

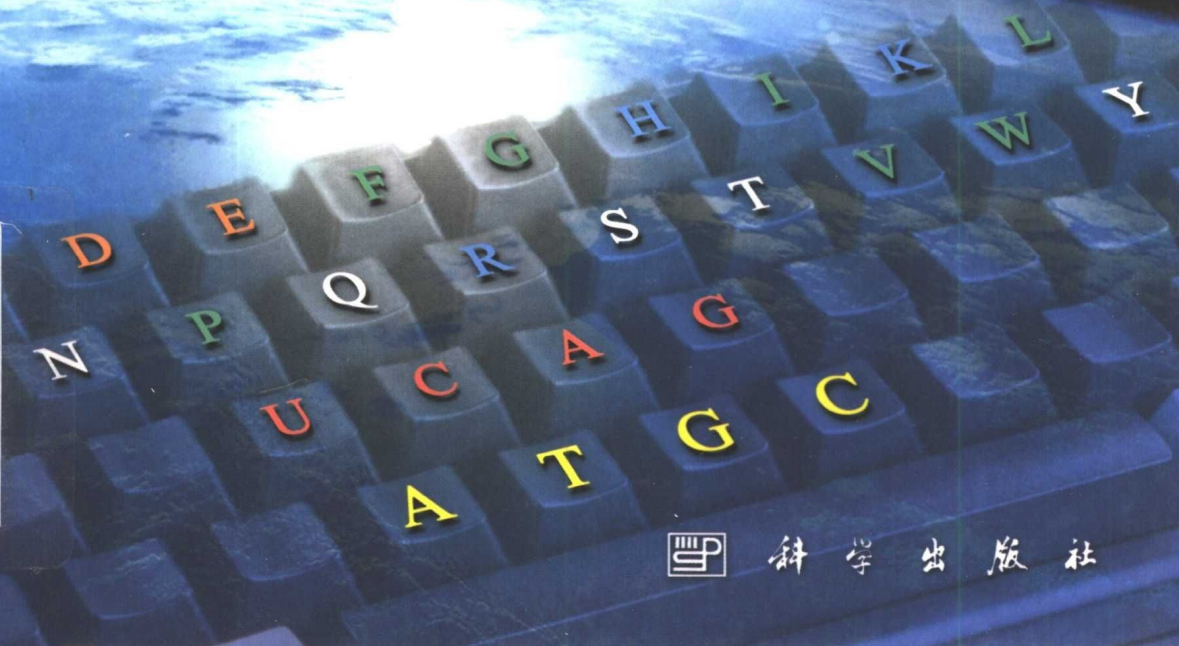
现 代 遗 传 学 丛 书

TIG 遗传命名指南

Trends in Genetics Genetic Nomenclature Guide

[英] 遗传学进展编辑部编

王金发 等 译



科学出版社

内 容 简 介

本书译自国际权威杂志《遗传学进展》编辑部编辑的《TIG 遗传命名指南》。书中信息源自国际遗传学会、国际微生物学会、各物种国际命名委员会和遗传收藏中心等权威机构和人类及各物种基因组计划网点。内容包括植物、人类及脊椎无脊椎动物、酵母与丝状真菌、原生动物及细菌等 18 种生物的基因与等位基因、基因型、表现型和蛋白质产物和突变体符号,质粒、转座子和叶绿体、线粒体基因符号,转基因生物书写方式、染色体及其畸变符号,品系、菌株和噬菌体命名,以及各物种命名的信息来源、网点、有关出版物和参考文献。

本书是一本简明、实用、极具指导性的工具书,可供广大从事生物学、医学、农学的教学与科研人员、大学生和研究生参考与使用。

This edition is translated from the English language edition of the *Trends in Genetics Genetic Nomenclature Guide* by Elsevier Science Ltd. © 1998 by Elsevier Science Ltd. All rights reserved.

图字:01-1999-3435

图书在版编目(CIP)数据

TIG 遗传命名指南/(英)遗传学进展编辑部编;王金发等译. —北京:科学出版社,2000.8

(现代遗传学丛书/李汝祺、谈家桢主编)

书名原文:Trends in Genetics Genetic Nomenclature Guide

ISBN 7-03-008118-8

I. T… II. ① 遗… ② 王 III. 生物-遗传-命名法-指南 IV. Q311

中国版本图书馆 CIP 数据核字(2000)第 04262 号

科学出版社 出版

北京东黄城根北街 16 号

邮政编码:100717

科地亚印刷厂印刷

科学出版社发行 各地新华书店经销

*

2000 年 8 月第 一 版 开本:787×1092 1/16

2000 年 8 月第一次印刷 印张:6

印数:1—3 000 字数:124 000

定价: 18.00 元

(如有印装质量问题,我社负责调换(新欣))

本指南译者名单

翻译:(按贡献排序)

王金发 陈中健 杨 琳

王宏斌 何海琼 何小玲

易厚富 蔡小钿

校订:王金发

《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》,此书出版后曾经过修订再版,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童克中、刘良式、盛志廉、陈瑶生和孟金陵等教授分别撰写的《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。今年内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》、顾万春教授主编的《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作出应有的贡献。

李家瑞
2000/4/19

译者的话

生物学家研究的对象是宏观和微观生物界的个体及群体,包括人类自己。五彩缤纷的生命现象的本质是基因活动的结果,即基因型决定的,而表现出来的表现型(包括生理的和生化的表型)就是基因的表(现)型。

遗传学的发展突飞猛进,特别是基因工程技术的建立,使人们对遗传现象和本质的认识越来越深入。由于研究手段的更新,新技术的创立,遗传命名不再仅是对基因、基因型、以及遗传表型的命名。一些新的研究遗传的分子工具,如 RFLP、EST、STS,以及用这些工具发现的基因和基因座的命名都需要有相应的命名规则,各方面的专家根据自己的研究领域,对不同生物遗传命名规则不断进行整理、总结和创新。生物学研究的科学家们必须掌握的基本知识,不仅仅是遗传学家,生物学科各方面的研究人员都应懂得遗传命名规则,这不仅有利于学术交流,更有利于我们自己的研究工作。

在长期的教学和科学研究中,发现我们的学生(包括研究生)在这方面存在很大的不足,同时也发现有部分研究人员对遗传命名规则不甚明了,因而在论文中常常出现一些不应有的错误,甚至是笑话。

由《遗传学进展》编辑部编辑的《遗传命名指南和网络信息》(Genetic Nomenclature Guide with information on websites),简称《TIG 遗传命名指南》(TIG Genetic Nomenclature Guide)的 1998 年版本总结了当今遗传命名的重要规则,可为科学研究中常用的模式生物提供标准化的遗传命名指导,同时该命名指南还收集了大量相关的网络信息,因此是生物学家必备的一本案头书和信息资源库。经编者同意,我们将此命名规则译成中文,以便让更多的中国学者了解和掌握现行的遗传命名规则。

在翻译过程中,我们参照了该指南 1995 年版本的内容,并对原文中个别笔误进行了修改。译文的格式与原文有所不同,初衷是便于中国学者使用。参与本指南翻译工作的主要是研究生,组织研究生集体翻译,目的是让他们有一个系统学习遗传命名规则的机会。

推动本指南翻译出版的是科学出版社、中国科学院遗传研究所童克中先生、中山大学刘良式先生。他们在出版《基因及其表达》(童克中著)、《植物分子遗传学》(刘良式主编)等现代遗传学丛书的分册中,努力进行遗传命名标准化的尝试,并向我们提供了该指南的两个版本,促使我们萌生翻译之念,在翻译过程中给予了诸多良好建议,在此向他们致以诚挚的谢意。

感谢童克中先生为现代遗传学丛书的封面提出了具有学科特色的、激发思维和想象力的、赋有时代感的创意。

由于译校者水平有限,加上时间仓促,不妥之处在所难免,欢迎读者批评指正。

王金发

1999-12-于中山大学

前 言

编辑出版《TIG 遗传命名指南》(TIG Genetic Nomenclature Guide)的目的是总结遗传命名的重要规则,同时为科学研究中常用的模式生物提供遗传命名指导。生物学家为了进行正确的学术交流,不仅要熟悉自己所研究生物的遗传命名规则,也应对其他生物的命名规则有所了解。1998 年的版本在第一版的基础上进行了更新和拓展,为遗传家与发育生物学家提供了研究工作中常用的 18 种模式生物的命名规则和指南。有些生物的命名规则最近进行了修订,由于尚未定论,这一部分内容仅作为一个暂时规则,请读者留意相关学会发布的最新公告。为使读者能了解命名规则的最新进展,命名指南的所有章节后均附有至少一个撰稿人的通讯地址,读者可以直接与他们联系。

由于不可能对遗传学家们研究的所有生物进行介绍,因此选择了一些典型的模式生物,例如植物中的拟南芥(*Arabidopsis*)、玉米(maize)、豌豆(pea)、衣藻(*Chlamydomonas*),脊椎动物中的人(human)、小鼠(mouse)、大鼠(rat)、鸡(chick),无脊椎动物中的果蝇(*Drosophila*)、秀丽新小杆线虫(*Caenorhabditis elegans*),酵母中的裂殖酵母(*Schizosaccharomyces pombe*)、酿酒酵母(*Saccharomyces cerevisiae*),丝状真菌中的构巢曲霉(*Aspergillus nidulans*)、链孢霉(*Neurospora crassa*),原生动物门的盘基网柄菌属盘基亚目(*Dictyostelium discoideum*),细菌中重点介绍大肠杆菌(*Escherichia coli*)和枯草杆菌(*Bacillus subtilis*)。其他几种被研究的热门生物,如非洲爪蟾(*Xenopus*)、犬(dog)等的命名规则正在制定,将编辑在以后的版本中。

每种生物的命名条目分为三部分:① 概要:简要介绍有关基因座、等位基因、蛋白质和表型的基本规则;② 命名细则:对概要中的基本规则进行拓展叙述以及提供诸如细胞遗传命名规则、菌株命名规则;③ 资料来源:提供命名规则和基因组计划的相关信息、主要网点和文献,并在每个主标题下给出一个关键网点。

命名指南采用表格形式便于读者快速查寻(命名规则列在左边,相应的例子列在右边),同时关键词用黑体标出便于浏览相关信息。每种生物的命名规则分成若干个不同的部分进行分述,并按相同的顺序排列。对于某些生物,由于某些部分与之无关或尚未定论,因此缺省该部分内容。同时注意命名规则的广泛性以及必要的交叉参考信息。

编委会尽可能使命名指南中的信息正确和完整,欢迎提出批评和建议,请读者将意见寄至 TIG 编辑部(tig@elsevier.co.uk)。

Rachel Wood

Assistant Editor *Trends in Genetics*

目 录

细菌 (bacteria)	(1)
原生动物 (protozoan)	(7)
网柄菌 (<i>Dictyostelium discoideum</i>)	(7)
酵母 (yeast)	(10)
裂殖酵母 (<i>Schizosaccharomyces pombe</i>)	(10)
酿酒酵母 (<i>Saccharomyces cerevisiae</i>)	(14)
丝状真菌 (filamentous fungi)	(18)
构巢曲霉 (<i>Aspergillus nidulans</i>)	(18)
链孢霉 (<i>Neurospora crassa</i>)	(21)
植物 (plants)	(24)
已测序植物基因的标准命名法	(24)
莱哈衣藻 (<i>Chlamydomonas reinhardtii</i>)	(27)
玉米 (maize)	(30)
豌豆 (pea)	(34)
拟南芥 (<i>Arabidopsis thaliana</i>)	(37)
无脊椎动物 (invertebrates)	(40)
秀丽新小杆线虫 (<i>Caenorhabditis elegans</i>)	(40)
果蝇 (<i>Drosophila melanogaster</i>)	(43)
脊椎动物 (vertebrates)	(49)
斑马鱼 (zebrafish)	(49)
鸡 (chick)	(52)
大鼠 (rat)	(57)
小鼠 (mouse)	(61)
人 (human)	(67)
索引 (index)	(72)
译者增编附录: 网址索引	(75)

细菌(bacteria)

E. coli Genetic Stock Center(CGSC)

<http://cgsc.biology.yale.edu>

概要

基因(基因座)

基因符号由三个小写斜体字母组成。具有相同表型的不同基因座突变用斜体大写字母后缀相区别,如:*uvrA*、*uvrB*(若有出入,参照细则部分)。

等位基因

等位基因用紧随基因座名称后的一系列特定的数字表示,如:*lacY1*、*araA1*、*araA2*。

如果发生突变的精确基因座不能确定,并且等位基因的名称只有三个字母,则基因座字母由连字符代替,如:*lac-23*、*eda-1*。

在一类特定的基因座中,等位基因的编号是不同的,如:*lacY1*、*lacY2*(注:当*lac-1*被定名为*lacY1*后,不能再将它定名为*lacZ1*。大肠杆菌等位基因的命名在*E. coli*遗传收藏中心都有注册)。

蛋白质

蛋白质产物的名称与对应的基因或等位基因相同,但不用斜体,且首字母大写,如:*UvrA*、*UvrB*。

表型

表型的命名由三个正体字母组成,且首字母大写,如:*Ara*。

(a) 野生型和突变表型分别在名称后面加上标“+”和“-”,如*His⁺*、*Ara⁻*;(b) 可以用小写或大写字母的上标对链霉素抗性等表型进行更细致的描述,如:*Str^s*或*Str^r*。

命名细则

基因

命名基因:通常用反映基因的某些性状(如突变表型)的缩写符号命名基因(最早由Demerec^[1]等提出)。

在基因符号后加上适当的后缀表示**基因座**。建议用后缀*p*和*o*表示**启动子**(promoter)和**操纵子**(operator)位点{参考文献2~4},**终止子**(terminators)及其他位点的命名也应

遵循这样的原则,如:*lacZp*、*lacAt*、*lacZo*。然而,这个系统并没有被广泛接受,主要原因是担心同质粒的命名系统混淆(见下文)。

枯草芽孢杆菌(*Bacillus subtilis*)中许多通过阻遏孢子形成突变而确定的基因座的命名与 Demerec 等^[1]提出的命名规则{详见参考文献 5}存在很大差别。其中包括:①在 0 期阻遏孢子形成的枯草芽孢杆菌突变,如:*spo0A*、*spo0B*;②在 II 期阻遏孢子形成的枯草芽孢杆菌突变,如:*spoII A*、*spoII B* 等。随后发现许多这样的基因座都包含多个基因(通常形成一个操纵子),因此在基因座后加上一个字母表示基因,如:*spoII AA*、*spoII AB*、*spoII AC*。

有很多实例(如 *pur* 基因座的命名)都表明,枯草芽孢杆菌的基因命名不仅逐渐与大肠杆菌基因的命名相一致(详见参考文献 5),而且历来都是以大肠杆菌与沙门氏菌中相应基因的命名作为标准。

未定性的可读框(ORFs):Kenneth Rudd 提出了对 *E. coli* 和 *Salmonella* 中 ORFs 的命名规则。该命名规则不仅被用于 SWISS-PROT,并且被美国微生物学会的出版物所采用。如:染色体 0~10 分钟区段的 ORF 命名规则为 $y_a[a \sim j]A \sim Z$,例如 *yabK*;染色体 10~20 分钟区段 ORF 的命名规则为 $y_b[a \sim j]A \sim Z$,例如 *ybaC*(如果在任何分钟内的命名用完了 26 个字母,则要进入下一个 10 分钟开始新的循环,如:*yaaZ* 之后是 *ykaA*、*ykaB*。

从对基因座的遗传和表型描述过渡到分子描述时偶尔会出现问题,尤其是在对编码次级代谢物的基因簇进行命名时。*Streptomyces coelicolor* 中放线菌紫素合成基因的命名就反映了该系统有关实验的发展史,如:*actII - ORF4*。

等位基因

本命名指南在对等位基因进行命名时,尽量避免使用上标和下标。只有在表示**野生型等位基因**时,才用上标‘+’,如:*ara⁺*或 *ara A⁺*。

突变体参照各种属命名,不用‘-’号,如:*ara* 或 *araA*。

突变的类型可以在等位基因后的括号内用适当的非斜体缩写表示,如:组成型突变(Con)、温度敏感型突变(Ts)、冷敏感型突变(Cs)。

抑制基因和抑制的突变体命名要在括号中表明其性质:①用 AS 表示琥珀 UAG 抑制基因,如:*serU67(AS)*;②用 OS 表示赭石 UAA 抑制基因,如:*lysT46(OS)*;③用 Am 表示琥珀抑制基因的等位基因,如:*lacZ82(Am)*。

当抑制基因的性质未知时,常用 sup 对其命名。上述提到的抑制基因原先也是用 sup 或 Su 命名的,后来发现它们是 tRNA 基因,才改用相应的系统命名。

下标在某些情况下可以使用:①区别不同物种或株系中同名的基因,如:*his A_{E. coli}*;②区别多重启动子等不同的元件,如:*glnAp₁*、*glnAp₂*。

结构突变

基因组的缺失、融合、转座和倒位分别用 Δ 、 Φ 、TP 和 IN 表示,后接所影响的基因或区域,然后是突变的编号。如果重排延伸到基因外,则在括号中标明受到影响的区域。

对于含有多个基因座的**缺失**,只给一个特定的编号,对于基因座内多个位点的缺失也

只给一个特定的编号,如: $\Delta trpE5$ 是 *trp*-5 突变的独特编号, $\Delta(aroP-aceF)73$ 是特定的第 73 号缺失突变。

倒位是很少见的,用 IN(区域)表示,其后跟特定的编号,如:IN(*rrnD-rrnE*)1。

转座出现的概率很少,下面的例子是转座命名的一种方法:TP(*lacI-purE*)33 表示 *lacI* 和 *purE* 间的区域(包括这两个基因)被插入染色体的某个位点(这里指 *proB* 和 *proA* 之间)。

由于大多数的转座都涉及 F' 因子的插入,因此可以用另一种方法来命名,该方法用 TOR 加编号表示。TOR 表示位于 Hfr 上的转座位点的起点,转座的编号是特定的,用于表示某 DNA 片段转座的方向和位点,如:TOR13。

融合用 Φ 表示,如: $\Phi(ara-lac)95$ 表示 *ara* 和 *lac* 之间的一种融合。

插入的命名是在基因座名后加上双冒号,然后是插入的基因或元件的符号,如:*aroA273::Tn 10*(注意:有时也用双冒号代替 Φ 表示体外构建的融合体。所有的缺失、倒位和转座的等位基因编号都在收藏中心注册,从而避免重复使用。对于复杂的转座元件的插入,如多重插入的命名,以及如何描述元件的方向,详见参考文献 6。

如果插入或融合所涉及的基因被缩短,则用“'”号表示,如: $\Phi(araB'-lacZ)96$ 。

其他来源的基因整合到细菌基因组中,这些基因的命名与其他插入物的命名规则相同,而且所用的符号应与原来的一致,如:*lacZ::gfp*。

菌株基因型

菌株的完整基因型包括菌株的交配型,性别以及所有的突变(用空格或逗号相隔),并在括号中对附属元件(如质粒或原噬菌体)进行遗传说明。如果附属元件本身是一个具有基因型的外源元件,且基因型必须被标出,则染色体与外源遗传因子之间用斜线隔开,如:(a) F^- , λ^- *thyA748::Tn10 rph-1 deo-77* 代表一个 F^- 菌株;(b) *lacZ53(Am) galK35* λ^- *pyrD34 trp-90::Tn 5 rpsL125(strR) malT1* (λ -resistant) *mtlA2 thr-1 valR119/F106-3* (*aroA273::Tn10*) 是一个 F' 菌株的基因型。斜线后的 *F106-3* 表示 F' 因子,括号中的 *aroA* 等位基因表示 F' 因子携带突变。

菌株命名

通过突变或分离得到的新菌株必须给一个新的命名,命名符号由一个或多个非斜体的字母(一般反映该菌株来源的研究者或实验室)和一个系列号组成,如:C600(注:为了避免重复,菌株名的前缀应清楚地表示出对应的收藏中心)。

转座子

细菌的转座元件包括 IS 元件(简单插入序列,通常小于 2kb,只含有与插入功能有关的基因)、转座子(相对 IS 元件而言,序列较长,结构较复杂)和附加体(包括 *E. coli* 中可以整合或独立存在的 λ 噬菌体和 F 质粒)。

IS 元件和转座子的命名^[6]分别用符号 IS 和 Tn 表示,其后为斜体的编号,如:IS 3、Tn9。对 Tn 元件有一个专门的命名注册处^[7],所有新的 Tn 元件都必须经注册处核实。

噬菌体

噬菌体基因的命名用1~3个斜体的字母,大写或小写均可,如:*N*、*cI*、*int*。

当几个基因突变产生类似表型时,可加上罗马数字编号,如:*cI*、*cII*、*cIII*;或者,其基因产物可用gp加上基因名称表示,如:gp43;或是在基因名称后加上“蛋白”二字,如:*cI*蛋白。

表型与基因产物的命名不用斜体,但首字母要大写,如:*N*、*Int*。

插入噬菌体基因组的宿主DNA命名时要写在方括号中,但一些最常见的例子并没有遵循这个原则,如:*λgal*、*λbio*。研究中常用的噬菌体(如:*λ*噬菌体、T4噬菌体和T7噬菌体)的基因符号在文献8中均有列出。溶源性噬菌体在细菌染色体中的附着部位通常命名为*att*位点,其后是所对应噬菌体的名称,如:*attλ*、*attP4*、*attHK022*。

质粒

新发现或改造的质粒命名由小写的前缀“p”加上非斜体字母和编号组成,如:pBR322。

对天然存在的质粒仍使用原有的命名,如:Col E1、F、R100、SCP1等。

插入质粒的基因或质粒突变体的命名根据Demerec等^[1]的命名规则进行,并用方括号或圆括号括起,前缀表示质粒的F型,如:F128-20[*proBA*⁺, *lacI*p-4000(*lacQ*)Δ*lacZ*(*M15*)],此例中*proBA*⁺是一种缩写,表示插入序列包括*proB*和*proA*基因座,并且从该处一直扩展到*lac*基因座。

质粒中的缺失和其他类型重排的命名与细菌基因组中的缺失和重排的命名规则相同,如:缺失了基因*cad*与*asa*的质粒命名为pI258Δ(*cad-asa*)7。参考文献7中有对各种不同类型质粒的命名以及对重组质粒命名的建议。

资料来源

命名信息

细菌遗传命名规则首先由Demerec等^[1]提出并得到一致认同。上述的命名概要是以Demerec命名系统细则为基础,并从后来陆续发表的文章以及美国微生物学会(American Society for Microbiology)出版的一些杂志(例如*The journal of Bacteriology*, *Molecular and Cellular Biology*)的征稿细则中选取了一些建议。本命名手册也对一些与命名细则不同的情况进行了介绍。

对于革兰氏阴性菌*E. coli* K-12菌株的命名细则可以参见参考文献2~7,而对于革兰氏阳性菌*B. subtilis*,可以参见参考文献5。这些参考文献中引用的一些书籍也包括其他一些细菌,如*S. typhimurium*、*Bacillus*和*Staphylococcus*的某些种、*Streptococcus*、*Clostridium*和*Streptomyces*等的命名。

网点

CGSC(*E. coli* Genetic Stock Center,大肠杆菌遗传收藏中心)提供有关细菌基因型、

菌株、基因名称、基因特性、蛋白质和 RNA 基因产物、突变体和连锁图谱的资料。此外, CGSC 还是官方正式负责等位基因、菌株名前缀、缺失体和其他结构突变体及 F' 因子的命名注册机构。CGSC 的网址是:

CGSC: <http://cgsc.biology.yale.edu>

TIGR 微生物数据库(The TIGR Microbial Database)提供进入全世界所有细菌基因组测序计划的途径,包括已完成的和正在进行中的。网址是:

TIGR: <http://www.tigr.org/tdb/mdb/mdb.html>

Sanger 中心(The Sanger Centre)提供许多细菌基因组测序计划的信息。网址是:

Sanger: <http://www.sanger.ac.uk/Projects/Microbes>。

EBI 和 NCBI 均提供具有注释的全基因组序列数据库信息。NCBI 同时还是 GenBank DNA 序列数据库中所有生物的分类浏览器。网址是:

EBI: <http://www.ebi.ac.uk>

NCBI: <http://www.ncbi.nlm.nih.gov>

Colibri 和 SubtiList 采用相同的图表界面显示 *E. coli* 和 *B. subtilis* 基因组数据,其中包括基因图谱、DNA 和蛋白质序列。网址是:

Colibri: <http://www.pasteur.fr/Bio/Colibri.html>

SubtiList: <http://bioweb.pasteur.fr/GeneoList/SubtiList>

ExPASy 分子生物学服务器(ExPASy Molecular Biology Server) 位于瑞士生物信息学院,提供一份酶的命名手册——ENZYME。网址是:

ExPASy: <http://www.expasy.ch>

MycDB 是一个综合性的分枝杆菌数据库(可以从 AGIS 中获得,即农业基因组信息服务部,the Agricultural Genome Information Server)。网址是:

MycDB: <http://www.kiev.physchem.kth.se>

AGIS: <http://probe.nalusda.gov:8000/all dbs.html>

基因组计划

12 个细菌基因组(以及 4 个古细菌基因组)已完全测序,例如: *Escherichia coli*、*Haemophilus influenzae*、*Helicobacter pylori* 和 *Methanococcus jannaschii* (详见 TIGR 和 Sanger 中心的网点)。

收藏中心

***E. coli* 遗传收藏中心**:主要是负责网上查询和索取菌株,详见 CGSC 网点。

***Salmonella* 遗传收藏中心**:联系人是 Kenneth Sanderson,其电子邮箱是:

kesander@acs.ucalgary.ca

***Bacillus* 遗传收藏中心**:联系人是 Daniel Zeigler,其电子邮箱是:

dzeigler@magnus.acs.ohio-state.edu

撰稿人

Mary Berlyn(mary.berlyn@yale.edu): *E. coli* Genetic Stock Center, Department of

Biology, Yale University, USA.

Kenn Rudd(rudd@ecogene.med.miami.edu): Department of Biochemistry and Molecular Biology, University of Miami, USA.

Keith Chater(keith.chater@bbsrc.ac.uk): John Innes Centre, Norwich, UK.

参 考 文 献

1. Demerec, M. et al. A proposal for a uniform nomenclature in bacterial genetics. *Genetics*, 1966. 54: 61~76
2. Berlyn, M.K.B. Linkage map of *Escherichia coli* K-12, edition 9; the traditional map. *Microbiol. Mol. Biol. Rev.*, 1998. 62: 814~984
3. Rudd, K.E. Linkage map of *Escherichia coli* K-12, edition 9; the physical map. *Microbiol. Mol. Biol. Rev.*, 1998. 62: 985~1019
4. Berlyn, M.K.B. et al. Integrated linkage map of *Escherichia coli* K-12 (Edn 9) in *Escherichia coli* and *Salmonella*: *Cellular and Molecular Biology*, (2nd edn) (Neidhart, R. et al., eds), ASM Press, 1996. 1715~1902
5. Agnostopoulos, C. et al. The genetic map of *Bacillus subtilis* in *Bacillus subtilis* and *Other Gram-Positive Bacteria* (Sonenshein, A.L., Hoch, J.A. and Losick, R., eds), American Society for Microbiology, 1993. 425~461
6. Campbell, A. et al. Nomenclature of transposable elements in prokaryotes. *Gene*, 1979. 5: 197~206
7. 质粒咨询中心(The Plasmid Reference Center)作为命名注册处可以避免出现重名的现象,该中心建立了无争议的质粒前缀名称的档案。新的前缀以及 IS 和 Tn 序号系列的分配均可以通过书面形式申请注册,联系方式:Plasmid Reference Center, Department of Microbiology and Immunology, 5402, Stanford University School of Medicine, Stanford, CA 94305-2499, USA
Tel: (+) 415 723 1772 Fax: (+) 415 725 6757
8. O'Brien, S.J. (ed.) *Genetic Maps* (6th edn), Cold Spring Harbor Laboratory Press
9. Novick, R.P. et al. Uniform nomenclature for plasmids; a proposal. *Bacteriol. Rev.*, 1976. 40: 168~189

网柄菌 (*Dictyostelium discoideum*)

Dictyostelium discoideum Database(DictyDB)

<http://www-biology.ucsd.edu/others/dsmith/dictydb.html>

概要

基因(基因座)

基因符号由三个小写斜体字母组成。突变后具有相同表型或序列相近的不同基因座可用大写斜体字母后缀加以区别,如:*acpA*、*acpB*、*carC*、*dscA*、*lipA*(注:目前仍然存在与 Demerec^[1]命名系统不一致的基因命名符号,如 *mrpC1*。为了避免与等位基因命名混淆,建议统一采用 Demerec 的等位基因命名系统,详见“等位基因”)。

等位基因

等位基因用不同的系列号进行命名,如:*aggA1*、*aggA2*。

发生突变的基因座在确定之前,以连字符代替基因座名称中的斜体大写字母,但仍保留系列号,如:*agg-3*。

突变体与野生型等位基因符号分别用“-”上标和“+”上标区分,如:*agg⁻*、*agg⁺*。

目前还没有建立关于区分显性等位基因与隐性等位基因的命名系统。

蛋白质

蛋白质采用相应的基因符号,不用斜体,但首字母大写,如:*AcpA*、*AcpB*、*CarC*、*DscA*、*LipA*、*MrpC1*。偶有蛋白质的命名不遵循这种规则的情况,如:由 *rgaA* 基因编码的蛋白质命名为 RasGAP1,而由 *piaA* 基因编码的蛋白质命名为 Pianissimo(按照 Demerec 命名系统,应避免这种情况的发生)。

表型

表型用正体全称表示,如: finger、slugger、rapidly developing(偶尔也用符号表示,如 *agg⁻*)。

命名细则

基因

命名基因:基因符号用反映基因某些特性(例如突变表型或蛋白质产物)的缩写表示

(按 Demerec^[1]规则)。

未定性的**可读框**(ORF)与**cDNA**基因在命名时分别加上前缀 ORF 和 cDNA,如: ORF1010、cDNA1010。

染色体

6个连锁群已经确定,它们分属的6个染色体分别用正体阿拉伯数字表示,即1到6(注:由于尚不清楚第五连锁群是否独立存在,因此原来的第七连锁群命名为染色体5号,以避免染色体序号不连续。预计不久,根据 HAPPY 的作图结果^[2]会弄清这一问题),网柄菌属没有性染色体,所有染色体均有一个假定的近末端着丝粒。

质粒

盘基网柄菌属中存在大量的**天然质粒**,并衍生了许多转化载体。质粒名称由表示属和种(Dd表示 *D. discoideum*; Dp表示 *D. purpureum*)的字母作为前缀,随后是表示质粒的小写字母 p 和阿拉伯数字组成,采用正体书写方式,如: Ddp1、Dpp1。相同质粒的多个分离群用大写字母后缀标明,唯一的例外是 pDG1。

菌株

每种菌株均有一个特定的名称,一般由两到三个大写字母(表示分离出该菌株的实验室)以及系列号组成,如: HM44、NP12、HTY217。

资料来源

命名信息

目前尚未成立命名委员会,但20年前网柄菌属研究学会正式同意使用 Demerec^[1]命名系统。所有在已发表图谱中定位的基因座均按这一命名系统重新命名。虽然网柄菌属研究学会并未正式讨论过蛋白质的命名方式,但仍倾向于采用 Demerec 命名系统。

网点

DictyDB: 可通过该网点查询研究人员通讯录、cDNA 基因、REMI 基因、遗传基因座和物理图谱等。网址是:

DictyDB: <http://www-biology.ucsd.edu/others/dsmith/dictydb.html>

网柄菌属万维网服务器(Dictyostelium WWW Server): 可以查询研究人员资料库、盘基网柄菌属研究人员电子邮箱通讯录、正在印刷的文章摘要、CSM 新闻简报、载体序列、基因序列、Dicty 出版物的 Franke 数据库、密码子偏爱表、实验室交流和实验方法。网址是:

WWW Server: <http://dicty.cmb.nwu.edu/dicty/dicty.html>

Tsukaba cDNA 计划: 提供克隆简介、克隆表、序列、性染色体 cDNA 文库、发育各阶段 cDNA 文库。网址是:

Tsukaba; <http://www.csm.biol.tsukuba.ac.jp/cDNAproject.html>

基因组计划

现在有三个盘基网柄菌属基因组计划:由 Angelika Noegel 在德国 Jena 发起并由 André Rosenthal 负责的基因组计划;由 Bart Barrell、Jeff William 和 Rob Kay 在欧洲发起的基因组计划和由 Adam Kuspa、Bill Loomis 和 Richard Gibbs 在美国发起的基因组计划。

撰稿人

Rob Kay (rrk@mrc-lmb.cam.ac.uk): MRC Laboratory of Molecular Biology, Cambridge, UK.

Bill Loomis (wloomis@ucsd.edu): Department of Biology, University of California, San Diego, USA.

Peter Devreotes (pnd@welchlink.welch.jhu.edu): Department of Biological Chemistry, The Johns Hopkins School of Medicine, USA.

参 考 文 献

1. Demerec, M. *et al.* A proposal for a uniform nomenclature in bacterial genetics. *Genetics*, 1966. 54, 61~76
2. Dear, P.H. and Cook P.R. Happy mapping: linkage mapping using a physical analogue of meiosis. *Nucleic Acids Res.*, 1993. 21, 13~20
3. Metz, B.A. *et al.* Identification of an endogenous plasmid in *Dictyostelium discoideum*. *EMBO J.*, 1983. 2, 515~519
4. Orii, *et al.* A new type of plasmid from a wild isolate of *Dictyostelium* species: the existence of closely situated long inverted repeats. *Nucleic Acids Res.*, 1987. 15, 1097~1107

裂殖酵母(*Schizosaccharomyces pombe*)

Schizosaccharomyces pombe biology

<http://www2.bio.uva.nl/pombe>

概要

基因(基因座)

基因符号由三个小写斜体字母与一个阿拉伯数字组成(裂殖酵母不存在全基因名称),如: *arg1*、*leu2*、*cdc25*、*rad21*。

等位基因

对于同一基因座上不同等位基因,则是在基因座符号后用连字符加上等位基因特异的后缀来命名,如: *ade6-M26*、*ade6-469*。

① 特定基因的野生型等位基因用上标“+”表示,如: *arg1*⁺;②在描述基因型时,野生型等位基因用“+”表示。基因符号不反映基因的显隐性关系。

蛋白质

蛋白质用相应基因的基因符号表示,正体书写,并且首字母大写,如: Arg1、Leu2、Cdc25、Rad21。按照酿酒酵母的命名习惯,裂殖酵母研究协会也逐步采用了“p”后缀表示蛋白的表示方法,如: Arg1p、Leu2p、Cdc25p、Rad21p。

表型

表型名称通常用相应基因名称的三个正体字母表示。

野生型和突变体 分别采用“+”和“-”上标表示,如: *arg*⁻、*arg*⁺。

命名细则

基因

多功能基因的互补群命名时在基因座符号后标注一个大写字母,如: *trp1A*。

线粒体基因符号必要时用方括号括起以便与核基因相区别,如: [*cox1*]。(注含有质粒的菌株采用类似的命名方法,详见“基因型”,但这种用法有一定的限制。)

通过克隆鉴定获得的基因的命名规则与通过突变、或是通过基因产物(但不是表型)鉴定的基因命名原则相同,或者以“*orf*”作为临时性的命名,如: *orf39*。

以 ORF 临时命名的基因最终要根据鉴定的结果按照基因的命名规则进行重新命名,