



(12) 发明专利申请

(10) 申请公布号 CN 103451284 A

(43) 申请公布日 2013. 12. 18

(21) 申请号 201310370253. 3

(22) 申请日 2013. 08. 22

(71) 申请人 中国科学院生物物理研究所
地址 100101 北京市朝阳区大屯路 15 号

(72) 发明人 刘光慧 顾颖
胡安·卡洛斯·伊斯毕华·贝尔蒙特
曲静 杨济平 张维琦 任若通

(74) 专利代理机构 北京纪凯知识产权代理有限公司 11245

代理人 关畅

(51) Int. Cl.

C12Q 1/68 (2006. 01)

C12Q 1/04 (2006. 01)

C12N 15/12 (2006. 01)

权利要求书2页 说明书12页 附图6页

(54) 发明名称

一组人类心肌细胞的新型分子标记物及其应用

(57) 摘要

本发明公开了一组人类心肌细胞的新型分子标记物及其应用。该组标记物由如下 29 种基因组成: POPDC2、ALPK2、TRIM55、ASPH、FILIP1、HSPB7、NPPA、KBTBD10、MYL4、CRYAB、LRRN4、MYOZ2、OBSCN、SMPX、NRK、ABLM1、SYNPO2L、ANKRD1、TNNI1、MYH6、ABRA、TPM1、C15orf52、MYOM1、MASP1、MYBPC3、TNS1、FLNC、F2RL2。本发明的标记物在心肌细胞的筛选和鉴定上具有重要作用,并且在人类心脏早期发育研究、心脏疾病模型构建、心脏疾病的治疗以及心脏疾病药物靶点的开发或药物研发中都具有潜在的作用与极大的意义。

1. 一组用于鉴定或辅助鉴定人类心肌细胞的标记物,由如下 29 种基因组成:POPDC2、ALPK2、TRIM55、ASPH、FILIP1、HSPB7、NPPA、KBTBD10、MYL4、CRYAB、LRRN4、MYOZ2、OBSCN、SMPX、NRK、ABLIM1、SYNP02L、ANKRD1、TNNI1、MYH6、ABRA、TPM1、C15orf52、MYOM1、MASP1、MYBPC3、TNS1、FLNC、F2RL2;

所述 29 种基因满足如下条件:与人类胚胎多能干细胞和/或人类神经干细胞相比,在人类心肌细胞中该 29 种基因的启动子区 CpG 岛甲基化水平均显著降低、且在人类心肌细胞中该 29 种基因的 mRNA 表达量均显著提高。

2. 根据权利要求 1 所述的标记物,其特征在于:所述该 29 种基因的启动子区 CpG 岛甲基化水平均显著降低是指在人类心肌细胞中每种基因的启动子区 CpG 岛甲基化的量均比人类胚胎多能干细胞和/或人类神经干细胞中相同基因的启动子区 CpG 岛甲基化的量降低 5% 以上;

所述该 29 种基因的 mRNA 表达量均显著提高是指在人类心肌细胞中每种基因的 mRNA 表达量均是人类胚胎多能干细胞和/或人类神经干细胞中相同基因 mRNA 表达量的 2 倍以上。

3. 权利要求 1 或 2 所述标记物在鉴定人心肌细胞中的应用,所述应用为非疾病治疗或诊断方法;

或,权利要求 1 或 2 所述标记物在制备、开发或设计鉴定人心肌细胞的产品中的应用。

4. 权利要求 1 或 2 所述标记物在从人心肌细胞、人胚胎多能干细胞和人神经干细胞中鉴定出人心肌细胞中的应用;

或,权利要求 1 或 2 所述标记物在从人心肌细胞和人胚胎多能干细胞中鉴定出人心肌细胞中的应用。

5. 权利要求 1 或 2 所述标记物在制备、开发或设计具有如下功能的产品中的应用:从人心肌细胞、人胚胎多能干细胞和人神经干细胞中鉴定出人心肌细胞;

或,权利要求 1 或 2 所述标记物在制备、开发或设计具有如下功能的产品中的应用:从人心肌细胞和人胚胎多能干细胞中鉴定出人心肌细胞。

6. 根据权利要求 3-5 任一所述的应用,其特征在于:所述人心肌细胞是由所述人胚胎多能干细胞分化而来的;所述人神经干细胞是由所述人胚胎多能干细胞分化而来的。

7. 检测 29 种基因的启动子区 CpG 岛甲基化水平的物质和检测 29 种基因的 mRNA 表达量的物质在制备鉴定人类心肌细胞的试剂盒中的应用;

所述 29 种基因为 POPDC2、ALPK2、TRIM55、ASPH、FILIP1、HSPB7、NPPA、KBTBD10、MYL4、CRYAB、LRRN4、MYOZ2、OBSCN、SMPX、NRK、ABLIM1、SYNP02L、ANKRD1、TNNI1、MYH6、ABRA、TPM1、C15orf52、MYOM1、MASP1、MYBPC3、TNS1、FLNC、F2RL2。

8. 检测 29 种基因的启动子区 CpG 岛甲基化水平的物质和检测 29 种基因的 mRNA 表达量的物质在制备具有如下功能的试剂盒中的应用:从人心肌细胞、人胚胎多能干细胞和人神经干细胞中鉴定出人心肌细胞;

所述 29 种基因为 POPDC2、ALPK2、TRIM55、ASPH、FILIP1、HSPB7、NPPA、KBTBD10、MYL4、CRYAB、LRRN4、MYOZ2、OBSCN、SMPX、NRK、ABLIM1、SYNP02L、ANKRD1、TNNI1、MYH6、ABRA、TPM1、C15orf52、MYOM1、MASP1、MYBPC3、TNS1、FLNC、F2RL2。

9. 检测 29 种基因的启动子区 CpG 岛甲基化水平的物质和检测 29 种基因的 mRNA 表达

量的物质在制备具有如下功能的试剂盒中的应用：从人心肌细胞和人胚胎多能干细胞中鉴定出人心肌细胞；

所述 29 种基因为 POPDC2、ALPK2、TRIM55、ASPH、FILIP1、HSPB7、NPPA、KBTBD10、MYL4、CRYAB、LRRN4、MYOZ2、OBSCN、SMPX、NRK、ABLIM1、SYNP02L、ANKRD1、TNNI1、MYH6、ABRA、TPM1、C15orf52、MYOM1、MASP1、MYBPC3、TNS1、FLNC、F2RL2。

10. 根据权利要求 7-9 任一所述的应用，其特征在于：所述人心肌细胞是由所述人胚胎多能干细胞分化而来的；所述人神经干细胞是由所述人胚胎多能干细胞分化而来的。

一组人类心肌细胞的新型分子标记物及其应用

技术领域

[0001] 本发明涉及一组人类心肌细胞的新型分子标记物及其应用。

背景技术

[0002] 心脏病是威胁人类健康的杀手,每年都有上千万人因罹患各类心脏疾病失去生命,给家庭和公共卫生事业带来了极大的压力和负担。无论在发达国家还是发展中国家,心脏病的发病率和病死率都居高不下,然而面对如此严峻的考验,目前心脏疾病的治疗却不容乐观,药物治疗停留在缓解症状的阶段,手术治疗如搭桥手术等开支较大且存在一定的风险和可能的副作用,介入治疗也仍处于探索阶段,鲜有标本兼治的治疗方法问世。究其原因,很大一方面是人类对心脏以及心脏疾病认识的不足。心脏疾病的病因追根溯源是心肌细胞的缺失或功能的紊乱。因而未来根治心脏疾病的一个思路便是研究心脏病的分子致病机理,寻找影响心脏纤维化或者心肌细胞增殖的决定性因子,实现对紊乱的心肌细胞进行人为调控使其恢复正常功能的目的(1)。

[0003] 近年来,随着诱导型多能干细胞(induced Pluripotency Stem Cells, iPSCs)技术的产生,人们能够在体外获得 iPSCs 并分化为病人特异的心肌细胞,通过基因矫正或其他手段得到“健康”的病人来源的心肌细胞(1);利用转分化(transdifferentiation)技术可以不经 iPSCs 过程由其他便于获得的细胞类型直接获得心肌细胞(2),这些技术的发展为将来实现在病人体内的细胞移植治疗提供了可能。

[0004] 1. Mordwinkin NM, Burrige PW, Wu JC. 2013. A review of human pluripotent stem cell-derived cardiomyocytes for high-throughput drug discovery, cardiotoxicity screening, and publication standards. *Journal of cardiovascular translational research*6:22-30.

[0005] 2. Wada R, Muraoka N, Inagawa K, Yamakawa H, Miyamoto K, Sadahiro T, Umei T, Kaneda R, Suzuki T, Kamiya K, Tohyama S, Yuasa S, Kokaji K, Aeba R, Yozu R, Yamagishi H, Kitamura T, Fukuda K, Ieda M. 2013. Induction of human cardiomyocyte-like cells from fibroblasts by defined factors. *Proceedings of the National Academy of Sciences*110 :12667-12672.

发明内容

[0006] 本发明的目的是提供一组人类心肌细胞的新型分子标记物及其应用。

[0007] 本发明提供的一组用于鉴定或辅助鉴定人类心肌细胞的标记物,由如下 29 种基因组成:POPDC2、ALPK2、TRIM55、ASPH、FILIP1、HSPB7、NPPA、KBTBD10、MYL4、CRYAB、LRRN4、MYOZ2、OBSCN、SMPX、NRK、ABLIM1、SYNPO2L、ANKRD1、TNNI1、MYH6、ABRA、TPM1、C15orf52、MYOM1、MASP1、MYBPC3、TNS1、FLNC、F2RL2。

[0008] 所述 29 种基因满足如下条件:与人类胚胎多能干细胞和/或人类神经干细胞相比,在人类心肌细胞中该 29 种基因的启动子区 CpG 岛甲基化水平均显著降低、且在人类心

肌细胞中该 29 种基因的 mRNA 表达量均显著提高。

[0009] 上述标记物中,所述该 29 种基因的启动子区 CpG 岛甲基化水平均显著降低是指在人类心肌细胞中每种基因的启动子区 CpG 岛甲基化的量均比人类胚胎多能干细胞和 / 或人类神经干细胞中相同基因的启动子区 CpG 岛甲基化的量降低 5% 以上。

[0010] 所述该 29 种基因的 mRNA 表达量均显著提高是指在人类心肌细胞中每种基因的 mRNA 表达量均是人类胚胎多能干细胞和 / 或人类神经干细胞中相同基因 mRNA 表达量的 2 倍以上。

[0011] 上述任一所述标记物在鉴定人心肌细胞中的应用也属于本发明的保护范围。

[0012] 上述应用中,所述应用为非疾病治疗或诊断方法。

[0013] 上述任一所述标记物在制备、开发或设计鉴定人心肌细胞的产品中的应用也属于本发明的保护范围。

[0014] 上述任一所述标记物在从人心肌细胞、人胚胎多能干细胞和人神经干细胞中鉴定出人心肌细胞中的应用也属于本发明的保护范围。

[0015] 上述任一所述标记物在从人心肌细胞和人胚胎多能干细胞中鉴定出人心肌细胞中的应用也属于本发明的保护范围。

[0016] 上述任一所述标记物在制备、开发或设计具有如下功能的产品中的应用也属于本发明的保护范围:从人心肌细胞、人胚胎多能干细胞和人神经干细胞中鉴定出人心肌细胞。

[0017] 上述任一所述标记物在制备、开发或设计具有如下功能的产品中的应用也属于本发明的保护范围:从人心肌细胞和人胚胎多能干细胞中鉴定出人心肌细胞。

[0018] 上述任一所述的应用中,所述人心肌细胞是由所述人胚胎多能干细胞分化而来的;所述人神经干细胞是由所述人胚胎多能干细胞分化而来的。

[0019] 检测 29 种基因的启动子区 CpG 岛甲基化水平的物质和检测 29 种基因的 mRNA 表达量的物质在制备鉴定人类心肌细胞的试剂盒中的应用也属于本发明的保护范围。

[0020] 所述 29 种基因为 POPDC2、ALPK2、TRIM55、ASPH、FILIP1、HSPB7、NPPA、KBTBD10、MYL4、CRYAB、LRRN4、MYOZ2、OBSCN、SMPX、NRK、ABLIM1、SYNPO2L、ANKRD1、TNNT1、MYH6、ABRA、TPM1、C15orf52、MYOM1、MASP1、MYBPC3、TNS1、FLNC、F2RL2。

[0021] 检测 29 种基因的启动子区 CpG 岛甲基化水平的物质和检测 29 种基因的 mRNA 表达量的物质在制备具有如下功能的试剂盒中的应用也属于本发明的保护范围:从人心肌细胞、人胚胎多能干细胞和人神经干细胞中鉴定出人心肌细胞。

[0022] 所述 29 种基因为 POPDC2、ALPK2、TRIM55、ASPH、FILIP1、HSPB7、NPPA、KBTBD10、MYL4、CRYAB、LRRN4、MYOZ2、OBSCN、SMPX、NRK、ABLIM1、SYNPO2L、ANKRD1、TNNT1、MYH6、ABRA、TPM1、C15orf52、MYOM1、MASP1、MYBPC3、TNS1、FLNC、F2RL2。

[0023] 检测 29 种基因的启动子区 CpG 岛甲基化水平的物质和检测 29 种基因的 mRNA 表达量的物质在制备具有如下功能的试剂盒中的应用也属于本发明的保护范围:从人心肌细胞和人胚胎多能干细胞中鉴定出人心肌细胞。

[0024] 所述 29 种基因为 POPDC2、ALPK2、TRIM55、ASPH、FILIP1、HSPB7、NPPA、KBTBD10、MYL4、CRYAB、LRRN4、MYOZ2、OBSCN、SMPX、NRK、ABLIM1、SYNPO2L、ANKRD1、TNNT1、MYH6、ABRA、TPM1、C15orf52、MYOM1、MASP1、MYBPC3、TNS1、FLNC、F2RL2。

[0025] 上述任一所述应用中,所述人心肌细胞是由所述人胚胎多能干细胞分化而来的;

所述人神经干细胞是由所述人胚胎多能干细胞分化而来的。

[0026] 上述任一所述应用中,所述由人胚胎多能干细胞分化而来的人心肌细胞中和所述由人胚胎多能干细胞分化而来的人神经干细胞中,所述人胚胎多能干细胞为同一株系的人胚胎多能干细胞。

[0027] 由于 H9-ESC, H9-NSC 及 H9-CM 具有完全相同的基因型与遗传背景,将 H9-ESC 或 H9-NSC 作为对照,进行转录组及全基因组 DNA 甲基化分析具有严格的无偏见性。通过大规模比较 H9-ESC, H9-NSC 及 H9-CM 的基因表达谱和全基因组 DNA 甲基化谱,筛选出 29 种新型分子标记物。本发明涉及的新型分子标记物在心肌细胞的筛选和鉴定上具有重要作用,并且在心脏疾病模型构建、心脏疾病的治疗以及心脏疾病药物靶点开发或药物研发中都具有潜在的作用与极大的意义。

附图说明

[0028] 图 1 为人类心肌细胞的鉴定。

[0029] 图 2 为人类胚胎多能干细胞、人类神经干细胞和人类心肌细胞基因表达热图以及 qPCR 验证。

[0030] 图 3 为人类胚胎多能干细胞、人类神经干细胞和人类心肌细胞中全基因组启动子区 CpG 岛甲基化水平热图。

[0031] 图 4 为编码心肌结构蛋白与心脏转录因子的基因的 DNA 去甲基化与基因表达水平正相关性分析。

[0032] 图 5 为双核心稳定调控网络。

[0033] 图 6 为功能相关性网络分析。

[0034] 图 7 为基因 - 疾病网络分析。

具体实施方式

[0035] 下述实施例中所使用的实验方法如无特殊说明,均为常规方法。

[0036] 下述实施例中所用的材料、试剂等,如无特殊说明,均可从商业途径得到。

[0037] 人类胚胎多能干细胞系 H9 购自美国 WiCell Research Institute, 货号为 WA09(H9)-DL-7 ;

[0038] 丝裂霉素购自美国 Sigma 公司, 货号为 M0503 ;

[0039] 小鼠胚胎成纤维细胞购自美国 Millipore 公司, 货号为 PMEF-CFL ;

[0040] 细胞外基质 (qualified-Matrigel) 购自美国 BD Biosciences, 货号为 354277 ;

[0041] DMEM/F12 培养基购自 Invitrogen 公司, 货号为 11320-033 ;

[0042] CDF12 培养基配方如下 :

[0043] DMEM/F12(购自 Invitrogen, 货号为 11320-033)、0.1mM 非必需氨基酸 (购自 Invitrogen, 货号为 11140-050)、1mM GlutaMAX™ 二肽 (购自 Invitrogen, 货号为 35050-061)、20%Knockout 血清替代物 (购自 Invitrogen, 货号为 N10828-028)、1%青霉素 / 链霉素 (购自 Invitrogen, 货号为 15070-063)、55 μ M β - 巯基乙醇 (购自 Invitrogen, 货号为 21985-023) 和 10ng/ml Human FGF2(购自 Joint Protein Central) ;

[0044] RPMI1640 培养基购自 Invitrogen 公司, 货号为 11875119 ;

- [0045] B27 添加剂购自 Invitrogen 公司, 货号为 0080085SA ;
- [0046] B27 (不含 Insulin) 添加剂购自 Invitrogen 公司, 货号为 0050129SA ;
- [0047] mTeSR 培养基购自美国 StemCell Technologies ;
- [0048] ROCK 抑制剂 Y-27632 购自 Sigma-Aldrich, 货号为 Y0503 ;
- [0049] CHIR99021 购自 Selleck ;
- [0050] Wnt 抑制剂 IWP4 购自 Stemgent ;
- [0051] Triton X-100 购自 Sigma, 货号 T-8787 ;
- [0052] 鼠抗 cTnT 购自 Lab Vision ;
- [0053] 鼠抗 MF20 购自 Developmental Studies Hybridoma Bank ;
- [0054] 兔抗 MLC2v 购自 ProteinTech Group ;
- [0055] 鼠抗 α -Actinin 购自 Sigma-Aldrich ;
- [0056] 二抗 Alex Fluor488goat anti-mouse IgG 购自 Life Technologies ;
- [0057] 二抗 Alex Fluor568goat anti-rabbit IgG 购自 Life Technologies ;
- [0058] 人类胚胎多能干细胞系 H9 来源的神经干细胞(hNSCs) 及分化方案在文献“Liu GH, Qu J, Suzuki K, Nivet E, Li M, Montserrat N, Yi F, Xu X, Ruiz S, Zhang W, Wagner U, Kim A, Ren B, Li Y, Goebel A, Kim J, Soligalla RD, Dubova I, Thompson J, Yates J, 3rd, Esteban CR, Sancho-Martinez I, Izpisua Belmonte JC. 2012. Progressive degeneration of human neural stem cells caused by pathogenic LRRK2. Nature491:603-607.” 中公开过, 公众可从中国科学院生物物理研究所获得人类胚胎多能干细胞来源的神经干细胞(hNSCs) ;
- [0059] RNeasy Mini Kit 购自 QIAGEN ;
- [0060] AffymetrixGeneChipPrimeView Human Gene Expression Arrays 购自 Affymetrix, 货号为 901837。
- [0061] 实施例 1、人类胚胎多能干细胞的培养
- [0062] 将人类胚胎多能干细胞系 H9 利用如下方法进行培养 :
- [0063] 一、将 H9 细胞接种至预先培养了经过丝裂霉素灭活的小鼠胚胎成纤维细胞的培养板中, 使用人类多能干细胞培养基 (CDF12 培养基) 与小鼠胚胎成纤维细胞共同培养。
- [0064] 二、用 DMEM/F12 培养基将细胞外基质稀释至体积分数为 1% 的浓度, 并用稀释后的细胞外基质包被培养板。
- [0065] 三、将步骤一培养的 H9 细胞接种至预先用体积分数为 1% 的细胞外基质包被的培养板中, 使用 mTeSR 培养基培养。
- [0066] 实施例 2、人类胚胎多能干细胞向人类心肌细胞分化
- [0067] 人类胚胎多能干细胞系 H9 向人类心肌细胞分化方案基于 Palecek 方案 (Lian X, Hsiao C, Wilson G, Zhu K, Hazeltine LB, Azarin SM, Raval KK, Zhang J, Kamp TJ, Palecek SP. 2012. Robust cardiomyocyte differentiation from human pluripotent stem cells via temporal modulation of canonical Wnt signaling. Proceedings of the National Academy of Sciences of the United States of America109:E1848-1857.) 进行改进, 改进后的方案进一步提高了人类心肌细胞的纯度, 该方案的具体步骤如下 :

[0068] 一、用 DMEM/F12 培养基将细胞外基质稀释至体积分数为 1% 的浓度,用稀释后的细胞外基质包被培养板。

[0069] 二、将人类胚胎多能干细胞系 H9 消化为单细胞悬液,以 10^5 个细胞 / cm^2 的密度接种到经体积分数为 1% 的细胞外基质包被的培养板中,添加含有 $10 \mu\text{M}$ ROCK 抑制剂 Y-27632 的 mTeSR 培养基。37°C、5% 二氧化碳条件下培养 24 小时,然后改用 mTeSR 培养基继续培养 48 小时。

[0070] 三、起始心肌细胞的分化。

[0071] (一)在第 1 天换用含有 $12 \mu\text{M}$ CHIR99021 的 RPMI/B27-insulin 培养基培养 24 小时,24 小时后更换为 RPMI/B27-insulin 培养基继续培养。

[0072] (二)第 3 天换用含有 $5 \mu\text{M}$ Wnt 抑制剂 IWP4 的 RPMI/B27-insulin 培养基培养 48h,期间不换液。

[0073] (三)从第 5 天起每隔 2-3 天更换 RPMI/B27 培养基继续培养。

[0074] (四)第 15 天,将肉眼可见的可自主收缩的人类心肌细胞薄片或团块通过机器分选或人工挑取进行收集。

[0075] (五)将收集的细胞重新接种到经体积分数为 1% 的细胞外基质包被的培养板中,用 RPMI/B27 培养基进行培养。

[0076] 实施例 3、人类心肌细胞的鉴定

[0077] 人类胚胎多能干细胞来源的人类心肌细胞不仅在形态上表现出心肌细胞特有的收缩功能,并且可以进一步通过免疫荧光及流式细胞术对心肌细胞特异性的生物学标记进行鉴定。

[0078] 一、免疫荧光鉴定

[0079] (一)将实施例 2 得到的人类心肌细胞分为四组,各组分别经 4% 多聚甲醛固定,0.3% Triton X-100/PBS 通透。然后四组分别与四种一抗(鼠抗 cTnT, 鼠抗 MF20, 兔抗 MLC2v, 鼠抗 α -Actinin)4°C 孵育过夜。将一抗清洗后于室温进行二抗孵育 1 小时,其中第一组使用的二抗为 Alex Fluor488goat anti-mouse IgG,第二组使用的二抗为 Alex Fluor488goat anti-mouse IgG,第三组同时使用二抗 Alex Fluor488goat anti-mouse IgG 和二抗 Alex Fluor568goat anti-rabbit IgG 进行孵育。

[0080] (二)细胞核用 DAPI 进行负染。激光共聚焦显微镜采集图像进行分析。

[0081] cTnT 为心肌肌钙蛋白 T (cardiac troponin T);

[0082] MF20 为肌球蛋白重链;

[0083] MLC2v 为肌球蛋白轻链 2v (myosin light chain2V);

[0084] α -Actinin 为肌动蛋白。

[0085] 结果如图 1A 所示。

[0086] 图 1A 中,1 (a-c)表示 cTnT 在不同放大倍率下的染色结果,2 (a-c)表示 MF20 在不同放大倍率下的染色结果,3a 表示 α -Actinin 的染色结果,3b 表示 MLC2v 的染色结果,3c 为 3a 和 3b 重叠的结果。

[0087] 结果表明,人类胚胎多能干细胞来源的人类心肌细胞能够表达心肌特异的标记物心肌肌钙蛋白 T (cardiac troponin T)和肌球蛋白,通过对肌球蛋白和肌动蛋白的免疫荧光染色可观察到心肌纤维结构。

[0088] 二、流式细胞术鉴定

[0089] 结果如图 1B 所示。

[0090] 图 1B 表明,分化得到的人类心肌细胞群中大约 96% 是 cTnT 阳性细胞,大约 91% 是 MF20 阳性的细胞。

[0091] 实施例 4、转录组分析

[0092] 一、设置三组样品:人类胚胎多能干细胞(hESCs)H9(以下简称为 H9-ESC),人类胚胎多能干细胞系 H9 来源的神经干细胞(hNSCs)(以下简称为 H9-NSC)以及实施例 3 鉴定出的人类胚胎多能干细胞系 H9 来源的人类心肌细胞(hCMs)(以下简称为 H9-CM),每组细胞样品共设三个平行。

[0093] 二、使用 Trizol 法提取各组细胞各样品的总 RNA,并使用 RNeasy Mini Kit 进一步纯化。

[0094] 三、将三组细胞的样品用“AffymetrixGeneChipPrimeView Human Gene Expression Arrays”基因芯片处理。

[0095] 四、表达差异比较

[0096] 分析比较各组细胞中不同基因在 H9-ESC、H9-NSC 及 H9-CM 中的表达情况。由 AffymetrixGeneChip Scanner30007G 采集基因芯片的表达信号。利用 R/Bioconductor 处理表达信号,根据 RMA 算法归一化后转换为对应的基因表达值,并绘制 H9-ESC、H9-NSC 和 H9-CM 三种细胞各 3 份样品的表达热图(heatmap),结果如图 2A 所示。

[0097] 图 2A 中,

[0098] hESC # 1、hESC # 2 和 hESC # 3 分别为 H9-ESC 的三个平行样品;

[0099] hNSC # 1、hNSC # 2 和 hNSC # 3 分别为 H9-NSC 的三个平行样品;

[0100] hCM # 1、hCM # 2 和 hCM # 3 分别为 H9-CM 的三个平行样品;

[0101] 1/4x 表示基因的表达水平下调到 4 倍;

[0102] 1/2x 表示基因的表达水平下调到 2 倍;

[0103] 1x 表示基因的表达水平无显著变化;

[0104] 2x 表示基因的表达水平上调到 2 倍;

[0105] 4x 表示基因的表达水平上调到 4 倍。

[0106] 图中是以人胚胎干细胞 H9-ESC 作为基因上调和下调的标准。

[0107] 经过分析,与基因在 H9-ESC 和 H9-NSC 中的表达量相比,共有 695 个基因在 H9-CM 中的表达量同时是 H9-ESC 和 H9-NSC 中的表达量的 2 倍以上,共有 401 个基因在 H9-CM 中的表达量均不足 H9-ESC 和 H9-NSC 中的表达量的一半。

[0108] 五、实时荧光定量 PCR 验证关键基因的表达结果,表 1 为各关键基因的引物。其中部分结果如图 2B 所示,验证结果与芯片分析得到的结果一致。

[0109] 表 1 验证引物

[0110]

基因	正向引物序列 (5' -3')	反向引物序列 (5' -3')
ACTA2	GTGTTGCCCTGAAGAGCAT	GCTGGGACATTGAAAGTCTCA

CASQ2	GGCAGAAGAGGGGCTTAATTT	GAAGACACCGGCTCATGGTAG
CREM	ACACCACCTAGTATTGCTACCA	GGATTGTTCCACCTTGGGCTAT
CSRP3	CCTGTGAAAAGACCGTCTACC	GTCGTGCTGTCAAGAGCCT
DTNA	TACCCACGGAAGTTTTGGAGG	GGATCTGACATAAGCGTGTCC
EYA4	CTTCTTGCAGTCAAACAGAGC	GTGGATAGGGCTTGAAGGAT
FBX032	GCCTTTGTGCCTACAACCTGAA	CTGCCCTTTGTCTGACAGAAT
GATA4	CGACACCCCAATCTCGATATG	GTTGCACAGATAGTGACCCGT
GATA5	CTTCGTGTCCGACTTCTTGGA	CCGAGGCATTCCTTGTGGA
GATA6	CTCAGTTCCTACGCTTCGCAT	GTCGAGGTCAGTGAACAGCA
HAND2	CGCCGACACCAAACCTCTCC	TCGCCATTCTGGTCGTCT
HECW2	AAATCCCAGATGCGGTACAC	CGGCTCTCAGAAGTCACCA
ISL1	GCGGAGTGTAATCAGTATTTGGA	GCATTTGATCCCGTACAACCT
KCNA5	GCGGTCCACATCAACATCTC	GGTAGAAGCGTATCTCGTCCG
LBH	GCCCCGACTATCTGAGATCG	GCGGTCAAAATCTGACGGGT
LDB3	CTATCTCCCGATCACACCAG	GTGAGGCTCAAGTTGTAGCTG
LEFTY2	TGGACCTCAGGGACTATGGAG	CCGAGGCGATACACTGTCTG
MARCH11	CCGGGAGAGTCTCCACA	GCCTCATCCGATTGAACAGGT
MEF2A	ACTACAGACCTCACAGTGCCA	GCCTAAGCTATTTGCACCAGT
MEF2C	CCAACTTCGAGATGCCAGTCT	GTCGATGTGTTACACCAGGAG

[0111]

MIR21	CTGCCTGACTGTCTGCTTGTTF	AGATTCAACAGTCAACATCAGTCTGA
MITF	CAGTCCGAATCGGGGATCG	TGCTCTTCAGCGGTTGACTTT
MYBPC3	CCATGACGCTGAGGTCAAATG	GCTTGGCACCAGTGGACTC
MYOCD	ACGGATGCTTTTGCCTTTGAA	AACCTGTCCAAGGGGTATCTG

MYOZ2	CCAGGCTATTTAAGATGCGTCA	CCTTCCAAGTTACTTCCATCCAC
NEXN	ACTGTGAAGGGTAGATTTGCTG	TTCTGCGTTTTTCGTTCCCTCCT
NKX2	CCAAGGACCCTAGAGCCGAA	ATAGGCGGGGTAGGCGTTAT
NPPA	CAACGCAGACCTGATGGATTT	AGCCCCGCTTCTTCATTC
NPPB	TGGAAACGTCCGGGTACAG	CTGATCCGGTCCATCTTCCT
PCGF5	AGCCAACAACAGTGACGGAAT	TGAACTTGGTTGCCACACCTT
PLN	ACCTCACTCGCTCAGCTATAA	CATCACGATGATACAGATCAGCA
PPAPDC3	AGCTGAACCCCTCCTTCAAG	CCGATGACAAAGCCGGAGA
PPARGC1A	TCTGAGTCTGTATGGAGTGACAT	CCAAGTCGTTACATCTAGTTCA
PRICKLE1	TTTGCTTGCTTACCAGAGGAAA	ACTGGCAATACCGTACCTCAT
RARB	TCCGAAAAGCTCACCAGGAAA	GGCCAGTTCACTGAATTTGTCC
SMAD6	GCTACCAACTCCCTCATCACT	CGTACACCGCATAGAGGGG
SMAD7	TTCTCCGCTGAAACAGGG	CCTCCCAGTATGCCACCAC
SNTA1	TGCTCCTCTACTTGTCTCTCC	TCTGCATCGTAGGGCACTGA
STAT4	TGTTGGCCCAATGGATTGAAA	GGAAACACGACCTAACTGTTTAT
TBX2	GCTGACGATTGCCGCTATAAG	GGCTGTCTGGGTGGATGTA
TCEA3	AAGAGCACGGACATGAAGTACC	CTCTGCCGTCATCTTGGCTA
TNNT2	GGAGGAGTCCAAACCAAAGCC	TCAAAGTCCACTCTCTCTCCATC
TRIM63	CTTCCAGGCTGCAAATCCCTA	ACACTCCGTGACGATCCATGA
TTN	CCCCATCGCCATAAGACAC	CCACGTAGCCCTCTTGCTTC
XPO4	ATTTACAGCGTTACCTTGCACT	CAGCACATTTTGCAGGATTC
ZFPM2	GGCCTGAAAATCTGAGCTGC	CAGTCGTCTGTCTCAACTCCA

[0112]

[0113] 实施例 5、DNA 甲基化分析

[0114] 一、设置三组样品：H9-ESC, H9-NSC 以及 H9-CM, 每组细胞样品共设三个平行。

[0115] 二、提取各组各样品细胞的 DNA, 利用基于重亚硫酸盐挂锁探针(bisulfite padlock probes)的 DNA 甲基化测序技术以及甲基化水平计算方法(Diep D, Plongthongkum N, Gore A, Fung HL, Shoemaker R, Zhang K. 2012. Library-free methylation sequencing with bisulfite padlock probes. Nature methods 9:270-272.) 分析 H9-ESC, H9-NSC 以及 H9-CM 的全基因组 DNA 甲基化水平, 部分基因检测结果如图 3 所示。

[0116] 图 3 中,

[0117] hESC 为 H9-ESC ;

[0118] hNSC 为 H9-NSC ;

[0119] hCM 为 H9-CM ;

[0120] -10% 表示基因的启动子区 CpG 岛甲基化水平下调 10% ;

[0121] -5% 表示基因的启动子区 CpG 岛甲基化水平下调 5% ;

[0122] 0 表示基因的启动子区 CpG 岛甲基化无显著变化 ;

[0123] 5% 表示基因的启动子区 CpG 岛甲基化水平上调 5% ;

[0124] 10% 表示基因的启动子区 CpG 岛甲基化水平上调 10%。

[0125] 图中是以人胚胎干细胞 H9-ESC 作为甲基化上调和下调的标准。

[0126] 图 3 表明, 与 H9-ESC 和 H9-NSC 相比, H9-CM 具有最高的全基因组 DNA 甲基化水平。

[0127] 经过分析, 与基因启动子区 CpG 岛在 H9-ESC 和 H9-NSC 中的甲基化水平相比, 共有 985 个基因启动子区 CpG 岛在 H9-CM 中的甲基化水平同时比该基因启动子区 CpG 岛在 H9-ESC 和 H9-NSC 中的甲基化水平有 5% 以上的提高, 共有 195 个基因启动子区 CpG 岛在 H9-CM 中的甲基化水平同时比该基因启动子区 CpG 岛在 H9-ESC 和 H9-NSC 中的甲基化水平有 5% 以上的降低。

[0128] 将 DNA 甲基化分析与实施例 4 的转录组分析相结合得到基因的甲基化水平与表达水平的结果。部分基因的结果如图 4 所示, 图 4 表明, 心脏结构与心脏转录因子相关基因的 DNA 去甲基化与基因表达水平成正相关性。

[0129] 将 DNA 甲基化分析与实施例 4 的转录组分析相结合得出 : 与 H9-ESC 和 H9-NSC 相比, 在 H9-CM 中, 鉴别出基因的表达水平提高到 2 倍以上且启动子区 CpG 岛甲基化水平降低 5% 以上的 29 种基因。这些基因可以作为人类心肌细胞的分子标记物, 如表 2 所示。

[0130] 表 2 29 种基因的 DNA 甲基化水平和表达水平

[0131]

基因名	GeneBank ID	DNA 甲基化水平		基因表达水平	
		人心肌细胞 VS 人胚胎多能干细 胞	人心肌细胞 VS 人神经干细胞	人心肌细胞 VS 人胚胎多能干细 胞	人心肌细胞 VS 人神经干细胞
POPC2	64091	0.57	0.56	8.402616293	7.562352603
ALPK2	115701	0.32	0.42	8.069528368	3.647151766
TRIM55	84675	0.38	0.51	7.373137025	7.267597143
ASPH	444	0.64	0.57	4.281290866	3.122582294
FILIP1	27145	0.59	0.72	6.920428571	7.794390913
HSPB7	27129	0.58	0.72	6.447867991	6.295762763
NPEA	4878	0.78	0.80	7.983519001	8.03271611
KBTBD10	10324	0.80	0.79	2.910404942	2.670527553
MYL4	4635	0.70	0.79	9.22112583	9.002175408
CRYAB	1410	0.72	0.80	5.935066651	6.305402948
LRRN4	164312	0.78	0.75	2.491450066	2.533850351

[0132]

MYOZ2	51778	0.86	0.86	9.48535322	9.510284971
OBSCN	84033	0.81	0.85	2.598541342	2.401037422
SMPX	23676	0.79	0.86	9.91336154	9.678290914
NRK	203447	0.79	0.75	2.403506377	4.248352806
ABLIM1	3983	0.88	0.88	2.257898689	2.653017499
SYNPO2L	79933	0.91	0.88	9.25673506	9.420535272
ANKRD1	27063	0.40	0.56	9.714009013	8.638574866
TNNI1	7135	0.78	0.89	6.705106323	6.794815577
MYH6	4624	0.79	0.91	9.518535857	9.794227963
ABRA	137735	0.88	0.92	5.464788018	5.340692628
TPM1	7168	0.91	0.80	3.637236226	6.143185497
C15orf52	388115	0.92	0.88	5.192978911	5.918630472
MYOM1	8736	0.91	0.93	5.928375111	5.942358498
MASP1	5648	0.80	0.88	7.240842303	6.569487061
MYBPC3	4607	0.87	0.93	6.01967711	6.115402912
TNS1	7145	0.77	0.92	2.641474521	2.546919208
FLNC	2318	0.78	0.56	5.199968343	4.370411535
F2RL2	2151	0.27	0.77	3.218976947	3.106845011

[0133] 实施例 6、生物信息分析

[0134] 对表达差异最显著的 50 个基因进行基因 - 疾病网络分析,发现有 27 种基因至少与已知的一种心血管疾病有关。其中包括上述鉴别出的 29 种人类心肌细胞分子标记物中的 6 种,它们分别是 NPPA, MYL4, MYOZ2, ANKRD1, TNNI1, MYH6。

[0135] 实施例 7、表达差异显著基因的生物信息分析

[0136] 对实施例 4 中表达水平存在显著差异的基因进行进一步的生物信息分析,分析包括基因聚类分析(Gene Ontology, GO),功能相关性网络分析以及基因 - 疾病网络分析等。

[0137] 一、基因聚类分析

[0138] 使用 Cytoscape 软件的 BiNGO 插件进行聚类分析,分析结果显示显著上调的基因多数被归入了心脏功能相关的 GO 类别中,该类别包括心肌收缩,心脏发育,肌节结构等。相反,显著下调的基因多数被归入有丝分裂期,细胞核分裂和有丝分裂等类别,提示了心肌细胞中细胞的有丝分裂受到强烈的抑制。此外,通过分析基因调控网络上下游中受调控基因的显著性可鉴定出 hESCs 向 hCMs 分化最小的决定性基因组合,通过计算模型确定了一个双核心的稳定调控网络,如图 5 所示。这两个核心分别与多潜能性和心肌细胞特性相关。人为改变双核心中的基因表达水平能够引发调控的链式反应,使得 hESCs 向 hCMs 分化。对双核心中部分基因有目的的调控同样可能实现其他细胞类型向 hCMs 的转分化。

[0139] 二、功能相关性网络分析

[0140] 为揭示实施例 4 发现的 695 个 hCMs 特异基因的功能相关性,利用 Cytoscape 软件的 GeneMANIA 插件产生了基于相关性的网络,包括共表达,共定位,相互作用等。此外,将上述 GO 聚类分析中被归入心肌收缩,心脏发育,心肌转录调控等类别的基因群分别产生次级网络,如图 6 所示。

[0141] 图 6A 为心肌收缩相关基因网络。

[0142] 图 6B 为心肌转录调控网络。

[0143] 图 6C 为心脏发育相关基因网络。

[0144] 图 6 所示的网络为心肌基因新功能、新相关性的预测等提供了支持。

[0145] 三、基因 - 疾病网络分析

[0146] 对 hCMs 特异基因中表达差异最显著的 50 个基因利用 Cytoscape 软件的 DisGeNET 插件进行基因 - 疾病网络分析。其中,27 个基因至少与已知的一种心血管疾病有关。此外,有 5 种基因(NPPB, TNNT2, NPPA, RYR2 和 PLN)与超过 10 种不同种类的心血管疾病有联系。超过 140 种疾病相关的基因突变被证明与这些表达差异最显著的 hCMs 特异基因有关,部分结果如图 7 所示。基因 - 疾病网络的分析结果进一步强调了这些 hCMs 特异基因对维持正常心脏发育的重要作用,为揭示心脏病机制乃至治疗心脏病提供了参考和依据。

[0147] 实施例 5 得到的 29 个基因在聚类分析中被归入肌肉收缩,肌肉结构组成,肌节结构等类别,为心脏结构相关的基因。可见,与心脏结构相关的基因主要受 DNA 甲基化调控。而其余未在表达水平上产生显著改变的去甲基化基因可能在其它水平上接受调控。此外,多数与心脏结构相关的基因在 hESCs 和 hNSCs 中表现出 CpG 岛的高甲基化,提示了在非心肌细胞的细胞谱系中,这些基因的启动子区 DNA 甲基化是最主要的抑制信号。

[0148] 此外,存在一些心肌细胞特异的转录因子(如 NKX2.5, GATA6, GATA4, MYOCD, HAND2, TBX5, TBX18 等)在 hESCs, hNSCs 及 hCMs 中均表现类似的低甲基化水平,但在 hCMs 中表达水平远高于 hESCs 和 hNSCs。先前的研究也发现了这些转录因子在 hESCs 向 hCMs 的分化过程中抑制性的组蛋白修饰 H3K27me3 水平发生了降低。提示了对这些基因,组蛋白修饰才是最主要的表观遗传调控因素。

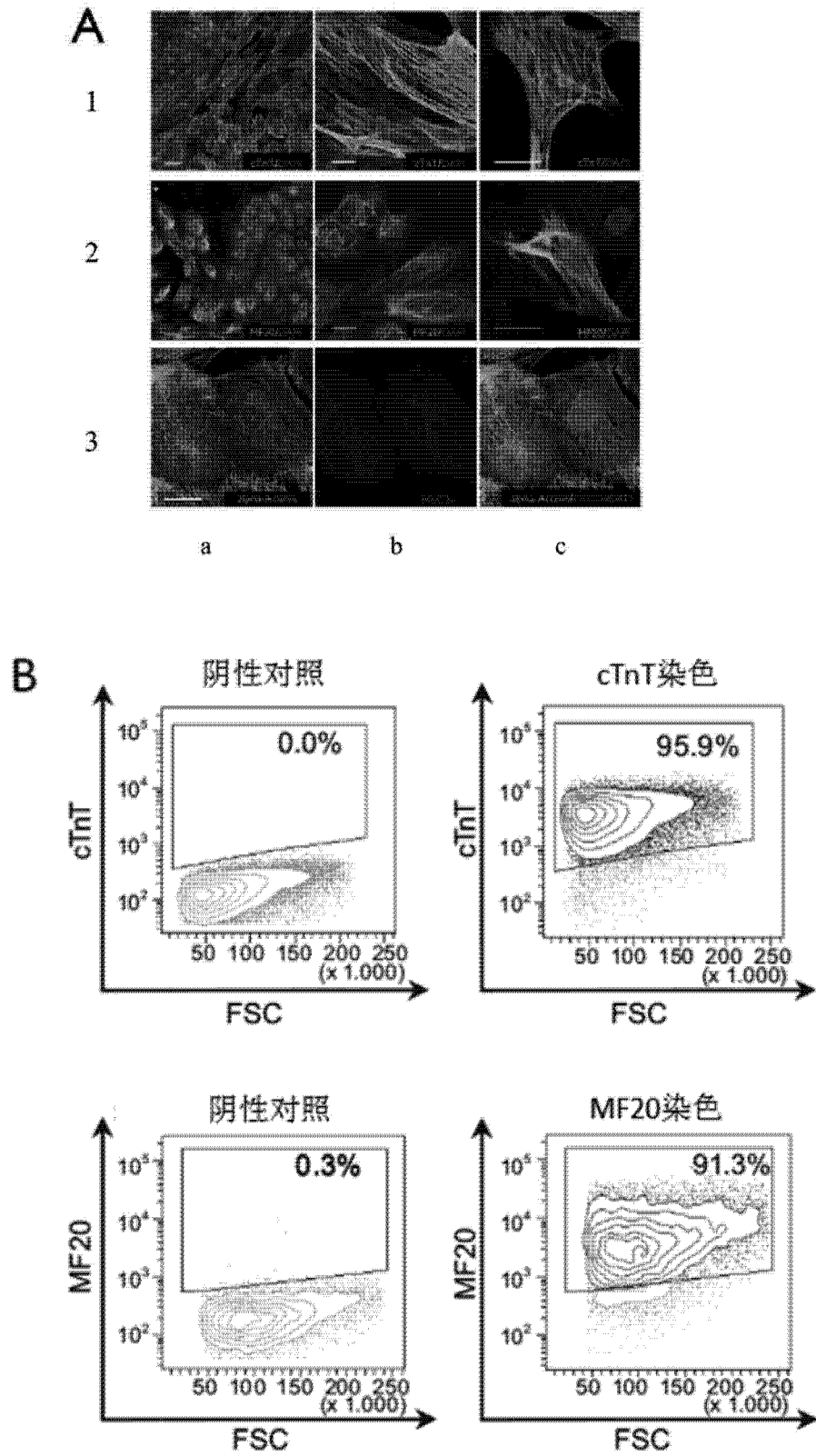


图 1

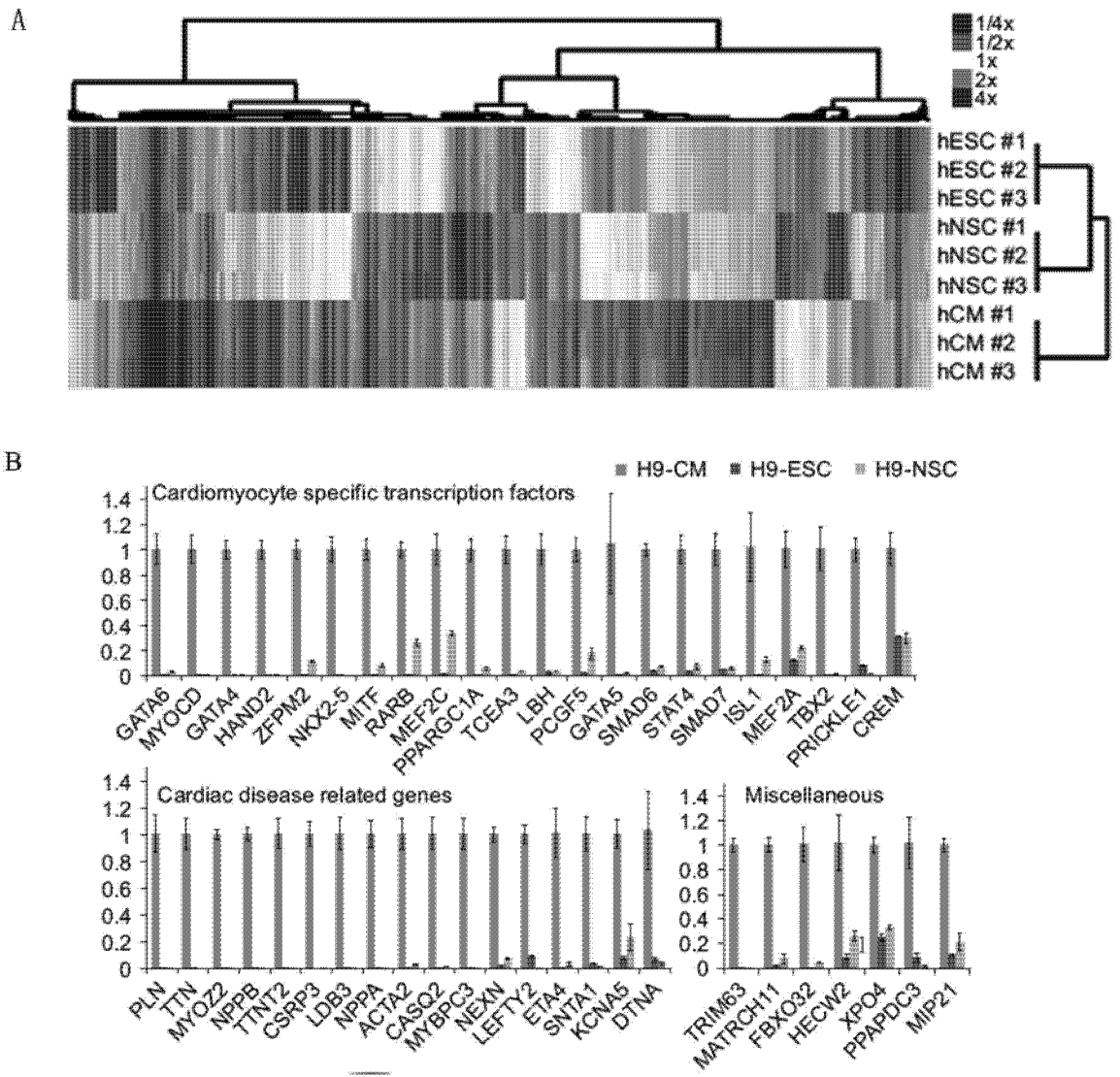


图 2

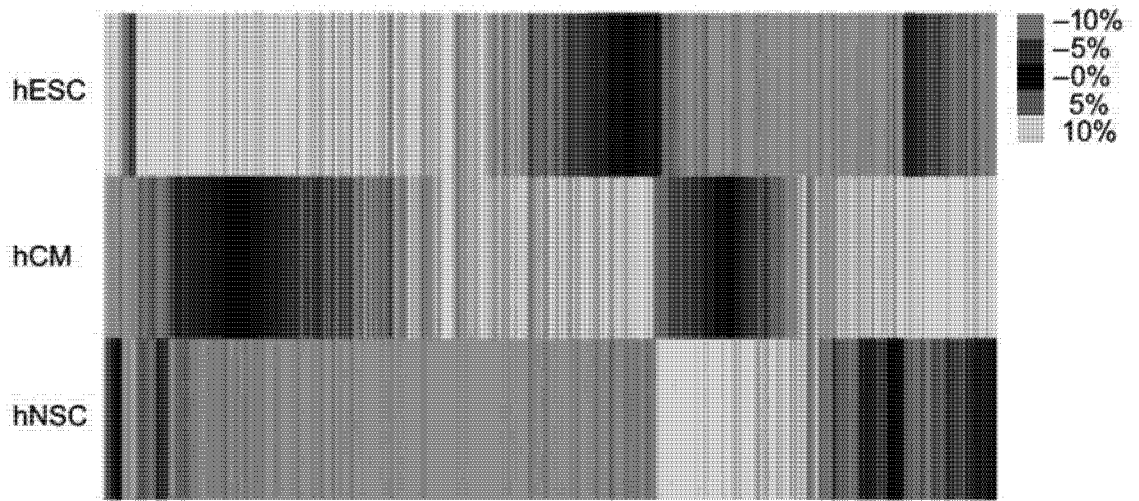


图 3

编码心肌结构蛋白的基因

	DNA methylation			Gene expression	
	ES-mCpG	NSC-mCpG	CM-mCpG	CM vs ES	CM vs NSC
TRIM55	0.89	0.67	0.34	7.17	7.27
ASPH	0.83	0.93	0.53	4.28	3.12
KBTBD10	0.90	0.91	0.72	2.91	2.67
MYL4	0.92	0.81	0.64	3.73	3.00
CRYAB	0.94	0.85	0.68	5.94	5.41
MYOZ2	0.97	0.96	0.83	3.25	3.12
OBSCN	0.90	0.86	0.73	2.60	2.40
ABLIM1	0.83	0.83	0.73	2.26	2.65
SYNPO2L	0.90	0.93	0.82	3.37	3.62
ANKRD1	0.25	0.18	0.10	3.73	3.64
TNNI1	0.79	0.70	0.62	6.73	6.79
MYH6	0.87	0.76	0.69	3.92	3.79
ABRA	0.97	0.92	0.85	5.46	5.34
TPM1	0.77	0.88	0.70	3.64	6.18
MYOM1	0.92	0.90	0.84	5.93	5.96
MYBPC3	0.89	0.83	0.77	6.02	6.32
TNS1	0.74	0.62	0.57	2.64	2.55
FLNC	0.23	0.32	0.18	5.20	4.37

心肌转录因子

Gene Name	DNA methylation			Gene expression	
	ES-mCpG	NSC-mCpG	CM-mCpG	CM vs ES	CM vs NSC
GATA6	0.01	0.01	0.01	6.01	6.05
GATA4	0.01	0.01	0.01	7.09	7.09
MYOCD	0.05	0.03	0.04	6.70	6.91
HAND2	0.05	0.07	0.14	6.43	6.41
NKX2-5	0.03	0.04	0.03	5.28	5.32
TBX5	0.03	0.02	0.18	5.11	5.15
TBX18	0.01	0.01	0.03	4.45	3.54



图 4

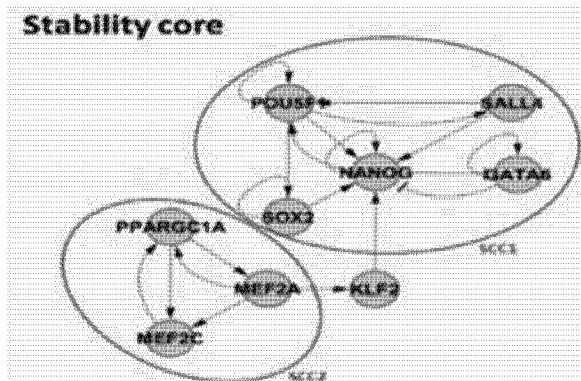
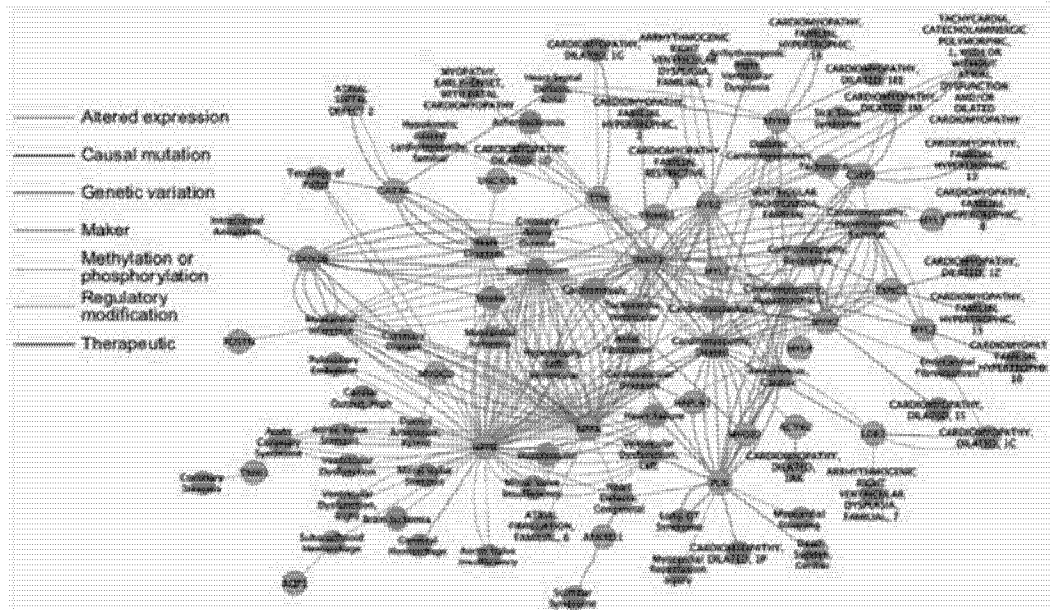


图 5

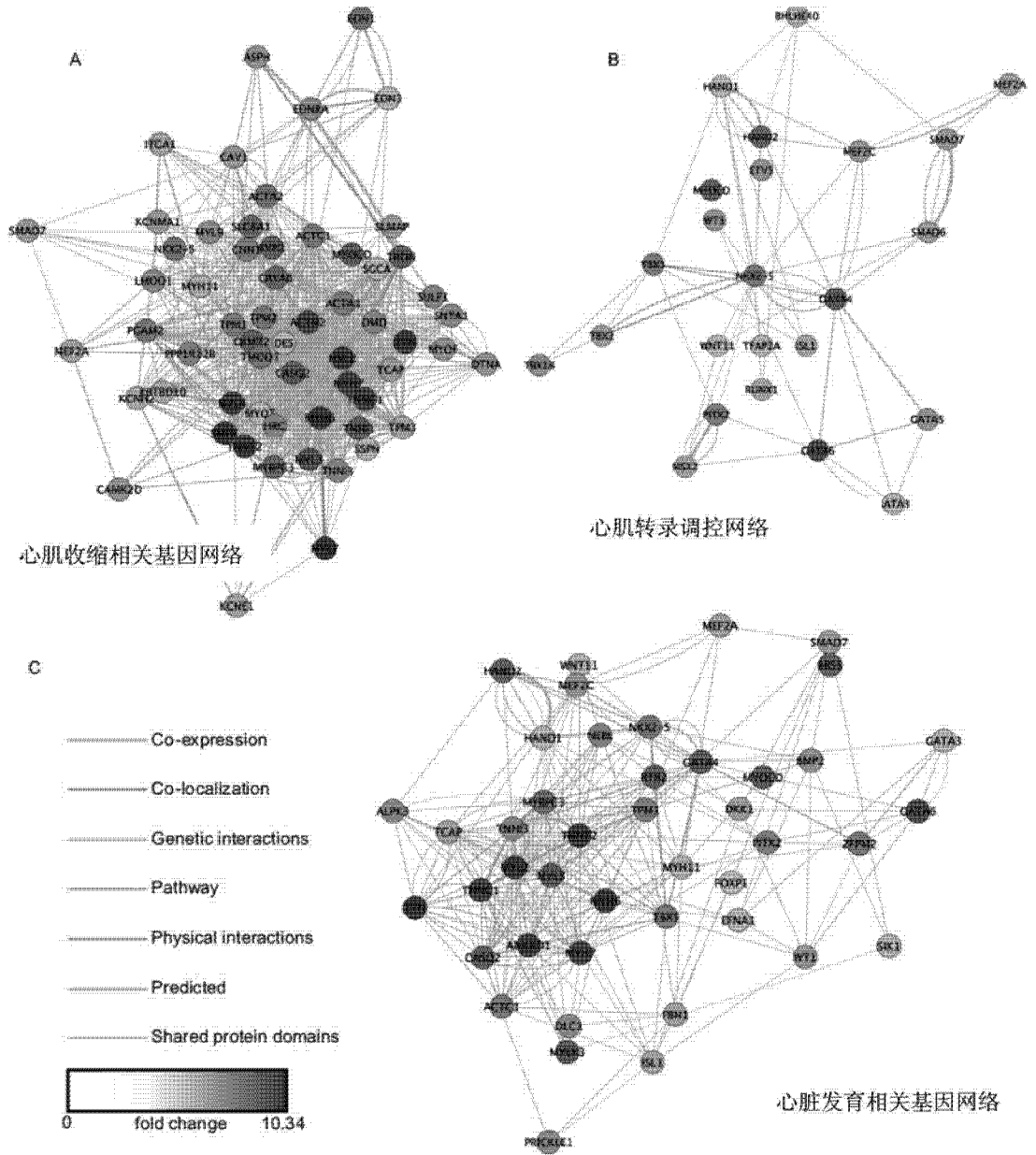


图 6

Gene	Diseases	Mutations in DNA or protein sequence
NPPB	HCM	D175N
	EH	Variable number of tandem repeat of 5'-flanking region
	DCM	I79N, R141W, K210del
TNNT2	HCM	I79N, I90M, R92C, R92W, F110I, K124N, R130C, R278C, G287A, Δexon 14, a five-base (CTTCT) insertion/deletion in intron 3
	RCM	I79N, Q96del
NPPA	Ischemic Stroke	G86A
	AF	c.436-457delAA
	EH LVH	V7M, -664G -2843A>G
TNNC1	HCM	A8V, L29Q, A31S, C84Y, E134D, D145E
	DCM	Y5H, M103I, D145E, I148V
ACTN2	DCM	Q9R
MYL3	HCM	D778E
GATA4	CHD	A6V, S52F, P163S, E216D, E359K, S377G, P407Q, S429T, A442V, IVS4+55C>A
		c.1074delC
RYR2	CPVT	R414C, E1724K, A2254V, R2267M, A2394G, V2475F, F4020L, E4076K, N4104I, N4104K, H4108N, H4108Q, R4497C, S4565R, G4662S, H4762P, V4771I, P4902S, genomic deletion involved RYR2 exon-3
MYOCD	CHD	K259R
TRIM83	HCM	A48V, I130M, Q247del
Gene	Diseases	Mutations in DNA or protein sequence
PLN	HCM	-42C>G, -77A>G
	DCM	R9C, R14del
MYH6	HCM or DCM	R795Q, P830L, A1004S, Q1065H, E1457K
	ASD	I820N
MYOZ2	HCM	S48P, I246M
MYL2	HCM	A13T, F18L, E22K, N47K, R58Q, K103E, IVS6-1G>C
CSRFP3	HCM	L44P, S54R/E55G, C58G, I72T>G
	DCM	W4R
TTN	HCM	R740L, S3799Y
	DCM	V54M, A743V, W930R, R25618Q, Q4053ter, S4465D, 2-bp insertion in exon 326
LDB3	DCM	D117N, K136M, S196L, T213I, I352M, D626N
	LVNC	D117N, K136M, S196L
MYH7	DCM	P211L, L390V, R403W, R403Q, Y501C, H576R, R653C, Q734P, I736T, G741W, A797T, R870H, A901G, D928N, E930K, R1053Q, E1358K, A1379T, A1454T, R1500W, S1776G
	DCM	I201T, Y412N, A550V, T1019N, R1193S, E1426K, R1634S
	LVNC	M531R
ANKTD1	TAPVR	10;21 translocation
DCM	Dilated Cardiomyopathy	
HCM	Hypertrophic Cardiomyopathy	
RCM	Restrictive Cardiomyopathy	
ASD	Atrial Septal Defect	
LVNC	Left Ventricular Non-Compaction	
TAPVR	Total Anomalous Pulmonary Venous Return	
EH	Essential Hypertension	
AF	Atrial Fibrillation	
LVH	Left Ventricular Hypertrophy	
CPVT	Catecholaminergic Polymorphic Ventricular Tachycardia	
CHD	Congenital Heart Disease	
ARVC	Arrhythmogenic Right Ventricular Cardiomyopathy	

图 7