

DOI: 10.3724/SP.J.1005.2011.00100

# 全基因组关联研究的深度分析策略

权晟, 张学军

安徽医科大学皮肤病研究所, 合肥 230032

**摘要:** 2005 年至今, 全基因组关联研究(Genome-wide association study, GWAS)发现了大量复杂疾病/性状相关变异。近来, 科学家们关注的焦点又集中在了如何利用 GWAS 数据进行深入分析, 期待发现更多复杂疾病/性状的易感基因。一些新的策略和方法已经被尝试应用到复杂疾病/性状 GWAS 的后续研究中, 例如深入分析 GWAS 数据; 鉴定新的复杂疾病/性状易感基因/位点; 国际合作和 Meta 分析; 易感区域精细定位及测序; 多种疾病共同易感基因研究; 以及基因型填补, 基于通路的关联分析, 基因-基因、基因-环境交互作用和上位研究等。这些策略和方法的应用弥补了经典 GWAS 的一些不足之处, 进一步推动了人类对复杂疾病/性状遗传机制的认识。文章对上述研究的策略、方法以及所面临的问题和挑战进行了综述, 为读者描绘了 GWAS 后期工作的一个简要框架。

**关键词:** 全基因组关联研究; 复杂疾病; 性状; 易感基因

## Research strategies for the next step of genome-wide association study

QUAN Cheng, ZHANG Xue-Jun

*Institute of Dermatology of Anhui Medical University, Hefei 230032, China*

**Abstract:** Since 2005, genome-wide association studies (GWAS) have yielded an unprecedented number of complex diseases/traits-associated variants. Recently, scientists have focused on performing further analysis by utilizing the genome-wide genotyping data to identify more susceptibility genes of complex diseases/traits. Many strategies and methods have been applied in the following GWAS, such as screening other new susceptibility genes/loci for complex diseases/traits, international collaboration and meta-analysis, fine mapping and resequencing, studies on shared susceptibility genes in different diseases, imputation methods, pathway analysis, gene-gene and gene-environment interaction, and epistasis study and so on. The application of these strategies and methods compensates the limitation of the traditional GWAS and provides new insights into genetics basis of complex diseases/traits. We reviewed these strategies and methods, as well as their difficulty and challenge. Meanwhile, we presented a brief framework of GWAS next step to readers.

**Keywords:** genome-wide association studies; complex diseases; traits; susceptibility gene

从 2005 年至今, 全世界众多研究小组开展了大量复杂疾病/性状易感基因的全基因组关联研究(Genome-wide association study, GWAS), 发现了近 3 000 个疾病/表型相关变异, 即单核苷酸多态性

(Single nucleotide polymorphism, SNP), 同时构建了海量的基因分型数据库。如何采取有效措施充分利用现有 GWAS 数据, 提高发现易感基因/位点的效能, 明确致病的序列变异与疾病/表型的关系是新一轮

收稿日期: 2010-08-16; 修回日期: 2010-10-19

基金项目: 国家“863”计划项目(编号: 2007AA02Z161)资助

作者简介: 权晟, 男, 主治医师, 博士, 研究方向: 复杂疾病遗传学。E-mail: quan5797635@163.com

通讯作者: 张学军, 男, 主任医师, 教授, 博士, 博士生导师, 研究方向: 复杂疾病遗传学。E-mail: ayzxj@vip.sina.com

GWAS 面临的挑战。

## 1 GWAS数据深度挖掘的策略和方法

### 1.1 深入分析GWAS数据,进行多种群大样本量验证

最初的GWAS多采用两个阶段的设计:首先采用覆盖整个基因组的高通量SNP分型芯片对一批样本进行分型和分析,然后筛选出最显著差异的SNP(如 $P < 10^{-7}$ )进行验证。GWAS两阶段研究设计减少了基因分型的工作量和花费,同时通过重复实验降低了研究的假阳性率。然而这种两阶段的研究设计却存在着另一个问题:第一阶段通常在较少样本中对全基因组数量庞大的SNPs进行分析(可达100万个),因而没有足够的检验效能发现所有可能与疾病关联的SNPs。因此,为了发现更多的易感基因/位点,目前常采用的方法就是在扩大GWAS样本量的同时,适当放宽第一阶段筛选SNPs的标准、扩大验证范围。以类风湿性关节炎为例,2007年,Wellcome Trust Case Control Consortium (WTCCC)在7种重大复杂疾病的GWAS仅发现PTPN22和MHC与疾病相关联<sup>[1]</sup>。随后,Thomson等<sup>[2]</sup>在前期基础上对 $5 \times 10^{-7} < P < 10^{-5}$ 的SNPs进行验证,发现一个新的易感位点6q23(TNFAIP3)。此后,Barton等<sup>[3]</sup>进一步选择 $10^{-5} < P < 10^{-4}$ 的SNPs进行验证,又发现3个新的易感位点(10p15、12q13和22q13)。此外,在前列腺癌的验证研究中,Eeles等<sup>[4]</sup>从第一阶段GWAS的结果选出约46,000个SNPs在另外2个独立人群(分别包括7600与31000样本)进行了广泛和深入的验证,共发现7个新的前列腺癌易感位点。最近,本课题组利用其他多种疾病的GWAS数据作为参考对照增加第一阶段的检验效能,对银屑病前期GWAS数据进行深入分析,并在中国以及欧美多个人群进行验证,继发现LCE易感基因后又新发现了6个新的银屑病易感位点<sup>[5]</sup>。

随着对同一种疾病易感位点发现数量的增多,进一步搜寻其余效应相对微弱变异的难度不断增大,因此如何提高从GWAS数据中甄别真正疾病关联变异是一个比较棘手的问题。虽然验证人群不断增多,但每个独立研究群体的种族、表型界定多不一致,如何恰当处理各验证群体间的异质性,也是目前一个亟待解决的问题。

### 1.2 国际合作和Meta分析

目前的GWAS样本量多为1000病例/1000对照左右,发现的疾病相关SNPs基本上都属于常见(MAF>0.2)且具有中度效力(OR=1.2)的变异;而对于频率低(MAF<0.05)、效力弱(OR值接近1)的SNPs,由于统计把握度较低很难被检出,因此增加GWAS样本量是提高检验效能最直接和有效的方式。但是鉴于GWAS巨大的工作量以及昂贵的实验费用,单个GWAS样本量可能难以达到研究需要。由于Meta分析可以通过合并多个研究数据的方式增加样本含量,提高统计把握度,有利于发现新的易感位点,目前受到众多研究者的青睐。DIAGRAM协会进行了第一个GWAS的Meta分析,他们合并分析3个GWAS数据共计10000欧洲样本,新发现了6个1型糖尿病的易感位点<sup>[6]</sup>。此后,多种疾病GWAS的Meta研究相继开展,如:克隆病<sup>[7]</sup>、多发性硬化<sup>[8]</sup>、类风湿性关节炎<sup>[9]</sup>、直肠癌<sup>[10]</sup>等,发现了大量新的易感位点。在新发现位点中不仅包括大量常见变异,还有少部分遗传效力较强的低频变异,如:克隆病易感位点(包括LRRK2和MUC19)的一个疾病相关SNP rs11175593的风险等位基因在对照中的频率仅为0.017<sup>[7]</sup>;多发性硬化易感基因TNFRSF1A的一个疾病相关SNP rs4149584的风险等位基因频率仅为0.02<sup>[8]</sup>。这些研究表明,GWAS的Meta分析不仅可以检出常见变异,同样可以发现一些强效的罕见变异,是发掘现有GWAS数据,鉴定更多疾病易感基因的一种经济和高效的方法。

Meta分析需要对多个研究的检验统计量、标准误差或者p值进行合并,由于所涉及的研究设计、人群、表型界定多不一致,因此多项研究间的异质性是影响分析结果准确性的最大因素。根据异质性的程度,Meta分析采用不同的分析模型:当异质性来自随机误差(弱),采用固定效应模型;反之,当异质性来自人群分层(强),则采用随机效应模型<sup>[11]</sup>。目前常用的统计分析工具,如METAL、Comprehensive Meta-analysis、Stata/s.e.m. 9.1等均可实现GWAS的Meta分析。

尽管GWAS的Meta分析在多项研究中取得了很大成功,但也面临着一些问题和挑战:1. 随着研究样本量的增多,需要采用更加完善的异质性检验和数据质控方法以提升Meta分析的可靠性。2. 随

着测序技术的发展, 序列变异的数量和种类也在不断增多, 这时就需要一种新的统计方法将测序数据与已有 SNP 芯片数据进行 Meta 分析。3. 目前的 Meta 分析主要是针对 SNP 本身, 将来则需要新的统计模型可以对一个区域内多个 SNPs 进行 Meta 分析。

### 1.3 根据HapMap数据进行基因型填补(Imputation)

已经报道的GWAS所采用的基因分型芯片大多可以同时检测 50 万个SNPs位点或低于这个水准, 远远少于HapMap中 1 000 万个SNPs位点, 一些真正与疾病相关的位点可能会被遗漏。尽管全基因组DNA测序可以解决这一问题, 但是对于大多数研究者而言费用过于昂贵。Imputation方法在一定程度上可以解决这一难题。Imputation可以根据已检测位点基因型信息和HapMap提供的单体型结构, 推断未检测位点基因型, 进而对该位点与疾病的关联性进行分析。BEAGLE、IMPUTE、MACH和PLINK等多种工具均可实现这一目标<sup>[12]</sup>。由于该方法具有经济和高效的特点, 在目前开展的GWAS中应用已经十分常见。此外, Imputation还可以对一些分型失败位点的基因型进行推测和校正; 更重要的是, GWAS的Meta分析需要将不同分型平台的数据合并分析, 由于不同平台所检测的SNPs位点也不尽相同, 因此Imputation也是合并数据过程的关键步骤<sup>[13]</sup>。目前Imputation主要使用HapMap数据作为参考对照, 但由于HapMap计划所分型的样本数量相对较少(尤其是非欧洲人群), 因此参考对照的基因分型数据并不能完全准确代表各人群的真实情况。随着“千人基因组计划”的完成, 将能够提供更加全面和准确的DNA序列变异数据, 此时Imputation的准确性必将得到显著提升。

### 1.4 多种疾病共同易感基因研究

GWAS所发现的各类疾病易感位点和(或)区域大多具有疾病特异性, 但是随着GWAS的广泛开展, 也发现部分多种疾病所共有的易感区域。如: 8q24 最早确定为前列腺癌易感区域, 随后陆续发现该区域与多种肿瘤(如: 结直肠癌、乳腺癌、膀胱癌等)的易感性存在关联<sup>[14]</sup>; 同样, 6q27 区域与克隆病<sup>[7]</sup>、类风湿性关节炎<sup>[15]</sup>及白癜风等疾病都相关<sup>[16, 17]</sup>, 而 *TERT-CLPTMIL* 与多种癌症易感性相关<sup>[18]</sup>。这表明不同疾病可能存在共同遗传易感机制, 此类共同位

点多位于不同疾病发病网络的衔接点和(或)中枢点, 在疾病的发生和发展过程中起到至关重要的作用, 因此针对此类位点的研究有利于促进对疾病发病机制的深入理解。例如: 8q24 区域SNP rs6983267 的危险等位基因可以作为转录增强子对该区域的原癌基因*MYC*起到远端调控作用, 进而影响肿瘤的易感性<sup>[19, 20]</sup>; 6q27 区域*CCR6* 基因的 5 端一个三等位基因二核苷酸多态可以影响*CCR6* 基因转录, 并且与患者血清中IL-17水平相关, 提示*CCR6* 基因可能是通过IL-17途径参与多种自身免疫性疾病的发病<sup>[15]</sup>。因此, 将同类疾病的GWAS数据进行合并分析是很有意义的尝试, 可以提高发现共同易感位点的效能, 对揭示疾病的潜在机制具有特殊的意义<sup>[21]</sup>。

### 1.5 易感位点的精细定位和测序分析

精细定位是通过增加易感区域内遗传标记的检测密度, 找到与疾病关联程度最强的一组变异, 进而确定易感基因或者缩小易感区域。通过GWAS发现的关联信号通常来自常见变异, 然而目前常见SNPs目录并不完整, 即使第2代的HapMap数据库也仅能提供全部基因组 30%的常见SNPs<sup>[22]</sup>。因此, 要想获得一个区域完整的SNPs信息, 对一定数量的参考对照进行测序是必须要做的工作。通过对该区域的测序还可能发现一些新的拷贝数变异(Copy number variations, CNV)或者小的插入-缺失变异, 建立SNPs与这些DNA序列变异间的关系, 有利于确定与疾病相关的功能性变异。例如: 与克隆病相关的一个非编码SNP与*IRGM*基因上游一个 20kb缺失有很强的连锁关系。由于缺失变异可以影响基因表达, 因此这个 20 kb缺失可能是真正与疾病易感性相关的变异<sup>[23]</sup>。

尽管通过GWAS发现的关联信号通常来自常见变异, 但是高外显性的少见变异对疾病的作用也不容忽视。许多疾病相关的少见变异在普通人群中十分罕见, 或者为近代新出现的突变, 所以不能够通过SNPs间的连锁不平衡关系进行分析。因此, 对于患者的疾病易感基因候选区域进行深度测序是有效地补充手段。由于全基因组测序的费用极其昂贵, 所以目前多针对GWAS或连锁分析发现的候选区域进行测序。例如: 对 1 型糖尿病的 10 个候选区域测序发现*IFIH1* 基因的 4 个少见变异与疾病相关<sup>[24]</sup>,

Johansen等<sup>[25]</sup>通过对高甘油三酯血症的易感基因(*APOA5*、*GCKR*、*LPL*和 *APOB*)进行深度测序,发现438名患者中存在154个低频的错义或者无义突变,而327名正常对照中仅有53个突变( $P = 6.2 \times 10^{-8}$ ),进一步证实少见变异在复杂疾病易感性的作用。

### 1.6 基于通路的GWAS

GWAS产生了海量的分型数据,如何利用这些数据构建基因调控网络也成为研究热点。利用相关的软件平台实现对基因调控网络的构建和分析,在生物信息学、系统生物学、生物医学等领域起着重要的作用。将GWAS所有分型的SNPs按照不同的生物学通路排列,然后比较各个通路在病例/对照间的差异,这就是基于通路GWAS的研究思路。当前, Gene-set enrichment analysis (GSEA) 和 Ingenuity Pathways Analysis(IPA)是最为常用的基因通路分析软件。2005年, Wang等<sup>[26]</sup>对2个帕金森氏病GWAS数据进行分析,成功鉴定出疾病的易感基因和相关通路。还有学者利用WTCCC的GWAS数据进行通路关联分析,发现三种自身免疫病(克隆病、1型糖尿病和类风湿性关节炎)的一些相关病理通路。TLR通路中多个基因与克隆病相关; JAK-STAT信号通路以及T细胞激活相关通路与克隆病、1型糖尿病和类风湿性关节炎相关<sup>[27]</sup>。同样, Jia等<sup>[28]</sup>发现了多个通路(如:谷氨酸盐代谢通路、TGF- $\beta$ 通路、TNFR1通路等)与精神分裂症的易感性相关; Menashe等<sup>[29]</sup>发现了3个与乳腺癌易感性相关的通路(多配体蛋白聚糖-1介导信号通路、肝细胞生长因子受体信号通路、生长激素信号通路)和一个信号级联反应。上述研究表明基于通路的GWAS和基于单个位点的GWAS是两种互补的分析方法,将两者有效地结合可以发现更多的疾病易感基因或者相关生物学通路。由于疾病相关的生物学通路能够更准确反应疾病的发病机制,对于新药的开发和临床诊疗技术的提高具有更大价值。

然而,目前的基因通路研究中也存在一些亟待解决的问题:(1)主要针对基因区域内最显著的SNPs进行分析,很可能会忽略其它具有中度显著水平SNPs,但是纳入通路分析后显著性会明显增高的SNPs。(2)虽然对多种不同类型变异的累加效应进行分析可以增强通路分析的效能,然而这对方法学

提出了更高的要求。(3)没有考虑基因-基因交互作用的因素。

### 1.7 基因-基因、基因-环境交互作用研究

疾病易感基因之间以及易感基因与危险环境因素之间的相互作用对疾病易感性有着重要作用。经典GWAS的研究对象主要是单个SNP或者一群SNPs与疾病的关系,而较少考虑到在人群的不同亚群中可能存在一些非常重要的疾病相关基因<sup>[30]</sup>。针对各种类型GWAS分型数据,应用不同的统计工具可以进行多种上位效应(epistasis effect)分析,如: PLINK进行标准的logistic回归分析, Random Jungle进行随机森林(Random Forest)分析, BEAM进行贝叶斯分隔分析。尤其是PLINK,其具有数据质控和处理模块,应用起来非常方便,是目前最常用的GWAS分析软件<sup>[31]</sup>。例如, Tang等<sup>[32]</sup>对老年黄斑变性GWAS数据进行多位点间交互作用的分析,发现了2个位点间的交互作用与疾病易感性相关,并通过功能学实验进行了证实。他们采用类似的方法对帕金森病GWAS数据进行分析,结果发现了7个易感位点。基因-环境交互作用分析还可以发现对某一环境危害敏感的人群,如:红色肉类(杂环胺)饮食对结直肠癌的易感性仅有轻度影响,而 *CYP1A2* 与 *NAT2* 基因可以影响杂环胺的代谢。基因-环境交互作用分析发现携带 *CYP1A2* 与 *NAT2* 易感基因型的人群,食用高红色肉类饮食,其患病风险显著增高( $OR=8.8$ )<sup>[33]</sup>。同样,有研究发现携带 *NAT2* 基因易感基因型的个体中,吸烟可以显著增加患膀胱癌的风险<sup>[34]</sup>。

尽管目前有多种统计工具可以进行不同模型的交互作用分析,但是交互作用研究同样面临着许多的问题和挑战,如:样本量与统计效能的确定、人群异质性的处理、环境因素暴露的评估等。更重要的是,目前的交互作用分析仅仅是基于统计分析层面的交互作用,即使其显著性水平超过了多重校正的阈值,仍然可能存在着假阳性,因此通过生物学水平的验证是检验交互作用的金标准<sup>[31,35]</sup>。分析基因-基因、基因-环境的交互作用有助于阐明基因与环境因素对疾病的影响。只有综合考虑基因与环境因素交互作用的预测模型,才能对公共健康和个体化用药起到准确的指导作用。



## 2 结语与展望

上述研究策略和方法的应用弥补了经典GWAS的一些不足之处,发现了更多的复杂疾病/性状相关变异,进一步推进了人们对于复杂疾病/性状遗传机制的认识。尽管如此,目前这种基于HapMap设计的GWAS还存在很多的局限性:第一、通过GWAS发现的疾病相关变异多为常见变异(MAF>5%),平均频率在36%左右,而对低频变异的检出效能不够<sup>[36]</sup>。主要原因在于常见变异具有更高的统计学效能,因此在既定的样本量下比罕见变异更容易被检测。就目前的基因分型手段而言,无论是直接分型还是间接连锁不平衡(Linkage disequilibrium, LD)推算,对于罕见SNPs的覆盖都不够充分。第二、GWAS对SNPs的选择是根据HapMap数据库中LD关系,这种间接设计原则决定GWAS更容易发现疾病相关的区域(位点),而对真正致病的变异检出效能不够。有研究对目前GWAS所发现的疾病关联SNPs进行分类,结果提示大部分变异位于基因间(43%)和内含子区域(45%);只有少部分位于基因的功能区域<sup>[35]</sup>,这些非功能区的变异很难明确其在疾病发生发展中的作用。第三、GWAS研究的对象主要是SNPs,对SNPs以外的其它变异检出效能非常微弱。人类基因组的变异除了SNPs外,还包括小片段的缺失、串联重复序列、拷贝数变异和其他结构变异等。这些非SNPs变异通常包含了多个核苷酸,往往可以影响到基因的表达,因此具有更高的生物学意义,可以解释相当一部分复杂疾病/性状的遗传易感性<sup>[37]</sup>。第四、GWAS前期识别的变异,仅仅是在DNA水平观察到的关联而已,仍然可能存在着假阳性,而根本的解决办法将是功能上进行实验验证。利用样本组织或者细胞系观测变异位点对基因功能的影响,或者通过在细胞或动物模型中的基因操作(如基因敲除或过表达等方法)对基因与疾病的关联性进行证实,将有助于明确基因的功能,以及该基因在疾病发生中的作用。

如果将目前基于HapMap设计的GWAS称为狭义GWAS,那么作为一种在全基因组水平进行的关联分析研究方法,随着新技术的不断发展其内涵也必将随之不断扩展。广义的GWAS概念应该是在全基因组范围内,利用关联分析的原理和方法进行的

疾病基因组学研究,不仅包括SNPs,还包括突变、拷贝数变异、基因表达、表观遗传修饰等。因此,在未来几年内GWAS研究将会涉及到疾病基因组研究的几个方面:(1)全基因组拷贝数变异关联分析;(2)全基因组测序研究;(3)全基因组外显子测序研究;(4)全基因组转录组和表达谱研究;(5)全基因组表观遗传学研究。

总之,随着科学技术的持续发展,新的检测技术不断出现,将为疾病基因组学研究带来空前的机遇。GWAS作为一种研究方法具有与时俱进的特点,可以不断吸收和利用这些新技术,并将其应用到复杂疾病的研究中,为疾病预警、临床诊断以及个体化治疗奠定理论基础<sup>[38]</sup>。

## 参考文献(References):

- [1] the Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*, 2007, 447(7145): 661–678. [\[DOI\]](#)
- [2] Thomson W, Barton A, Ke XY, Eyre S, Hinks A, Bowes J, Donn R, Symmons D, Hider S, Bruce IN, Wilson AG, Marinou I, Morgan A, Emery P, YEAR Consortium, Carter A, Steer S, Hocking L, Reid DM, Wordsworth P, Harrison P, Strachan D, Worthington J. Rheumatoid arthritis association at 6q23. *Nat Genet*, 2007, 39(12): 1431–1433. [\[DOI\]](#)
- [3] Barton A, Thomson W, Ke XY, Eyre S, Hinks A, Bowes J, Plant D, Gibbons LJ, Wellcome Trust Case Control Consortium, YEAR Consortium, BIRAC Consortium, Wilson AG, Bax DE, Morgan AW, Emery P, Steer S, Hocking L, Reid DM, Wordsworth P, Harrison P, Worthington J. Rheumatoid arthritis susceptibility loci at chromosomes 10p15, 12q13 and 22q13. *Nat Genet*, 2008, 40(10): 1156–1159. [\[DOI\]](#)
- [4] Eeles RA, Kote-Jarai Z, Al Olama AA, Giles GG, Guy M, Severi G, Muir K, Hopper JL, Henderson BE, Haiman CA, Schleutker J, Hamdy FC, Neal DE, Donovan JL, Stanford JL, Ostrander EA, Ingles SA, John EM, Thibodeau SN, Schaid D, Park JY, Spurdle A, Clements J, Dickinson JL, Maier C, Vogel W, Dörk T, Rebbeck TR, Cooney KA, Cannon-Albright L, Chappuis PO, Hutter P, Zeegers M, Kaneva R, Zhang HW, Lu YJ, Foulkes WD, English DR, Leongamornlert DA, Tymrakiewicz M, Morrison J, Arder-Jones AT, Hall AL, O'Brien LT, Wilkinson RA,

- Saunders EJ, Page EC, Sawyer EJ, Edwards SM, Dearnaley DP, Horwich A, Huddart RA, Khoo VS, Parker CC, Van As N, Woodhouse CJ, Thompson A, Christmas T, Ogden C, Cooper CS, Southey MC, Lophatananon A, Liu JF, Kolonel LN, Le Marchand L, Wahlfors T, Tammela TL, Auvinen A, Lewis SJ, Cox A, FitzGerald LM, Koopmeiners JS, Karyadi DM, Kwon EM, Stern MC, Corral R, Joshi AD, Shahabi A, McDonnell SK, Sellers TA, Pow-Sang J, Chambers S, Aitken J, Gardiner RA, Batra J, Kedda MA, Lose F, Polanowski A, Patterson B, Serth J, Meyer A, Luedeke M, Stefflova K, Ray AM, Lange EM, Farnham J, Khan H, Slavov C, Mitkova A, Cao GW, The UK Genetic Prostate Cancer Study Collaborators/British Association of Urological Surgeons' Section of Oncology, The UK ProtecT Study Collaborators, The PRACTICAL Consortium, Easton DF. Identification of seven new prostate cancer susceptibility loci through a genome-wide association study. *Nat Genet*, 2009, 41(10): 1116–1121. [\[DOI\]](#)
- [5] Sun LD, Cheng H, Wang ZX, Zhang AP, Wang PG, Xu JH, Zhu QX, Zhou HS, Ellinghaus E, Zhang FR, Pu XM, Yang XQ, Zhang JZ, Xu AE, Wu RN, Xu LM, Peng L, Helms CA, Ren YQ, Zhang C, Zhang SM, Nair RP, Wang HY, Lin GS, Stuart PE, Fan X, Chen G, Tejasvi T, Li P, Zhu J, Li ZM, Ge HM, Weichenthal M, Ye WZ, Zhang C, Shen SK, Yang BQ, Sun YY, Li SS, Lin Y, Jiang JH, Li CT, Chen RX, Cheng J, Jiang X, Zhang P, Song WM, Tang J, Zhang HQ, Sun L, Cui J, Zhang LJ, Tang B, Huang F, Qin Q, Pei XP, Zhou AM, Shao LM, Liu JL, Zhang FY, Du WD, Franke A, Bowcock AM, Elder JT, Liu JJ, Yang S, Zhang XJ. Association analyses identify six new psoriasis susceptibility loci in the Chinese population. *Nat Genet*. 2010, 42(11): 1005–1009.
- [6] Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, Hu T, de Bakker PI, Abecasis GR, Almgren P, Andersen G, Ardlie K, Boström KB, Bergman RN, Bonnycastle LL, Borch-Johnsen K, Burt NP, Chen H, Chines PS, Daly MJ, Deodhar P, Ding CJ, Doney ASF, Duren WL, Elliott KS, Erdos MR, Frayling TM, Freathy RM, Gianniny L, Gallett H, Grarup N, Groves CJ, Guiducci C, Hansen T, Herder C, Hitman GA, Hughes TE, Isomaa B, Jackson AU, Jørgensen T, Kong A, Kubalanza K, Kuruvilla FG, Kuusisto J, Langenberg C, Lango H, Lauritzen T, Li Y, Lindgren CM, Lyssenko V, Marvelle AF, Meisinger C, Midtjell K, Mohlke KL, Morken MA, Morris AD, Narisu N, Nilsson P, Owen KR, Palmer CN, Payne F, Perry JRB, Pettersen E, Platou C, Prokopenko I, Qi L, Qin L, Rayner NW, Rees M, Roix JJ, Sandbaek A, Shields B, Sjögren M, Steinthorsdottir V, Stringham HM, Swift AJ, Thorleifsson G, Thorsteinsdottir U, Timpson NJ, Tuomi T, Tuomilehto J, Walker M, Watanabe RM, Weedon MN, Willer CJ, Wellcome Trust Case control Consortium, Illig T, Hveem K, Hu FB, Laakso M, Stefansson K, Pedersen O, Wareham NJ, Barroso I, Hattersley AT, Collins FS, Groop L, McCarthy MI, Boehnke M, Altshuler D. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat Genet*, 2008, 40(5): 638–645. [\[DOI\]](#)
- [7] Barrett JC, Hansoul S, Nicolae DL, Cho JH, Duerr RH, Rioux JD, Brant SR, Silverberg MS, Taylor KD, Barmada MM, Bitton A, Dassopoulos T, Datta LW, Green T, Griffiths AM, Kistner EO, Murtha MT, Regueiro MD, Rotter JI, Schumm LP, Steinhart AH, Targan SR, Xavier RJ, the NIDDK IBD Genetics Consortium, Libioulle C, Sandor C, Lathrop M, Belaiche J, Dewit O, Gut I, Heath S, Laukens D, Mni M, Rutgeerts P, Van Gossum A, Zelenika D, Franchimont D, Hugot JP, de Vos M, Vermeire S, Louis E, the Belgian-French IBD Consortium, the Wellcome Trust Case control Consortium, Cardon LR, Anderson CA, Drummond H, Nimmo E, Ahmad T, Prescott NJ, Onnie CM, Fisher SA, Marchini J, Ghori J, Bumpstead S, Gwilliam R, Tremelling M, Deloukas P, Mansfield J, Jewell D, Satsangi J, Mathew CG, Parkes M, Georges M, Daly MJ. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat Genet*, 2008, 40(8): 955–962. [\[DOI\]](#)
- [8] De Jager PL, Jia XM, Wang J, de Bakker PIW, Ottoboni L, Aggarwal NT, Piccio L, Raychaudhuri S, Tran D, Aubin C, Briskin R, Romano S, International MS Genetics Consortium, Baranzini SE, McCauley JL, Pericak-Vance MA, Haines JL, Gibson RA, Naeglin Y, Uitdehaag B, Matthews PM, Kappos L, Polman C, McArdle WL, Strachan DP, Evans D, Cross AH, Daly MJ, Compston A, Sawcer SJ, Weiner HL, Hauser SL, Hafler DA, Oksenberg JR. Meta-analysis of genome scans and replication identify *CD6*, *IRF8* and *TNFRSF1A* as new multiple sclerosis susceptibility loci. *Nat Genet*, 2009, 41(7): 776–782. [\[DOI\]](#)
- [9] Raychaudhuri S, Remmers EF, Lee AT, Hackett R, Guiducci C, Burt NP, Gianniny L, Korman BD, Padyukov L, Kurreeman FAS, Chang M, Catanese JJ, Ding B, Wong S, van der Helm-van Mil AHM, Neale BM, Cobylin J, Cui J, Tak PP, Wolbink GJ, Crusius JBA, van der Horst-Bruinsma IE, Criswell LA, Amos CI, Seldin MF, Kastner DL, Ardlie KG, Alfredsson L, Costenbader KH,

- Altshuler D, Huizinga TWJ, Shadick NA, Weinblatt ME, de Vries N, Worthington J, Seielstad M, Toes REM, Karlsson EW, Begovich AB, Klareskog L, Gregersen PK, Daly MJ, Plenge RM. Common variants at *CD40* and other loci confer risk of rheumatoid arthritis. *Nat Genet*, 2008, 40(10): 1216–1223. [\[DOI\]](#)
- [10] Houlston RS, Webb E, Broderick P, Pittman AM, Di Bernardo MCC, Lubbe S, Chandler I, Vijayakrishnan J, Sullivan K, Penegar S, Colorectal Cancer Association Study, Carvajal-Carmona L, Howarth K, Jaeger E, Spain SL, Walther A, Barclay E, Martin L, Gorman M, Domingo E, Teixeira AS, CorGI Consortium, Kerr D, Cazier JBB, Niittymäki I, Tuupanen S, Karhu A, Aaltonen LA, Tomlinson IP, Farrington SM, Tenesa A, Prendergast JG, Barnetson RA, Cetnarskyj R, Porteous ME, Pharoah PD, Koessler T, Hampe J, Buch S, Schafmayer C, Tepl J, Schreiber S, Völzke H, Chang-Claude J, Hoffmeister M, Brenner H, Zanke BW, Montpetit A, Hudson TJ, Gallinger S, International Colorectal Canetic Association Consortium, Campbell H, Dunlop MG. Meta-analysis of genome-wide association data identifies four new susceptibility loci for colorectal cancer. *Nat Genet*, 2008, 40(12): 1426–1435. [\[DOI\]](#)
- [11] Ioannidis JP, Patsopoulos NA, Evangelou E. Heterogeneity in meta-analyses of genome-wide association investigations. *PLoS One*, 2007, 2(9): e841. [\[DOI\]](#)
- [12] Pei YF, Li J, Zhang L, Papasian CJ, Deng HW. Analyses and comparison of accuracy of different genotype imputation methods. *PLoS One*, 2008, 3(10): e3551. [\[DOI\]](#)
- [13] Browning SR. Missing data imputation and haplotype phase inference for genome-wide association studies. *Hum Genet*, 2008, 124(5): 439–450. [\[DOI\]](#)
- [14] Stadler ZK, Thom P, Robson ME, Weitzel JN, Kauff ND, Hurley KE, Devlin V, Gold B, Klein RJ, Offit K. Genome-wide association studies of cancer. *J Clin Oncol*, 2010, 28(27): 4255–4267. [\[DOI\]](#)
- [15] Kochi Y, Okada Y, Suzuki A, Ikari K, Terao C, Takahashi A, Yamazaki K, Hosono N, Myouzen K, Tsunoda T, Kamatani N, Furuichi T, Ikegawa S, Ohmura K, Mimori T, Matsuda F, Iwamoto T, Momohara S, Yamanaka H, Yamada R, Kubo M, Nakamura Y, Yamamoto K. A regulatory variant in *CCR6* is associated with rheumatoid arthritis susceptibility. *Nat Genet*, 2010, 42(6): 515–519. [\[DOI\]](#)
- [16] Jin Y, Birlea SA, Fain PR, Mailloux CM, Riccardi SL, Gowan K, Holland PJ, Bennett DC, Wallace MR, McCormack WT, Kemp EH, Gawkrödger DJ, Weetman AP, Picardo M, Leone G, Taïeb A, Jouary T, Ezzedine K, van Geel N, Lambert J, Overbeck A, Spritz RA. Common variants in *FOXP1* are associated with generalized vitiligo. *Nat Genet*, 2010, 42(7): 576–578. [\[DOI\]](#)
- [17] Quan C, Ren YQ, Xiang LH, Sun LD, Xu AE, Gao XH, Chen HD, Pu XM, Wu RN, Liang CZ, Li JB, Gao TW, Zhang JZ, Wang XL, Wang J, Yang RY, Liang L, Yu JB, Zuo XB, Zhang SQ, Zhang SM, Chen G, Zheng XD, Li P, Zhu J, Li YW, Wei XD, Hong WS, Ye Y, Zhang Y, Wu WS, Cheng H, Dong PL, Hu DY, Li Y, Li M, Zhang X, Tang HY, Tang XF, Xu SX, He SM, Lv YM, Shen M, Jiang HQ, Wang Y, Li K, Kang XJ, Liu YQ, Sun L, Liu ZF, Xie SQ, Zhu CY, Xu Q, Gao JP, Hu WL, Ni C, Pan TM, Yao S, He CF, Liu YS, Yu ZY, Yin XY, Zhang FY, Yang S, Zhou Y, Zhang XJ. Genome-wide association study for vitiligo identifies susceptibility loci at 6q27 and the MHC. *Nat Genet*, 2010, 42(7): 614–618. [\[DOI\]](#)
- [18] Rafnar T, Sulem P, Stacey SN, Geller F, Gudmundsson J, Sigurdsson A, Jakobsdottir M, Helgadóttir H, Thorlacius S, Aben KKH, Blöndal T, Thorgeirsson TE, Thorleifsson G, Kristjansson K, Thorisdóttir K, Ragnarsson R, Sigurgeirsson B, Skuladóttir H, Gudbjartsson T, Isaksson HJ, Einarsson GV, Benediktsdóttir KR, Agnarsson BA, Olafsson K, Salvarsdóttir A, Bjarnason H, Asgeirsdóttir M, Kristinsson KT, Matthiasdóttir S, Sveinsdóttir SG, Polidoro S, Höiom V, Bottella-Estrada R, Hemminki K, Rudnai P, Bishop DT, Campagna M, Kellen E, Zeegers MP, de Verdier P, Ferrer A, Isla D, Vidal MJ, Andres R, Saez B, Juberias P, Banzo J, Navarrete S, Tres A, Kan DH, Lindblom A, Gurzau E, Koppova K, de Vegt F, Schalken JA, van der Heijden HFM, Smit HJ, Termeer RA, Oosterwijk E, van Hooij O, Nagore E, Porru S, Steineck G, Hansson J, Buntinx F, Catalona WJ, Matullo G, Vineis P, Kiltie AE, Mayordomo JI, Kumar R, Kiemeny LA, Frigge ML, Jonsson T, Saemundsson H, Barkardóttir RB, Jonsson E, Jonsson S, Olafsson JH, Gulcher JR, Masson G, Gudbjartsson DF, Kong A, Thorsteinsdóttir U, Stefansson K. Sequence variants at the *TERT-CLPTMIL* locus associate with many cancer types. *Nat Genet*, 2009, 41(2): 221–227. [\[DOI\]](#)
- [19] Pomerantz MM, Ahmadiyah N, Jia L, Herman P, Verzi MP, Doddapaneni H, Beckwith CA, Chan JA, Hills A, Davis M, Yao KL, Kehoe SM, Lenz HJ, Haiman CA, Yan CL, Henderson BE, Frenkel B, Barretina J, Bass A, Tabernero J, Baselga J, Regan MM, Manak JR, Shivdasani R, Coetzee GA, Freedman ML. The 8q24 cancer risk variant rs6983267 shows long-range interaction with *MYC* in colorectal cancer. *Nat Genet*, 2009, 41(8): 882–884. [\[DOI\]](#)
- [20] Tuupanen S, Turunen M, Lehtonen R, Hallikas O, Vanha-

- ranta S, Kivioja T, Björklund M, Wei GH, Yan J, Niitty-mäki I, Mecklin JP, Järvinen H, Ristimäki A, Di-Bernardo M, East P, Carvajal-Carmona L, Houlston RS, Tomlinson I, Palin K, Ukkonen E, Karhu A, Taipale J, Aaltonen LA. The common colorectal cancer predisposition SNP rs6983267 at chromosome 8q24 confers potential to enhanced Wnt signaling. *Nat Genet*, 2009, 41(8): 885–890. [\[DOI\]](#)
- [21] Burton PR, Clayton DG, Cardon LR, Craddock N, Deloukas P, Duncanson A, Kwiatkowski DP, McCarthy MI, Ouwehand WH, Samani NJ, Todd JA, Donnelly Chair P, Barrett JC, Davison D, Donnelly P, Easton D, Evans DM, Leung HT, Marchini JL, Morris AP, Spencer CC, Tobin MD, Cardon LR, Clayton DG, Attwood AP, Boorman JP, Cant B, Everson U, Hussey JM, Jolley JD, Knight AS, Koch K, Meech E, Nutland S, Prowse CV, Stevens HE, Taylor NC, Walters GR, Walker NM, Watkins NA, Winzer T, Todd JA, Ouwehand WH, Jones RW, McArdle WL, Ring SM, Strachan DP, Pembrey M, Breen G, Clair DS, Caesar S, Gordon-Smith K, Jones L, Fraser C, Green EK, Grozeva D, Hamshire ML, Holmans PA, Jones IR, Kirov G, Moskvina V, Nikolov I, O'Donovan MC, Owen MJ, Craddock N, Collier DA, Elkin A, Farmer A, Williamson R, McGuffin P, Young AH, Ferrier IN, Ball SG, Balmforth AJ, Barrett JH, Bishop TD, Iles MM, Maqbool A, Yuldasheva N, Hall AS, Braund PS, Burton PR, Dixon RJ, Mangino M, Stevens S, Tobin MD, Thompson JR, Samani NJ, Bredin F, Tremelling M, Parkes M, Drummond H, Lees CW, Nimmo ER, Satsangi J, Fisher SA, Forbes A, Lewis CM, Onnie CM, Prescott NJ, Sanderson J, Matthew CG, Barbour J, Mohiuddin MK, Todhunter CE, Mansfield JC, Ahmad T, Cummings FR, Jewell DP, Webster J, Brown MJ, Clayton DG, Lathrop MG, Connell J, Dominiczak A, Samani NJ, Marcano CAB, Burke B, Dobson R, Gungadool J, Lee KL, Munroe PB, Newhouse SJ, Onipinla A, Wallace C, Xue MZ, Caulfield M, Farrall M, Barton A, Bruce IN, Donovan H, Eyre S, Gilbert PD, Hilder SL, Hinks AM, John SL, Potter C, Silman AJ, Symmons DP, Thomson W, Worthington J, Stevens HE, Dunger DB, Widmer B, Tidd JA, Frayling TM, Freathy RM, Lango H, Perry JR, Shields BM, Weedon MN, Hattersley AT, Hitman GA, Walker M, Elliott KS, Groves CJ, Lindgren CM, Rayner NW, Timpson NJ, Zeggini E, McCarthy MI, Newport M, Sirugo G, Lyons E, Vannberg F, Hill AVS, Bradbury LA, Farrar C, Pointon JJ, Wordsworth P, Brown MA, Franklyn JA, Heward JM, Simmonds MJ, Gough SC, Seal S, Stratton MR, Rahman N, Ban M, Goris A, Sawcer SJ, Compston A, Conway D, Jallow M, Newport M, Sirugo G, Rockett KA, Kwiatkowski DP, Bumpstead SJ, Chaney A, Downes K, Ghori MJ, Gwilliam R, Hunt SE, Inouye M, Keniry A, King E, McGinnis R, Potter S, Ravindrarajah R, Whittaker P, Widdens C, Withers D, Deloukas P, Leung HT, Nutland S, Stenens HE, Walker NM, Todd JA, Easton D, Tobin MD, Barrett JC, Evans DM, Morris AP, Cardon LR, Cardin NJ, Davison D, Ferreira T, Pereira-Gale J, Hallgrimsdottir IB, Howie BN, Marchini JL, Spencer CC, Su Z, Teo YY, Vukcevic D, Donnelly P, Bentley D, Brown MA, Caulfield M, Clayton DG, Compston A, Craddock N, Deloukas P, Donnelly P, Farrall M, Gough SC, Hall AS, Hattersley AT, Mitchell SL, Newby PR, Brand OJ, Carr-Smith J, Pearce SH, Gough SCL, McGinnis R, Keniry A, Deloukas P, Reveille JD, Deloukas P, Reveille JD, Zhou XD, Bradbury LA, Sims AM, Dowling A, Taylor J, Doan T, Cardon LR, Davis JC, Pointon JJ, Savage L, Ward MM, Learch TL, Weisman MH, Wordsworth P, Brown MA. Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. *Nat Genet*, 2007, 39(11): 1329–1337. [\[DOI\]](#)
- [22] Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, Belmont JW, Boudreau A, Hardenbol P, Leal SM, Pasternak S, Wheeler DA, Willis TD, Yu F, Yang H, Zeng C, Gao Y, Hu H, Hu W, Li C, Lin W, Liu S, Pan H, Tang X, Wang J, Wang W, Yu J, Zhang B, Zhang Q, Zhao H, Zhou J, Gabriel SB, Barry R, Blumenstiel B, Camargo A, Defelice M, Faggart M, Goyette M, Gupta S, Moore J, Nguyen H, Onofrio RC, Parkin M, Roy J, Stahl E, Winchester E, Ziaugra L, Altshuler D, Shen Y, Yao Z, Huang W, Chu X, He Y, Jin L, Liu Y, Shen Y, Sun W, Wang H, Wang Y, Xiong X, Xu L, Waye MM, Tsui SK, Xue H, Wong JT, Galver LM, Fan JB, Gunderson K, Murray SS, Oliphant AR, Chee MS, Montpetit A, Chagnon F, Ferretti V, Leboeuf M, Olivier JF, Phillips MS, Roumy S, Sallé C, Verner A, Hudson TJ, Kwok PY, Cai D, Koboldt DC, Miller RD, Pawlikowska L, Taillon-Miller P, Xiao M, Tsui LC, Mak W, Song YQ, Tam PK, Nakamura Y, Kawaguchi T, Kitamoto T, Morizono T, Nagashima A, Ohnishi Y, Sekine A, Tanaka T, Tsunoda T, Deloukas P, Bird CP, Delgado M, Dermitzakis ET, Gwilliam R, Hunt S, Morrison J, Powell D, Stranger BE, Whittaker P, Bentley DR, Daly MJ, de Bakker PI, Barrett J, Chretien YR, Maller J, McCarroll S, Patterson N, Pe'er I, Price A, Purcell S, Richter DJ, Sabeti P, Saxena R, Schaffner SF, Sham PC, Varilly P, Stein LD, Krishnan L, Smith AV, Tello-Ruiz MK,



- Thorisson GA, Chakravarti A, Chen PE, Cutler DJ, Kashuk CS, Lin S, Abecasis GR, Guan W, Li Y, Munro HM, Qin ZS, Thomas DJ, McVean G, Auton A, Bottolo L, Cardin N, Eyheramendy S, Freeman C, Marchini J, Myers S, Spencer C, Stephens M, Donnelly P, Cardon LR, Clarke G, Evans DM, Morris AP, Weir BS, Mullikin JC, Sherry ST, Feolo M, Skol A, Zhang H, Matsuda I, Fukushima Y, Macer DR, Suda E, Rotimi CN, Adebamowo CA, Ajayi I, Aniagwu T, Marshall PA, Nkwodimmah C, Royal CD, Leppert MF, Dixon M, Peiffer A, Qiu R, Kent A, Kato K, Niikawa N, Adewole IF, Knoppers BM, Foster MW, Clayton EW, Watkin J, Muzny D, Nazareth L, Sodergren E, Weinstock GM, Yakub I, Birren BW, Wilson RK, Fulton LL, Rogers J, Burton J, Carter NP, Clee CM, Griffiths M, Jones MC, McLay K, Plumb RW, Ross MT, Sims SK, Willey DL, Chen Z, Han H, Kang L, Godbout M, Wallenburg JC, L'Archevêque P, Bellemare G, Saeki K, An D, Fu H, Li Q, Wang Z, Wang R, Holden AL, Brooks LD, McEwen JE, Guyer MS, Wang VO, Peterson JL, Shi M, Spiegel J, Sung LM, Zacharia LF, Collins FS, Kennedy K, Jamieson R, Stewart J. A second generation human haplotype map of over 3.1 million SNPs. *Nature*, 2007, 449(7164): 851–861. [\[DOI\]](#)
- [23] McCarroll SA, Huett A, Kuballa P, Chlewicki SD, Landry A, Goyette P, Zody MC, Hall JL, Brant SR, Cho JH, Duerr RH, Silverberg MS, Taylor KD, Rioux JD, Altshuler D, Daly MJ, Xavier RJ. Deletion polymorphism upstream of *IRGM* associated with altered *IRGM* expression and Crohn's disease. *Nat Genet*, 2008, 40(9): 1107–1112. [\[DOI\]](#)
- [24] Nejentsev S, Walker N, Riches D, Egholm M, Todd JA. Rare variants of *IFIH1*, a gene implicated in antiviral responses, protect against type 1 diabetes. *Science*, 2009, 324(5925): 387–389. [\[DOI\]](#)
- [25] Johansen CT, Wang J, Lanktree MB, Cao HN, McIntyre AD, Ban MR, Martins RA, Kennedy BA, Hassell RG, Visser ME, Schwartz SM, Voight BF, Elosua R, Salomaa V, O'Donnell CJ, Dallinger-Thie GM, Anand SS, Yusuf S, Huff MW, Kathiresan S, Hegele RA. Excess of rare variants in genes identified by genome-wide association study of hypertriglyceridemia. *Nat Genet*, 2010, 42(8): 684–687. [\[DOI\]](#)
- [26] Wang K, Li MY, Bucan M. Pathway-based approaches for analysis of genomewide association studies. *Am J Hum Genet*, 2007, 81(6): 1278–1283. [\[DOI\]](#)
- [27] Eleftherohorinou H, Wright V, Hoggart C, Hartikainen AL, Jarvelin MR, Balding D, Coin L, Levin M. Pathway analysis of GWAS provides new insights into genetic susceptibility to 3 inflammatory diseases. *PLoS One*, 2009, 4(11): e8068. [\[DOI\]](#)
- [28] Jia P, Wang L, Meltzer HY, Zhao Z. Common variants conferring risk of schizophrenia: a pathway analysis of GWAS data. *Schizophr Res*, 2010, 122(1/3): 38–42. [\[DOI\]](#)
- [29] Menashe I, Maeder D, Garcia-Closas M, Figueroa JD, Bhattacharjee S, Rotunno M, Kraft P, Hunter DJ, Chanock SJ, Rosenberg PS, Chatterjee N. Pathway analysis of breast cancer genome-wide association study highlights three pathways and one canonical signaling cascade. *Cancer Res*, 2010, 70(11): 4453–4459. [\[DOI\]](#)
- [30] Engelman CD, Baurley JW, Chiu YF, Joubert BR, Lewinger JP, Maenner MJ, Murcray CE, Shi G, Gauderman WJ. Detecting gene-environment interactions in genome-wide association data. *Genet Epidemiol*, 2009, 33(Suppl 1): S68–S73. [\[DOI\]](#)
- [31] Cantor RM, Lange K, Sinsheimer JS. Prioritizing GWAS results: a review of statistical methods and recommendations for their application. *Am J Hum Genet*, 2010, 86(1): 6–22. [\[DOI\]](#)
- [32] Tang WW, Wu XB, Jiang R, Li YD. Epistatic module detection for case-control studies: a Bayesian model with a Gibbs sampling strategy. *PLoS Genet*, 2009, 5(5): e1000464. [\[DOI\]](#)
- [33] Le Marchand L, Hankin JH, Wilkens LR, Pierce LM, Franke A, Kolonel LN, Seifried A, Custer LJ, Chang W, Lum-Jones A, Donlon T. Combined effects of well-done red meat, smoking, and rapid N-acetyltransferase 2 and CYP1A2 phenotypes in increasing colorectal cancer risk. *Cancer Epidemiol Biomarkers Prev*, 2001, 10(12): 1259–1266. [\[DOI\]](#)
- [34] García-Closas M, Malats N, Silverman D, Dosemeci M, Kogevinas M, Hein DW, Tardón A, Serra C, Carrato A, García-Closas R, Lloreta J, Castaño-Vinyals G, Yeager M, Welch R, Chanock S, Chatterjee N, Wacholder S, Samanic C, Torà M, Fernández F, Real FX, Rothman N. NAT2 slow acetylation, GSTM1 null genotype, and risk of bladder cancer: results from the Spanish Bladder Cancer Study and meta-analyses. *Lancet*, 2005, 366(9486): 649–659. [\[DOI\]](#)
- [35] Thomas D. Gene-environment-wide association studies: emerging approaches. *Nat Rev Genet*, 2010, 11(4): 259–272. [\[DOI\]](#)
- [36] Hindorf LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci USA*, 2009, 106(23): 9362–9367. [\[DOI\]](#)
- [37] Ku CS, Loy EY, Pawitan Y, Chia KS. The pursuit of genome-wide association studies: where are we now? *J Hum Genet*, 2010, 55(4): 195–206. [\[DOI\]](#)
- [38] 韩建文, 张学军. 全基因组关联研究现状. *遗传*, 2011, 33(1): 25–35.