



复杂性状遗传 CC 小鼠信息资源及研究应用

孔琪, 夏霞宇, 秦川*

(中国医学科学院医学实验动物研究所, 北京协和医学院比较医学中心; 卫生部人类疾病比较医学重点实验室;
国家中医药管理局人类疾病动物模型三级实验室, 北京 100021)

【摘要】 复杂性状遗传 CC 小鼠(CC 小鼠)由不同遗传背景的小鼠品系杂交后,近交培育而成,用于研究人类复杂性状疾病和疾病的遗传差异。CC 小鼠能体现不同小鼠亚种的遗传差异,其单核苷酸多态性是传统近交系小鼠的四倍。CC 小鼠在生命科学与医学研究领域的应用越来越广泛。本文通过信息检索等方式对 CC 小鼠相关信息资源进行梳理,阐述了 CC 小鼠的起源、数据库、应用工具,以及研究领域等,以推动 CC 小鼠资源在中国得到广泛应用。

【关键词】 复杂性状遗传 CC 小鼠,小鼠品系,资源,数据库,人类疾病

【中图分类号】 R-332 **【文献标识码】** A **【文章编号】** 1671-7856(2016)07-0011-05

doi: 10.3969/j.issn.1671-7856.2016.07.002

The Collaborative Cross mice resource information and application

KONG Qi, XIA Xia-yu, QIN Chuan*

(Institute of Laboratory Animal Science, Chinese Academy of Medical Sciences & Comparative Medical Center,
Peking Union Medical College, Key Laboratory of Human Disease Comparative Medicine, Ministry of Health;
Key Laboratory of Human Diseases Animal Model, State Administration of Traditional Chinese Medicine, Beijing 100021, China)

【Abstract】 Collaborative Cross mice (CC mice) are series of inbred mice strains generated from hybrid strains of mice with different genetic background which used for human complex diseases and genetic diversity diseases studies. Genetic diversity of CC mice can reflect different mouse subspecies, the single nucleotide polymorphism is four times than traditional inbred mice. CC mice are more and more widely used in the field of life science and medical research. Based on information retrieval of CC mice, we introduced the related information resources of CC mice origin, database, application tools, and research results, to promote CC mice resources to be used widely in China.

【Key words】 Collaborative Cross mice; Mouse strain; Resource; Database; Human disease

小鼠(mice, *Mus musculus*)是研究人类疾病理想而且常用的模式动物。随着基因工程技术的发展,使得小鼠品系资源愈加丰富,在人类疾病研究中的应用也更广泛。复杂性状疾病,即多基因病,是指由多个基因共同参与,且和环境因素相互作用决定表型的遗传病,如高血压、冠心病、糖尿病、哮

喘和某些癌症等。

采用单一遗传背景的小鼠已无法有效复制人类多基因相互作用引起的复杂性状疾病。复杂性状遗传 CC 小鼠(Collaborative Cross mice, CC 小鼠)应运而生,成为模拟人类复杂性状疾病的利器,具有更丰富的性状差异,更丰富的遗传多态性,能够

[基金项目] 中央科研院所基本业务费(DWS200709, DWS201208, DWS201512); 传染病科技重大专项(2009ZX10004-503)。

[作者简介] 孔琪(1978-), 助理研究员, 博士, 研究方向: 实验动物学。E-mail: infor@cnilas.org。

[通讯作者] 秦川, 研究员, 博士生导师, 研究方向: 实验病理学。E-mail: qinchuan@pumc.edu.cn。

体现不同人群对病因敏感的差异性,可应用于精准医疗、基因功能发现、疾病模型建立和人类复杂性状疾病等方面研究。

本文通过信息检索、文献查阅、数据库搜索等方式对 CC 小鼠相关信息资源进行梳理,阐述了 CC 小鼠的概述、数据资源、应用工具,以及研究应用等,以使我国科研人员深入了解和更好的使用 CC 小鼠,支撑人类复杂性状疾病研究和精准医疗等相关领域发展。

1 概述

CC 小鼠来源于复杂性状联盟 (Complex Trait Community, CTC) 启动的一个项目。其目的是发展一种更适用于研究复杂性状或复杂病因疾病的研究工具,以模拟人群基因多样性进行复杂性状相关研究^[1]。包括数百种不同基因型的小鼠品系,来自多个原始种系。截止到 2016 年 3 月美国北卡大学 (University of North Carolina, UNC) 共对外提供 72 个 CC 小鼠品系^[2]。

研究人员采用 CC 小鼠进行正常肝组织的表达数量性状基因座 (expression Quantitative Trait Loci, *eQTL*)^[3,4], 过敏性鼻炎 (allergic airway inflammation)^[5], 黑色素瘤易感性^[6], 流行性感与埃博拉等病毒易感性^[7,8] 基因型与性状相关研究。CC 小鼠名声大噪源于 Science 杂志在线发表的一篇《Host genetic diversity enables Ebola hemorrhagic fever pathogenesis and resistance》文章^[8]。作者在 47 个 CC 小鼠品系中测试了埃博拉病毒引发的宿主应答。病毒感染对不同品系小鼠造成的影响不同,揭示了人类对埃博拉易感性不同源于其存在的遗传学差异。与此同时,以 CC 小鼠为基础的数据库以及研究工具也逐步建立起来,极大程度地丰富了科研人员的研究资源。

2 数据资源

2.1 UNC CC 小鼠数据库

UNC 数据库介绍了 CC 小鼠项目的发展历程,并提供了当前发布的 72 个 CC 小鼠品系的相关数据,包括名称、八种初始品系的基因组组成、杂合比例、健康状况、饲养环境、图片等基础信息,毛色、眼睛色、繁殖能力等生理数据,以及相应的基因型、单倍型等多种数据,供研究人员下载使用。

UNC 同时也提供了一系列的可视化工具对 CC 小鼠的基因型等数据进行展示。通过 UNC 的 CC 小鼠浏览器 2.0 (<http://csbio.unc.edu/CCstatus/>

<http://csbio.unc.edu/gecco/>), 研究人员可以获取其基因型、单倍型及系统进化等信息。通过 CC 小鼠表达谱浏览器 (<http://csbio.unc.edu/gecco/>), 研究人员可直观了解不同 CC 小鼠中基因表达水平异同。除发布的 CC 小鼠品系信息,UNC 还整合了部分其它相关数据。

UNC 系统遗传学中心负责 CC 小鼠项目后期的繁殖,保种与信息发布。其数据库包含了最全面的 CC 小鼠品系相关信息,并通过官网 (<http://csbio.unc.edu/CCstatus/index.py>) 提供给相关科研人员。

2.2 Sanger 小鼠基因组数据库

CC 小鼠作为研究复杂性状疾病分子机制的工具,通过其基因图谱定位可以找出决定性状的关键因素。而根据 CC 小鼠的来源,可以将其基因组近似表征为八种纯系小鼠基因组的嵌合体。因此,这八种纯系小鼠的基因组数据是所有后续相关研究的基础。

这些基因组数据均可以通过 Sanger 小鼠基因组项目获得。自 2011 年开始,Sanger 小鼠基因组数据库 (<http://www.sanger.ac.uk/resources/mouse/genomes/>) 便发布了所有八种纯系小鼠以及其它九种常见的实验小鼠品系的高通量全基因组重测序数据^[9]。与此同时发布的还有各实验小鼠品系与小鼠参考基因组 (C57BL/6J, CRCm38 版) 的比对及变异,包括单核苷酸多态性 (Single nucleotide polymorphism, SNP)、短片段插入/删除 (Insertion/Deletion, Indel), 以及结构变异 (Structural variation) 等数据信息。其中,SNP 及短片段 Indel 均已通过 Ensembl Variant Effect Predictor 加以注释^[10], 研究人员可通过基因组位置、基因名称、小鼠品系、变异类型及预测功能等关键词进行相关信息的检索及可视化浏览。截止 2016 年 3 月,数据库中所包含的实验小鼠品系数已增加至 36 种,其中有 19 种 (包括构建 CC 小鼠所采用的八种) 已整合进 UNC 数据库中。

2.3 MDA 小鼠基因型数据库

美国 Jackson 实验室建立的小鼠 MDA 基因型数据库 (<http://cgd.jax.org/datasets/diversityarray.shtml>) 已收录了包括 UNC CC 小鼠在内的 1,902 份实验小鼠的原始及处理后芯片数据,可供研究人员下载使用。

小鼠多样性芯片 (Mouse Diversity Array, MDA) 是最早针对实验小鼠发明的高密度基因分型芯片^[11]。MDA 芯片基于 Affymetrix 平台,包含了 60 余万种实验小鼠全基因组范围的 SNP,及另外 90 余

万种用于检测拷贝数变异的功能性保守标签信息^[11]。虽然不如全基因组测序精度高,MDA 小鼠分型芯片仍被广泛应用于 CC 前期(G2;F1 代)及后期(前 CC)基因分型及功能研究中^[3, 5, 7, 12]。

2.4 MUGA 及 MegaMUGA 基因型数据库

MUGA 和 MegaMUGA 是另外两种专门针对 CC 小鼠所设计的基因分型芯片,基于 Illumina Infinium 平台,主要用于确定各 CC 小鼠品系中八个纯系小鼠基因组的贡献,以及繁殖过程中不同 CC 小鼠品系间的剩余杂合率^[13]。MUGA 芯片包含 7,851 种 SNP 分型标记,其升级版 MegaMUGA(第二代小鼠基因分型芯片)则是其 10 倍。MUGA 及 MegaMUGA 芯片是 UNC 推荐使用的小鼠基因分型工具。

CC 小鼠构建过程中部分样本的 MUGA 或 MegaMUGA 基因型检测数据均已有 UNC 收录。其中,MUGA 数据共计 626 份,包括 168 份八种纯系及 F1 代的参考基因型数据,以及另 458 份其它样本检测数据;MegaMUGA 数据共计 792 份,包括 72 种成熟 CC 小鼠品系基因型数据,以及 720 份其它样本的检测数据。MUGA/MegaMUGA 数据库,以及芯片相关的分析工具,均可通过 UNC 官方网站获取。

2.5 小鼠表型数据库

美国 Jackson 实验室建立的小鼠表型数据库(Mouse Phenome Database,MPD)旨在为研究人员提供一个近交系实验小鼠基因组变异信息分类集的 Web 接口^[14]。其 Genotype 模块(<http://phenome.jax.org/db/q? rtn = snp/home>)包含了 Sanger 实验小鼠基因组及 Jax-UNC 小鼠的 MDA 变异信息,以及部分其它的结构变异数据集。

研究人员可通过基因组位置、基因名称、小鼠品系、变异类型及预测的功能等关键词进行相关信息的检索。此外,Genotype 模块还提供了不同小鼠品系的两两基因型比较,基因型数据获取等功能。

2.6 小鼠基因型重估数据库

经典近交系小鼠通常由单一品系近交培育而成,遗传相似度在 99% 以上。每个近交系小鼠个体的基因组可近似认为是由祖先株单倍型片段组成的嵌合体。Yang 等人对 100 种经典小鼠品系的 MDA 基因型数据进行了四配子检验,以界定家鼠(*M. m. domesticus*)的单倍型域^[15]。随后,Wang 等人结合这些单倍型域,以及 12 个品系的全基因组重测序数据,对 1.2×10^{10} 个位点进行了高置信度基因型重估^[16]。所有单倍型域及重估数据可以通过 UNC 的模块(<http://msub.csbio.unc.edu/>和

<http://csbio.unc.edu/imputation/>)获取。这些结果对于后续 CC 小鼠相关的数量性状遗传位点(Quantitative trait location, QTL)分析研究工作至关重要。

3 应用工具

3.1 单倍型重构工具

CC 小鼠是由多个纯系小鼠重组而来,其基因组中不同祖先株的单倍型域组成极大程度影响了后续的 QTL 分析与研究。单倍型重构是应用 CC 小鼠进行 QTL 分析研究的第一步。以 CC 小鼠基因型芯片数据为基础,研究人员通常采用 HAPPY^[17] 与 GAIN^[18] 软件包来实现。

HAPPY 与 GAIN 均以隐马尔可夫模型(Hidden markov model,HMM)为基础,并根据基因型芯片数据的观测值估算每个位点祖先株相应的概率^[19]。HAPPY 是一个 R 软件包,起源较早,包括一系列针对封闭群进行 QTL 分析的函数。GAIN 函数主要是针对 CC 小鼠所设计,通过结合 HMM 算法与系谱相关知识有效推断祖先株的概率。HAPPY 软件包及相关数据可通过 Wellcome Trust 人类遗传中心网站(<http://mus.well.ox.ac.uk/CC/>)在线获取。

所有已发布的 CC 小鼠品系的单倍型重构数据信息可以通过 CC 小鼠浏览器(<http://csbio.unc.edu/CCstatus/index.py? run = CCV>)进行浏览与下载。

3.2 QTL 定位工具

BAGPIPE(<http://valdarlab.unc.edu/software/bagpipe>)^[20] 及 BAGPHENOTYPE(<http://valdarlab.unc.edu/bagphenotype.html>)^[21] 是两种针对 CC 小鼠进行 QTL 定位研究的软件包。两个软件包均以 HAPPY 的固定效应模型(fixed-effects model)为基础,模拟加性效应与显性效应,通过八种原始株的单倍型概率矢量对表型值进行拟合。其它如实验(如:批次)及生物学(如:性别)等协变量可采用固定或随机效应模型。统计显著性程度可通过非限制性排列进行估算。

BAGPIPE 及 BAGPHENOTYPE 均以小鼠的单倍型概率矩阵为输入值,通过命令行进行操作。BAGPIPE 通常适用于群体结构未知的正态分布性状的单位点定位,而 BAGPHENOTYPE 则采用了重采样模型平均及多位点模型。此外,BAGPHENOTYPE 也可以根据广义线性模型对非正态分布性状(如:枚举、二元性状等)进行定位。

3.3 高通量测序分析工具

高通量测序读段与参考基因组的比对是测序相关研究的第一步。其保真度与有效性随着测序物种与参考基因组的基因距离增大而降低,将导致杂合样本的研究产生误差。理想的参考基因组可以尽可能地融合更多的待测样本信息,包括二倍性等,以减少误差。以此为核心,美国北卡大学的 McMillan 团队,以及美国 Jackson 实验室的 Churchill 团队先后研发了两个等位基因特异性比对流程^[22, 23]。两种方法的原理较为相似,均首先以一个参考基因组及另一个个体特异性的变异位点(相对于参考基因组)列表为输入值,构建一个改进的、重构的、二倍体参考基因组(伪基因组),测序读段继而与伪基因组进行比对。对比质量改进,以及等位基因特异性,有利于后续的分析。

4 研究应用

以“Collaborative cross mice”为关键词对 Web of ScienceTM核心合集数据库进行检索(截止 2016 年 3 月),共得到 206 条记录,其中研究类文献共计 152 篇,出版年份为 2006 年至今,呈现逐年递增趋势。其中,121 篇(79.6%)研究文献来源于美国的研究机构。

对研究方向进行统计分析表明,超过 50.0% (80 篇)CC 小鼠相关的科学研究均为基因遗传领域,其次为分子生物学(37, 24.3%)、生物技术应用微生物学(31, 21.4%),神经生物学(11, 7.2%)等。此外,CC 小鼠在细胞生物学、行为科学、微生物学、毒理学及生理学等领域均有所应用。

5 总结

CC 小鼠项目自 2004 年正式启动至今已经经历了 12 年,三个繁殖中心已建立起了百余种 CC 小鼠品系,并逐渐开始广泛应用于生命科学研究中。与此同时,相关的小鼠基因组与遗传数据也逐渐得以积累,在一定程度上推动了相关数据库与分析工具的构建。

UNC 等数据库已整合了多种 CC 小鼠相关的数据资源,供科研人员浏览和下载使用。诸如 HAPPY、GAIN 等针对 CC 小鼠而发展的分析工具也已逐步发展。对 Web of ScienceTM数据库中收录的 CC 小鼠为主题的科研论文进行统计分析也表明近年来呈现出了高速增长的态势。可以预期,未来随着人们对复杂性状疾病的日益重视,以及 CC 小鼠项目在中国医学科学院医学实验动物研究所落地,将进一步推动其在中国范围内的广泛应用。

参考文献:

- [1] Churchill GA, Airey DC, Allayee H, *et al.* The Collaborative Cross, a community resource for the genetic analysis of complex traits [J]. *Nat Genet.* 2004,36:1133 – 1137.
- [2] UNC Systems Genetics. [OL][2016 – 03 – 26]. <http://www.csbio.unc.edu/CCstatus/index.py>.
- [3] Aylor DL, Valdar W, Foulds-Mathes W, *et al.* Genetic analysis of complex traits in the emerging Collaborative Cross [J]. *Genome Res.* 2011, 21:1213 – 1222.
- [4] Weiser M, Mukherjee S, Furey TS, *et al.* Novel distal eQTL analysis demonstrates effect of population genetic architecture on detecting and interpreting associations [J]. *Genetics* 2014, 198: 879 – 893.
- [5] Kelada SNP, Carpenter DE, Aylor DL, *et al.* Integrative genetic analysis of allergic inflammation in the murine lung [J]. *Am J Respir Cell Mol Biol.* 2014, 51:436 – 445.
- [6] Ferguson B, Ram R, Handoko HY, *et al.* Melanoma susceptibility as a complex trait: genetic variation controls all stages of tumor progression [J]. *Oncogene.* doi: 10.1038/onc.2014.227.
- [7] Ferris MT, Aylor DL, Bottomly D, *et al.* Modeling host genetic regulation of influenza pathogenesis in the Collaborative Cross [J]. *PLoS Pathog.* 2013,9:e1003196.
- [8] Rasmussen AL, Okumura A, Ferris MT, *et al.* Host genetic diversity enables Ebola hemorrhagic fever pathogenesis and resistance [J]. *Science.* 2014, doi:10.1126/science.1259595
- [9] Keane TM, Goodstadt L, Danecek P, *et al.* Mouse genomic variation and its effect on phenotypes and gene regulation [J]. *Nature.* 2011, 477:289 – 294.
- [10] McLaren W, Pritchard B, Rios D, *et al.* Deriving the consequences of genomic variants with the Ensembl API and SNP effect predictor [J]. *Bioinformatics.* 2010, 26:2069 – 2070.
- [11] Yang H, Ding Y, Hutchins LN, *et al.* A customized and versatile high-density genotyping array for the mouse [J]. *Nat Methods.* 2009, 6:663 – 666.
- [12] Liu EY, Morgan AP, Chesler EJ, *et al.* High-resolution sex-specific linkage maps of the mouse reveal polarized distribution of crossovers in male germline [J]. *Genetics.* 2014, 197:91 – 106.
- [13] Collaborative Cross Consortium, The Genome Architecture of the Collaborative Cross Mouse Genetic Reference Population [J]. *Genetics.* 2012, 190, 2:389 – 402.
- [14] Maddatu TP, Grubb SC, Bult CJ, Bogue MA. Mouse Phenome Database (MPD) [J]. *Nucleic Acids Res.* 2012, Jan; 40 (Database issue):D887 – 94.
- [15] Yang H, Wang JR, Didion JP, *et al.* Subspecific origin and haplotype diversity in the laboratory mouse [J]. *Nat Genet.* 2011, 43:648 – 655.
- [16] Wang J, Moore KJ, Zhang Q, *et al.* Genome-wide compatible SNP intervals and their properties. In: *Proceedings of the first a CM international conference on bioinformatics and computational biology—bCB’10* [M]. ACM Press, New York. 2010. 43

- recordings using a miniaturised matrix electrode: a new technique for small animals [J]. *J Neurosci Methods*, 2000, 97(1):69 – 75.
- [45] Canu MH, Garnier C, Lepoutre FX, *et al.* A 3D analysis of hindlimb motion during treadmill locomotion in rats after a 14 – day episode of simulated microgravity [J]. *Behav Brain Res*, 2005, 157(2):309 – 321.
- [46] Metz GA, Whishaw IQ. Cortical and subcortical lesions impair skilled walking in the ladder rung walking test: a new task to evaluate fore- and hindlimb stepping, placing, and co-ordination [J]. *J Neurosci Methods*, 2002, 115(2):169 – 179.
- [47] Riek – Burchardt M, Henrich-Noack P, Metz GA, *et al.* Detection of chronic sensorimotor impairments in the ladder rung walking task in rats with endothelin – 1 – induced mild focal ischemia [J]. *J Neurosci Methods*, 2004, 137(2):227 – 233.
- [48] Chou KH, Lin WC, Lee PL, *et al.* Structural covariance networks of striatum subdivision in patients with Parkinson’s disease [J]. *Hum Brain Mapp*, 2015, 36(4):1567 – 1584.
- [49] 余文娟, 彭蓉. 帕金森病冻结步态的研究进展[J]. *华西医学*, 2014, 29(1):155 – 159.
- [50] Nanhoe – Mahabier w, Snijders AH, Delval A, *et al.* Split-belt locomotion in Parkinson’s disease with and without freezing of gait [J]. *Neuroscience*, 2013, 236:110 – 116.
- [51] Tang WL, McDowell K, Limsam M, *et al.* Locomotion analysis of Sprague-Dawley rats before and after injecting 6 – OHDA [J]. *Behav Brain Res*, 2010, 210(1):131 – 133.
- [52] Madete JK, Klein A, Dunnett SB, *et al.* Three-dimensional motion analysis of postural adjustments during over-ground locomotion in a rat model of Parkinson’s disease [J]. *Behav Brain Res*, 2011, 220(1):119 – 125.
- [53] 王军. 6 – 羟基多巴胺致偏侧帕金森大鼠的 CatWalk 行为学研究[D]. 深圳:南方医科大学, 2012:3 – 4.
- [54] Lakes EH, Allen KD. Gait analysis methods for rodent models of arthritic disorders: reviews and recommendations [J]. *Osteoarthritis Cartilage*, 2016, 3:1 – 13.
- [55] Chao OY, Pum ME, Li JS, *et al.* The grid-walking test: assessment of sensorimotor deficits after moderate or severe dopamine depletion by 6 – hydroxydopamine lesions in the dorsal striatum and medial forebrain bundle [J]. *Neuroscience*, 2012, 202:318 – 325.
- [56] Chao OY, Wang AL, Nikolaus S. NK 3 receptor agonism reinstates temporal order memory in the hemiparkinsonian rat [J]. *Behav. Brain Res*, 2015, 285:208 – 212.
- [57] Asakawa T, Fang H, Sugiyama K, *et al.* Animal behavioral assessments in current research of Parkinson’s disease [J]. *Neurosci Biobehav Rev*, 2016, 65:63 – 94.
- [58] 潘钰, 郑淑燕. 脊髓损伤神经功能定量评价研究进展[J]. *中国康复医学杂志*, 2014, 29(4):379 – 385.
- [59] Oliveri RS, Bello S, Biering-Sørensen F. Mesenchymal stem cells improve locomotor recovery in traumatic spinal cord injury: Systematic review with meta-analyses of rat models [J]. *Neurobiol Dis*, 2014, 62:338 – 353.
- [60] Wu M, Landry JM, Schmit BD, *et al.* Robotic resistance treadmill training improves locomotor function in human spinal cord injury: a pilot study [J]. *Arch Phys Med Rehabil*, 2012, 93(5):782 – 789.
- [61] Jin Y, Bouyer J, Shumsky JS, *et al.* Transplantation of neural progenitor cells in chronic spinal cord injury [J]. *Neuroscience*, 2016, 21(320):69 – 82.
- [62] Sun Y, Liu D, Su P, *et al.* Changes in autophagy in rats after spinal cord injury and the effect of hyperbaric oxygen on autophagy [J]. *Neurosci Lett*, 2016, 618:139 – 145.
- [63] Lee SH, Kim Y, Rhew D, *et al.* Effect of the combination of mesenchymal stromal cells and chondroitinase ABC on chronic spinal cord injury [J]. *Cytherapy*, 2015, 17(10):1374 – 1383.
- [64] Redondo – Castro E, Torres-Espín A, García-Álías G, *et al.* Quantitative assessment of locomotion and interlimb coordination in rats after different spinal cord injuries [J]. *J Neurosci Methods*, 2013, 213(2):165 – 178.
- [65] Knuttgen HG, Vogel JA, Poortmans JR, *et al.* *Biochemistry of Exercise* [M]. Boston: Human Kinetics Publishers, 1983. 63 – 75.
- [66] 田野, 高铁群. 大鼠运动性疲劳模型的建立[J]. *北京体育大学学报*, 1995, 18(4):49 – 53.

[修回日期]2016-04-18

(上接第 14 页)

- [17] Mott R, Talbot CJ, Turri MG, *et al.* A method for fine mapping quantitative trait loci in outbred animal stocks [J]. *Proc Natl Acad Sci USA*. 2000, 97:12649 – 12654.
- [18] Liu EY, Zhang Q, McMillan L, *et al.* Efficient genome ancestry inference in complex pedigrees with inbreeding [J]. *Bioinformatics*. 2010, 26:i199 – i207.
- [19] Baum, L. E.; Petrie, T. *Statistical Inference for Probabilistic Functions of Finite State Markov Chains* [J]. *The Annals of Mathematical Statistics*. 1966, 37(6):1554 – 1563.
- [20] Valdar W, Solberg LC, Gauguier D, *et al.* Genome-wide genetic association of complex traits in heterogeneous stock mice [J]. *Nat Genet*. 2006, 38:879 – 887.
- [21] Valdar W, Holmes CC, Mott R, Flint J, *et al.* Mapping in structured populations by resample model averaging [J]. *Genetics*. 2009, 182:1263 – 1277.
- [22] Huang S, Holt J, Kao C-Y, *et al.* A novel multi-alignment pipeline for high-throughput sequencing data [J]. *Database* 2014:bau057.
- [23] Munger SC, Raghupathy N, Choi K, *et al.* RNA-seq alignment to individualized genomes improves transcript abundance estimates in multiparent populations [J]. *Genetics*. 2014, 198:59 – 73.

[修回日期]2016-04-05