



(12) 发明专利申请

(10) 申请公布号 CN 105002286 A

(43) 申请公布日 2015. 10. 28

(21) 申请号 201510460957. 9

(22) 申请日 2015. 07. 30

(71) 申请人 中国医学科学院阜外心血管病医院
地址 100037 北京市阜城门外北礼士路 167 号

(72) 发明人 顾东风 鲁向锋 黄建凤 李宏帆
杨学礼 陈恕凤

(74) 专利代理机构 北京三友知识产权代理有限公司 11127

代理人 韩蕾 姚亮

(51) Int. Cl.

C12Q 1/68(2006. 01)

C12N 15/11(2006. 01)

权利要求书2页 说明书17页 附图11页

(54) 发明名称

与高血压和 / 或心血管疾病发病风险相关的多个单核苷酸多态性位点及相关应用

(57) 摘要

本发明提供了与高血压和 / 或心血管疾病发病风险相关的多个单核苷酸多态性位点及相关应用。具体而言, 本发明提供了一种检测多个单核苷酸多态性位点的试剂在制备用于预测高血压和 / 或心血管疾病发病风险的检测制剂或检测装置中的应用, 还提供了一种用于预测高血压和 / 或心血管疾病发病风险的检测装置, 本发明中, 通过多个单核苷酸多态性检测预测高血压和心血管疾病发病风险, 整合影响亚洲人群血压变异相关的 22 个单核苷酸多态性位点, 在 26262 例研究对象中进行基因分型, 并对研究对象进行随访研究, 在传统危险因素模型中, 加入遗传风险评分 GRS, 能提高对高血压和心血管疾病风险的预测能力。

1. 检测多个单核苷酸多态性位点的试剂在制备用于预测高血压和 / 或心血管疾病发病风险的检测制剂或检测装置中的应用, 其中, 所述多个单核苷酸多态性位点包括以下 22 个单核苷酸多态性位点:

CASZ1 区域 rs880315、MOV10 区域的 rs10745332 和 rs17030613、FIGN 区域 rs16849225、SLC4A7 区域 rs820430、ULK4 区域 rs9815354、CACNA1D 区域 rs9810888、FGF5 区域 rs1902859、ENPEP 区域 rs6825911、GUCY1A3 区域 rs13143871、NPR3 区域 rs1173766、HFE 区域 rs1799945、HLA-B 区域 rs9266359、CYP21A2 区域 rs2021783、CYP17A1 区域 rs4409766、SOX6 区域 rs4757391、ATP2B1 区域 rs17249754、ALDH2 区域 rs11066280、TBX3-TBX5 区域 rs1991391、TBX3 区域 rs35444、MED13L 区域 rs11067763 和 JAG1 区域 rs1887320。

2. 根据权利要求 1 所述的应用, 其中, 用于预测高血压和 / 或心血管疾病发病风险的检测制剂或检测装置包括检测单元以及数据分析单元;

所述检测单元是用于检测待测个体携带所述 22 个单核苷酸多态性位点的危险等位基因情况, 获得检测结果;

所述数据分析单元包括用于对检测单元的检测结果进行分析处理的处理单元, 获得遗传风险评分; 其中, 遗传风险评分的高低与待测个体患高血压和 / 或心血管疾病发病风险的高低成正相关。

3. 根据权利要求 2 所述的应用, 其中, 所述 22 个单核苷酸多态性位点的危险等位基因分别是: rs880315 多态性位点为 C; rs17030613 多态性位点为 C; rs10745332 多态性位点为 A; rs16849225 多态性位点为 C; rs820430 多态性位点为 A; rs9815354 多态性位点为 A; rs9810888 多态性位点为 G; rs1902859 多态性位点为 C; rs6825911 多态性位点为 C; rs13143871 多态性位点为 T; rs1173766 多态性位点为 C; rs1799945 多态性位点为 G; rs9266359 多态性位点为 C; rs2021783 多态性位点为 C; rs4409766 多态性位点为 T; rs4757391 多态性位点为 C; rs17249754 多态性位点为 G; rs11066280 多态性位点为 T; rs1991391 多态性位点为 G; rs35444 多态性位点为 A; rs11067763 多态性位点为 A; rs1887320 多态性位点为 A。

4. 根据权利要求 2 所述的应用, 所述遗传风险评分符合按照以下方法计算得到的数值:

遗传风险评分 = $\sum \beta_i \times N_i$, 其中 β_i 指第 i 个 SNP 的收缩压和舒张压效应值的平均值, N_i 指研究对象所携带第 i 个 SNP 的危险等位基因数目。

5. 根据权利要求 1 ~ 4 任一项所述的应用, 其中, 所述的待测个体为中国汉族人。

6. 根据权利要求 1 ~ 4 任一项所述的应用, 其中, 待测样品来自待测个体的血液、尿、唾液、胃液、头发或活组织检查。

7. 一种用于预测高血压和 / 或心血管疾病发病风险的检测装置, 该检测装置包括检测单元以及数据分析单元;

所述检测单元包括用于检测待测个体携带多个单核苷酸多态性位点的危险等位基因情况的检测单元, 获得检测结果; 其中, 所述多个单核苷酸多态性位点包括以下 22 个单核苷酸多态性位点: CASZ1 区域 rs880315、MOV10 区域的 rs10745332 和 rs17030613、FIGN 区域 rs16849225、SLC4A7 区域 rs820430、ULK4 区域 rs9815354、CACNA1D 区域 rs9810888、FGF5 区域 rs1902859、ENPEP 区域 rs6825911、GUCY1A3 区域 rs13143871、NPR3 区域 rs1173766、HFE

区域 rs1799945、HLA-B 区域 rs9266359、CYP21A2 区域 rs2021783、CYP17A1 区域 rs4409766、SOX6 区域 rs4757391、ATP2B1 区域 rs17249754、ALDH2 区域 rs11066280、TBX3-TBX5 区域 rs1991391、TBX3 区域 rs35444、MED13L 区域 rs11067763 和 JAG1 区域 rs1887320；

所述数据分析单元包括用于对检测单元的检测结果进行分析处理的处理单元，获得遗传风险评分；其中，遗传风险评分的高低与待测个体患高血压和 / 或心血管疾病发病风险的高低成正相关。

8. 根据权利要求 7 所述的检测装置，其中，所述 22 个单核苷酸多态性位点的危险等位基因分别是：rs880315 多态性位点为 C；rs17030613 多态性位点为 C；rs10745332 多态性位点为 A；rs16849225 多态性位点为 C；rs820430 多态性位点为 A；rs9815354 多态性位点为 A；rs9810888 多态性位点为 G；rs1902859 多态性位点为 C；rs6825911 多态性位点为 C；rs13143871 多态性位点为 T；rs1173766 多态性位点为 C；rs1799945 多态性位点为 G；rs9266359 多态性位点为 C；rs2021783 多态性位点为 C；rs4409766 多态性位点为 T；rs4757391 多态性位点为 C；rs17249754 多态性位点为 G；rs11066280 多态性位点为 T；rs1991391 多态性位点为 G；rs35444 多态性位点为 A；rs11067763 多态性位点为 A；rs1887320 多态性位点为 A。

9. 根据权利要求 7 或 8 所述的检测装置，其中：

所述检测单元还包括用于检测待测个体传统危险因素的检测单元；其中，高血压相关传统危险因素包括年龄、性别、体重指数、吸烟、饮酒、脉搏、收缩压和舒张压；心血管疾病相关传统危险因素包括性别、年龄、体重指数，吸烟、饮酒状况、糖尿病、总胆固醇、高密度脂蛋白胆固醇、心血管病家族史、收缩压和服用降压药情况；

所述数据分析单元还包括用于对检测单元的传统危险因素检测结果进行分析处理的处理单元，获得综合遗传风险评分和传统危险因素的高血压和 / 或心血管疾病发病风险预测评分。

10. 根据权利要求 9 所述的检测装置，其中，所述综合遗传风险评分和传统危险因素的高血压和 / 或心血管疾病发病风险预测评分符合按照以下模型计算得到的数值：

高血压发病预测模型：

高血压发病风险 = $1/(1+\exp(-Z))$ ； $Z = -14.808 + 0.142 \times G2 + 0.193 \times G3 + 0.265 \times G4 + 0.339 \times G5 + 0.165 \times \text{性别} + 0.034 \times \text{年龄} + 0.063 \times \text{体重指数} + (-0.013) \times \text{是否吸烟} + 0.159 \times \text{是否饮酒} + 0.01 \times \text{脉搏} + 0.061 \times \text{收缩压} + 0.04 \times \text{舒张压}$ ；

心血管疾病事件发生预测模型：

心血管疾病发病风险 = $1/(1+\exp(-Z))$ ； $Z = -8.558 + 0.126 \times G2 + 0.159 \times G3 + 0.297 \times G4 + 0.256 \times G5 + (-0.273) \times \text{性别} + 0.056 \times \text{年龄} + 0.035 \times \text{体重指数} + 0.19 \times \text{是否吸烟} + (-0.475) \times \text{是否饮酒} + 0.358 \times \text{有无糖尿病} + 0.006 \times \text{总胆固醇} + (-0.028) \times \text{高密度脂蛋白胆固醇} + 0.643 \times \text{是否有心血管疾病家族史} + 0.015 \times \text{收缩压} + 0.309 \times \text{是否服用降压药}$ 。

与高血压和 / 或心血管疾病发病风险相关的多个单核苷酸多态性位点及相关应用

技术领域

[0001] 本发明涉及与高血压和 / 或心血管疾病发病风险相关的多个单核苷酸多态性位点及相关应用,具体涉及检测多个单核苷酸多态性位点的试剂在制备用于预测高血压和 / 或心血管疾病发病风险的检测制剂或检测装置中的应用,还涉及一种用于预测高血压和 / 或心血管疾病发病风险的检测装置,通过影响亚洲人群血压变异相关的 22 个单核苷酸多态性位点检测预测高血压和心血管疾病发病风险。

背景技术

[0002] 心血管疾病是全球人类健康的首要疾病负担。我国心血管病的发病率和死亡率呈快速上升趋势。据估计,我国心血管疾病现患人数至少 2.3 亿,每年约 350 万人死于心血管病。高血压是最常见的心血管疾病之一,也是心血管病的主要危险因素。2013 年中国心血管病报告显示,约 70% 的脑卒中和 50% 的心肌梗死的发生与血压升高有关,至少一半的心血管病死亡与高血压有关。根据历次全国大规模高血压患病率调查,从上世纪五六十年代到 2007-2009 年,高血压患病率已由 5.1% 上升到 26.6%,上升趋势明显,已成为我国重要疾病负担。因此,高血压的防治已经成为预防心血管疾病、保障国民健康的重要需求。

[0003] 高血压和心血管疾病是复杂疾病,由多个微效基因与环境因素长期相互作用所致。鉴定出与高血压和 / 或血压升高相关的易感基因和 / 或致病基因,在人群中筛选增加疾病风险的易感基因确定易感个体,将有助于高血压和心血管疾病事件(冠心病和脑卒中)的发病风险预测、新药开发、诊断和个体化治疗。

[0004] 遗传风险评分(genetic risk score, GRS)是基于遗传易感基因建立个体患病风险的预测评分模型,在个体水平利用个体遗传风险评分结合传统危险因素进行模型预测是识别高血压和心血管疾病发病高危人群的有效方法。早期识别高危人群,采取针对性的生活方式干预,能够有效地减少高血压及心血管疾病的发生,遏制医疗费用增长、延长国民期望寿命。

发明内容

[0005] 本发明的一个目的在于提供一种通过多个单核苷酸多态性检测预测高血压和 / 或心血管疾病发病风险的方法。

[0006] 本发明的另一目的在于提供检测多个单核苷酸多态性位点的试剂在制备用于预测高血压和 / 或心血管疾病发病风险的检测制剂或检测装置中的应用。

[0007] 本发明的另一目的在于提供一种用于预测高血压和 / 或心血管疾病发病风险的检测装置。

[0008] 本发明中,通过整合影响亚洲人群血压变异相关的多个单核苷酸多态性位点,在大量样本的研究对象中进行基因分型,并对研究对象进行多年的随访研究,计算遗传风险评分(GRS)。在传统危险因素的模型中,加入遗传风险评分,提高对高血压和心血管事件风

险的预测能力。

[0009] 根据本发明的具体实施方案,所述亚洲人群血压变异的相关的多个单核苷酸多态性位点包括以下 22 个单核苷酸多态性位点: CASZ1 区域 rs880315、MOV10 区域的 rs10745332 和 rs17030613、FIGN 区域 rs16849225、SLC4A7 区域 rs820430、ULK4 区域 rs9815354、CACNA1D 区域 rs9810888、FGF5 区域 rs1902859、ENPEP 区域 rs6825911、GUCY1A3 区域 rs13143871、NPR3 区域 rs1173766、HFE 区域 rs1799945、HLA-B 区域 rs9266359、CYP21A2 区域 rs2021783、CYP17A1 区域 rs4409766、SOX6 区域 rs4757391、ATP2B1 区域 rs17249754、ALDH2 区域 rs11066280、TBX3-TBX5 区域 rs1991391、TBX3 区域 rs35444、MED13L 区域 rs11067763 和 JAG1 区域 rs1887320。通过检测这些单核苷酸多态性位点,可以预测高血压和 / 或心血管疾病发病风险。

[0010] 从而,一方面,本发明提供了检测多个单核苷酸多态性位点的试剂在制备用于预测高血压和 / 或心血管疾病发病风险的检测制剂或检测装置中的应用,其中,所述多个单核苷酸多态性位点包括以下 22 个单核苷酸多态性位点:

[0011] CASZ1 区域 rs880315、MOV10 区域的 rs10745332 和 rs17030613、FIGN 区域 rs16849225、SLC4A7 区域 rs820430、ULK4 区域 rs9815354、CACNA1D 区域 rs9810888、FGF5 区域 rs1902859、ENPEP 区域 rs6825911、GUCY1A3 区域 rs13143871、NPR3 区域 rs1173766、HFE 区域 rs1799945、HLA-B 区域 rs9266359、CYP21A2 区域 rs2021783、CYP17A1 区域 rs4409766、SOX6 区域 rs4757391、ATP2B1 区域 rs17249754、ALDH2 区域 rs11066280、TBX3-TBX5 区域 rs1991391、TBX3 区域 rs35444、MED13L 区域 rs11067763 和 JAG1 区域 rs1887320。

[0012] 在本发明的大量实验结果基础上,本发明的进一步研究发现,这些 SNPs 的危险等位基因分别是: rs880315 多态性位点为 C; rs17030613 多态性位点为 C; rs10745332 多态性位点为 A; rs16849225 多态性位点为 C; rs820430 多态性位点为 A; rs9815354 多态性位点为 A; rs9810888 多态性位点为 G; rs1902859 多态性位点为 C; rs6825911 多态性位点为 C; rs13143871 多态性位点为 T; rs1173766 多态性位点为 C; rs1799945 多态性位点为 G; rs9266359 多态性位点为 C; rs2021783 多态性位点为 C; rs4409766 多态性位点为 T; rs4757391 多态性位点为 C; rs17249754 多态性位点为 G; rs11066280 多态性位点为 T; rs1991391 多态性位点为 G; rs35444 多态性位点为 A; rs11067763 多态性位点为 A; rs1887320 多态性位点为 A。

[0013] 可用本领域中任何的可行技术在 DNA 水平、RNA 水平检测本发明所述的多个单核苷酸多态性位点。例如:可以采用直接测序的方法,通过 DNA 直接测序可以直接揭示对照基因和携带突变基因之间的序列差异,具体可以是传统的使用商业测序试剂盒或自动测序仪对 DNA 直接进行测序,或是近年来发展的焦磷酸测序 (Pyrosequencing)、微测序 (SNaPshot) 等。也可以采用基于杂交的方法,具体包括 Taqman 探针法, DNA 芯片法等。还可以采用基于引物延伸的方法,如基质辅助激光解析离子飞行时间质谱 (MALDI-Tof-MS)。还可以采用基于构象的方法,具体例如限制性片段长度多态性 (RFLP) 分析、单链构象多态性 (single-strand conformational polymorphism, SSCP) 分析、变性梯度凝胶电泳 (denaturing gradient gel electrophoresis, DGGE) 分析、变性高效液相色谱技术 (denaturing high performance liquid chromatography, dHPLC) 等分析技术。还可以采用高分辨率溶解曲线分析技术 (HRM)。在具体实施时,本领域的技术人员可以根据实

实际情况选择上述的任一种技术体外检测本发明所述多个单核苷酸多态性位点,也可以采用多种技术的组合来体外检测所述多个单核苷酸多态性位点。本发明中所述的检测多个单核苷酸多态性位点的试剂可以是这些任何可行的检测所述多个单核苷酸多态性位点的技术中所用到的试剂材料和/或仪器设备等。在本发明的一个具体实施方式中,应用的是 iPLEX Sequenom MassARRAY 分型系统对所述多个单核苷酸多态性位点进行检测,其中将 PCR 扩增后的延伸产物用 MassARRAY Nanodispenser RS1000 点样仪 (Sequenom) 移至 SpectroCHIP (Sequenom) 芯片上,使用 MALDI-TOP (matrix-assisted laser desorption/ionization-time of flight, 基质辅助飞行时间质谱) 分析,并使用 TYPER 4.0 软件 (Sequenom) 分型并输出结果。

[0014] 根据本发明的具体实施方案,所述的待测个体可以为亚洲人群,优选为中国汉族人群。具体检测时,待测样品可来自待测个体的血液、尿、唾液、胃液、头发或活组织检查。

[0015] 在本发明的具体实施方案中,本发明在中国汉族人群中选取 26262 例正常人群研究对象,平均随访观察 6.9 年,采集空腹血并收集性别、年龄、血压水平、血脂水平、身高、体重、吸烟、饮酒和疾病史等相关信息。应用 iPLEX Sequenom MassARRAY 分型系统,对 22 个血压相关单核苷酸多态性 (SNP) 位点进行基因分型实验 (CASZ1 区域 rs880315、MOV10 区域的 rs10745332 和 rs17030613、FIGN 区域 rs16849225、SLC4A7 区域 rs820430、ULK4 区域 rs9815354、CACNA1D 区域 rs9810888、FGF5 区域 rs1902859、ENPEP 区域 rs6825911、GUCY1A3 区域 rs13143871、NPR3 区域 rs1173766、HFE 区域 rs1799945、HLA-B 区域 rs9266359、CYP21A2 区域 rs2021783、CYP17A1 区域 rs4409766、SOX6 区域 rs4757391、ATP2B1 区域 rs17249754、ALDH2 区域 rs11066280、TBX3-TBX5 区域 rs1991391、TBX3 区域 rs35444、MED13L 区域 rs11067763 和 JAG1 区域 rs1887320)。使用 logistic 和 cox 回归分析评估每个 SNP 位点与年平均血压增加、高血压发病以及心血管病事件发生 (冠心病和脑卒中) 的关系,计算危险等位基因的效应值 (β)、相对危险度 (Odds ratio, OR) 或风险比 (Hazard ratio, HR)。综合每个个体携带 22 个 SNP 位点的危险等位基因情况,对每个研究对象进行遗传风险评分,评分方法是利用前期血压全基因组研究获得的风险等位基因效应值 (β)。将研究对象携带危险等位基因的数目与危险等位基因效应值 (β) 相乘获得每个 SNP 位点遗传效应。然后对 22 个 SNP 的遗传效应求和得到该研究对象的遗传风险评分 (GRS),计算公式:遗传风险评分 (GRS) = $\sum \beta_i \times N_i$,其中 β_i 指第 i 个 SNP 的收缩压和舒张压效应值的平均值, N_i 指研究对象所携带第 i 个 SNP 的危险等位基因数目。将遗传风险评分作为连续型变量,计算遗传风险评分的标准差 (SD)。同时,根据所有研究对象遗传风险评分将研究对象等分为 5 组。分别为第 1 组 7.765 ~ 13.465 (G1)、第 2 组 13.466 ~ 14.616 (G2)、第 3 组 14.617 ~ 15.545 (G3)、第 4 组 15.545 ~ 16.654 (G4) 和第 5 组 16.655 ~ 22.125 (G5)。与遗传风险评分最低的 1 组相比,计算其他各组发病风险的高低。结果显示遗传风险评分是血压升高、高血压以及心血管事件发生的独立危险因素。

[0016] 在本发明的具体实施方案中,本发明进一步建立了高血压和心血管疾病的发病风险预测模型。高血压预测模型中包括基因型信息 (遗传风险评分分组) 和传统危险因素包括性别、年龄、体重指数、吸烟、饮酒、脉搏、收缩压和舒张压。心血管疾病发病风险预测模型中包括基因型信息 (遗传风险评分分组) 和传统危险因素包括性别、年龄、体重指数、吸烟、饮酒状况、糖尿病、总胆固醇、高密度脂蛋白胆固醇、心血管疾病家族史、收缩压和服用降压

药情况。综合上述遗传和环境因素建立高血压和心血管疾病发病风险预测模型：

[0017] 高血压发病风险 = $1/(1+\exp(-Z))$ 。 $Z = -14.808 + 0.142 \times G2 + 0.193 \times G3 + 0.265 \times G4 + 0.339 \times G5 + 0.165 \times \text{性别} + 0.034 \times \text{年龄} + 0.063 \times \text{体重指数} + (-0.013) \times \text{是否吸烟} + 0.159 \times \text{是否饮酒} + 0.01 \times \text{脉搏} + 0.061 \times \text{收缩压} + 0.04 \times \text{舒张压}$ 。其中,性别:男=1,女=0;是否吸烟:吸烟=1,不吸烟=0;是否饮酒:饮酒=1,不饮酒=0;脉搏项:次/分;收缩压:mmHg;舒张压:mmHg。

[0018] 心血管疾病发病风险 = $1/(1+\exp(-Z))$ 。 $Z = -8.558 + 0.126 \times G2 + 0.159 \times G3 + 0.297 \times G4 + 0.256 \times G5 + (-0.273) \times \text{性别} + 0.056 \times \text{年龄} + 0.035 \times \text{体重指数} + 0.19 \times \text{是否吸烟} + (-0.475) \times \text{是否饮酒} + 0.358 \times \text{有无糖尿病} + 0.006 \times \text{总胆固醇} + (-0.028) \times \text{高密度脂蛋白胆固醇} + 0.643 \times \text{是否有心血管疾病家族史} + 0.015 \times \text{收缩压} + 0.309 \times \text{是否服用降压药}$ 。其中,性别:男=1,女=0;是否吸烟:吸烟=1,不吸烟=0;是否饮酒:饮酒=1,不饮酒=0;有无糖尿病:有=1,无=0;总胆固醇:mg/dl;高密度脂蛋白胆固醇:mg/dl;是否有心血管病家族史:有=1,无=0;收缩压:mmHg;是否服用降压药:是=1,否=0。

[0019] 上述模型中, G2、G3、G4、G5 为前述根据研究对象遗传风险评分的分组,第1组 7.765 ~ 13.465(G1)、第2组 13.466 ~ 14.616(G2)、第3组 14.617 ~ 15.545(G3)、第4组 15.545 ~ 16.654(G4) 和第5组 16.655 ~ 22.125(G5)。当待测个体遗传风险评分落入 G2、G3、G4、G5 任一组时,其相应的组别 G 值为 1,其余取值为 0。例如,当待测个体遗传风险评分落入 G2 组,G2 取值为 1,G3、G4、G5 均取值 0(当待测个体遗传危险因素评分落入 G1 组,G2、G3、G4、G5 均取值 0)。

[0020] 在本发明的具体实施方案中,本发明将遗传风险评分分组和环境因素暴露信息放入风险预测模型计算出每名研究对象的高血压和心血管疾病的发病风险。采用作受试者工作特征曲线(ROC)和模型拟合曲线(Calibration)评估高血压和心血管疾病发生风险预测模型。根据 ROC 曲线下面积评价该预测方法的价值。结果显示,高血压和心血管疾病发生的 ROC 下面积为分别为 0.776 和 0.813,表明该模型预测性能良好,可以很好的识别高危人群。模型拟合曲线显示高血压和心血管事件发生风险的预测值和观察值无显著差异,P 值分别为 0.882 和 0.24,模型拟合良好。

[0021] 从而,另一方面,本发明还提供了一种用于预测高血压和/或心血管疾病发病风险的检测装置,该检测装置包括检测单元以及数据分析单元;

[0022] 所述检测单元包括用于检测待测个体携带多个单核苷酸多态性位点的危险等位基因情况的检测单元(可包括检测所用的试剂材料和/或仪器设备等),获得检测结果;其中,所述多个单核苷酸多态性位点包括以下 22 个单核苷酸多态性位点: CASZ1 区域 rs880315、MOV10 区域的 rs10745332 和 rs17030613、FIGN 区域 rs16849225、SLC4A7 区域 rs820430、ULK4 区域 rs9815354、CACNA1D 区域 rs9810888、FGF5 区域 rs1902859、ENPEP 区域 rs6825911、GUCY1A3 区域 rs13143871、NPR3 区域 rs1173766、HFE 区域 rs1799945、HLA-B 区域 rs9266359、CYP21A2 区域 rs2021783、CYP17A1 区域 rs4409766、SOX6 区域 rs4757391、ATP2B1 区域 rs17249754、ALDH2 区域 rs11066280、TBX3-TBX5 区域 rs1991391、TBX3 区域 rs35444、MED13L 区域 rs11067763 和 JAG1 区域 rs1887320;

[0023] 所述数据分析单元包括用于对检测单元的检测结果进行分析处理的处理单元,获得遗传风险评分;其中,遗传风险评分的高低与待测个体患高血压和/或心血管疾病发病

风险的高低成正相关。

[0024] 根据本发明的具体实施方案,本发明的用于预测高血压和 / 或心血管疾病发病风险的检测装置,其中:

[0025] 所述检测单元还包括用于检测待测个体传统危险因素的检测单元(包括检测用的试剂材料和 / 或仪器设备等);其中,高血压相关传统危险因素包括性别、年龄、体重指数、吸烟、饮酒、脉搏、收缩压和舒张压;心血管疾病相关传统危险因素包括性别、年龄、体重指数、吸烟、饮酒状况、糖尿病、总胆固醇、高密度脂蛋白胆固醇、心血管疾病家族史、收缩压和服用降压药情况;

[0026] 所述数据分析单元还包括用于对检测单元的传统危险因素检测结果进行分析处理的处理单元,获得综合遗传风险评分和传统危险因素的高血压和 / 或心血管疾病发病风险预测评分。

[0027] 本发明所述的用于预测高血压和 / 或心血管疾病发病风险的检测装置,可以是虚拟装置,只要能实现所述检测单元以及数据分析单元的功能即可。所述的检测单元可以是包括各种检测试剂材料和 / 或检测仪器设备等;所述的数据分析单元可以是任何可以实现对检测单元的检测结果进行分析处理而得出待测个体遗传风险评分(或综合遗传风险评分与传统危险因素的结果)的运算仪器、模块或是虚拟设备,例如可以是预先按照上述遗传风险评分公式而制定的数据图表,将检测单元的检测结果对照该数据图表即能得出遗传风险评分数值。

[0028] 根据本发明的具体实施方案,本发明的用于预测高血压和 / 或心血管疾病发病风险的检测装置,其中,所述综合遗传风险评分和传统危险因素的高血压和 / 或心血管疾病发病风险预测评分符合按照以下模型计算得到的数值:

[0029] 高血压发病风险预测模型:

[0030] 高血压发病风险 = $1/(1+\exp(-Z))$ 。 $Z = -14.808 + 0.142 \times G2 + 0.193 \times G3 + 0.265 \times G4 + 0.339 \times G5 + 0.165 \times \text{性别} + 0.034 \times \text{年龄} + 0.063 \times \text{体重指数} + (-0.013) \times \text{是否吸烟} + 0.159 \times \text{是否饮酒} + 0.01 \times \text{脉搏} + 0.061 \times \text{收缩压} + 0.04 \times \text{舒张压}$ 。其中,性别:男 = 1,女 = 0;是否吸烟:吸烟 = 1,不吸烟 = 0;是否饮酒:饮酒 = 1,不饮酒 = 0;脉搏:次/分;收缩压:mmHg;舒张压:mmHg。

[0031] 心血管疾病发病风险预测模型:

[0032] 心血管疾病发病风险 = $1/(1+\exp(-Z))$ 。 $Z = -8.558 + 0.126 \times G2 + 0.159 \times G3 + 0.297 \times G4 + 0.256 \times G5 + (-0.273) \times \text{性别} + 0.056 \times \text{年龄} + 0.035 \times \text{体重指数} + 0.19 \times \text{是否吸烟} + (-0.475) \times \text{是否饮酒} + 0.358 \times \text{有无糖尿病} + 0.006 \times \text{总胆固醇} + (-0.028) \times \text{高密度脂蛋白胆固醇} + 0.643 \times \text{是否有心血管疾病家族史} + 0.015 \times \text{收缩压} + 0.309 \times \text{是否服用降压药}$ 。其中,性别:男 = 1,女 = 0;是否吸烟:吸烟 = 1,不吸烟 = 0;是否饮酒:饮酒 = 1,不饮酒 = 0;有无糖尿病:有 = 1,无 = 0;总胆固醇:mg/dl;高密度脂蛋白胆固醇:mg/dl;是否有心血管病家族史:有 = 1,无 = 0;收缩压:mmHg;是否服用降压药:是 = 1,否 = 0。

[0033] 上述模型中 G2、G3、G4、G5 取值参照前述。

[0034] 应用本发明的技术,根据模型计算获得该研究对象高血压和 / 或心血管事件发生的风险,可以判断该研究对象是否为高危人群。完成个体化健康指导报告,报告高血压和心血管疾病事件发生风险的高低。个体化健康指导报告是在基因型结果的基础上,结合生活

方式危险因素评估研究对象高血压和心血管疾病事件发生的风险,并制定出针对研究对象的个性化健康行动方案。可见,本发明的技术在高血压和 / 或心血管疾病的预测和预防中有重要的应用前景。

附图说明

[0035] 图 1- 图 9 :rs4757391、rs9810888、rs1902859、rs4409766、rs10745332、rs17249754、rs 880315、rs 9815354 和 rs11066280 基因分型结果散点图 (各 2 块 384 板)。

[0036] 图 10 :遗传风险评分分组和环境因素暴露信息预测高血压和心血管病发病风险的 ROC 曲线。

[0037] 图 11 :遗传风险评分分组和环境因素暴露信息预测高血压和心血管病发病风险的模型拟合曲线。

[0038] 图 12 :血压相关位点危险等位基因遗传效应速查图表。

[0039] 图 13 :遗传风险评分速查图表。

具体实施方式

[0040] 为了更清楚地理解本发明,现参照下列实施例及附图进一步描述本发明。实施例仅用于解释而不以任何方式限制本发明。实施例中未注明具体条件的实验方法为所属领域熟知的常规方法和常规条件,或按照制造商所建议的条件。

[0041] 实施例一、高血压和心血管疾病的发病风险预测模型的建立

[0042] 一、病例对照样本入选标准

[0043] 首先在中国汉族人群中选取正常人群研究对象,研究对象来自中国心血管病多中心协作研究 (InterAsia in China) 和中国心血管健康研究多个北方研究中心,研究方案经各调查中心当地研究机构伦理委员会批准,所有参加者都签署了知情同意书。中国心血管病多中心协作研究于 2000-2001 年开始,2008 年进行了随访。中国心血管健康研究是 2006-2007 年开始调查,2013-2014 年进行了随访。平均随访观察 6.9 年,采集空腹血并收集性别、年龄、血压水平、血脂水平、身高、体重、吸烟、饮酒和疾病史等相关信息。

[0044] 血压测量方法:研究对象血压由培训合格的调查员根据美国心脏病学会推荐血压测量标准测量三次。血压测量前静坐 5 分钟;并且血压测量前 30 分钟内避免饮酒、吸烟、饮用咖啡或茶饮料和体力活动。对于服用降压药物的患者,血压分别按收缩压、舒张压测量者各增加 10 和 5mmHg 计算。高血压定义为:收缩压 ≥ 140 mm Hg 和 / 或舒张压 DBP ≥ 90 mm Hg 或自我报告以治疗高血压为目的服用降压药。正常血压定义为不服用任何降压药,且收缩压 <140 mm Hg 同时舒张压 <90 mm Hg。

[0045] 吸烟定义为:一生中吸烟支数是否超过 100 支 (是 1/ 否 0)。

[0046] 饮酒定义为:过去一年研究是否超过 12 次 (是 1/ 否 0)。

[0047] 糖尿病定义为:医院临床诊断 (有 1/ 无 0)。

[0048] 体重指数 (BMI) 计算公式:体重指数 (kg/m^2) = 体重 (kg) / 身高 (m)²

[0049] 随访阶段:心血管终点事件定义为发生心肌梗死、心绞痛、血运重建、卒中以及死亡归因为冠状动脉疾病和卒中。随访内容包括跟踪调查对象及其监护人当前居住地进行入户调查以确定疾病状态等重要信息。如果调查对象报告在随访期间有住院或者急诊留观,

经过培训的调查员使用标准调查表,通过问卷提取其病史、体检结果、实验室检查结果和出院诊断等信息。所有随访期间的死亡报告都通过从当地卫生部门或公安机关获得死亡证明进行核实。由中国医学科学院(北京)的心脏病专家、神经科医师和临床流行病学家组成的终点评估委员会通过回顾所有医院记录和死亡证明以确定调查终点事件或最终诊断死亡原因。心血管终点事件由两个委员会成员的独立核实诊断,如诊断结果不同,则由更多的委员会成员讨论确定。

[0050] 本发明中,分析选取了随访资料和遗传因素评分都有完整数据的参加者,并排除了基线阶段发生心血管病事件个体,共 26262 名个体纳入分析(表 1),在随访期间 378 例冠心病和 749 例脑卒中发生,共计 1078 例心血管病事件。7724 名基线血压正常且有随访血压数据的调查对象纳入血压升高和高血压发生的分析。

[0051] 表 1. 研究人群的基本特征

[0052]

总样本量	26262
男性 / 女性 (% 男性)	11736/14526 (44.69)
年龄 (标准差)	50.86 (8.92)
BMI (kg/m ²) (标准差)	24.24 (3.65)
收缩压 (SBP, mmHg) (标准差)	130.69 (22.78)
舒张压 (DBP, mmHg) (标准差)	81.68 (12.57)
高血压患病率 (%)	37.79
总胆固醇 (TC, mmol/l) (标准差)	4.71 (0.97)
高密度脂蛋白胆固醇 (HDL, mmol/l) (标准差)	1.35 (0.40)
甘油三脂 (TG, mmol/l) (标准差)	1.60 (1.20)
饮酒 (%)	20.99
吸烟 (%)	31.02
心血管疾病家族史 (%)	9.69

[0053] 二、SNP 选择、基因分型

[0054] 本发明中,选择了在东亚人群和中国人群全基因组关联研究中与收缩压或者舒张压相关的 22 个单核苷酸多态性 (SNP) 位点,应用 iPLEX Sequenom Mass ARRAY 平台进行基因分型。所述 22 个 SNP 为 :CASZ1 区域 rs880315、MOV10 区域的 rs10745332 和 rs17030613、FIGN 区域 rs16849225、SLC4A7 区域 rs820430、ULK4 区域 rs9815354、CACNA1D 区域 rs9810888、FGF5 区域 rs1902859、ENPEP 区域 rs6825911、GUCY1A3 区域 rs13143871、

NPR3 区域 rs1173766、HFE 区域 rs1799945、HLA-B 区域 rs9266359、CYP21A2 区域 rs2021783、CYP17A1 区域 rs4409766、SOX6 区域 rs4757391、ATP2B1 区域 rs17249754、ALDH2 区域 rs11066280、TBX3-TBX5 区域 rs1991391、TBX3 区域 rs35444、MED13L 区域 rs11067763 和 JAG1 区域 rs1887320。

[0055] 基因分型应用 iPLEX Sequenom MassARRAY 分型系统 (Complete iPLEX[®] Gold Genotyping Reagent Set 384 ;Sequenom ;Cat No. 10148-2) 完成。其原理为 :将包含 SNP 位点区域的 DNA 技术扩增,应用 MassARRAY iPLEX 单碱基延伸技术,将特异的延伸引物与 PCR 产物进行单碱基延伸反应。延伸产物不同的末端碱基将导致延伸后的产物分子量的差多态性引起的碱基差异通过分子量的差异而体现,通过基质辅助激光解吸检测延伸产物分子量的大小,应用专用的分析软件,通 SNP 分型检测。

[0056] 基因分型的具体实验步骤 :

[0057] 1) 应用 Sequenom 公司 Genotyping Tools 及 MassARRAY Assay Design 软件设计待测 SNP 位点的 PCR 扩增引物及单碱基延伸引物。

[0058] 2) 采集受试者外周血,提取 DNA 样本,并将 DNA 浓度调整为 $55 \mu\text{g}/\mu\text{l}$ ($\pm 5 \mu\text{g}/\mu\text{l}$)。

[0059] 3) PCR 反应 :

[0060] 仪器 :Gene Amp PCR System 9700 (Dural 384), Amplied Biosystems

[0061] 反应体系 ($5/\mu\text{l}$) :

[0062]

10×PCR Buffer	0.5 μl
MgCl ₂ (25mM)	0.4 μl
dNTPs(25mM)	0.1 μl
HotStar Taq(5U/ μl)	0.1 μl
PCR Primer Mix(10pmol/ μl)	1 μl
Genomic DNA (55ng/ μl)	1 μl
ddH ₂ O	1.9 μl

[0063] PCR 条件 :

[0064]

Step1 94°C 4min;

Step2 94°C 20sec,

Step3 56°C 30sec,

Step4 72°C 1min;

Goto step2 44 times;

Step5 72°C 3min;

Step6 4°C ∞。

[0065] 4) PCR 产物碱性磷酸酶处理 :在 PCR 反应结束后,将 PCR 产物用 SAP(shrimp

alkaline phosphatase, 虾碱性磷酸酶) 处理, 以去除体系中游离的 dNTPs。

[0066] 反应体系:

[0067]

SAP Buffer (10×)	0.17μl
SAP Enzyme (1.7U/μl)	0.3μl
PCR 产物	5μl
ddH ₂ O	1.53μl

[0068] 反应条件: 37°C 40min; 85°C 5min; 4°C ∞。

[0069] 5) 单碱基延伸

[0070] 将单碱基延伸反应液 2 μl, 加入 4) 对应的 384 孔中, 于 PCR 仪上进行如下反应:

[0071]

Step 1	94°C	30sec
Step2	94°C	5sec
Step3	52°C	5sec
Step4	80°C	5sec
Goto step 3 for 4 times		
Goto step 2 for 39 times		
step 5	72°C	3min
step6	4°C	∞

[0072] 单碱基延伸反应液配比 (2 μl):

[0073]

Extend primer mix	0.94μl
iPLEX buffer plus	0.2μl
iPLEX terminator	0.2μl
iPLEX enzyme	0.041μl
ddH ₂ O	0.619μl。

[0074] 6) 树脂纯化

[0075] 将 Clean Resin 树脂平铺到 6mg 的树脂板中, 加 16 μl 水到延伸产物的对应孔内; 将干燥后的树脂倒入延伸产物板中, 封膜; 低速垂直旋转 30 应物充分接触; 离心使树脂沉入孔底部。

[0076] 7) 芯片点样

[0077] 启动 MassARRAY Nanodispenser RS1000 点样仪 (Sequenom), 将树脂纯化后的延伸产物移至 384 孔 SpectroCHIP (Sequenom) 芯片上。

[0078] 8) 质谱检测

[0079] 将点样后的 SpectroCHIP 芯片使用 MALDI-TOP (matrix-assisted laser desorption/ionization - time of flight, 基质辅助飞行时间质谱) 分析, 检测结果使用

TYPER 4.0 软件 (Sequenom) 分型并输出结果。

[0080] 9) 分型结果

[0081] 图 1 至图 9 分别是 rs4757391、rs9810888、rs1902859、rs4409766、rs10745332、rs17249754、rs 880315、rs 9815354 和 rs11066280 基因分型结果散点图 (各 2 块 384 板)。

[0082] 三、遗传风险评估

[0083] 采用血压的遗传风险评估 (GRS) 评价 22 个 SNP 的累积效应, 根据全基因组关联研究结果, 确定用于评价每个 SNP 对收缩压 (SBP) 和舒张压 (DBP) 的平均效应大小 (β 加权系数)。22 个位点信息见表 2。用每个 SNP 的 β 加权系数乘以该位点的危险等位基因数并将结果相加, 计算遗传风险评估 (GRS), 计算公式: 遗传风险评估 = $\sum \beta_i \times N_i$, 其中 β_i 指第 i 个 SNP 的收缩压和舒张压效应值的平均值, N_i 指研究对象所携带第 i 个 SNP 的危险等位基因数目。将遗传风险评估作为连续型变量或进行 5 分位分组, 分析与高血压和心血管病事件发生的关系。

[0084] 统计方法应用 Logistic 回归和 Cox 比例风险回归分析单个 SNPs 以及它们构成的 GRS 与血压增高、高血压发生以及心血管病事件发生的相关性, 并在模型中校正传统危险因素。比较在有 GRS 的模型中的 C 指数 (C-index), 以评价应用遗传风险评估对高血压和心血管事件发生风险预测的改进。采用模型拟合曲线 (Calibration) 评估高血压和心血管疾病发生风险预测模型拟合是否良好。

[0085] 表 2. 22 个血压相关的位点及其效应

[0086]

附近基因	SNP	染色体	基因组位置	危险/非危险等位基因	危险等位基因频率	收缩压效应 β SBP*	舒张压效应 β DBP*	平均效应 (β 加权系数)
CASZ1	rs880315	1	10719453	C/T	0.61	0.97	0.46	0.715
MOV10	rs17030613	1	112971190	C/A	0.46	0.49	0.38	0.435
MOV10	rs10745332	1	112990576	A/G	0.82	0.96	0.53	0.745
FIGN	rs16849225	2	164615066	C/T	0.61	0.45	0.10	0.275
SLC4A7	rs820430	3	27523904	A/G	0.32	0.76	0.27	0.515
ULK4	rs9815354	3	41887655	A/G	0.18	0.10	0.43	0.265
CACNA1D	rs9810888	3	53610635	G/T	0.39	0.53	0.39	0.46
FGF5	rs1902859	4	81376727	C/T	0.40	1.34	0.71	1.025
ENPEP	rs6825911	4	111,601,087	C/T	0.47	0.6	0.39	0.495
GUCY1A3	rs13143871	4	156838654	T/C	0.79	0.96	0.49	0.725
NPR3	rs1173766	5	32,840,285	C/T	0.65	0.63	0.36	0.495
HFE	rs1799945	6	26199158	G/C	0.04	0.95	0.88	0.915
HLA-B	rs9266359	6	31440718	C/T	0.60	0.44	0.29	0.365
CYP21A2	rs2021783	6	32152829	C/T	0.78	0.68	0.49	0.585
CYP17A1	rs4409766	10	104606653	T/C	0.70	1.24	0.59	0.915

[0087]

SOX6	rs4757391	11	16259515	C/T	0.28	0.67	0.36	0.515
ATP2B1	rs17249754	12	88584717	G/A	0.65	1.03	0.52	0.775
ALDH2	rs11066280	12	111302166	T/A	0.79	0.96	0.62	0.79
TBX3-TBX5	rs1991391	12	113837049	G/A	0.84	0.6	0.21	0.405
TBX3	rs35444	12	114036820	A/G	0.76	0.83	0.36	0.595
MED13L	rs11067763	12	114682724	A/G	0.61	0.81	0.51	0.66
JAG1	rs1887320	20	10913998	A/G	0.52	0.78	0.43	0.605

[0088] * 效应值来源于东亚人群和中国人群开展的高血压全基因组关联研究 (Nat Genet. 2011 ;43:531-538, Hum Mol Genet. 2015 ;24:865-874)。

[0089] 22 个血压位点除 rs6825911 外, 每年升高收缩压 0.01 - 0.11mm Hg 和舒张压 0.004 - 0.07mm Hg。其中 20 个 SNPs 增加高血压和心血管疾病事件的发生 ($P = 6.06 \times 10^{-5}$) (表 3)。6 个 SNP (rs17030613, rs10745332、rs9810888、rs1902859、rs11066280、rs35444 与血压变化或高血压发病相关达到显著水平 ($P < 0.05$)。5 个 SNP (rs880315、rs820430、rs13143871、rs11066280 和 rs1887320) 与心血管病事件发生相关性达到显著水平 ($P < 0.05$)。

[0090] 表 3. 2 2 位点对血压变化、高血压以及心血管疾病事件发生的影响

[0091]

SNP	附近基因	染色体	基因组位置	危险/非危险等位基因	收缩压升高/年		舒张压升高/年		高血压发生		心血管事件	
					效应 Beta(se)	P 值	效应 Beta(se)	P 值	效应 Beta(se)	P 值	效应 Beta(se)	P 值
rs880315	<i>CASZ1</i>	1	10719453	C/T	0.05(0.04)	0.19	0.04(0.02)	0.09	0.04(0.04)	0.30	0.11(0.05)	0.02
rs17030613	<i>MOV10</i>	1	112971190	C/A	0.10(0.03)	2.77×10^{-3}	0.05(0.02)	0.02	0.10(0.04)	6.54×10^{-3}	0.05(0.07)	0.50
rs10745332	<i>MOV10</i>	1	112990576	A/G	0.11(0.05)	0.01	0.03(0.03)	0.25	0.13(0.05)	7.79×10^{-3}	-0.03(0.06)	0.61
rs16849225	<i>FIGN</i>	2	164615066	C/T	0.06(0.04)	0.08	0.04(0.02)	0.10	0.06(0.04)	0.11	0.01(0.04)	0.81
rs820430	<i>SLC4A7</i>	3	27523904	A/G	0.02(0.04)	0.57	0.01(0.02)	0.66	0.05(0.04)	0.17	0.09(0.05)	0.04
rs9815354	<i>ULK4</i>	3	41887655	A/G	0.02(0.05)	0.61	0.04(0.03)	0.13	0.06(0.05)	0.17	0.01(0.06)	0.89
rs9810888	<i>CACNA1D</i>	3	53610635	G/T	0.06(0.04)	0.11	0.02(0.02)	0.37	0.09(0.04)	0.01	0.08(0.04)	0.05
rs1902859	<i>FGF5</i>	4	81376727	C/T	0.05(0.04)	0.15	0.06(0.02)	0.01	0.14(0.04)	7.12×10^{-3}	0.03(0.04)	0.48
rs6825911	<i>ENPEP</i>	4	111601087	C/T	-0.02(0.03)	0.64	-0.03(0.02)	0.24	-0.02(0.04)	0.65	0.03(0.07)	0.67
rs13143871	<i>GUCY1A3</i>	4	156838654	T/C	0.01(0.04)	0.81	0.004(0.03)	0.88	0.08(0.04)	0.09	0.12(0.06)	0.04
rs1173766	<i>NPR3</i>	5	32840285	C/T	0.03(0.04)	0.34	0.01(0.02)	0.83	0.05(0.04)	0.17	0.04(0.07)	0.60
rs1799945	<i>HFE</i>	6	26199158	G/C	0.11(0.09)	0.25	0.02(0.06)	0.78	0.09(0.09)	0.32	0.03(0.11)	0.79
rs9266359	<i>HLA-B</i>	6	31440718	C/T	0.04(0.04)	0.32	0.01(0.02)	0.75	0.01(0.04)	0.82	0.02(0.04)	0.60
rs2021783	<i>CYP21A2</i>	6	32152829	C/T	0.05(0.04)	0.28	0.04(0.03)	0.20	0.04(0.04)	0.31	0.08(0.05)	0.14
rs4409766	<i>CYP17A1</i>	10	104606653	T/C	0.05(0.04)	0.18	0.03(0.03)	0.25	0.07(0.04)	0.06	0.07(0.05)	0.15
rs4757391	<i>SOX6</i>	11	16259515	C/T	0.07(0.04)	0.09	0.05(0.03)	0.08	0.07(0.04)	0.07	0.01(0.05)	0.82
rs17249754	<i>ATP2B1</i>	12	88584717	G/A	0.03(0.04)	0.46	0.01(0.02)	0.57	0.01(0.04)	0.86	0.04(0.05)	0.34
rs11066280	<i>ALDH2</i>	12	111302166	T/A	0.01(0.04)	0.76	0.06(0.03)	0.03	-0.003(0.04)	0.95	0.14(0.06)	0.02
rs1991391	<i>TBX3-TBX5</i>	12	113837049	G/A	0.04(0.05)	0.43	0.04(0.03)	0.23	0.07(0.05)	0.18	0.04(0.06)	0.52
rs35444	<i>TBX3</i>	12	114036820	A/G	0.11(0.04)	0.01	0.07(0.03)	5.82×10^{-3}	0.06(0.04)	0.17	-0.02(0.05)	0.66
rs11067763	<i>MED13L</i>	12	114682724	A/G	0.05(0.04)	0.17	0.03(0.02)	0.16	0.03(0.04)	0.49	0.06(0.04)	0.15

[0092]

rs1887320	<i>JAG1</i>	20	10913998	A/G	0.02(0.04)	0.68	0.03(0.02)	0.24	0.04(0.04)	0.33	0.10(0.04)	0.02
-----------	-------------	----	----------	-----	------------	------	------------	------	------------	------	------------	------

[0093] 由上述可以看出,单个遗传位点的作用相对较弱,本发明整合了 22 个位点的累积效应,可以显著增加遗传因素的预测能力。按照遗传风险评分值将研究对象 5 分位法分为 5 组,分别为第 1 组 7.765 ~ 13.465 (G1)、第 2 组 13.47 ~ 14.616 (G2)、第 3 组 14.62 ~ 15.545 (G3)、第 4 组 15.545 ~ 16.654 (G4) 和第 5 组 16.655 ~ 22.125 (G5)。与遗传风险评分最低的 1 组相比,计算其他各组发病风险的高低。随着遗传风险评分分值的增加,收缩压和舒张压水平也显著增加。(见表 4)。遗传风险评分分值从低到高的各组高血压累计发病率分别为 29.4%、31.6%、32.4%、36.2% 和 36.8%。在模型中加入性别、年龄、体重指数,吸烟、饮酒状况,脉搏、收缩压和舒张压,与遗传风险评分最低的分组个体相比,处在第二、三、四和第五个五分位法分组个体发生高血压风险分别增加 15%,21%,30%,和 40% (趋势检验 $P = 3.82 \times 10^{-5}$)。遗传风险评分从低到高各组心血管疾病累计发病率分别为 3.1%、3.5%、3.7%、4.2% 和 4.4%。在模型中加入性别、年龄、体重指数,吸烟、饮酒状况、糖尿病、总胆固醇、高密度脂蛋白胆固醇、心血管病家族史、收缩压和服用降压药情况,与遗传风险评分最低的分组个体相比,处在第二、三、四和第五个五分位法分组个体发生心血管疾病事

件发生风险分别增加 11%、15%、32%、和 26% (趋势检验 $P = 4.57 \times 10^{-3}$)。上述研究结果显示遗传风险评分是血压升高、高血压以及心血管疾病事件发生的独立危险因素。本发明的遗传风险评分可以用于预测血压升高、高血压以及心血管疾病事件发生。

[0094] 表 4. 遗传风险评分与血压升高、高血压以及心血管事件发生危险

[0095]

	第 2 分位数组对第 1 分位数组	第 3 分位数组对第 1 分位数组	第 4 分位数组对第 1 分位数组	第 5 分位数组对第 1 分位数组	趋势检验 P 值
年平均收缩压增加 (mmHg)	0.17(0.02~0.33)	0.22(0.07~0.38)	0.31(0.16~0.46)	0.33(0.18~0.49)	4.66×10^{-6}
年平均舒张压增加 (mmHg)	0.07(-0.03~0.16)	0.09(-0.01~0.19)	0.14(0.05~0.24)	0.22(0.12~0.32)	3.10×10^{-6}
高血压发病相对危险度 (OR 值)	1.15(0.97~1.37)	1.21(1.02~1.44)	1.30(1.1~1.55)	1.40(1.18~1.66)	3.82×10^{-5}
心血管疾病事件发生风险比 (HR 值)	1.11(0.9~1.31)	1.15(0.95~1.35)	1.32(1.13~1.52)	1.26(1.06~1.45)	4.57×10^{-3}

[0096] 高血压校正性别、年龄、体重指数,吸烟、饮酒状况,脉搏、收缩压和舒张压。

[0097] 心血管病事件发生校正性别、年龄、体重指数,吸烟、饮酒状况、糖尿病、总胆固醇、高密度脂蛋白胆固醇、心血管病家族史、收缩压和服用降压药。

[0098] 遗传风险评分在高血压和心血管事件发生风险预测模型的作用评估:

[0099] 在传统危险因素模型中,加入遗传风险评分 GRS,能提高对高血压和心血管事件发生风险的预测能力,C-指数变化为 0.2%,P 值都均 < 0.05 (表 5)。

[0100] 表 5. 传统危险因素加入遗传风险评分后的预测模型评估

[0101]

	模型 (不含遗传风险评分)	模型 (包含遗传风险评分)	P 值
	C 指数(95%CI)	C 指数 (95%CI)	
高血压预测模型			
传统危险因素包括年龄、性别、体重指数、吸烟、饮酒、脉搏、收缩压和舒张压	0.774(0.763-0.785)	0.776(0.765-0.787)	0.012
心血管事件发生预测模型			
传统危险因素包括性别、年龄、体重指数、吸烟、饮酒状况、糖尿病、总胆固醇、高密度脂蛋白胆固醇、心血管病家族史、收缩压和服用降压药情况	0.811(0.799-0.823)	0.813(0.801-0.825)	0.043

[0102] 四、建立高血压和心血管疾病的发病风险预测模型

[0103] 高血压预测模型中包括基因型信息 (遗传风险评分分组) 和传统危险因素包括年龄、性别、体重指数、吸烟、饮酒、脉搏、收缩压和舒张压,相关效应见表 6。心血管疾病发病风险预测模型中包括基因型信息 (遗传风险评分分组) 和传统危险因素包括性别、年龄、体重指数、吸烟、饮酒状况、糖尿病、总胆固醇、高密度脂蛋白胆固醇、心血管疾病家族史、收缩压和服用降压药情况,各个变量的效应见表 7。综合上述遗传和环境因素建立高血压和心血管

疾病发病风险预测模型：

[0104] 高血压发病风险预测模型

[0105] 高血压发病风险 = $1/(1+\exp(-Z))$ 。 $Z = -14.808+0.142 \times G2+0.193 \times G3+0.265 \times G4+0.339 \times G5+0.165 \times \text{性别}+0.034 \times \text{年龄}+0.063 \times \text{体重指数}+(-0.013) \times \text{是否吸烟}+0.159 \times \text{是否饮酒}+0.01 \times \text{脉搏}+0.061 \times \text{收缩压}+0.04 \times \text{舒张压}$ 。

[0106] 心血管疾病发病风险预测模型：

[0107] 心血管疾病发病风险 = $1/(1+\exp(-Z))$ 。 $Z = -8.558+0.126 \times G2+0.159 \times G3+0.297 \times G4+0.256 \times G5+(-0.273) \times \text{性别}+0.056 \times \text{年龄}+0.035 \times \text{体重指数}+0.19 \times \text{是否吸烟}+(-0.475) \times \text{是否饮酒}+0.358 \times \text{有无糖尿病}+0.006 \times \text{总胆固醇}+(-0.028) \times \text{高密度脂蛋白胆固醇}+0.643 \times \text{是否有心血管疾病家族史}+0.015 \times \text{收缩压}+0.309 \times \text{是否服用降压药}$ 。

[0108] 表 6. 高血压发生风险预测模型变量

[0109]

	效应 Beta	95%置信区间下限	95%置信区间上限
遗传风险评分分组			
第 2 分位 (G2) vs 第 1 分位 (G1)	0.142	-0.03	0.314
第 3 分位 (G3) vs 第 1 分位 (G1)	0.193	0.021	0.365

[0110]

第 4 分位 (G4) vs 第 1 分位 (G1)	0.265	0.095	0.435
第 5 分位 (G5) vs 第 1 分位 (G1)	0.339	0.169	0.509
常数项	-14.808	-15.764	-30.409
年龄	0.034	0.028	0.057
性别 (1=男, 0=女)	0.165	0.011	0.1
体重指数	0.063	0.047	0.101
饮酒 (1=饮酒, 0=不饮酒)	0.159	0.02	0.111
吸烟 (1=吸烟, 0=不吸烟)	-0.013	-0.166	-0.247
脉搏 (次/分)	0.01	0.001	0.006
收缩压 (mmHg)	0.061	0.055	0.111
舒张压 (mmHg)	0.04	0.031	0.064

[0111] 表 7. 心血管病事件发生风险预测模型变量

[0112]

	效应 Beta	95%置信区间下限	95%置信区间上限
遗传风险评分分组			
第2分位 (G2) vs 第1分位 (G1)	0.126	-0.085	-0.058
第3分位 (G3) vs 第1分位 (G1)	0.159	-0.049	0.009
第4分位 (G4) vs 第1分位 (G1)	0.297	0.094	0.288
第5分位 (G5) vs 第1分位 (G1)	0.256	0.054	0.209
常数项	-8.558	-9.356	-17.93
年龄	0.056	0.048	0.098
性别 (1=男, 0=女)	-0.273	-0.444	-0.783
体重指数	0.035	0.017	0.043
饮酒 (1=饮酒, 0=不饮酒)	-0.475	-0.665	-1.207
吸烟 (1=吸烟, 0=不吸烟)	0.19	0.012	0.115
糖尿病 (0=患有糖尿病, 0=无糖尿病)	0.358	0.041	0.243
总胆固醇(mg/dl)	0.006	0.004	0.009
高密度脂蛋白胆固醇(mg/dl)	-0.028	-0.033	-0.063
心血管病家族史 (1=有, 0=无)	0.643	0.467	1.004
收缩压 (mmHg)	0.015	0.012	0.025
是否服用降压药 (1=是, 0=否)	0.309	0.153	0.379

[0113] 将遗传风险评分分组和环境因素暴露信息放入风险预测模型计算出每名研究对象的高血压和心血管疾病的发病风险。采用受试者工作特征曲线 (ROC) 和模型拟合曲线 (Calibration) 评估高血压和心血管疾病发生风险预测模型。根据 ROC 曲线下面积评价该预测方法的价值。如图 10 所示, 高血压和心血管疾病发生的 ROC 下面积为分别为 0.776 和 0.813, 表明该模型预测性能良好, 可以很好的识别高危人群。模型拟合曲线显示高血压和心血管事件发生风险的预测值和观察值无显著差异, P 值分别为 0.882 和 0.24, 模型拟合良好, 见图 11。

[0114] 实施例二、待测个体高血压和心血管疾病发病风险评估实例

[0115] 根据模型计算获得该研究对象高血压和心血管疾病事件发生的风险, 判断该研究对象是否为高危人群。完成个体化健康指导报告, 报告高血压和心血管疾病事件发生风险的高低。个体化健康指导报告是在基因型结果的基础上, 结合生活方式危险因素评估研究对象高血压和心血管疾病事件发生的风险, 并制定出针对研究对象的个体化健康行动方案。

[0116] 根据实施例一中所检测的 26262 名个体 22 个多态性位点 (rs880315、rs17030613、rs10745332、rs16849225、rs820430、rs9815354、rs9810888、rs1902859、rs6825911、rs13143871、rs1173766、rs1799945、rs9266359、rs2021783、rs4409766、rs4757391、rs17249754、rs11066280、rs1991391、rs35444、rs11067763 和 rs1887320) 情况, 平均 6.9 年的随访观察结果, 制定血压相关位点危险等位基因遗传效应速查图表 (图 12) 以及遗传风险评分速查图表 (图 13)。遗传风险评分最低组到最高组, 高血压和心血管疾病发病风险逐渐升高, 存在明显的剂量效应关系。

[0117] 待测个体王某, 中国汉族人, 男性, 73 岁, 利用本发明的用于评估待测个体高血压

和 / 或心血管疾病发病风险的检测装置评估其患高血压和 / 或心血管疾病发病风险高低, 所述的用于评估待测个体高血压和 / 或心血管疾病发病风险的检测装置包括检测单元、数据分析单元、评估单元, 主要按照如下步骤进行:

[0118] (一) 采集空腹血并收集性别、年龄、血压水平、血脂水平、身高、体重、吸烟、饮酒和疾病史等相关信息, 填入下表 (表 8)。

[0119] 表 8 待测个体信息登记表

[0120]

姓名	王某	年龄	73
性别 (1=男, 0=女)			1
脉搏 (次/分)	70	体重 (kg)	73
身高 (m)	1.61	BMI (kg/m ²)	28
饮酒 (1=饮酒, 0=不饮酒)			0
吸烟 (1=吸烟, 0=不吸烟)			0
糖尿病 (1=有, 0=无)			1
是否有心血管病家族史 (1=有, 0=无)			0
是否服用降压药 (1=是, 0=否)			0
总胆固醇(mg/dl)			299
高密度脂蛋白胆固醇(mg/dl)			48
收缩压 SBP (mmHg)			135
舒张压 DBP (mmHg)			80

[0121] (二) 分离待测个体抗凝血中 DNA, 利用检测单元检测 22 个位点的基因型, 该实施例中, 检测单元主要包括 iPLEX Sequenom MassARRAY 分型系统, 具体操作参照说明书的记载进行。

[0122] 经检测, 王某的 22 个位点的基因型分别为: rs880315 为 CC; rs17030613 为 AC; rs10745332 为 AG; rs16849225 为 CC; rs820430 为 AG; rs9815354 为 AG; rs9810888 为 TT; rs1902859 为 CT; rs6825911 为 CC; rs13143871 为 CC; rs1173766 为 CC; rs1799945 为 CC; rs9266359 为 CC; rs2021783 为 CC; rs4409766 为 TT; rs4757391 为 TT; rs17249754 为 GA; rs11066280 为 AT; rs1991391 为 GG; rs35444 为 AA; rs11067763 为 AG; rs1887320 为 AG; 其风险等位基因的数目分别为: 2、1、1、2、1、1、0、1、2、0、2、0、2、2、0、1、1、2、2、1 和 1。

[0123] (三) 对检测结果进行分析处理:

[0124] (1) 利用数据分析单元对检测结果进行分析处理, 其中包括图 12 所示血压相关位点危险等位基因遗传效应速查表将所述 22 个单核苷酸多态性位点的检测结果对照该图表, 查找出各位点相应的危险等位基因遗传贡献, 相加, 得到遗传风险评分, 具体如下:

[0125] 遗传风险评分 = $\sum \beta_i \times N_i$, 其中 β_i 指第 i 个 SNP 的收缩压和舒张压效应值的平均值,

[0126] N_i 指研究对象所携带第 i 个 SNP 的危险等位基因数目。

[0127] 王某高血压发病遗传风险评分为 15.505, 查阅图 13, 处于遗传风险评分第 3 组。

[0128] (2) 利用高血压发病预测模型, 计算高血压发病风险:

[0129] 高血压发病风险 = $1 / (1 + \exp(-Z))$ 。 $Z = -14.808 + 0.142 \times G_2 + 0.193 \times G_3 + 0.265 \times G_4 + 0.339 \times G_5 + 0.165 \times \text{性别} + 0.034 \times \text{年龄} + 0.063 \times \text{体重指数} + (-0.013) \times \text{是否吸烟}$

+0.159× 是否饮酒 +0.01× 脉搏 +0.061× 收缩压 +0.04× 舒张压。具体为：

[0130] $Z = -14.808 + 0.142 \times 0 + 0.193 \times 1 + 0.265 \times 0 + 0.339 \times 0 + 0.165 \times 1 + 0.034 \times 73 + 0.063 \times 28 + (-0.013) \times 0 + 0.159 \times 0 + 0.01 \times 70 + 0.061 \times 135 + 0.04 \times 80 = 1.931$ 。

[0131] 高血压发病风险 = $1 / (1 + \exp(-Z)) = 0.87$ 。

[0132] 王某高血压发病风险为 0.87, 即 87%。

[0133] (3) 利用心血管疾病发病风险预测模型, 计算心血管疾病发病风险：

[0134] 心血管疾病发病风险 = $1 / (1 + \exp(-Z))$ 。 $Z = -8.558 + 0.126 \times G2 + 0.159 \times G3 + 0.297 \times G4 + 0.256 \times G5 + (-0.273) \times \text{性别} + 0.056 \times \text{年龄} + 0.035 \times \text{体重指数} + 0.19 \times \text{是否吸烟} + (-0.475) \times \text{是否饮酒} + 0.358 \times \text{有无糖尿病} + 0.006 \times \text{总胆固醇} + (-0.028) \times \text{高密度脂蛋白胆固醇} + 0.643 \times \text{是否有心血管疾病家族史} + 0.015 \times \text{收缩压} + 0.309 \times \text{是否服用降压药}$ 。具体为：

[0135] $Z = -8.558 + 0.126 \times 0 + 0.159 \times 1 + 0.297 \times 0 + 0.256 \times 0 + (-0.273) \times 1 + 0.056 \times 73 + 0.035 \times 28 + 0.19 \times 0 + (-0.475) \times 0 + 0.358 \times 1 + 0.006 \times 299 + (-0.028) \times 48 + 0.643 \times 0 + 0.015 \times 135 + 0.309 \times 0 = -0.771$ 。心血管疾病发病风险 = $1 / (1 + \exp(-Z)) = 0.31$ 。

[0136] 王某心血管疾病发病风险为 0.31, 即 31%。

[0137] (三) 对王某进行高血压和心血管疾病发病风险评估：

[0138] 通过检测结果的分析处理, 王某高血压发病风险为 87%。

[0139] 按照 Framingham 研究对心血管疾病发生风险分类 (0-10% 之间为低危险, 10-20% 之间为中度危险, 大于 20% 为高危人群), 王某心血管疾病发病风险为 31%, 属于高危人群。

[0140] 通过遗传和环境因素预测确定王某的高血压和心血管疾病发生风险较高, 建议其除采取健康生活方式外, 注意控制血糖、血脂, 定期进行健康体检, 如有异常及时就医。本发明提供了一种根据多个单核苷酸多态性多态性预测个体高血压和 / 或心血管疾病发病风险的模型, 能够很好的识别高血压和心血管疾病高危人群, 从而针对性的进行干预, 达到早期发现、早期预防高血压和 / 或心血管疾病发生的目的。利用研究对象基因型结果计算遗传风险评分, 在传统危险因素模型中加入遗传风险评分, 完成待测个体基因型检测报告和个体化健康指导报告, 并报告其发生高血压和 / 或心血管疾病风险的高低。

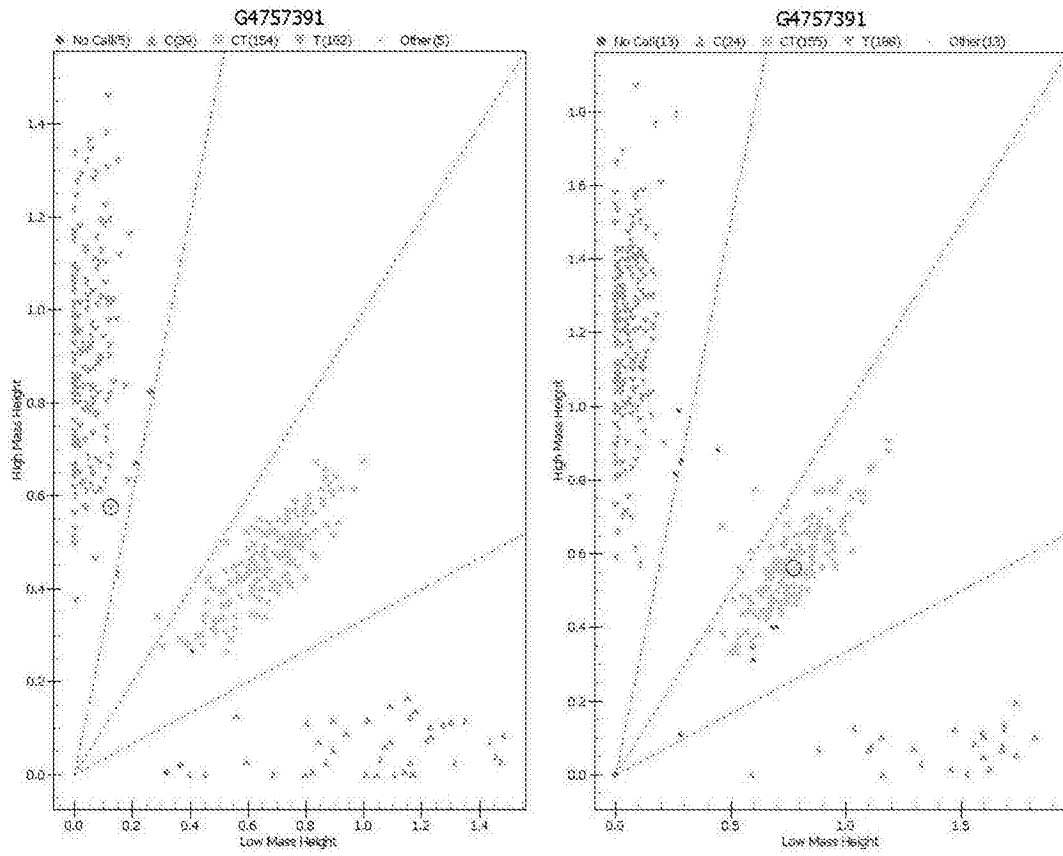


图 1

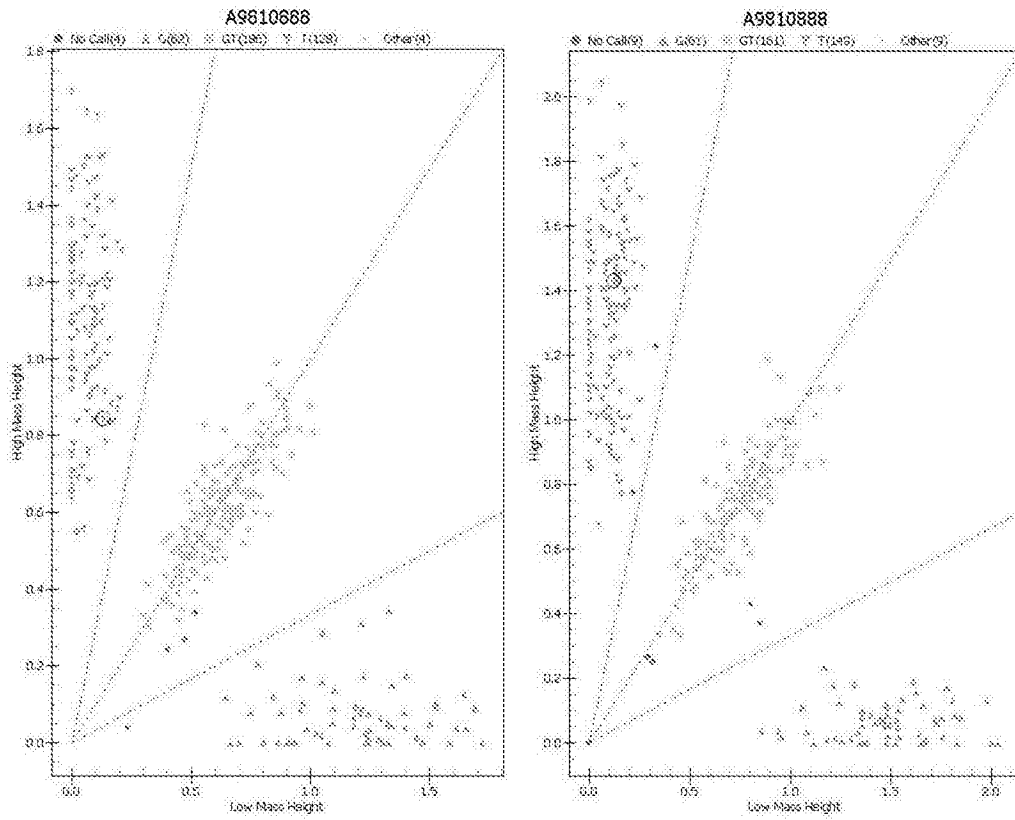


图 2

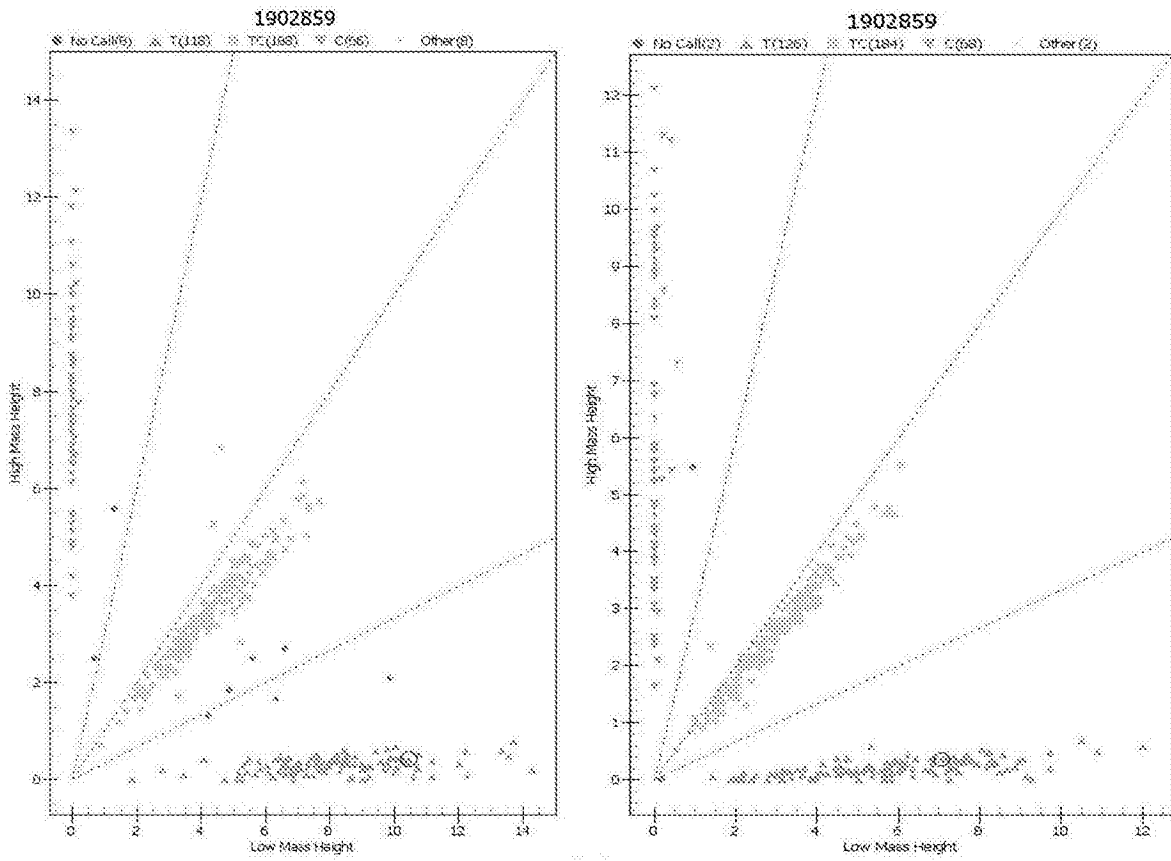


图 3

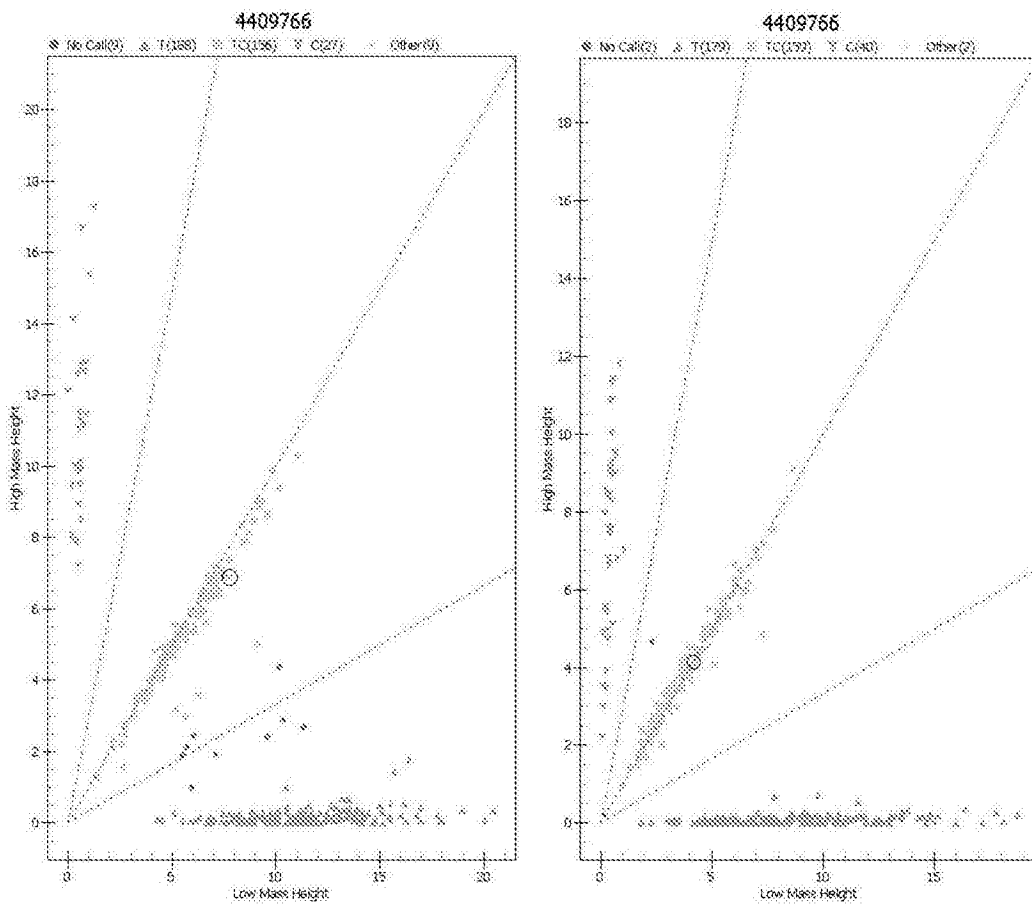


图 4

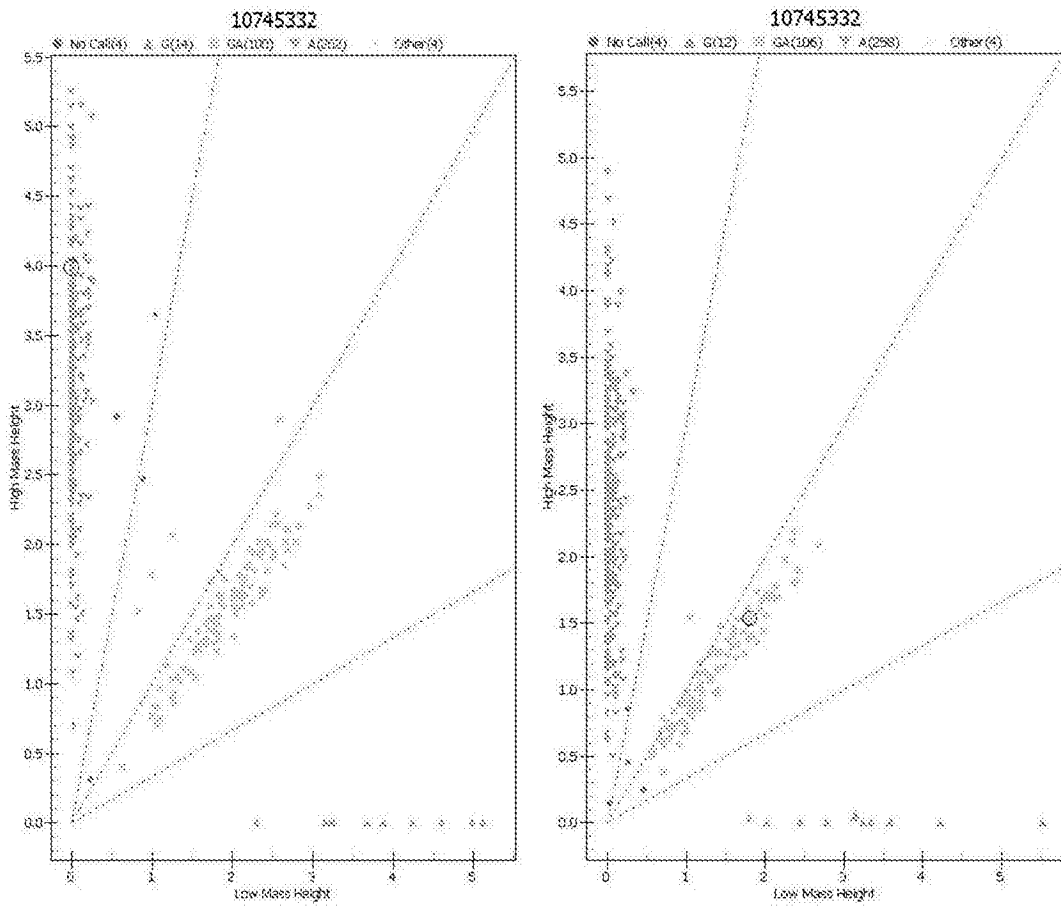


图 5

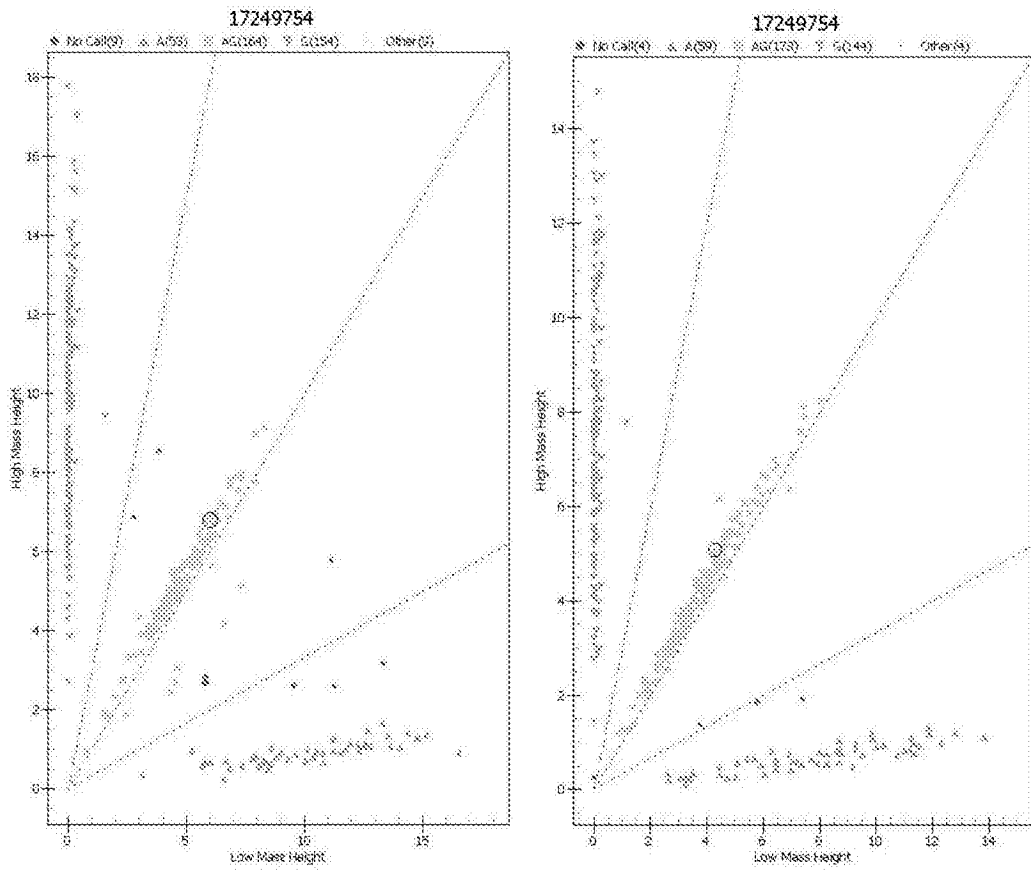


图 6

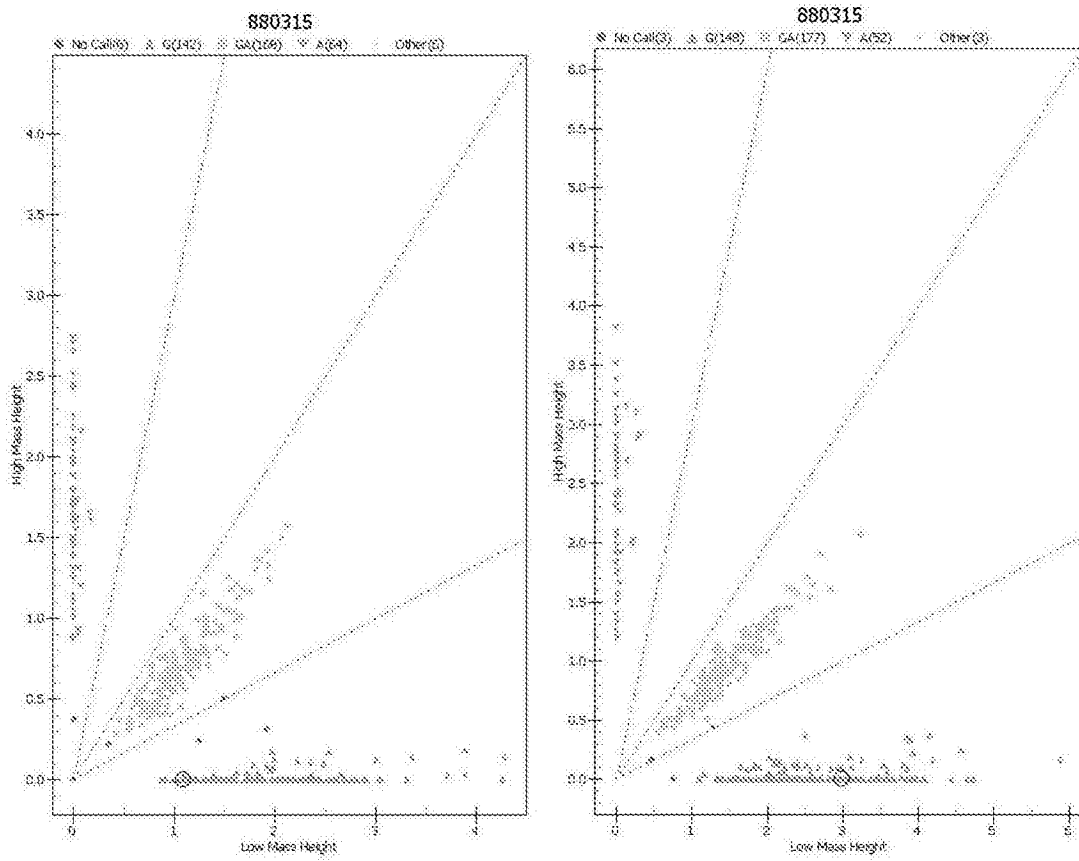


图 7

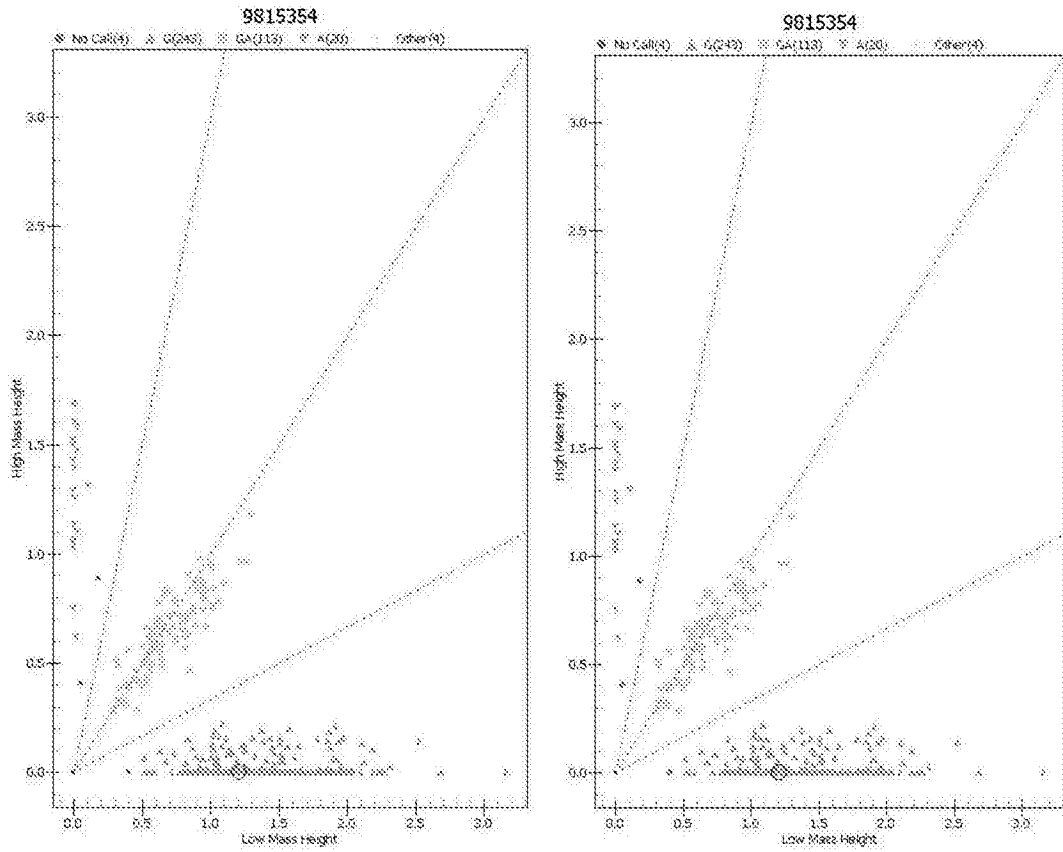


图 8

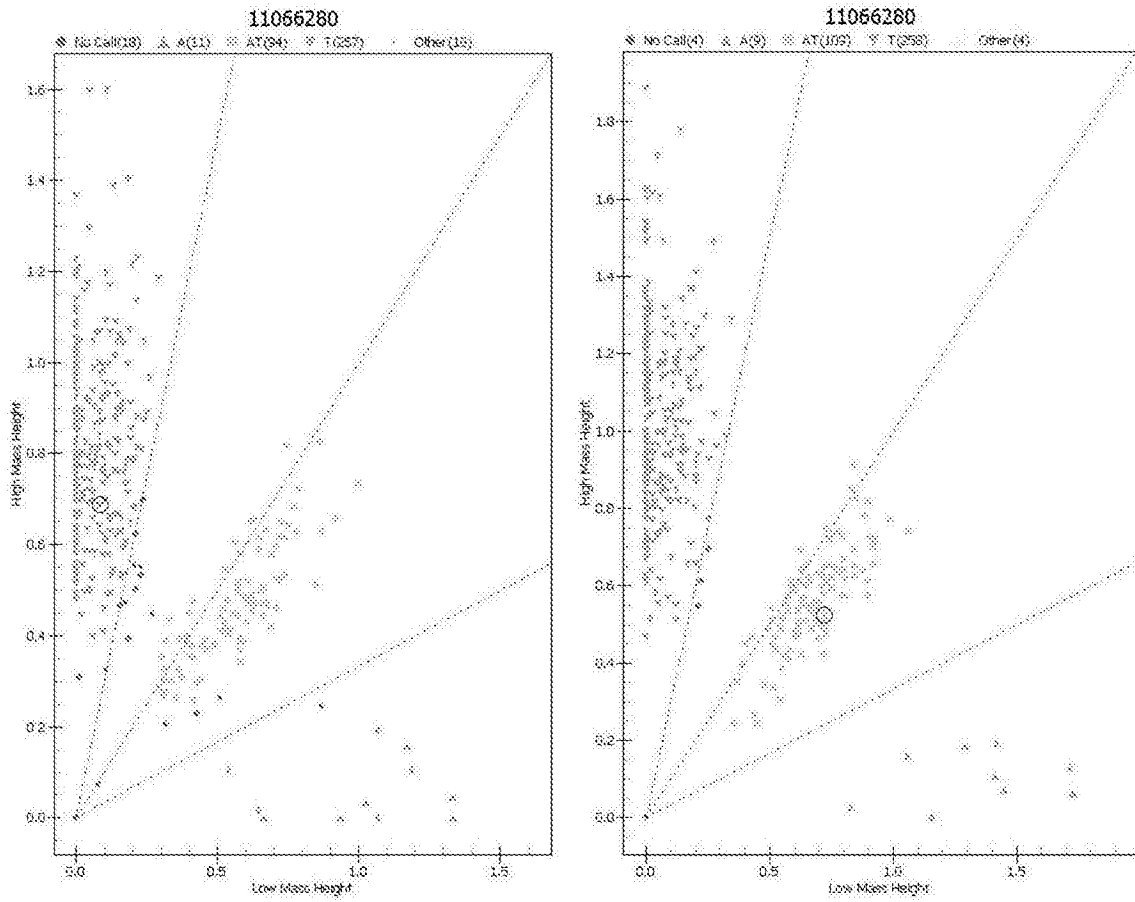


图 9

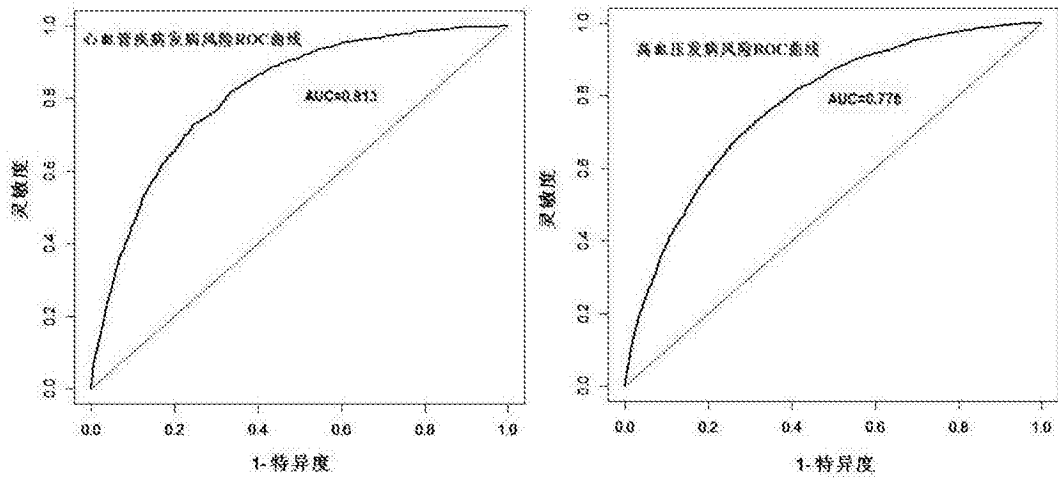


图 10

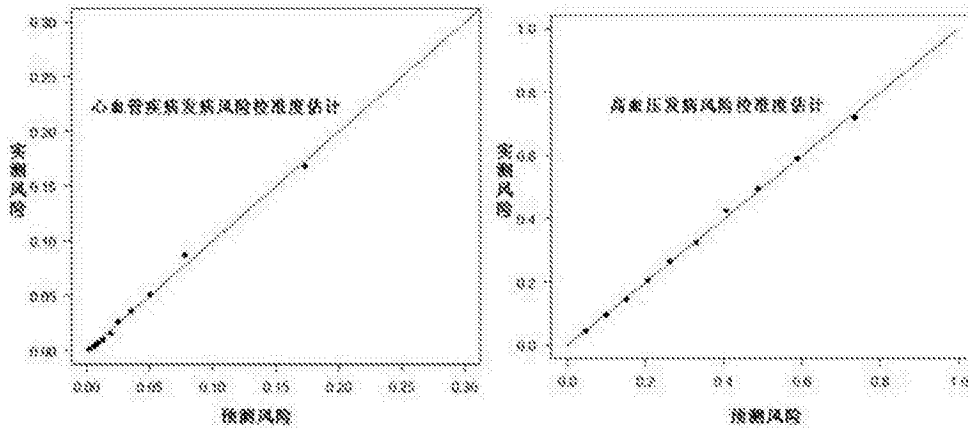


图 11

★血压相关位点危险等位基因遗传效应速查图表★

SNP	危险/非危险等位基因	收缩压效应 β SBP*	舒张压效应 β DBP*	平均效应 β 加权系数	基因型	危险等位基因数 N1	血压效应值 β ₁	基因型	危险等位基因数 N1	血压效应值 β ₁	基因型	危险等位基因数 N1	血压效应值 β ₁
rs889315	C/T	0.97	0.46	0.715	CC	2	1.43	CT	1	0.715	TT	0	0
rs17030613	C/A	0.49	0.33	0.435	CC	2	0.87	CA	1	0.435	AA	0	0
rs10745373	A/G	0.96	0.53	0.745	AA	2	1.49	AG	1	0.745	GG	0	0
rs16349225	C/T	0.45	0.1	0.275	CC	2	0.85	CT	1	0.275	TT	0	0
rs820430	A/G	0.76	0.27	0.515	AA	2	1.03	AG	1	0.515	GG	0	0
rs9815354	A/G	0.1	0.43	0.265	AA	2	0.53	AG	1	0.265	GG	0	0
rs9810688	G/T	0.53	0.39	0.46	GG	2	0.92	GT	1	0.46	TT	0	0
rs1902859	C/T	1.34	0.71	1.035	CC	2	2.05	CT	1	1.035	TT	0	0
rs6835911	C/T	0.6	0.39	0.495	CC	2	0.99	CT	1	0.495	TT	0	0
rs13143371	T/C	0.96	0.49	0.725	TT	2	1.45	TC	1	0.725	CC	0	0
rs1173766	C/T	0.63	0.36	0.495	CC	2	0.99	CT	1	0.495	TT	0	0
rs1799945	G/C	0.95	0.38	0.915	GG	2	1.63	GC	1	0.915	CC	0	0
rs9266359	C/T	0.44	0.29	0.365	CC	2	0.73	CT	1	0.365	TT	0	0
rs2031783	C/T	0.68	0.49	0.585	CC	2	1.17	CT	1	0.585	TT	0	0
rs4409766	T/C	1.24	0.59	0.915	TT	2	1.63	TC	1	0.915	CC	0	0
rs4757391	C/T	0.67	0.36	0.515	CC	2	1.05	CT	1	0.515	TT	0	0
rs17249754	G/A	1.03	0.52	0.775	GG	2	1.55	GA	1	0.775	AA	0	0
rs11066359	T/A	0.96	0.62	0.79	TT	2	1.58	TA	1	0.79	AA	0	0
rs1991391	G/A	0.6	0.21	0.405	GG	2	0.81	GA	1	0.405	AA	0	0
rs35444	A/G	0.33	0.36	0.595	AA	2	1.19	AG	1	0.595	GG	0	0
rs11067763	A/G	0.81	0.51	0.66	AA	2	1.32	AG	1	0.66	GG	0	0
rs1887329	A/G	0.78	0.43	0.605	AA	2	1.21	AG	1	0.605	GG	0	0

图 12

★遗传风险评分（GRS）速查图表★					
分组（G）	G1	G2	G3	G4	G5
GRS	7.765~13.465	13.466~14.616	14.617~15.545	15.545~16.654	16.655~22.125

图 13