	迁移率; 群体间配偶相关系数	\hat{y}, Z	人工选择收获量函数
N_m	基因流	$\Delta G, \Delta G(t)$	遗传增益, 对应的良种增益(ΔB)
F, f, f_{ST}	近交率, 近交系数, F 分布	$K(\hat{i})$	选择指数
C	系数; 有效生殖系数; 收缩系数	\bar{y}_i	性状的期望值
L, GI	遗传负荷	EGA	生态梯度
CR	相关响应	CR, E, E_r	(人工)选择效率, MAS 选择效率
$u(p)$	基因固定率	χ^2	常用的卡方(适合性)检验
$H, \bar{H}, H_e, H_o, H_i$	杂合度(基因杂合度)	θ, cM	遗传图距, 厘摩
J, J_s, J_T	基因一致度; 纯合度; 遗传相似性系数		
D^N, d	基因多样性; 遗传距离		
N	总体数量; 群体实际大小		
N_e	群体有效大小(个数); 基因(型)多样性		
\tilde{N}_e	群体有效繁殖大小(个数)		

目 录

《现代遗传学丛书》序	
序	
前言	
统计遗传学主要参考符号	
第一章 绪论	1
第一节 统计遗传学	1
第二节 频率与概率	7
第三节 木本植物群体与个体	11
第四节 选择的效应式	15
第二章 基因频率与选择作用	17
第一节 基因频率的概率过程	17
第二节 选择情况下基因频率的随机变化	21
第三节 Hardy-Weinberg 定律	23
第四节 自然选择与适合度的基因频率	26
第五节 人工选择的基因频率	31
第三章 基因频率的估算	37
第一节 基因频率的估算和检验	37
第二节 样本及平衡定律离差的估算	43
第三节 近交系数与杂合度的估算	46
第四节 突变基因频率的估算	51
第五节 两种分布模型选择强度的参数估算	54
第四章 选择与适合度	57
第一节 自然选择的基本定理与选择系数	57
第二节 适应的量化描述	62
第三节 选择引起适合度的变化	64
第四节 适合度与数量性状	68
第五节 遗传负荷	73
第五章 分布、迁移与群体剖分的频率和概率	78
第一节 增长模型与群体分布	78
第二节 主要分布模型的基因频率和概率	82
第三节 迁移和群体剖分	89
第四节 亚群体与随机交配指数	92

第五节 迁移对群体遗传结构的影响	94
第六节 木本植物群体的基因流	98
第六章 近亲交配和性状方差	102
第一节 近亲关系的协方差	102
第二节 已知亲缘关系的协方差	104
第三节 近亲交配对数量性状的影响	106
第四节 群体剖分引起的方差变化	108
第五节 群体有效大小与杂合度变化率	110
第六节 群体剖分与性状方差	112
第七章 繁殖群体与瓶颈效应	116
第一节 物种与群体	116
第二节 繁殖结构	118
第三节 群体有效大小	121
第四节 交配系统	123
第五节 “瓶颈效应”的定量模型	127
第六节 群体有效大小与瓶颈效应	133
第八章 遗传漂变和样本策略	138
第一节 遗传漂变及概率估算	138
第二节 遗传漂变的影响及调控	140
第三节 群体有效大小与子群体样本	144
第四节 家系的样本	150
第五节 基因型与配子体的样本	153
第六节 “核心种质”的样本策略	154
第七节 运用“捕获曲线”制定核心种质的样本	160
第九章 遗传多样性参数	165
第一节 表型多样性参数	165
第二节 基因频率与基因种类的参数	171
第三节 基因多样度和多态位点百分率	175
第四节 杂合度与一致度	179
第五节 亚群体基因多样度	184
第六节 遗传距离	186
第七节 遗传分化系数(G_{ST})与固定指数(F_{ST})	191
第八节 DNA 信息参数的符号	196
第十章 基因梯度与变异	199
第一节 梯度与生态遗传	199
第二节 基因频率的梯度	201
第三节 群体性状和 RFLP 分子标记的梯度变异	205

第四节	生态适应指数(EQ)评价表型梯度变异	209
第五节	梯度变异的动态评价与预测	213
第十一章	生态遗传与生态梯度	216
第一节	生态梯度[EGA(PC _r)]与梯度变异	216
第二节	生态梯度[EGA(CA)]用于生态遗传研究	223
第三节	性状值、环境指数与生态梯度的耦合	230
第四节	中国生态梯度与应用式	235
第十二章	概率与方差	241
第一节	数值的概率与方差	241
第二节	数量性状与基因频率方差	246
第三节	基因平均效应与方差	250
第四节	显性方差与上位方差的频率估算	254
第五节	基因型值与育种值	260
第六节	基因频率与效应对遗传方差的影响	262
第七节	环境方差和 GE 交互作用	268
第十三章	重复率、遗传力与配合力	273
第一节	重复率	273
第二节	遗传力	278
第三节	遗传力的估算	282
第四节	遗传力估算方法的比较	285
第五节	配合力	292
第六节	调查估算遗传参数	298
第十四章	遗传相关、回归与通径	305
第一节	遗传协方差	305
第二节	遗传相关	309
第三节	遗传、环境及表型的相关	311
第四节	GE 互作与遗传相关	314
第五节	回归与遗传通径	318
第六节	模糊数学方法估算隔代遗传相似性	321
第十五章	性状的遗传稳定性和生长适应性	324
第一节	GE 互作及遗传稳定性和生长适应性	324
第二节	遗传稳定性和生长适应性的评价参数	325
第三节	分析评价技术	329
第四节	结构模型的评价参数 $\hat{\alpha}_i$ 与 $\hat{\lambda}_i$	334
第五节	GE 互作分析评价参数的简评	337
第十六章	选择效率	338
第一节	选择效应预测式	338

第二节	选择差和选择强度	340
第三节	选择效应预测与实际遗传力	342
第四节	选择效率	343
第五节	木本植物群体大小与群体选择效率	348
第十七章	选择指数	351
第一节	常用选择指数的种类	351
第二节	选择指数的构建与适合度权重	356
第三节	选择指数的基本模型	359
第四节	选择指数的运用	363
第五节	纸浆材育种的选择指数	369
第六节	选择指数(RESI 程序)的变量与结果评价	372
第十八章	分子标记辅助选择(MAS)实验的统计遗传	378
第一节	MAS 与统计遗传	378
第二节	QTL 实验精密性与有效性	381
第三节	实验样本与相对效率	384
第四节	QTL 与 MAS 的参数估算	387
第五节	MAS 选择指数和效率评价	390
第六节	MAS 技术中有关问题	395
附录一	数学模型的期望值	398
第一节	基本描述	398
第二节	单点试验的数学模型及 EMS	398
第三节	多点试验或多年试验的数学模型及 EMS	399
第四节	多点、多年试验的数学模型及 EMS	402
附录二	回归相关与组内相关	405
第一节	回归分析	405
第二节	回归的检验与预测的运用	408
第三节	相关系数分析与检验	410
第四节	组内相关	412
附录三	梯度研究的数学方法	415
主要参考文献		420

第一章 绪 论

20 世纪初以来,人类发展与科技进步,促进了遗传学的建立与完善。经过“实践、归纳、演绎和逻辑”认识论升华,遗传学已经成为生命科学的重要学科。随着理论指导实践的双向循环,先后建立了群体遗传学、数量遗传学、生态遗传学、发育遗传学和分子遗传学等诸多遗传学分支学科,深化和拓展了遗传学理论与方法论。

统计遗传学的产生是基于植物遗传学研究与遗传改良的需要,运用频率与概率的归纳逻辑,讨论遗传变异量化规律、遗传参数及其相互关系,较为完整地阐述群体与个体的遗传学联系。

第一节 统计遗传学

统计遗传学(statistic genetics),是统计、归纳遗传规律与参数的逻辑关系,并与遗传改良相结合的遗传学分支学科。

统计遗传学依托遗传学已有知识,运用统计学与遗传学的概率与频率的基础原理,将群体与个体遗传学规律联系在一起,解析遗传学特征与量化逻辑关系,贯通遗传学基础理论与遗传改良实践的有机联系。适用于植物遗传多样性与遗传改良的研究需要。

一、统计遗传学运用归纳逻辑

统计遗传学对实验数据的数学处理,运用归纳逻辑,把数集简化为少数简单的常数或参数,评定这些常数或参数的显著性或重要性。其实验的目的是由特殊去论证一般,涉及更多的是归纳逻辑方法,而非演绎逻辑方法。

概率是统计遗传学的基本内容之一。实验数据收集与分析,建立在概率基础上,从计算结果的几个可能抉择中选出 1 个作为结论,包括参数方法与非常数方法。

在 n 个事件中,有 m 个特殊事件的概率(P):

$$P = m/n \quad (1.1)$$

两个独立事件同时发生的概率,等于各个事件自发生概率的积。

频率是频次的分布量值,是统计遗传学中的基本统计量与参数之一。在群体中分组(类)个体计量,或遗传单元的位点与谱带的计量,或数量性状观测值大小的组阶计量,每组的中值为中价,得到频次分布。譬如,正态分布概率的二项式展开,有如下公式。

求 m 为 x 组的频率,如 μ 为均值、 σ 为标准差、 σ^2 为方差、 e 为自然对数的底、 π 为圆周率,则正态分布曲线公式是:

$$m = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (1.2)$$

式(1.2)勾画成“钟形”曲线。曲线中心线为平均值 μ ，钟形曲线斜率最大点至曲线中心线的最短距离即是标准差(σ)。

二、统计遗传学重视群体/个体遗传逻辑关系

大多数物种在自然界通常都是以群体方式生活着，多年生木本植物也不例外。种以下群体的分布状况和增长速率，取决于群体的遗传结构、群体大小或密度、交配机制或体系以及群体所处的生态环境。群体内的基因与基因型的种类和频率的变异，规定了群体遗传结构及其变化规律。

生命对环境反应的单元是个体，适应与进化单元是群体和物种。群体由个体组成，群体具有超越个体的遗传学特性。多年生木本植物群体可以表现出种内遗传单元构建层次上系统等级、各个等级内与等级间直至个体之间的相互作用以及其他各种遗传特征。

表 1.1 统计遗传学研究群体/个体的对应学科与相关遗传因子

自然群体 群体遗传理论与模式参数	人工群体 群体遗传评价与利用	(亚群体)谱系与个体 数量性状遗传的研究与改良	群体与个体 遗传信息测定与利用
群体遗传学 生态遗传学 统计遗传学	群体遗传学 生态遗传学 统计遗传学	数量遗传学(育种学) 发育遗传学 统计遗传学	分子遗传学 发育遗传学 统计遗传学
群体遗传系统:	繁殖群体遗传系统:	目的性状(组)遗传变异及系统:	遗传系统的分子标记 基因组的组合:
遗传结构 遗传传递方式 交配系统 遗传分化	遗传结构 繁殖结构 遗传传递方式 遗传分化	遗传变异规律 有性和无性繁殖 繁殖结构 遗传分化	遗传信息组成 酶标记 DNA 标记(定位、标记) 遗传分化
影响因子: 频率 p . 概率 P	影响因子: 频率 p . 概率 P	影响因素: 频率 p . 概率 P	影响因素: 频率 p . 概率 P
总体——参数; 样本——统计量。 数据分布型, 概率, 精度。			
选择(S) 适合度(W) 突变(μ) 迁移(m) 遗传漂变与群体大小 (N , N_e , \tilde{N}_e) 近交(F) 随机交配指数(P) 遗传分化(f_{ST} 等)	分布模型(M) 选择强度(i , SD) 适合度(W) 群体有效大小 (N_e , \tilde{N}_e) 隔离与迁移(m) 自交与近交(F) 群体剖分 瓶颈效应	选择强度(i) 选择差(SD) 适合度(有性或无性)(w) 数量性状遗传参数 (R , h^2 , ΔG , I ...) 选择预测反应 $G \times E$ 优化配置 多世代育种 MAS	分子标记, 遗传多样性 核酸碱基配制, 基因定位 基因功能, 基因复制(克隆) 突变(诱变、导入) 核外基因(cpDNA, RNA) 基因互作, 连锁互换 核酸与蛋白质互作, 基因克隆与表达 QTL, MAS

种内遗传单元的等级，类同于生物系统学等级，包括：物种以下的群体(“产地”)/亚群体(林分)/家系/个体(包括无性系)。遗传单元等级是种内完整的遗传系统，具有统一

性与完整性。在实践中需要统筹安排,或者分段研究逻辑评价。通常情况下,各个遗传单元样本大小直接制约频率大小与概率分布,同时,也直接影响遗传特性或表型特征参数的可信程度与重要性。

植物遗传学研究及遗传改良工作,对自然群体、人工群体以及群体/个体的组合材料,遗传改良主要是群体选择利用,个体选择的群体构建利用或无性系利用。从群体到个体以及从个体到群体双向的遗传学逻辑关系,尤其是遗传参数所描述的规律,是统计遗传学研究的重点之一(表 1.1)。

三、统计遗传学表达遗传系统的完整性

遗传系统(genetic system),是指一个生物种的遗传物质、遗传结构及遗传传递(繁殖)方式等构建的组合系统。遗传系统是对物种遗传属性的描述,是统计遗传学研究的基本内容。比较草本植物及其他生物,木本植物遗传系统具有特有的属性。

木本植物中绝大多数物种野生性强,杂合度高,遗传背景复杂,所处生态环境多样,个体生长和发育周期长,群体繁衍进化历史缓慢。多数物种内天然群体残次性(濒危群体)较为普遍,少数树种的部分群体经过 1~3 代的个体选择人工遗传改良,仍保留自然群体的遗传结构及固有的遗传属性。种间与种内遗传系统表现出多样性。

统计遗传学研究遗传规律与参数的逻辑关系,需要了解种内遗传系统,包括群体遗传结构与遗传(繁殖)传递方式,对其数值分布的频率与概率运用归纳逻辑方法,比较全面地评价物种的遗传变异与遗传多样性及其参数的重要性。以利于制定适宜的遗传改良策略。

生物遗传系统的研究,通过“实验、归纳”数集和参数,在一定可靠性与精度(概率)的要求下,研究遗传物质组成、遗传结构及其传递方式。在遗传系统中,遗传结构是主要研究内容,它延伸到遗传传递方式。遗传结构包括基因种类、基因与基因型频率。基因种类存在物种遗传单元系统之中,用基因座上的等位(复等位)基因数量、功能和分布表示,大体上包括 3 个层次:

- 1) 区分质量性状遗传与数量性状遗传的主基因与微效多基因,包括 QTL 技术中的微效基因分成主效基因(大效应)与微小效基因(微效应);
- 2) 区分基因功能的加性基因、显(隐)性基因和上位基因;
- 3) 基因分布频率上分为:全域基因、广域基因、局域基因与稀有基因等。

在基因种类研究的基础上,结合基因(型)频率的量化表述,剖析生物系统学等级上的基因(型)种类与分布。基因频率(p)和基因型频率(P)是群体遗传结构的表征之一,遗传单元间基因(型)频率方差、方差分量及贡献率是检验遗传结构的重要参数。分子标记计数的基因频率或谱带频率,为二项分布,在大样本前提下接近于正态分布。表征群体遗传结构的遗传参数受概率($P=1-\alpha$)制约;同时,群体遗传结构受自然和人为的影响,在环境压力、进化动力和偶发因素的影响下动态地有序或无序地运动着(图 1.1)。

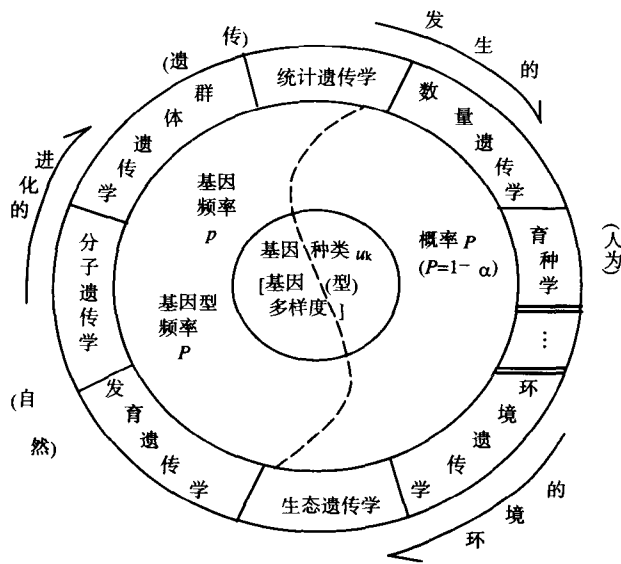


图 1.1 统计遗传学研究的内涵与外部联系

在人类创立的遗传学分支中，群体遗传学着眼于自然群体遗传系统的理论模型研究和验证；生态遗传学是从遗传学与生态学的关系出发，研究各类群体的遗传结构的生态反映；数量遗传学研究多基因控制的数量性状并对其收获量的数据进行处理与统计分析，并将其上升到遗传变异规律等(表 1.1)。生物的遗传现象及遗传系统本来是一个完整总体，物种/群体/家系/个体/细胞/DNA 分子之间密切联系、交织有序，体现了遗传多样性各级层次。

统计遗传学是运用统计学理论与方法解析遗传现象，属于遗传学交叉的学科分支。它通过(需要与可能的)统计分析 with 量化表达将遗传学各分支学科串联起来，拉近理论与实践、常规技术与高新技术的关系，对植物遗传改良更具有指导与应用价值。

四、统计遗传学强调群体遗传结构的量化关系

经典的群体遗传学、数量遗传学、生态遗传学和新近分子遗传学都十分重视群体遗传结构的研究(图 1.2)，但其学科分支侧重研究与运用的领域、范畴的内容有所区别，所用的遗传参数符号也不尽相同。为了说清楚统计遗传学与各个遗传学分支的关系，下面将主要遗传学科分支作简要介绍。

1. 群体遗传学

群体遗传学(population genetics)研究生物群体遗传结构及其传递规律的遗传分支。它是以自然群体为对象，根据遗传学原理，借助数学及统计学方法，分析研究群体基因种类、基因频率和基因型频率的变化规律及其进化规律的学科。多年来，群体遗传学吸取

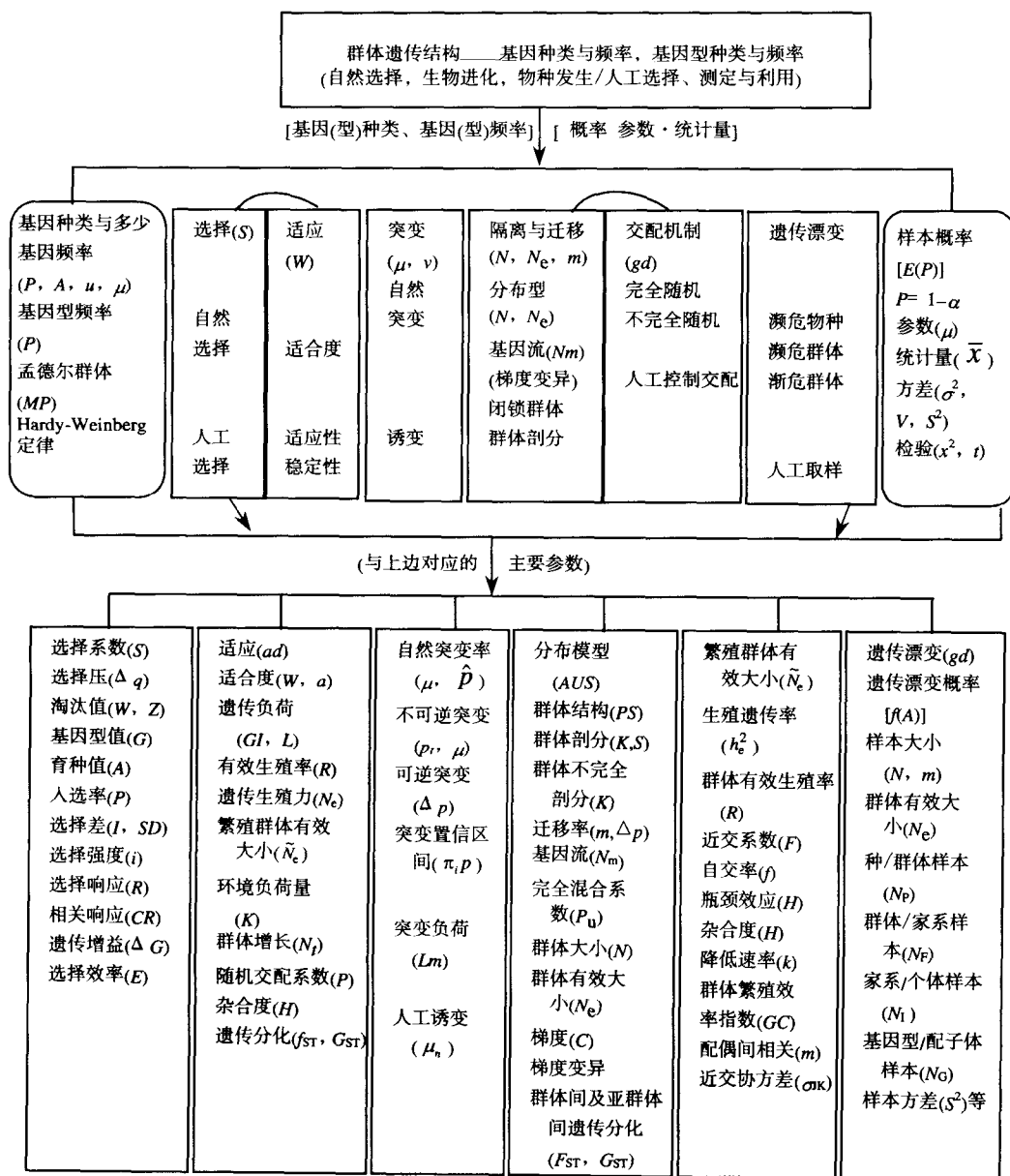


图 1.2 统计遗传学研究中影响群体遗传结构的相关因子简化示意图

人类遗传学和动植物遗传育种学的研究方法, 互为滋养, 发展成为现代遗传学含有多个分支的基础学科。与群体遗传学相关的群体生物学(population biology), 是研究生物在时间和空间上相互关系的学科, 包括生态学、分类学、群体遗传学、种质学以及其他主要研究生物相互关系的学科。1908年 H. Hardy 与 W. Weinberg 各自独立提出群体基因频率与基因型频率的变化规律, 成为研究群体遗传结构及其变化规律的理论基础, 在群体遗传学中确定了“哈特-温伯格”(H-W)基本定律。作为群体遗传学基础的数学理论, Fuahu