

DOI: 10.3724/SP.J.1005.2012.00793

## 原发性高血压全基因组关联研究进展

许睿玮<sup>1</sup>, 严卫丽<sup>2</sup>

1. 新疆医科大学公共卫生学院, 乌鲁木齐 830054;
2. 复旦大学附属儿科医院临床流行病学研究室, 上海 201102

**摘要:** 原发性高血压是一种由遗传与环境因素共同导致的复杂疾病, 具有高度的遗传异质性。自 2007 年首个高血压全基因组关联研究(Genome-wide association studies, GWAS)报道以来, 许多 GWAS 相继开展。文章首先对 2007 年 1 月至 2011 年 9 月期间报道的 24 篇血压/高血压易感基因的 GWAS 按人种与染色体位置对其结果进行汇总, 经统计位点 rs17249754、rs1378942 和 rs11191548 报道频数最多。其次介绍了 GWAS 方法学的研究进展, 包括选择高质量的数量表型和选择多阶段研究设计来增加研究发现阳性关联的机会。统计分析方面, 除强调了已经报道过的多重比较和重复(验证)研究等问题外, 文章还介绍了通过 Meta 分析对 GWAS 数据进行深度发掘, 并应用基因型填补法对缺失数据进行填补可以提高全基因组遗传标记的覆盖率的方法。尽管 GWAS 发现了许多我们未知的基因与疾病表型的关联, 为了解高血压的发病机制提供了更多线索, 但是目前 GWAS 发现的血压/高血压相关变异多为对人群血压的影响极其微弱的常见变异。因此今后的研究中可加强深度功能学研究对易感基因精细定位和外显子组测序技术的应用, 结合 GWAS 的成果进行生物信息学通路分析和表观遗传学机制研究等, 逐步揭示高血压的遗传机制。

**关键词:** 原发性高血压; 血压; 全基因组关联研究; 易感基因

## Advances in genome-wide association studies on essential hypertension

XU Rui-Wei<sup>1</sup>, YAN Wei-Li<sup>2</sup>

1. School of Public Health, Xinjiang Medical University, Urumqi 830054, China;
2. Department of Clinical Epidemiology, Children's Hospital of Fudan University, Shanghai 201102, China

**Abstract:** Since the first genome-wide association study was reported in 2007, hypertension has attracted numerous studies to identify its genetic basis. The first part of the current review summarizes the genetic loci associated with blood pressure/ hypertension identified by genome-wide association studies (GWAS) from January 2007 to September 2011, by race and chromosomal location. In the second part, we stress several important points in GWAS methodology, for example, selecting high-quality phenotypes and using multi-stage study design to increase the power studies to identify loci with minor effect. For statistical analysis, besides multiple testing correction and replication of the GWAS that have been intro-

收稿日期: 2011-12-13; 修回日期: 2012-02-17

作者简介: 许睿玮, 硕士在读, 研究方向: 复杂疾病遗传流行病学。E-mail: ruiwei526@gmail.com

通讯作者: 严卫丽, 教授, 博士, 研究方向: 复杂疾病遗传流行病学。Tel: 021-64931215; E-mail: yanwl@fudan.edu.cn

网络出版时间: 2012-5-31 02:27:18

URL: <http://www.cnki.net/kcms/detail/11.1913.R.20120531.1427.002.html>

duced in previous reviews, computer-based genotype imputation has been described for its advantages in compensating GWAS genotyping failures. Although GWAS identifies many unknown genetic variants and improves our understanding for the pathogenesis of hypertension, the loci related to blood pressure / hypertension are common sequence variations with minor effect. The association studies are difficult to be replicated in different populations. Further studies are expected including extensive functional studies and fine mapping using advanced techniques, such as whole genome exon sequencing and pathway analysis, as well as epigenetic study to elucidate the etiology of human essential hypertension.

**Keywords:** essential hypertension; blood pressure; genome-wide association study; susceptibility gene

高血压可分为原发性高血压(以下简称高血压)及继发性高血压两类,其中原发性高血压占总高血压患病率的 95%以上,即我们所称的高血压病。高血压以动脉血压持续升高为主要表现,是由多种因素导致的复杂疾病,具有高度遗传异质性,它作为一种独立疾病影响了全球 20%~30%人口的健康<sup>[1]</sup>,同时作为脑卒中、冠心病、心功能不全、肾脏疾病等多种疾病的重要危险因素,致使全球每年约 1 350 万人死于其相关性疾病<sup>[2]</sup>。既往流行病学研究提示,高血压的发病和遗传因素与环境因素的共同作用有关,其遗传度在家系研究中为 20%~40%,在双生子研究中则高达 60%<sup>[3,4]</sup>。

迄今为止,人类对以基因为单位和以区域为单位的高血压研究已有 20 年,经历了第一代限制性内切酶片段长度多态性到第二代短串联重复序列以至第三代单核苷酸多态(Single nucleotide polymorphism, SNP),已知的高血压候选基因不下 200 种,涉及肾素-血管紧张素-醛固酮系统、交感神经系统、下丘脑-垂体轴、内皮素、利钠肽、激肽释放酶-激肽系统、类固醇激素、前列腺素、生长因子和激素、骨架蛋白和粘附分子、细胞内信使、脂质代谢、糖代谢、载脂蛋白、离子通道或转运体等多个系统或功能,但是因为人群混杂等诸多因素影响使得高血压易感基因的研究在不同地区、不同人群中不易重复。

芯片技术的出现实现了遗传生物信息的高通量检测,2005 年基于 SNP 芯片技术的全基因组关联研究(Genome-wide association studies, GWAS)成为人类医学遗传研究史上的分水岭,为高血压易感基因的筛选开辟了新纪元。GWAS 基于连锁不平衡原理,同时选择几十万甚至上百万个多态位点代表基因组范围内的遗传变异,应用高通量的基因分型平台进行检测,分析全基因组范围内的遗传变异与所研究

疾病表型之间的关系<sup>[5]</sup>。GWAS 不需要事先提出研究假设,在全基因组水平上筛选疾病候选基因,研究设计多采用多阶段多中心的病例-对照研究,继而应用多个独立的研究进行重复验证,样本量较大,数据可靠且具有可重复性。近年来,世界范围内对肿瘤、心血管病、糖尿病、高血压、精神分裂症等复杂疾病的 GWAS 发展迅速,大量 GWAS 报道了高血压发病机制中遗传因素的作用,新发现了一系列血压水平相关基因序列变异。本文对高血压 GWAS 进展进行了综述。

## 1 高血压 GWAS 现状

### 1.1 高血压 GWAS 在欧洲人群中的研究结果

2007 年,The Wellcome Trust Case Control Consortium(WTCCC)<sup>[6]</sup>研究结果问世,该研究标志着高血压易感基因的研究进入 GWAS 的新时代。研究应用 Affymetrix 500K 基因芯片,入选了 2 000 例高血压病例和 3 000 例对照,然而校正多重比较后没有发现达到基因组显著性水平( $P < 5 \times 10^{-7}$ )的统计学关联。同年,美国研究者 Levy 等<sup>[7]</sup>在 Framingham Heart Study(FHS)基础上利用 Affymetrix 100K 基因芯片,对 FHS 研究中的 1 327 名参与者的 70 987 个 SNPs 进行分析,与收缩压(Systolic blood pressure, SBP)和舒张压(Diastolic blood pressure, DBP)关联最显著的位点分别为 1p31.3 区域的 rs10493340( $P = 1.7 \times 10^{-6}$ )和 8q13.3 区域的 rs1963982( $P = 3.3 \times 10^{-6}$ ),同样没有发现达到基因组显著性水平的位点。

2008 年,美国家庭血压计划(Family Blood Pressure Program, FBPP)<sup>[8]</sup>对 WTCCC 研究中最具显著性的 6 个位点( $P < 10^{-5}$ )在 11 433 名研究对象中进行了重复验证,发现在欧洲裔美国人和西班牙裔美国人

中,仅有rs1937506 G/C变异与高血压有关( $P=0.032$ 和 $P=0.005$ ),而在非洲裔美国人中没有发现任何位点与高血压存在阳性关联。

2009年,Wang等<sup>[9]</sup>应用Affymetrix 100 K芯片对来自阿米什人(Amish)的542名研究对象的DNA进行分析,发现位于9p21.3区域的rs4977950变异与SBP关联( $P=9.1\times 10^{-8}$ ),此外,2q24.3区域*STK39*基因的多个多态与血压水平关联( $P_{\max}=8.9\times 10^{-6}$ ),该基因能够调控肾脏盐运输蛋白质的表达,序列变异可能影响肾脏对钠盐的排泄过程而与血压升高有关。同年,Org等<sup>[10]</sup>应用Affymetrix 500 K芯片分别在第一阶段1644名、第二阶段1830名和第三阶段1823名欧洲裔研究对象中检测发现*CDH13*基因rs11646213变异与高血压呈显著相关( $P=5.3\times 10^{-8}$ ),该基因调控H-钙粘连蛋白的表达,参与血管壁重构和生成。同年Roslin等<sup>[11]</sup>应用Affymetrix 500 K芯片对来自FHS研究中的1659名研究对象进行研究,发现位于7p11区域的rs11976165 T/C变异与SBP水平呈显著相关( $P=6.88\times 10^{-8}$ ),此位点接近*TRIM60P16*基因。随后*Nature Genetics*杂志报道了美国CHARGE研究组<sup>[12]</sup>和欧洲Global BPgen协作组<sup>[13]</sup>的研究结果,*CYP17A1*、*CSK/CYP1A2/ULK3*、*SH2B3*等基因的遗传变异和血压水平显著相关,该结果在随后的两年被多个研究组重复验证。Levy等<sup>[12]</sup>在29136名白人中发现4个SBP相关基因变异(*ATP2B1*、*CYP17A1*、*PLEKHA7*和*SH2B3*),6个DBP相关变异(*ATP2B1*、*CACNB2*、*CSK-ULK3*、*SH2B3*、*TBX3-TBX5*和*ULK4*),1个高血压相关变异(*ATP2B1*)。Newton-Cheh等<sup>[13]</sup>采用两阶段法,在第一阶段的34433名欧洲裔美国人中,应用Affymetrix 500K芯片检测血压相关SNPs,然后在第二阶段不超过71225名欧洲裔美国人中和12889名印度-亚洲裔研究对象中进行验证研究,发现了8个SNP位点与SBP或DBP相关,相邻的基因分别为*CYP17A1*、*CYP1A2*、*FGF5*、*SH2B3*、*MTHFR*、*c10orf107*、*ZNF652*和*PLCD3*( $P<5\times 10^{-8}$ )。

2011年,Zhu等<sup>[14]</sup>采用“混合定位”(Admixture mapping)分析方法以混合人群这一特殊遗传学特征进行基因定位,发现rs295796、rs295813、rs1033225、rs1033224、rs2236611与SBP相关,rs4957217与DBP相关( $P_s<10^{-5}$ )。随后研究人员在一个多中心的大样本人群中( $N=11882$ )对这6个SNPs进行验证并进行

Meta分析,发现位于5p13.3区域的rs7726475与血压水平关联(SBP: $P=4.45\times 10^{-7}$ ;DBP: $P=7.52\times 10^{-7}$ ),由此认为rs7726475是一个新的高血压易感位点。Ho等<sup>[15]</sup>也报道了一个新的高血压易感位点,研究者应用Illumina's Infinium II分析平台,对来自女性基因组健康研究队列中的23019名女性进行分析,将2009年CHARGE研究组和Global BPgen协作组报道的16个SNPs( $P_s<5.0\times 10^{-8}$ )、18个具有潜在意义的位点( $5.0\times 10^{-8}<P<5.6\times 10^{-5}$ ),以及Wang等<sup>[9]</sup>和Org等<sup>[10]</sup>报道的两个血压水平相关位点(位于*STK39*和*CDH13*)进行了重复验证研究,并将结果与之前报道的结果进行Meta分析,发现*BLK-GATA*区域的rs2898290与SBP( $P=4.32\times 10^{-8}$ )相关。2011年9月,*Nature*杂志刊登了“血压全基因组关联研究国际协会组织”的研究成果<sup>[16]</sup>。研究人员采用多阶段法对20万名欧洲裔研究对象进行了全基因组扫描,新发现了17个血压相关遗传位点(rs2932538、rs13082711、rs4190763、rs13107325、rs13139571、rs1173771、rs11953630、rs1799945、rs805303、rs4373814、rs932764、rs7129220、rs633185、rs2521501、rs17608766、rs1327235和rs6015450),其中6个位点位于已知的高血压候选基因之中(*GUCYIA3*、*GUCYIB3*、*NPR3-C5orf23*、*ADM*、*FURIN-FES*、*GNAS-EDN3*和*GOSR2*),在东亚、南亚和非洲裔美国人群中也有类似发现。研究人员还发现这些变异与左心室壁厚度、中风和冠状动脉病有关。

## 1.2 高血压 GWAS 在美国人中的研究结果

2009年Dermitzakis等<sup>[17]</sup>采用Affymetrix 6.0芯片在1017名非洲裔美国人中发现6个SNPs的遗传变异与SBP水平存在关联,分别是位于2q32.2区域的*PMS1*基因rs5743185( $P=2.09\times 10^{-11}$ ),6p22.3区域的*AL365265.23*基因rs16877320( $P=3.42\times 10^{-9}$ ),4q32.12区域的*SLC24A4*基因rs11160059( $P=1.54\times 10^{-8}$ ),8q22.3区域的*YWHAZ*基因rs17365948( $P=1.59\times 10^{-8}$ ),11p15.4区域的*IPO7*基因rs12279202( $P=4.80\times 10^{-8}$ )以及16p13.3区域的*CACNA1H*基因rs3751664( $P=6.71\times 10^{-8}$ )。

2011年Fox等<sup>[18]</sup>应用Affymetrix 6.0芯片在8591名非洲裔美国人中发现*GPR98-ARRDC3*区域的rs10474346(DBP: $P=3.6\times 10^{-8}$ )和*C21orf91*基因

rs2258119(SBP : $P=4.7\times 10^{-8}$ )与血压水平相关。但是,这两个位点在另外两个样本量分别为 11 882(非洲裔美国人)和 69 899(欧洲裔美国人)的人群中都没有得到验证。此外,该研究还验证了欧洲裔美国人中发现的 8 个阳性位点,提示*SH2B3* 基因rs3184504 ( $P=0.009$ )、*TBX3-TBX5* 区域rs2384550( $P=0.03$ )和*CSK-ULK3* 区域rs6495112( $P=0.0004$ )与SBP水平相关。

### 1.3 高血压 GWAS 在中国人群中的研究结果

2010 年,我国学者Niu等<sup>[19]</sup>采用候选基因策略对CHARGE 研究组所报道的 8 个最具统计学显著性的血压水平相关位点在中国人群中进行了重复验证。研究选取了 548 名高血压患者和 560 名年龄、性别相匹配的对照。结果发现*FGF5* 基因rs16998073 ( $P=0.002$ )和*ZNF652* 基因rs16948048( $P=0.037$ )的频率在高血压组和对照组中的分布具有统计学差异。Liu等<sup>[20]</sup>也选择了 8 个CHARGE研究组和Global BPgen协作组报道的易感基因/区域进行验证研究( $N=3\ 210$ ),发现在中国人群中*FGF5* 基因rs16998073 (SBP : $P=6\times 10^{-4}$ ; DBP : $P=8\times 10^{-4}$ ; 高血压 : $P=1\times 10^{-4}$ ), *CYP17A1* 基因rs11191548(SBP : $P=5\times 10^{-3}$ ; DBP : $P=0.07$ ; 高血压 : $P=9\times 10^{-3}$ ), *CYP17A1* 基因rs1004467(SBP : $P=5\times 10^{-3}$ ; DBP : $P=0.01$ ; 高血压 : $P=9\times 10^{-4}$ ), 和*MTHFR*基因rs17367504(DBP : $P=0.02$ )与血压水平相关。

2011 年 9 月, Lin等<sup>[21]</sup>报道了在中国畲族人群中( $N=4\ 460$ )*CYP17A1*、*CACNB2* 和*PLEKHA7* 基因的遗传变异与血压/高血压相关。研究人员对欧洲人群中报道过的 7 个阳性位点进行验证(*CYP17A1*、*PLEKHA7*、*CACNB2*、*ATP2B1*、*TBX3-TBX5*、*CSK-ULK3* 和 *SH2B3*), 发现经年龄、性别和BMI调整后, *CYP17A1* 基因的rs11191548 与血压水平显著相关 (SBP : $P=0.002$ ; DBP : $P=0.002$ ), *CACNB2* 基因的rs11014166 与SBP水平和高血压显著相关 (SBP : $P=0.006$ ; 高血压 : $P=0.021$ ), 位于*PLEKHA7* 基因的rs381815 位点与高血压相关( $P=0.046$ )。

### 1.4 高血压 GWAS 在日本人群中的研究结果

2010 年, Tabara等<sup>[22]</sup>在 14 105 名日本人中选择了Global BPgen和CHARGE 两大协作组报道的 7 个MAF>0.1 的位点进行验证, 结果发现*ATP2B1* 基因

rs11105378 与高血压的发生呈显著相关 ( $P=4.1\times 10^{-11}$ ), 经与Global BPgen<sup>[12]</sup>研究结果进行Meta分析, 发现*FGF5* 基因rs1458038( $P=1.1\times 10^{-11}$ )、*CYP17A1* 基因rs1004467( $P=6.2\times 10^{-9}$ )和*CSK*基因rs1378942( $P=4.2\times 10^{-8}$ )与血压水平相关。随后, Takeuchi等<sup>[23]</sup>群中再次验证了上述 7 个位点与高血压的关系。研究人员首先应用Illumina 550 磁珠阵列分析平台对 1 526 名研究对象的 456 825 个SNPs进行全基因组扫描, 然后在一个大样本人群中对中国全基因组关联研究中已报道的 27 个位点进行二阶段的重重复验证研究, 结果发现位点rs880315 (位于1p36.22 区域的*CASZ1* 基因, SBP : $P=2.2\times 10^{-8}$ ; DBP : $P=4.9\times 10^{-12}$ )、rs17367504(位于 1p36.22 区域的*MTHFR* 基因, SBP : $P=0.03$ ; DBP : $P=0.05$ )、rs155524(位于 3p22.2 区域的*ITGA9* 基因, SBP : $P=0.05$  DBP : $P=0.07$ )、rs16998073(位于 4q21.21 区域的*FGF5* 基因, SBP : $P=3.1\times 10^{-14}$ ; DBP : $P=1.9\times 10^{-12}$ )、rs12413409(位于 10q24.32 区域*CYP17A1* 基因与*CNNM2* 基因之间, SBP : $P=1.4\times 10^{-14}$ ; DBP : $P=2.9\times 10^{-10}$ )、rs2681472(位于 12q21.33 区域的*ATP2B1* 基因, SBP : $P=1.5\times 10^{-7}$ ; DBP : $P=1.2\times 10^{-4}$ )和rs1378942(位于 15q24.1 区域*CSK*基因和*ULK3* 基因之间, SBP : $P=0.05$ ; DBP : $P=0.009$ )与血压水平相关。

2011 年, Miyaki等<sup>[24]</sup>名日本人中对已报道的 10 个基因的 14 个阳性多态进行验证, 发现*CYP17A1* 基因的rs11191548(SBP : $P=0.026$ ; DBP : $P=0.013$ ; 高血压 : $P=0.032$ )、*COMT* 基因的rs4680 (SBP : $P<0.001$ ; DBP : $P<0.001$ ; 高血压 : $P=0.001$ )和rs4633(SBP : $P<0.001$ ; DBP : $P<0.001$ ; 高血压 : $P=0.003$ )、*ATP2B1* 基因的rs17249754(SBP : $P=0.015$ ; 高血压 : $P=0.008$ )以及*CSK*基因的rs1378942(SBP : $P=0.009$ )与血压/高血压水平相关。

### 1.5 高血压 GWAS 在韩国人群中的研究结果

2009 年Cho等<sup>[25]</sup>项在韩国人群中展开的大规模全基因组关联研究(一阶段 $N=8\ 848$ , 二阶段  $N=7\ 861$ ), 发现邻近 12q21 区域的*ATP2B1* 基因rs17249754 遗传变异可能与SBP水平有关( $P=1.3\times 10^{-7}$ ), 该基因编码血管内皮细胞里泵出钙离子的蛋白质。

2010年, Park等<sup>[26]</sup>34名韩国人中发现12q21.33区域*ATP2B1*基因rs17249754(SBP:  $P=2.53 \times 10^{-10}$ ; DBP:  $P=1.30 \times 10^{-9}$ )和rs7136259(SBP:  $P=1.28 \times 10^{-8}$ ; DBP:  $P=6.41 \times 10^{-8}$ )与SBP和DBP的关联均达到基因组显著性水平。同年, Hong等<sup>[27]</sup>*ATP2B1*基因rs17249754、*CSK*基因、*ARSG*基因和*CSMD1*基因与血压水平相关。该研究首先在8 842名研究对象中进行了第一阶段的GWAS分析, 发现10个与高血压显著相关的阳性SNPs, 接着在7 861名研究对象中进行了二阶段的重复验证, 结果表明12q21.33区域*ATP2B1*基因rs17249754、15q24.1区域*CSK*基因rs1378942、17q24.2区域*ARSG*基因rs12945290位点的遗传变异与高血压存在关联, 其中证据最强的是rs17249754( $P=4.25 \times 10^{-9}$ )。将两阶段的数据合并进行Meta分析, 发现位于8p23.2区域*CSMD1*基因的rs995322显著增加高血压患病危险( $P=1.00 \times 10^{-4}$ )。

2011年, Hong等<sup>[28]</sup>Affymetrix 5.0芯片对7 551名韩国人的352 228个SNPs进行初筛, 然后将有意义的20个SNPs应用Affymetrix 6.0芯片在3 703名研究对象中进行验证, 发现*AKAP13*基因rs11638762

与血压水平相关, 经功能研究证实其参与血压水平调控(SBP:  $P=1.4 \times 10^{-5}$ ; DBP:  $P=6.3 \times 10^{-4}$ )。同年6月, 由中、日、韩等6个国家的研究人员联合在*Nature Genetics*杂志上报道了13个东亚地区人群中与血压有关的位点<sup>[29]</sup>6个为首次发现。该研究采用多阶段研究设计, 首先对国际上8个研究小组高血压GWAS的数据进行Meta分析(共计19 068名研究对象), 发现了约240个可能与血压水平有关的序列变异, 第二阶段对10 618名日本人的13个SNPs进行基因分型, 最后将一、二阶段的数据进行联合分析, 继而在独立的20 247名亚洲人中进行了第三阶段的进一步验证, 最终确认了6个与血压有关的SNPs(rs17030613、rs16849225、rs6825911、rs1173766、rs11066280和rs35444)。

对近年来高血压GWAS结果进行归纳与总结(表1、图1), 经统计rs17249754、rs1378942和rs11191548为3个报道频数最多的与SBP和DBP有

表1 2007年1月~2011年9月GWAS报道的高血压相关基因多态位点

表型	所在染色体位置	基因名	SNP	P值	参考文献
SBP	1p13.2	<i>MOV10</i>	rs2932538	$1.2 \times 10^{-9}$	[16]
		<i>CASZ1</i>	rs880315	$2.2 \times 10^{-8}$	[19]
	1p36.22	<i>MTHFR</i>	rs17367504	$8.7 \times 10^{-22}$	[16]
				0.03	[19]
	2q24.3	<i>STK39</i> <i>FIGN</i> 、 <i>GRB14</i>	cluster of SNPs rs16849225	$2 \times 10^{-13}$	[13]
				$8.9 \times 10^{-6}$	[9]
				$3.5 \times 10^{-11}$	[29]
	2q32.2	<i>PMS1</i>	rs5743185	$2.09 \times 10^{-11}$	[17]
	3p22.1	<i>ULK4</i>	rs3774372	$9.0 \times 10^{-14}$	[16]
	3p22.2	<i>ITGA9</i>	rs155524	0.05	[19]
	3q26.2	<i>MECOM</i>	rs419076	$1.8 \times 10^{-13}$	[16]
	4q21.21	<i>FGF5</i>	rs16998073	$6 \times 10^{-4}$	[20]
				$3.1 \times 10^{-14}$	[23]
	4q24	<i>SLC39A8</i>	rs1458038	$1.5 \times 10^{-23}$	[16]
	4q32.12	<i>SLC24A4</i> <i>GPR98-ARRDC3</i>	rs13107325 rs11160059	$3.3 \times 10^{-14}$	[16]
				$1.54 \times 10^{-8}$	[17]
				$4.45 \times 10^{-7}$	[14]
	5p13.3	<i>NPR3-C5orf23</i> <i>NPR3</i>	rs1173771	$1.8 \times 10^{-16}$	[16]
				$1.9 \times 10^{-8}$	[29]
	5q33.3	<i>EBF1</i>	rs11953630	$3.0 \times 10^{-11}$	[16]

表型	所在染色体位置	基因名	SNP	P 值	参考文献
	6p21.33	BAT2-BAT5	rs805303	$1.5 \times 10^{-11}$	[16]
SBP	6p21.33	CACNB2(3')	rs4373814	$4.8 \times 10^{-11}$	[16]
	6p22.2	HFE	rs1799945	$7.7 \times 10^{-12}$	[16]
	6p22.3	AL365265.23	rs16877320	$3.42 \times 10^{-9}$	[17]
	7p11.2	邻近 TRIM60P16	rs11976165	$6.88 \times 10^{-8}$	[11]
	8q22.3	YWHAZ	rs17365948	$1.59 \times 10^{-8}$	[17]
	8p23.1	BLK-GATA	rs2898290	$4.32 \times 10^{-8}$	[10]
	9p21.3	Not known	rs4977950	$9.1 \times 10^{-8}$	[9]
	10p12.31	CACNB2(3')	rs1813353	$2.6 \times 10^{-12}$	[16]
	10p12.32	CYP17A1-NT5C2	rs11191548	$6.9 \times 10^{-26}$	[16]
			rs1004467	$1 \times 10^{-10}$	[12]
	10q24.3	CYP17A1		$5 \times 10^{-3}$	[20]
			rs11191548	$5 \times 10^{-3}$	[20]
				0.026	[24]
				0.002	[21]
	10p12.31	CACNB2	rs11014166	0.006	[21]
	10p12.33	PLCE1	rs932764	$7.1 \times 10^{-16}$	[16]
	10q24.32	CYP17A1-CNNM2	rs12413409	$1.4 \times 10^{-14}$	[23]
				$2 \times 10^{-9}$	[12]
	11p15.1	PLEKHA7	rs381815	$5.3 \times 10^{-11}$	[16]
	11p15.4	IPO7	rs12279202	$4.80 \times 10^{-8}$	[17]
	11p15.4	ADM	rs7129220	$3.0 \times 10^{-12}$	[16]
	11q22.1	FLJ32810-TMEM133	rs633185	$1.2 \times 10^{-17}$	[16]
				$1.8 \times 10^{-18}$	[16]
				$1.3 \times 10^{-7}$	[25]
			rs17249754	0.015	[24]
	12q21.33	ATP2B1		$2.53 \times 10^{-10}$	[26]
			rs2681472	$1.5 \times 10^{-7}$	[23]
			rs7136259	$1.28 \times 10^{-8}$	[26]
			rs2681492	$4 \times 10^{-11}$	[12]
	12q24.12	SH2B3	rs3184504	$5 \times 10^{-9}$	[12]
				$3.8 \times 10^{-18}$	[16]
	12q24.13	RPL6, PTPN11, ALDH2	rs11066280	$7.9 \times 10^{-31}$	[29]
	12q24.21	TBX5-TBX3	rs10850411	$5.4 \times 10^{-8}$	[16]
				$1 \times 10^{-23}$	[13]
	15q24.1	CSK	rs1378942	$5.7 \times 10^{-23}$	[16]
				0.009	[24]
				0.05;	[23]
	15q25.3	AKAP13	rs11638762	$1.3 \times 10^{-5}$	[28]
	15q26.1	FURIN-FES	rs2521501	$5.2 \times 10^{-19}$	[16]
	16p13.3	CACNA1H	rs3751664	$6.71 \times 10^{-8}$	[17]
	17q21.31	PLCD3	rs12946454	$1 \times 10^{-8}$	[13]
	17q21.32	GOSR2	rs17608766	$1.1 \times 10^{-10}$	[16]
	17q21.33	ZNF652	rs12940887	$1.8 \times 10^{-10}$	[16]
	20q12.32	GNAS-EDN3	rs6015450	$3.9 \times 10^{-23}$	[16]
	21q21.1	C21orf91	rs2258119	$4.7 \times 10^{-8}$	[18]
	22q11.21	COMT	rs4680	<0.001	[24]

续表 1

			rs4633	<0.001	[24]	
续表 1						
表型	所在染色体位置	基因名	SNP	P 值	参考文献	
DBP	1p13.2	<i>MOV10</i>	rs2932538	$1.2 \times 10^{-9}$	[16]	
		<i>ST7L</i> 、 <i>CAPZA1</i>	rs17030613	$1.2 \times 10^{-8}$	[29]	
	1p36.22	<i>CASZ1</i>	rs880315	$4.9 \times 10^{-12}$	[19]	
				$3.5 \times 10^{-19}$	[16]	
		<i>MTHFR</i>	rs17367504	0.05	[19]	
				0.02	[20]	
	3p22.1	<i>ULK4</i>	rs9815354	$3 \times 10^{-9}$	[12]	
				$9.0 \times 10^{-14}$	[16]	
	3p22.2	<i>ITGA9</i>	rs155524	0.07	[19]	
	3p24.1	<i>SLC4A7</i>	rs13082711	$3.8 \times 10^{-9}$	[16]	
	3q26.2	<i>MECOM</i>		rs419076	$2.1 \times 10^{-12}$	[16]
					$1 \times 10^{-21}$	[13]
	4q21.21	<i>FGF5</i>		rs16998073	$9 \times 10^{-4}$	[20]
					$1.9 \times 10^{-12}$	[23]
				rs1458038	$8.5 \times 10^{-25}$	[16]
	4q24	<i>SLC39A8</i>		rs13107325	$2.3 \times 10^{-17}$	[16]
	4q25	<i>ENPEP</i>		rs6825911	$9.0 \times 10^{-9}$	[29]
	4q32.1	<i>GUCY1A3-GUCY1B3</i>		rs13139571	$2.2 \times 10^{-10}$	[16]
	5p13.3	<i>GPR98-ARRDC3</i>		rs7726475	$7.52 \times 10^{-7}$	[14]
	5p13.3	<i>NPR3-C5orf23</i>		rs1173771	$9.1 \times 10^{-12}$	[16]
	5q14.3	<i>GPR98-ARRDC3</i>		rs10474346	$3.6 \times 10^{-8}$	[18]
	5q33.3	<i>EBF1</i>		rs11953630	$3.8 \times 10^{-13}$	[16]
	6p21.33	<i>BAT2-BAT5</i>		rs805303	$3.0 \times 10^{-11}$	[16]
	6p21.33	<i>CACNB2(3')</i>		rs4373814	$4.4 \times 10^{-10}$	[16]
	6p22.2	<i>HFE</i>		rs1799945	$1.5 \times 10^{-15}$	[16]
	8q13.3	Not known		rs1963982	$3.3 \times 10^{-6}$	[7]
				rs11014166	$1 \times 10^{-8}$	[12]
	10p12.31	<i>CACNB2</i>			0.021	[21]
				rs1813353	$2.3 \times 10^{-15}$	[16]
	10p12.31	<i>C10orf107</i>		rs4590817	$4.0 \times 10^{-12}$	[16]
	10p12.32	<i>CYP17A1-NT5C2</i>		rs11191548	$9.4 \times 10^{-13}$	[16]
					0.07	[20]
				rs11191548	$7 \times 10^{-24}$	[13]
10q24.32	<i>CYP17A1</i>			0.013	[24]	
				0.002	[21]	
			rs1004467	0.01	[20]	
	<i>CYP17A1-CNNM2</i>		rs12413409	$2.9 \times 10^{-10}$	[23]	
11p15.4	<i>ADM</i>		rs7129220	$6.4 \times 10^{-8}$	[16]	
11q22.1	<i>FLJ32810-TMEM133</i>		rs633185	$2.0 \times 10^{-15}$	[16]	
				$5.3 \times 10^{-10}$	[16]	
12q24.13	<i>RPL6</i> 、 <i>PTPN11</i> 、 <i>ALDH2</i>		rs11066280	$1.3 \times 10^{-35}$	[29]	
12q24.21	<i>TBX3-TBX5</i>		rs2384550	0.03	[18]	
				$4 \times 10^{-8}$	[12]	
				$5.4 \times 10^{-10}$	[16]	

		<i>TBX3</i>	<i>rs35444</i>	$1.3 \times 10^{-10}$	[29]
续表 1					
表型	所在染色体位置	基因名	SNP	P 值	参考文献
高血压	12q21.33	<i>ATP2B1</i>	<i>rs17249754</i>	$1.2 \times 10^{-14}$	[16]
				0.008	[24]
			<i>rs2681472</i>	$1.30 \times 10^{-9}$	[26]
				$1.2 \times 10^{-4}$	[23]
	12q24.12	<i>SH2B3</i>	<i>rs7136259</i>	$1 \times 10^{-9}$	[12]
				$6.41 \times 10^{-8}$	[26]
			<i>rs3184504</i>	$3 \times 10^{-14}$	[12]
				$3.6 \times 10^{-25}$	[16]
	15q24.1	<i>CSK-ULK3</i>	<i>rs653178</i>	0.009	[18]
				$3 \times 10^{-18}$	[13]
			<i>rs1378942</i>	$2.7 \times 10^{-26}$	[16]
	15q25.3	<i>AKAP13</i>		0.009	[23]
			<i>rs6495112</i>	0.0004	[18]
	15q26.1	<i>FURIN-FES</i>	<i>rs6495122</i>	$2 \times 10^{-10}$	[12]
	17q21.33	<i>ZNF652</i>	<i>rs11638762</i>	$6.3 \times 10^{-4}$	[28]
	20p12.2	<i>JAG1</i>	<i>rs2521501</i>	$1.9 \times 10^{-15}$	[16]
	20q12.32	<i>GNAS-EDN3</i>	<i>rs12940887</i>	$2.3 \times 10^{-14}$	[16]
	22q11.21	<i>COMT</i>	<i>rs6015450</i>	$5.6 \times 10^{-23}$	[16]
			<i>rs4680</i>	<0.001	[24]
	1p36.22	<i>MTHFR</i>	<i>rs4633</i>	<0.001	[24]
			<i>rs17367504</i>	$2.3 \times 10^{-10}$	[16]
	4q21.21	<i>FGF5</i>	<i>rs1458038</i>	$1.1 \times 10^{-11}$	[22]
			<i>rs16998073</i>	0.002	[19]
	5p13.3	<i>NPR3-C5orf23</i>		$1 \times 10^{-4}$	[20]
	6p21.33	<i>BAT2-BAT5</i>	<i>rs1173771</i>	$3.2 \times 10^{-10}$	[16]
	6p22.2	<i>HFE</i>	<i>rs805303</i>	$3.0 \times 10^{-11}$	[16]
	8p23.2	<i>CSMD1</i>	<i>rs1799945</i>	$1.8 \times 10^{-10}$	[16]
	10p12.31	<i>CACNB2(3')</i>	<i>rs995322</i>	$1.0 \times 10^{-4}$	[27]
	10p12.31	<i>C10orf107</i>	<i>rs1813353</i>	$6.2 \times 10^{-10}$	[16]
			<i>rs4590817</i>	$9.8 \times 10^{-9}$	[16]
	10q24.3	<i>CYP17A1</i>	<i>rs1004467</i>	$6.2 \times 10^{-9}$	[12]
				$9 \times 10^{-4}$	[20]
			<i>rs11191548</i>	$9 \times 10^{-3}$	[20]
	11p15.1	<i>PLEKHA7</i>		0.032	[24]
	11q22.1	<i>FLJ32810-TMEM133</i>	<i>rs381815</i>	0.046	[21]
			<i>rs633185</i>	$5.4 \times 10^{-11}$	[16]
	12q21.33	<i>ATP2B1</i>	<i>rs11105378</i>	$4.1 \times 10^{-11}$	[22]
			<i>rs17249754</i>	$1.1 \times 10^{-14}$	[16]
	13q21.32	Not known	<i>rs2681472</i>	$2 \times 10^{-11}$	[12]
			<i>rs1937506</i>	0.032(EA)	[8]
15q24.1	<i>CSK</i>		0.005(AA)	[8]	
16q23.3	<i>CDH13</i>	<i>rs1378942</i>	$4.2 \times 10^{-8}$	[22]	
17q21.32	<i>ZNF652</i>	<i>rs11646213</i>	$5.3 \times 10^{-8}$	[10]	
20q12.32	<i>GNAS-EDN3</i>	<i>rs16948048</i>	0.037	[19]	
		<i>rs6015450</i>	$4.2 \times 10^{-14}$	[16]	



22q11.21

COMT

rs4680

0.001

[\[24\]](#)

注:EA:欧洲裔美国人;AA:非洲裔美国人。

关的 SNPs。Rs17249754 位于 12q21.33 的 *ATP2B1* 基因,编码 I 型-细胞膜钙离子转运酶,在维持细胞内钙动态平衡中发挥关键作用。Rs1378942 位于 15q24.1 区域的 *CSK* 基因,编码一种酪氨酸激酶 CSK,通过使 PP60C-SRC 激酶 C 端酪氨酸磷酸化而下调后者,使之活性上升,参与细胞的多种功能调控,PP60C-SRC 活性升高可活化 Ras 信号通路,促进细胞增殖。Rs11191548 位于 10q24.3 区域的 *CYP17A1* 基因,该基因编码细胞色素 P450c17 $\alpha$  酶,介导类固醇 17 $\alpha$  羟化酶和 17.20 裂解酶活性,参与胆固醇前体向雄激素或孕激素以及盐皮质激素或糖皮质激素的转化。

由表 2 可见,即使在遗传背景相对接近的亚洲人群中,不同人群的高血压的易感基因、易感位点可能并不一致。中国人群中 *CYP17A1* 基因 rs11191548 与血压水平相关,*ATP2B1* 基因 rs17249754 与韩国人群血压水平显著相关,而在日本人群中,以上 3 个基因的多态位点均与血压水平显著相关。

## 2 高血压 GWAS 方法学进展

### 2.1 表型

尽可能选择测量简单、准确和遗传度高的数量表型,例如多次重复测量血压计算平均值可降低测量误差,增加该数量表型变异可以由遗传因素解释的比例,提高发现关联的可能性<sup>[30]</sup>7 年,首个高血压 GWAS WTCCC 研究<sup>[6]</sup>了“高血压”这个质量表型分析其与遗传标记之间的关系,结果没有发现任何

遗传变异与高血压之间的关联达到基因组显著性水平。次年,美国 FBPP 研究组<sup>[8]</sup>压水平和高血压同时作为疾病表型对 WTCCC 研究中最具显著性的 6 个序列变异进行验证,数量表型 SBP 与 DBP 的应用增加了研究的效能,结果发现 rs1937506 与高血压相关。

### 2.2 设计类型

高血压 GWAS 在研究设计上经历了从单阶段到两阶段以及多阶段的发展历程。2007 年的 WTCCC 研究采用了单阶段设计,一次性在 17 000 名研究对象中进行基因分型(其中针对高血压易感基因研究包括 2 000 病例和 3 000 共用对照),然后分析每个 SNP 与疾病的关联,此种设计缺点为工作量太大,基因分型耗资巨大。从 2009 年以后,两阶段设计备受研究者的青睐。研究者首先采用覆盖整个基因组的高通量 SNP 分型技术在一阶段的小样本中对全基因组范围内的 SNPs 进行基因分型,筛选出少量的阳性 SNPs 进行重复验证,然后结合两个阶段的结果进行分析。例如 Newton-Cheh 等<sup>[13]</sup>采用了二阶段设计,首先在 34 433 名欧洲裔美国人中进行第一阶段的检测,然后在第二阶段对 71 225 名欧洲裔美国人和 12 889 名印度-亚洲裔研究对象中进行重复验证研究,发现了 8 个 SNPs 与血压水平相关。而 Org 等<sup>[10]</sup>了多阶段法,发现了 *CDH13* 基因 rs11646213 与高血压相关。两阶段或多阶段的研究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。然而第一阶段筛选

表 2 亚洲人群高血压全基因组关联研究结果

人种	表型	所在染色体位置	基因名	SNP	$\beta$	SE	P 值	参考文献
中国人群	SBP	1p36.22	<i>CASZ1</i>	rs880315	-	-	$2.2 \times 10^{-8}$	<a href="#">[19]</a>
			<i>MTHFR</i>	rs17367504	-	-	0.03	<a href="#">[19]</a>
		3p22.2	<i>ITGA9</i>	rs155524	-	-	0.05	<a href="#">[19]</a>
			<i>FGF5</i>	rs16998073	1.97	1.38 ~ 2.55	$6 \times 10^{-4}$	<a href="#">[20]</a>
			<i>CACNB2</i>	rs11014166	-0.19	-0.88 ~ 0.5	0.006	<a href="#">[21]</a>
	10q24.3	rs1004467	<i>CYP17A1</i>	rs11191548	1.62	1.05 ~ 2.19	$5 \times 10^{-3}$	<a href="#">[20]</a>
				-1.58	-2.09 ~ -1.07	0.002	<a href="#">[21]</a>	
		10q24.3	<i>CYP17A1</i>	rs11191548	1.74	1.12 ~ 2.36	$5 \times 10^{-3}$	<a href="#">[20]</a>
				rs11191548	-1.58	-2.09 ~ -1.07	0.002	<a href="#">[21]</a>
				rs11191548	-1.58	-2.09 ~ -1.07	0.002	<a href="#">[21]</a>
DBP	1q36.22	<i>CASZ1</i>	rs880315	-	-	$4.9 \times 10^{-12}$	<a href="#">[19]</a>	
		<i>MTHFR</i>	rs17367504	-	-	0.05	<a href="#">[19]</a>	

人种	表型	所在染色体位置	基因名	SNP	$\beta$	SE	P 值	参考文献
		3p22.2	<i>ITGA9</i>	rs155524	1.03	0.58 ~ 1.48	0.02	[20]
					-	-	0.07	[19]
续表 2								
日本人群	SBP	4q21.21	<i>FGF5</i>	rs16998073	0.95	0.66 ~ 1.24	$9 \times 10^{-4}$	[20]
				rs11191548	0.55	0.25 ~ 0.85	0.07	[20]
				rs11191548	-0.94	-1.25 ~ -0.63	0.002	[21]
		10q24.3	<i>CYP17A1</i>	rs1004467	0.70	0.42 ~ 0.98	0.01	[20]
				rs1004467	1.36	1.16 ~ 1.60	$1 \times 10^{-4}$	[20]
				rs1004467	1.31	1.12 ~ 1.53	$9 \times 10^{-4}$	[20]
		17q21.32	<i>ZNF652</i>	rs16948048	-	-	0.037	[19]
		4q21.21	<i>FGF5</i>	rs16998073	1.51	1.12 ~ 1.89	$3.1 \times 10^{-14}$	[23]
		10q24.3	<i>CYP17A1</i>	rs11191548	3.05	-	0.026	[24]
		10q24.32	<i>CYP17A1-CNNM2</i>	rs12413409	1.58	1.18 ~ 1.98	$1.4 \times 10^{-14}$	[23]
				rs17249754	2.29	-	0.015	[24]
		12q21.33	<i>ATP2B1</i>	rs2681472	0.99	0.62 ~ 1.35	$1.5 \times 10^{-7}$	[23]
				rs2681472	0.99	0.62 ~ 1.35	$1.5 \times 10^{-7}$	[23]
		15q24.1	<i>CSK</i>	rs1378942	-4.20	-	0.009	[24]
				rs1378942	0.44	-0.01 ~ 0.89	0.05;	[23]
				rs4680	4.16	-	<0.001	[24]
		22q11.21	<i>COMT</i>	rs4633	9.15	-	<0.001	[24]
	DBP	4q21.21	<i>FGF5</i>	rs16998073	0.82	0.59 ~ 1.05	$1.9 \times 10^{-12}$	[23]
				rs11191548	2.24	-	0.013	[24]
		10q24.32	<i>CYP17A1-CNNM2</i>	rs12413409	0.76	0.53 ~ 1.00	$2.9 \times 10^{-10}$	[23]
		12q21.33	<i>ATP2B1</i>	rs2681472	0.43	0.21 ~ 0.64	$1.2 \times 10^{-4}$	[23]
		15q24.1	<i>CSK-ULK3</i>	rs1378942	0.35	0.09 ~ 0.62	0.009	[23]
				rs4680	2.38	-	<0.001	[24]
		22q11.21	<i>COMT</i>	rs4633	5.96	-	<0.001	[24]
				rs17249754	2.12	1.22 ~ 3.70	0.008	[24]
				rs17249754	-1.064	-1.266 ~ -0.863	$1.3 \times 10^{-7}$	[25]
韩国人群	SBP	12q21.33	<i>ATP2B1</i>	rs17249754	-1.63	-2.13 ~ 1.13	$2.53 \times 10^{-10}$	[26]
				rs7136259	-1.55	-2.05 ~ 1.05	$1.28 \times 10^{-8}$	[26]
		15q25.3	<i>AKAP13</i>	rs11638762	-0.89	-1.09 ~ 0.69	$1.3 \times 10^{-5}$	[28]
	DBP	12q21.33	<i>ATP2B1</i>	rs17249754	-0.94	-1.26 ~ -0.61	$1.30 \times 10^{-9}$	[26]
				rs7136259	-0.88	-1.20 ~ -0.56	$6.41 \times 10^{-8}$	[26]
		15q25.3	<i>AKAP13</i>	rs11638762	-0.48	-0.62 ~ -0.34	$6.3 \times 10^{-4}$	[28]
东亚人群	SBP	2q24.3	<i>FIGN, GRB14</i>	rs16849225	0.75	0.64 ~ 0.86	$3.5 \times 10^{-11}$	[29]
		5p13.3	<i>NPR3</i>	rs1173766	0.63	0.52 ~ 0.74	$1.9 \times 10^{-8}$	[29]
		12q24.13	<i>RPL6, PTPN11, ALDH2</i>	rs11066280	1.56	1.43 ~ 1.69	$7.9 \times 10^{-31}$	[29]
	DBP	1p13.2	<i>ST7L, CAPZA1</i>	rs17030613	0.38	0.31 ~ 0.45	$1.2 \times 10^{-8}$	[29]
		4q25	<i>ENPEP</i>	rs6825911	0.39	0.32 ~ 0.46	$9.0 \times 10^{-9}$	[29]

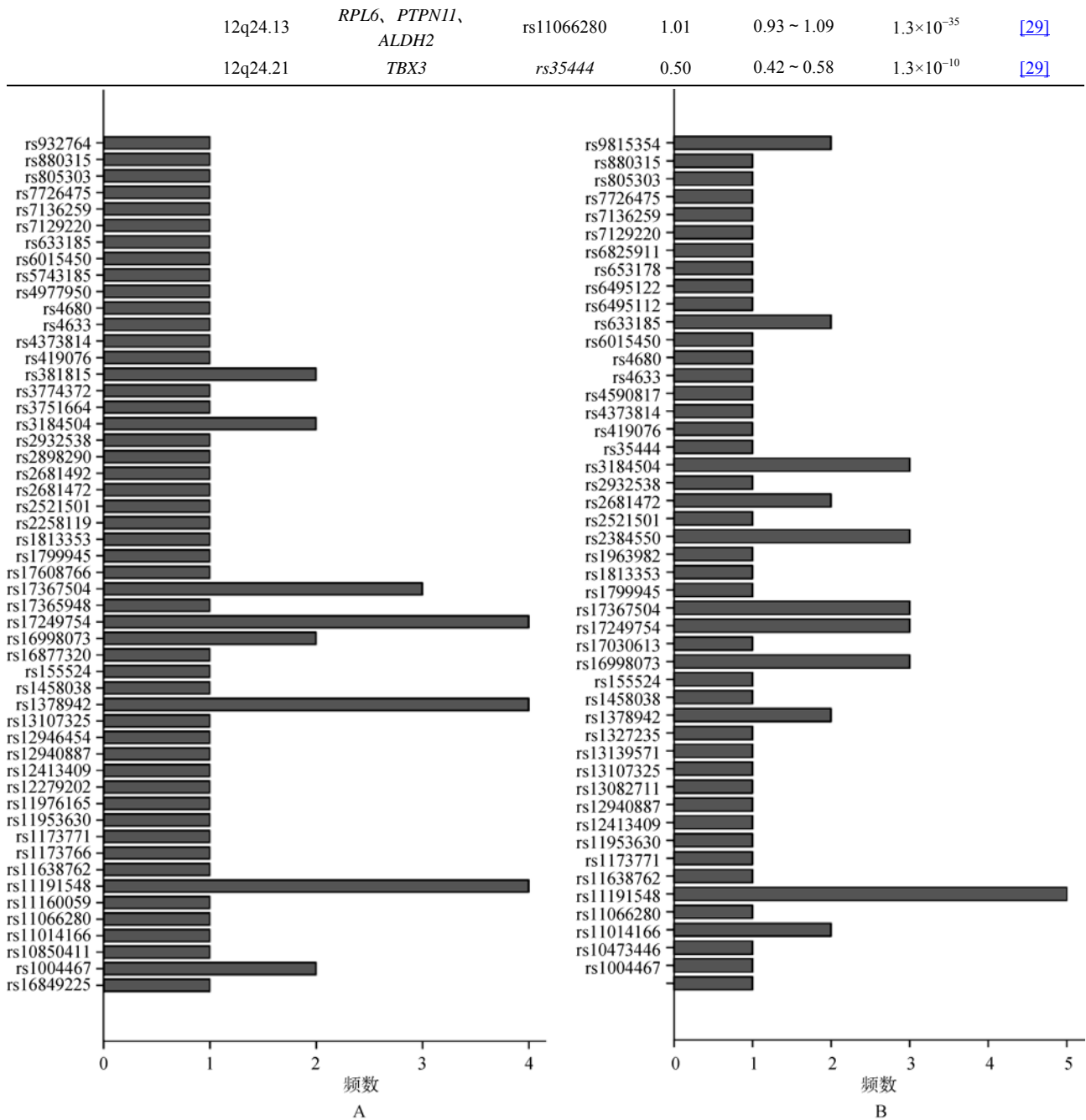


图 1 高血压全基因组关联研究 SBP(A)、DBP(B)阳性 SNP 报道频数分布图

通常样本量较小, 检验效能较低, 可能无法最大限度地发现与疾病或者表型关联的SNPs。为了解决这个问题, 在扩大GWAS样本量的同时, 可适当放宽第一阶段筛选SNPs的标准, 扩大筛选范围, 以此来发现更多的易感基因。例如 2011 年Zhu等<sup>[14]</sup>段中以  $P < 10^{-5}$  为检验水准, 发现 6 个位点与血压水平相关,

随后研究人员对这 6 个SNPs在一个多中心的大样本人群中验证, 发现位于 5p13.3 区域的rs7726475 位点与血压水平相关。

### 2.3 样本量与 GWAS 的 Meta 分析

早期的高血压GWAS在一阶段与二阶段的样本量多为 1 000~2 000 左右, 所发现的疾病相关SNPs

基本上属于常见SNPs(MAF>0.2), 对于低频的SNPs与疾病的关联(MAF<0.05), 由于统计分析检验效能较低较难被检出。因此通过增大样本量来提高检验效能是最直接和有效的方法。通过增加第一阶段基因分型的样本量, 不仅有助于发现基因频率较低的SNPs, 还有助于发现效应值较弱的SNPs。Hong等<sup>[27]</sup>和 Cho等<sup>[23]</sup>和二阶段均采用了较大的样本量(一阶段N=8 848, 二阶段 N=7 861), 同样Fox等<sup>[31]</sup>等<sup>[14]</sup>阶段都使用了较大的样本(一阶段N>6 000, 二阶段N>8 000), 发现了多个高血压相关的SNPs。除了增大单个研究样本量的方法, Meta分析也被研究者们广泛应用于高血压的易感基因研究<sup>[32]</sup>。Meta分析通过合并多个研究的效应指标来间接增加样本量, 是提高检验效能的有效方法。2009年Wang等<sup>[9]</sup>首个高血压GWAS的Meta分析, 他们首先在 542名研究对象中进行一阶段的分析, 然后将 6个人群的数据(N=7 125)合并后进行Meta分析, 发现STK39基因与血压水平显著关联。2009年, CHARGE和Global BpGen<sup>[12,13]</sup>GWAS分别报道了8个高血压相关的遗传变异, 研究者们对这两项研究的其他SNPs的统计量、标准误和P值进行Meta分析(N>60000), 又发现了5个SNPs与高血压相关。自此, Meta分析成为高血压遗传学研究的一种常用的高效且经济的方法。值得注意的是, 研究结果的相互合并分析需要基于一个前提, 即研究之间的同质性, 如研究对象、表型质量、测量方法、基因分型方法等, 因此对于国家间、地区间合作研究提出了更高的要求。

#### 2.4 多重比较

GWAS在全基因组范围内同时检测数以百万计的DNA序列变异与疾病的关联, 多重检验带来的假阳性概率增加是不可回避的问题。若每一次检验水准 $\alpha$ 仍然为5%, 犯I类错误的概率将以 $\alpha$ 的递增函数 $(1-(1-\alpha)^k)$ (k为比较的次数, 即为SNP的个数)增长。统计学上可以通过校正P值来降低假阳性率。常用的多重检验P值校正方法有Bonferroni法、递减调整法(Step-Down Adjustment)、模拟运算法(Permutation)和控制错误发现率法(False discovery rate, FDR), 4种方法的校正严格程度依次递减, 假阳性概率依次递增<sup>[30]</sup>、最保守的方法是Bonferroni法, 被应用于多个高血压GWAS中<sup>[11,14,20,23]</sup>真正的阳性位点失去关联。例

如2009年Wang等<sup>[9]</sup> Bonferroni校正使得原本存在关联的STK39基因不再具有统计学显著性。模拟运算方法先对未校正的P值排序, 然后依据基因之间结构上的关系, 通过反复抽样模拟运算, 分析P值的分布, 对所有的P值同时进行校正<sup>[33]</sup>。马尔可夫链蒙特卡罗(Markov chain Monte Carlo)算法和贝叶斯(Bayesian)方法也可以用来处理这个问题<sup>[34]</sup>。11年Hong等和Zhu等<sup>[14,28]</sup>了Bayesian方法进行分析。FDR法在4种方法中最为宽松, 通过控制错误发现率来调整P值, 首先将未校正的P值从小到大排序, 最大的P值保持不变, 其他的P值依次乘以系数(位点总数/该P值的位次)。例如Lin等<sup>[21]</sup> R法对研究数据进行分析, 在中国畲族人群中发现了CYP17A1基因、CACNB2基因和 PLEKHA7基因的遗传变异与血压/高血压相关。另有学者将Bonferroni法和FDR法结合使用<sup>[26]</sup> ATP2B1基因rs17249754位点和rs7136259位点与血压水平相关。

#### 2.5 验证研究

GWAS中为了控制假阳性, 常常采用重复研究来验证初期结果。验证研究所选人群可以和探索阶段的研究人群相一致, 也可以在不同人群或不同人种的人群中展开, 可以是内部验证, 也可是外部验证, 可以借助于已有的同类GWAS资料进行验证, 也可以是针对小规模或在某一DNA片段中更高密度的分型<sup>[35]</sup> RGE研究和Global BpGen研究在2009年报道的研究结果先后被6个研究组进行重复<sup>[15,19,20,22,23,31]</sup> 美国人, 其中MTHFR、FGF5、CYP17A1、CSK、ATP2B1基因均得到了验证。此外CHARGE和Global BpGen研究组还采用了另一种重复研究方法, 通过在第一阶段对各自的研究人群中所有SNP进行分析, 然后在第二阶段将对方的阳性结果在各自独立的人群中进行重复, 并且将两个样本合并进行Meta分析。这种做法最大限度控制了研究的假阴性和假阳性率, 具有较高的检验效能。

#### 2.6 基因型填补

基因型填补是指依据已检测位点基因型信息对数据缺失位点或未分型位点进行基因型推测的方法。GWAS中基因芯片分型的位点大约为10万~100万个, 芯片的分型缺失可能造成一些与疾病真正有关的位点被遗漏。对缺失数据进行填补可以提高

全基因组遗传标记的覆盖率,筛查出更多与疾病关联的位点。目前常用的软件有 BEAGLE、IMPUTE、MACH和PLINK等<sup>[32]</sup>。在高血压GWAS中,Levy等<sup>[12]</sup>和Newton-Cheh等<sup>[13]</sup>在2009年首次在研究中使用了基因型填补法,2011年Zhu等<sup>[14]</sup>也应用了该方法对检测数据进行基因填补,发现了一个新的高血压相关位点rs772647。基因型填补法目前已在GWAS中得到广泛使用,提高了研究的效率,节约了研究成本。

### 3 展望

随着人类基因组计划和人类单体型图谱计划的完成、高通量基因分型技术和统计方法的不断发展,人类在探索高血压病遗传机制的过程中不断取得突破性的进展。尽管如此,GWAS仍然存在一些不足之处:(1)GWAS尚未能揭示高血压的遗传模式,无法提供新的治疗靶点;(2)通过GWAS发现的血压相关变异多为常见变异,对于少见变异则未能发现其作用;(3)GWAS发现的血压相关变异对人群血压的影响极其微弱,通常仅<1 mm Hg;(4)由于遗传背景的差异常常导致相同位点在不同人群中不能得到重复验证。

因此,在将来的高血压GWAS中,我们仍然可以在以下几个方面进行努力:(1)科学规划,建立国际性或地区性协作组,对数据进行共享进行大规模的Meta分析,通过扩大样本量来增加统计效能;(2)在不同人种、人群中开展GWAS,丰富亚洲人群和非洲人群的数据;(3)结合表达分析、体内体外等功能研究,确定真正的致病位点;(4)充分结合生物信息学方法学的最新进展,如通路分析等,拓展GWAS结果的生物学意义;(5)进行易感位点的精细定位和外显子组深度测序;(6)继续深入研究基因-基因相互作用、基因-环境相互作用;(7)重视表观遗传在高血压发病机制中的作用,积极开展相应的研究。

GWAS时代的到来为我们研究复杂遗传病的发病机制带来了新契机,推动了人类对复杂疾病的认识。目前我国高血压GWAS仍然处于发展阶段,因此我们期待广大科研工作者们共同努力,深入研究高血压的潜在分子机制,为高血压的预防和治疗提供依据。

### 参考文献(References):

- [1] Kearney PM, Whelton M, Reynolds K, Muntner P, Whelton PK, He J. Global burden of hypertension: Analysis of worldwide data. *Lancet*, 2005, 365(9455): 217–223. DOI
- [2] Lawes CMM, Hoorn SV, Rodgers A. Global burden of blood-pressure-related disease, 2001. *Lancet*, 2008, 371(9623): 1513–1518. DOI
- [3] Feinleib M, Garrison RJ, Fabsitz R, Christian JC, Hrubec Z, Borhani NO, Kannel WB, Rosenman R, Schwartz JT, Wagner JO. The NHLBI twin study of cardiovascular disease risk factors: Methodology and summary of results. *Am J Epidemiol*, 1977, 106(4): 284–285. DOI
- [4] Hunt SC, Hasstedt SJ, Kuida H, Stults BM, Hopkins PN, Williams RR. Genetic heritability and common environmental components of resting and stressed blood pressures, lipids, and body mass index in Utah pedigrees and twins. *Am J Epidemiol*, 1989, 129(3): 625–638. DOI
- [5] Pearson TA, Manolio TA. How to interpret a genome-wide association study. *JAMA J Am Med Assoc*, 2008, 299(11): 1335–1344. DOI
- [6] Burton PR, Clayton DG, Cardon LR, Craddock N, Deloukas P, Duncanson A, Kwiatkowski DP, McCarthy MI, Ouwehand WH, Samani NJ. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*, 2007, 447(7145): 661–678. DOI
- [7] Levy D, Larson MG, Benjamin EJ, Newton-Cheh C, Wang TJ, Hwang SJ, Vasan RS, Mitchell GF. Framingham heart study 100K project: Genome-wide associations for blood pressure and arterial stiffness. *BMC Med Genet*, 2007, 8(Suppl. 1): S3. DOI
- [8] Ehret GB, Morrison AC, O'Connor AA, Grove ML, Baird L, Schwander K, Weder A, Cooper RS, Rao DC, Hunt SC, Boerwinkle E, Chakravarti A. Replication of the wellcome trust genome-wide association study of essential hypertension: The family blood pressure program. *Eur J Hum Genet*, 2008, 16(12): 1507–1511. DOI
- [9] Wang Y, O'Connell JR, McArdule PF, Wade JB, Dorff SE, Shah SJ, Shi XL, Pan L, Rampersaud E, Shen HQ, Kim JD, Subramanya AR, Steinle NI, Parsa A, Ober CC, Welling PA, Chakravarti A, Weder AB, Cooper RS, Mitchell BD, Shuldiner AR, Chang YPC. Whole-genome association study identifies *STK39* as a hypertension susceptibility gene. *Proc Natl Acad Sci USA*, 2009, 106(1): 226–231. DOI
- [10] Org E, Eyheramendy S, Juhanson P, Gieger C, Lichtner P, Klopp N, Veldre G, Döring A, Viigimaa M, Söber S,

- Tomberg K, Eckstein G, Kelgo P, Rebane T, Shaw-Hawkins S, Howard P, Onipinla A, Dobson RJ, Newhouse SJ, Brown M, Dominiczak A, Connell J, Samani N, Farrall M, Caulfield MJ, Munroe PB, Illig T, Wichmann HE, Meitinger T, Laan M. Genome-wide scan identifies *CDH13* as a novel susceptibility locus contributing to blood pressure determination in two european populations. *Hum Mol Genet*, 2009, 18(12): 2288–2296. [DOI](#)
- [11] Roslin NM, Hamid JS, Paterson AD, Beyene J. Genome-wide association analysis of cardiovascular-related quantitative traits in the framingham heart study. *BMC Proc*, 2009, 3(Suppl. 7): S117. [DOI](#)
- [12] Levy D, Ehret GB, Rice K, Verwoert GC, Launer LJ, Dehghan A, Glazer NL, Morrison AC, Johnson AD, Aspelund T, Aulchenko Y, Lumley T, Köttgen A, Vasán RS, Rivadeneira F, Eiriksdóttir G, Guo XQ, Arking DE, Mitchell GF, Mattace-Raso FUS, Smith AV, Taylor K, Scharpf RB, Hwang SJ, Sijbrands EJG, Bis J, Harris TB, Ganesh SK, O'Donnell CJ, Hofman A, Rotter JI, Coresh J, Benjamin EJ, Uitterlinden AG, Heiss G, Fox CS, Witteman JCM, Boerwinkle E, Wang TJ, Gudnason V, Larson MG, Chakravarti A, Psaty BM, van Duijn CM. Genome-wide association study of blood pressure and hypertension. *Nat Genet*, 2009, 41(6): 677–687. [DOI](#)
- [13] Newton-Cheh C, Johnson T, Gateva V, Tobin MD, Bochud M, Coin L, Najjar SS, Zhao JH, Heath SC, Eyheramendy S, Papadakis K, Voight BF, Scott LJ, Zhang F, Farrall M, Tanaka T, Wallace C, Chambers JC, Khaw KT, Nilsson P, van der Harst P, Polidoro S, Grobbee DE, Onland-Moret NC, Bots ML, Wain LV, Elliott KS, Teumer A, Luan J, Lucas G, Kuusisto J, Burton PR, Hadley D, McArdle WL; Wellcome Trust Case Control Consortium, Brown M, Dominiczak A, Newhouse SJ, Samani NJ, Webster J, Zeggini E, Beckmann JS, Bergmann S, Lim N, Song K, Vollenweider P, Waeber G, Waterworth DM, Yuan X, Groop L, Orho-Melander M, Allione A, Di Gregorio A, Guarrera S, Panico S, Ricceri F, Romanazzi V, Sacerdote C, Vineis P, Barroso I, Sandhu MS, Luben RN, Crawford GJ, Jousilahti P, Perola M, Boehnke M, Bonnycastle LL, Collins FS, Jackson AU, Mohlke KL, Stringham HM, Valle TT, Willer CJ, Bergman RN, Morken MA, Döring A, Gieger C, Illig T, Meitinger T, Org E, Pfeufer A, Wichmann HE, Kathiresan S, Marrugat J, O'Donnell CJ, Schwartz SM, Siscovick DS, Subirana I, Freimer NB, Hartikainen AL, McCarthy MI, O'Reilly PF, Peltonen L, Pouta A, de Jong PE, Snieder H, van Gilst WH, Clarke R, Goel A, Hamsten A, Peden JF, Seedorf U, Syvänen AC, Tognoni G, Lakatta EG, Sanna S, Scheet P, Schlessinger D, Scuteri A, Dörr M, Ernst F, Felix SB, Homuth G, Lorbeer R, Reffelmann T, Rettig R, Völker U, Galan P, Gut IG, Herberg S, Lathrop GM, Zelenika D, Deloukas P, Soranzo N, Williams FM, Zhai G, Salomaa V, Laakso M, Elosua R, Forouhi NG, Völzke H, Uitterwaal CS, van der Schouw YT, Numans ME, Matullo G, Navis G, Berglund G, Bingham SA, Koener JS, Connell JM, Bandinelli S, Ferrucci L, Watkins H, Spector TD, Tuomilehto J, Altshuler D, Strachan DP, Laan M, Meneton P, Wareham NJ, Uda M, Jarvelin MR, Mooser V, Melander O, Loos RJ, Elliott P, Abecasis GR, Caulfield M, Munroe PB. Genome-wide association study identifies eight loci associated with blood pressure. *Nat Genet*, 2009, 41(6): 666–676. [DOI](#)
- [14] Zhu XF, Young JH, Fox E, Keating BJ, Franceschini N, Kang S, Tayo B, Adeyemo A, Sun YV, Li YL, Morrison A, Newton-Cheh C, Liu K, Ganesh SK, Kutlar A, Vasán RS, Dreisbach A, Wyatt S, Polak J, Palmas W, Musani S, Taylor H, Fabsitz R, Townsend RR, Dries D, Glessner J, Chiang CWK, Mosley T, Kardia S, Curb D, Hirschhorn JN, Rotimi C, Reiner A, Eaton C, Rotter JI, Cooper RS, Redline S, Chakravarti A, Levy D. Combined admixture mapping and association analysis identifies a novel blood pressure genetic locus on 5p13: Contributions from the care consortium. *Hum Mol Genet*, 2011, 20(11): 2285–2295. [DOI](#)
- [15] Ho JE, Levy D, Rose L, Johnson AD, Ridker PM, Chasman DI. Discovery and replication of novel blood pressure genetic loci in the Women's Genome Health Study. *J Hypertens*, 2011, 29(1): 62–69. [DOI](#)
- [16] International Consortium for Blood Pressure Genome-Wide Association Studies. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature*, 2011, 478(7367): 103–109. [DOI](#)
- [17] Adeyemo A, Gerry N, Chen GJ, Herbert A, Doumatey A, Huang HX, Zhou J, Lashley K, Chen YX, Christman M, Rotimi C. A genome-wide association study of hypertension and blood pressure in African americans. *PLoS Genet*, 2009, 5(7): e1000564. [DOI](#)
- [18] Fox ER, Young JH, Li YL, Dreisbach AW, Keating BJ, Musani SK, Liu K, Morrison AC, Ganes S, Kutlar A, Ramachandran VS, Polak JF, Fabsitz RR, Dries DL, Farlow DN, Redline S, Adeyemo A, Hirschhorn JN, Sun YV, Wyatt SB, Penman AD, Palmas W, Rotter JI, Townsend RR, Doumatey AP, Tayo BO, Mosley TH Jr, Lyon HN, Kang SJ, Rotimi CN, Cooper RS, Franceschini N, Curb JD, Martin LW, Eaton CB, Kardia SL, Taylor HA, Caulfield MJ, Ehret GB, Johnson T; International Consortium for Blood Pressure Genome-wide Association Studies (ICBP-

- GWAS), Chakravarti A, Zhu X, Levy D, Munroe PB, Rice KM, Bochud M, Johnson AD, Chasman DI, Smith AV, Tobin MD, Verwoert GC, Hwang SJ, Pihur V, Vollenweider P, O'Reilly PF, Amin N, Bragg-Gresham JL, Teumer A, Glazer NL, Launer L, Zhao JH, Aulchenko Y, Heath S, Söber S, Parsa A, Luan J, Arora P, Dehghan A, Zhang F, Lucas G, Hicks AA, Jackson AU, Peden JF, Tanaka T, Wild SH, Rudan I, Igl W, Milanecchi Y, Parker AN, Fava C, Chambers JC, Kumari M, Go MJ, van der Harst P, Kao WH, Sjögren M, Vinay DG, Alexander M, Tabara Y, Shaw-Hawkins S, Whincup PH, Liu Y, Shi G, Kuusisto J, Seielstad M, Sim X, Nguyen KD, Lehtimäki T, Matullo G, Wu Y, Gaunt TR, Onland-Moret NC, Cooper MN, Platou CG, Org E, Hardy R, Dahgam S, Palmen J, Vitart V, Braund PS, Kuznetsova T, Uiterwaal CS, Campbell H, Ludwig B, Tomaszewski M, Tzoulaki I, Palmer ND; CARDIoGRAM consortium; CKDGen consortium; KidneyGen consortium; EchoGen consortium; CHARGE-HF consortium, Aspelund T, Garcia M, Chang YP, O'Connell JR, Steinle NI, Grobbee DE, Arking DE, Hernandez D, Najjar S, McArdle WL, Hadley D, Brown MJ, Connell JM, Hingorani AD, Day IN, Lawlor DA, Beilby JP, Lawrence RW, Clarke R, Collins R, Hopewell JC, Ongen H, Bis JC, Kähönen M, Viikari J, Adair LS, Lee NR, Chen MH, Olden M, Pattaro C, Hoffman Bolton JA, Köttgen A, Bergmann S, Mooser V, Chaturvedi N, Frayling TM, Islam M, Jafar TH, Erdmann J, Kulkarni SR, Bornstein SR, Grässler J, Groop L, Voight BF, Kettunen J, Howard P, Taylor A, Guarrera S, Ricceri F, Emilsson V, Plump A, Barroso I, Khaw KT, Weder AB, Hunt SC, Bergman RN, Collins FS, Bonnycastle LL, Scott LJ, Stringham HM, Peltonen L, Perola M, Vartiainen E, Brand SM, Staessen JA, Wang TJ, Burton PR, Soler Artigas M, Dong Y, Snieder H, Wang X, Zhu H, Lohman KK, Rudock ME, Heckbert SR, Smith NL, Wiggins KL, Shriner D, Veldre G, Viigimaa M, Kinra S, Prabhakaran D, Tripathy V, Langefeld CD, Rosengren A, Thelle DS, Corsi AM, Singleton A, Forrester T, Hilton G, McKenzie CA, Salako T, Iwai N, Kita Y, Ogiwara T, Ohkubo T, Okamura T, Ueshima H, Umemura S, Eyheramendy S, Meitinger T, Wichmann HE, Cho YS, Kim HL, Lee JY, Scott J, Sehmi JS, Zhang W, Hedblad B, Nilsson P, Smith GD, Wong A, Narisu N, Stancáková A, Raffel LJ, Yao J, Kathiresan S, O'Donnell C, Schwartz SM, Ikram MA, Longstreth WT Jr, Seshadri S, Shrine NR, Wain LV, Morken MA, Swift AJ, Laitinen J, Prokopenko I, Zitting P, Cooper JA, Humphries SE, Danesh J, Rasheed A, Goel A, Hamsten A, Watkins H, Bakker SJ, van Gilst WH, Janipalli C, Mani KR, Yajnik CS, Hofman A, Mattace-Raso FU, Oostra BA, Demirkan A, Isaacs A, Rivadeneira F, Lakatta EG, Orru M, Scuteri A, Ala-Korpela M, Kangas AJ, Lyytikäinen LP, Soininen P, Tukiainen T, Würz P, Ong RT, Dörr M, Kroemer HK, Völker U, Völzke H, Galan P, Hercberg S, Lathrop M, Zelenika D, Deloukas P, Mangino M, Spector TD, Zhai G, Meschia JF, Nalls MA, Sharma P, Terzic J, Kumar MJ, Denniff M, Zukowska-Szczzechowska E, Wagenknecht LE, Fowkes FG, Charchar FJ, Schwarz PE, Hayward C, Guo X, Bots ML, Brand E, Samani N, Polasek O, Talmud PJ, Nyberg F, Kuh D, Laan M, Hveem K, Palmer LJ, van der Schouw YT, Casas JP, Mohlke KL, Vineis P, Raitakari O, Wong TY, Tai ES, Laakso M, Rao DC, Harris TB, Morris RW, Dominiczak AF, Kivimäki M, Marmot MG, Miki T, Saleheen D, Chandak GR, Coresh J, Navis G, Salomaa V, Han BG, Kooner JS, Melander O, Ridker PM, Bandinelli S, Gyllenstein UB, Wright AF, Wilson JF, Ferrucci L, Farrall M, Tuomilehto J, Pramstaller PP, Elosua R, Soranzo N, Sijbrands EJ, Altshuler D, Loos RJ, Shuldiner AR, Gieger C, Meneton P, Uitterlinden AG, Wareham NJ, Gudnason V, Rettig R, Uda M, Strachan DP, Witteman JC, Hartikainen AL, Beckmann JS, Boerwinkle E, Boehnke M, Larson MG, Järvelin MR, Psaty BM, Abecasis GR, Elliott P, van Duijn CM, Newton-Cheh C. Association of genetic variation with systolic and diastolic blood pressure among African americans: The candidate gene association resource study. *Hum Mol Genet*, 2011, 20(11): 2273–2284. DOI
- [19] Niu WQ, Zhang Y, Ji KD, Gu ML, Gao PJ, Zhu DL. Confirmation of top polymorphisms in hypertension genome wide association study among Han chinese. *Clin Chim Acta*, 2010, 411(19-20): 1491–1495. DOI
- [20] Liu C, Li HX, Qi QB, Lu L, Gan W, Loos RJF, Lin X. Common variants in or near FGF5, CYP17A1 and MTHFR genes are associated with blood pressure and hypertension in Chinese Hans. *J Hypertens*, 2011, 29(1): 70–75. DOI
- [21] Lin YH, Lai XL, Chen B, Xu Y, Huang BY, Chen ZC, Zhu SH, Yao J, Jiang QQ, Huang HB, Wen JP, Chen G. Genetic variations in CYP17A1, CACNB2 and PLEKHA7 are associated with blood pressure and/or hypertension in She ethnic minority of China. *Atherosclerosis*, 2011, 219(2): 709–714. DOI
- [22] Tabara Y, Kohara K, Kita Y, Hirawa N, Katsuya T, Ohkubo T, Hiura Y, Tajima A, Morisaki T, Miyata T, Nakayama T, Takashima N, Nakura J, Kawamoto R, Takahashi N, Hata A, Soma M, Imai Y, Kokubo Y, Okamura T, Tomoike H, Iwai N, Ogiwara T, Inoue I, Tokunaga K, Johnson T, Caulfield M, Umemura S, Ueshima H, Miki T. Common variants in the ATP2B1 gene are associated with suscepti-

- bility to hypertension: The Japanese Millennium Genome Project. *Hypertension*, 2010, 56(5): 973–980. [DOI](#)
- [23] Takeuchi F, Isono M, Katsuya T, Yamamoto K, Yokota M, Sugiyama T, Nabika T, Fujioka A, Ohnaka K, Asano H, Yamori Y, Yamaguchi S, Kobayashi S, Takayanagi R, Ogihara T, Kato N. Blood pressure and hypertension are associated with 7 loci in the Japanese population. *Circulation*, 2010, 121(21): 2302–2309. [DOI](#)
- [24] Miyaki K, Htun NC, Song Y, Ikeda S, Muramatsu M, Shimbo T. The combined impact of 12 common variants on hypertension in Japanese men, considering GWAS results. *J Hum Hypertens*, 2011, doi: 10.1038/jhh.2011.50. [DOI](#)
- [25] Cho YS, Go MJ, Kim YJ, Heo JY, Oh JH, Ban HJ, Yoon D, Lee MH, Kim DJ, Park M, Cha SH, Kim JW, Han BG, Min H, Ahn Y, Park MS, Han HR, Jang HY, Cho EY, Lee JE, Cho NH, Shin C, Park T, Park JW, Lee JK, Cardon L, Clarke G, McCarthy MI, Lee JY, Lee JK, Oh B, Kim HL. A large-scale genome-wide association study of Asian populations uncovers genetic factors influencing eight quantitative traits. *Nat Genet*, 2009, 41(5): 527–534. [DOI](#)
- [26] Park JW, Uhm SY, Shin C, Cho NH, Cho YS, Lee JY. Genome-wide association analyses on blood pressure using three different phenotype definitions. *Geno Inform*, 2010, 8(3): 108–115. [DOI](#)
- [27] Hong KW, Go MJ, Jin HS, Lim JE, Lee JY, Han BG, Hwang SY, Lee SH, Park HK, Cho YS, Oh B. Genetic variations in ATP2B1, CSK, ARSG and CSMD1 loci are related to blood pressure and/or hypertension in two Korean cohorts. *J Hum Hypertens*, 2010, 24(6): 367–372. [DOI](#)
- [28] Hong KW, Lim JE, Oh B. A regulatory SNP in AKAP13 is associated with blood pressure in Koreans. *J Hum Genet*, 2011, 56(3): 205–210. [DOI](#)
- [29] Kato N, Takeuchi F, Tabara Y, Kelly TN, Go MJ, Sim XL, Tay WT, Chen CH, Zhang Y, Yamamoto K, Katsuya T, Yokota M, Kim YJ, Ong RTH, Nabika T, Gu DF, Chang LC, Kokubo Y, Huang W, Ohnaka K, Yamori Y, Nakashima E, Jaquish CE, Lee JY, Seielstad M, Isono M, Hixson JE, Chen YT, Miki T, Zhou X, Sugiyama T, Jeon JP, Liu JJ, Takayanagi R, Kim SS, Aung T, Sung YJ, Zhang X, Wong TY, Han BG, Kobayashi S, Ogihara T, Zhu D, Iwai N, Wu JY, Teo YY, Tai ES, Cho YS, He J. Meta-analysis of genome-wide association studies identifies common variants associated with blood pressure variation in east Asians. *Nat Genet*, 2011, 43(6): 531–538. [DOI](#)
- [30] 严卫丽. 复杂疾病全基因组关联研究进展——遗传统计分析. *遗传*, 2008, 30(5): 543–549. [DOI](#)
- [31] Fox ER, Young JH, Li YL, Dreisbach AW, Keating BJ, Musani SK, Liu K, Morrison AC, Ganesh S, Kutlar A, Ramachandran VS, Polak JF, Fabsitz RR, Dries DL, Farlow DN, Redline S, Adeyemo A, Hirschorn JN, Sun YV, Wyatt SB, Penman AD, Palmas W, Rotter JI, Townsend RR, Doumatey AP, Tayo BO, Mosley TH Jr, Lyon HN, Kang SJ, Rotimi CN, Cooper RS, Franceschini N, Curb JD, Martin LW, Eaton CB, Kardia SL, Taylor HA, Caulfield MJ, Ehret GB, Johnson T; International Consortium for Blood Pressure Genome-wide Association Studies (ICBP-GWAS), Chakravarti A, Zhu X, Levy D, Munroe PB, Rice KM, Bochud M, Johnson AD, Chasman DI, Smith AV, Tobin MD, Verwoert GC, Hwang SJ, Pihur V, Vollenweider P, O'Reilly PF, Amin N, Bragg-Gresham JL, Teumer A, Glazer NL, Launer L, Zhao JH, Aulchenko Y, Heath S, Söber S, Parsa A, Luan J, Arora P, Dehghan A, Zhang F, Lucas G, Hicks AA, Jackson AU, Peden JF, Tanaka T, Wild SH, Rudan I, Igl W, Milaneschi Y, Parker AN, Fava C, Chambers JC, Kumari M, Go MJ, van der Harst P, Kao WH, Sjögren M, Vinay DG, Alexander M, Tabara Y, Shaw-Hawkins S, Whincup PH, Liu Y, Shi G, Kuusisto J, Seielstad M, Sim X, Nguyen KD, Lehtimäki T, Matullo G, Wu Y, Gaunt TR, Onland-Moret NC, Cooper MN, Platou CG, Org E, Hardy R, Dahgam S, Palmén J, Vitart V, Braund PS, Kuznetsova T, Uiterwaal CS, Campbell H, Ludwig B, Tomaszewski M, Tzoulaki I, Palmer ND; CARDIOGRAM consortium; CKDGen consortium; KidneyGen consortium; EchoGen consortium; CHARGE-HF consortium, Aspelund T, Garcia M, Chang YP, O'Connell JR, Steinle NI, Grobbee DE, Arking DE, Hernandez D, Najjar S, McArdle WL, Hadley D, Brown MJ, Connell JM, Hingorani AD, Day IN, Lawlor DA, Beilby JP, Lawrence RW, Clarke R, Collins R, Hopewell JC, Ongen H, Bis JC, Kähönen M, Viikari J, Adair LS, Lee NR, Chen MH, Olden M, Pattaro C, Hoffman Bolton JA, Köttgen A, Bergmann S, Mooser V, Chaturvedi N, Frayling TM, Islam M, Jafar TH, Erdmann J, Kulkarni SR, Bornstein SR, Grässler J, Groop L, Voight BF, Kettunen J, Howard P, Taylor A, Guarrera S, Ricceri F, Emilsson V, Plump A, Barroso I, Khaw KT, Weder AB, Hunt SC, Bergman RN, Collins FS, Bonnycastle LL, Scott LJ, Stringham HM, Peltonen L, Perola M, Vartiainen E, Brand SM, Staessen JA, Wang TJ, Burton PR, Soler Artigas M, Dong Y, Snieder H, Wang X, Zhu H, Lohman KK, Rudock ME, Heckbert SR, Smith NL, Wiggins KL, Shriner D, Veldre G, Viigimaa M, Kinra S, Prabhakaran D, Tripathy V, Langefeld CD, Rosengren A, Thelle DS, Corsi AM,



Singleton A, Forrester T, Hilton G, McKenzie CA, Salako T, Iwai N, Kita Y, Ogihara T, Ohkubo T, Okamura T, Ueshima H, Umemura S, Eyheramendy S, Meitinger T, Wichmann HE, Cho YS, Kim HL, Lee JY, Scott J, Sehmi JS, Zhang W, Hedblad B, Nilsson P, Smith GD, Wong A, Narisu N, Stancáková A, Raffel LJ, Yao J, Kathiresan S, O'Donnell C, Schwartz SM, Ikram MA, Longstreth WT Jr, Seshadri S, Shrine NR, Wain LV, Morken MA, Swift AJ, Laitinen J, Prokopenko I, Zitting P, Cooper JA, Humphries SE, Danesh J, Rasheed A, Goel A, Hamsten A, Watkins H, Bakker SJ, van Gilst WH, Janipalli C, Mani KR, Yajnik CS, Hofman A, Mattace-Raso FU, Oostra BA, Demirkan A, Isaacs A, Rivadeneira F, Lakatta EG, Orru M, Scuteri A, Ala-Korpela M, Kangas AJ, Lyytikäinen LP, Soininen P, Tukiainen T, Würz P, Ong RT, Dörr M, Kroemer HK, Völker U, Völzke H, Galan P, Hercberg S, Lathrop M, Zelenika D, Deloukas P, Mangino M, Spector TD, Zhai G, Meschia JF, Nalls MA, Sharma P, Terzic J, Kumar MJ, Denniff M, Zukowska-Szczechowska E, Wagenknecht LE, Fowkes FG, Charchar FJ, Schwarz PE, Hayward C, Guo X, Bots ML, Brand E, Samani N, Polasek O, Talmud PJ, Nyberg F, Kuh D, Laan M, Hveem K, Palmer LJ, van der Schouw YT, Casas JP, Mohlke KL, Vineis P, Raitakari O, Wong TY, Tai ES, Laakso M, Rao DC, Harris TB, Morris RW, Dominiczak AF, Kivimaki M, Marmot MG, Miki T,

Saleheen D, Chandak GR, Coresh J, Navis G, Salomaa V, Han BG, Kooner JS, Melander O, Ridker PM, Bandinelli S, Gyllensten UB, Wright AF, Wilson JF, Ferrucci L, Farrall M, Tuomilehto J, Pramstaller PP, Elosua R, Soranzo N, Sijbrands EJ, Altshuler D, Loos RJ, Shuldiner AR, Gieger C, Meneton P, Uitterlinden AG, Wareham NJ, Gudnason V, Rettig R, Uda M, Strachan DP, Witteman JC, Hartikainen AL, Beckmann JS, Boerwinkle E, Boehnke M, Larson MG, Järvelin MR, Psaty BM, Abecasis GR, Elliott P, van Duijn CM, Newton-Cheh C. Association of genetic variation with systolic and diastolic blood pressure among African Americans: The candidate gene association resource study. *Hum Mol Genet*, 2011, 20(11): 2273–2284. DOI

- [32] 权晟, 张学军. 全基因组关联研究的深度分析策略. 遗传, 2011, 33(2): 100–108. DOI
- [33] Westfall PH, Young SS. Resampling-based multiple testing: Examples and methods for p-value adjustment. New York: Wiley-Interscience, 1993.
- [34] 顾东风. 常见复杂性疾病的遗传学和遗传流行病学研究: 挑战和对策. 中国医学科学院学报, 2006, 28(2): 115–118. DOI
- [35] 陈峰, 柏建岭, 赵杨, 荀鹏程. 全基因组关联研究中的统计分析方法. 中华流行病学杂志, 2011, 32(4): 400–404. DOI

## •科学新闻•

### 植物小分子 RNA 整合分析软件 psRobot 开发完成

植物小分子 RNA, 主要包括 microRNA 和小干扰 RNA, 在基因的转录和转录后调控过程中具有重要作用。第二代测序技术的逐渐成熟和广泛应用极大地推动了小分子 RNA 相关研究的发展, 对不同组织和材料中的小分子 RNA 进行深度测序已成为研究小分子 RNA 的常规手段。因此, 针对这些数据的生物信息学分析成为了研究者们面临的日益突出的问题。

中国科学院遗传与发育生物学研究所王秀杰课题组的研究人员开发了一套在线的小分子 RNA 整合分析软件 psRobot。该软件仅需要用户提交小分子 RNA 的成熟体序列, 就可以借助多种预存的高通量数据系统的鉴定这些小分子 RNA 是否为 microRNA (或具有发夹结构前体的小 RNA) 以及它们的靶基因情况。该在线工具主要分为两个功能模块, 第一个功能模块提供与小分子 RNA 本身相关的信息, 包括小分子 RNA 在参考基因组上的定位以及茎环结构的位置、在不同小分子 RNA 合成蛋白突变体或 AGO 结合实验中的表达水平等。用户可以基于上述分析结果来判断小分子 RNA 的产生与作用过程。第二个功能模块进行小分子 RNA 的靶基因预测, 提供包括参考序列库中的所有靶基因位点列表、靶位点的多重性、靶位点的保守性以及降解组数据等生物实验数据信息。通过以上工具用户可以方便地获得所分析的小分子 RNA 的产生过程、作用机制以及具有的生物学功能等信息, 为系统地研究小分子 RNA 的功能提供了极大的便利。PsRobot 的网址为: <http://omicslab.genetics.ac.cn/psRobot/>。

该研究在线发表于英国生物信息学期刊《核酸研究》(Nucleic Acids Research)(DOI:10.1093/nar/gks554)上。王秀杰课题组的博士生吴华君和马英克为论文共同第一作者。该项研究得到了国家自然科学基金, 中科院和农业部项目等经费

资助。

引自中国科学院遗传与发育生物学研究所网站