

· 分子标志物 ·

基于全转录组的心力衰竭预测标志物

马 伟, 李丹丹, 张常建, 陈伟浩, 熊 鸣, 乔媛媛, 李 军

[摘要] 目的 挖掘基于转录组的预测心力衰竭(heart failure, HF)的标志物。方法 本课题组开发了一个基于全转录组的HF预测标志物,整理转录本对应的蛋白质的氧原子含量,然后计算转录本蛋白质氧原子含量与转录本表达水平的相关性,通过随机森林模型训练预测模型。结果 对于训练集,预测模型5折交叉验证的准确率、阳性预测值和阴性预测值分别为83.6%、85.8%和81.5%。训练集的准确率、阳性预测值和阴性预测值分别为77.8%、80.0%和75.0%。在另外一个单独寻找的测试集,准确率、阳性预测值和阴性预测值分别为86.8%、86.7%和87.5%。结论 本课题组开发的基于全转录组的HF预测标志物具有非常稳定的预测性能,而且准确率高,可以用于HF预测。

[关键词] 心力衰竭;标志物;最优密码子;蛋白质氧含量;随机森林

[中图分类号] R541.6

[文献标志码] A

[文章编号] 2095-3097(2021)06-0349-05

doi: 10.3969/j.issn.2095-3097.2021.06.001

A global transcriptome-based biomarker for the robust classification of heart failure

MA Wei¹, LI Dandan², ZHANG Changjian¹, CHEN Weihao³, XIONG Ming¹, QIAO Yuanyuan¹, LI Jun¹

(1. the Sixth Medical Center of PLA General Hospital, Beijing 100048, China; 2. SUN YAT-SEN Memorial Hospital SUN YAT-SEN University, Guangzhou Guangdong 510120, China; 3. Department of Medical Service, the Third Medical Center of PLA General Hospital, Beijing 100143, China)

[Abstract] **Objective** To explore a Heart failure (HF) prediction biomarker based on transcriptome. **Method** We presented a global-transcriptome based HF biomarker, which integrated the expression level of all the transcripts and the protein oxygen atom content (OAC) of the corresponding transcripts. We trained a random forest classifier using the correlation coefficients between gene expression level and OAC. **Result** We presented a global-transcriptome based HF biomarker, which integrated the expression level of all the transcripts and the protein oxygen atom content (OAC) of the corresponding transcripts. We trained a random forest classifier using the correlation coefficients between gene expression level and OAC. The accuracy, positive predictive value (PPV) and negative predictive value (NPV) are 83.6%, 85.8% and 81.5% from the five-fold cross-validation, and 77.8%, 80.0%, and 75.0% on the testing dataset, respectively. Moreover, on a totally independent testing dataset, the classifier achieved an accuracy of the 86.8%, with a PPV of 86.7% and a NPV of 87.5%. **Conclusion** The global transcriptome-based biomarker we developed show robust prediction performance across different datasets, which further suggests that this biomarker could be useful for the classification of HF.

[Key words] Heart failure; Biomarker; Optimal codon; Protein oxygen content; Random Forest

心力衰竭(heart failure, HF)是一个复杂的临床症状,其特点是心脏供血不足^[1]。在发展中国家, HF是患者住院的常见原因,其发病率和病死率也高^[2-5]。HF的发病率逐年增加,已经成为重要公共健康问题^[6-8]。在发展中国家, HF非常普遍,其患病率高达0.4%~2%^[9-10]。虽然现在的诊断和治疗水平已经改善了HF患者的预后,但是其5年生存率仍不足50%^[11-12]。HF越早发现,患者预后越好,因此寻找预测HF的标志物对于降低患者病死率以及提

高患者生活质量非常有帮助。目前也发现了一些HF的风险因子。例如, Onoue等^[13]发现肌肉减少症可以预测HF患者的不良事件,而肌肉减少症可以通过结合年龄、握力和小腿围来进行评估。Soyama等^[14]通过超声心动图监测心脏发现心脏舒张应变(diastolic wall strain, DWS)能够评估HF患者心脏的功能。他们发现低DWS水平患者,心功能比较好,而左心室功能受损的患者DWS值升高, DWS高的患者预后往往不好。近些年来,随着高

[基金项目] 国家自然科学基金(31500756);解放军总医院第六医学中心创新培育基金(CXPY202007)

[作者单位] 100048 北京,解放军总医院第六医学中心(马 伟,张常建,熊 鸣,乔媛媛,李 军);510120 广东 广州,中山大学孙逸仙纪念医院(李丹丹);100143 北京,解放军总医院第三医学中心卫勤部(陈伟浩)

[通讯作者] 乔媛媛, E-mail: qiaoyuan75@126.com; 李 军, E-mail: chinalijun@126.com

通量分子实验技术的发展,多种多样的基于血液或组织的分子标志物也出现了。Huang等^[15]发现外周血白细胞的线粒体DNA(mitochondrial DNA, mtDNA)的拷贝数变异(copy number variation, CNV)是HF的独立危险因素,能够预测HF患者病死率。Tumorigenicity-2 (sST2)能够为HF患者的后期处置提供帮助^[16]。Circulating cell-free DNA (cfDNA)也可以作为HF标志物^[17]。在失代偿期的HF患者,血清高水平内源性促红细胞生成素的长期的临床预后更差^[18]。HF标志物的概况参见这三篇综述^[19-21]。芯片技术的发展和二代测序及三代测序的发展,累积了很多全转录组数据,这也为开发新的多分子标志物带来了机遇和挑战^[22-23]。通过转录组分析,Maciejak发现*RNAS1*、*FMN1*和*JDP2*与HF患者的临床指标有显著的相关性。ST段太高的心梗第一天,这三个基因的表达值在随后发生HF的患者中高于随后未发生HF的患者^[24]。基于转录组的方法能够发现基因表达特征作为疾病的标志物^[25]。Heidecker基于转录组发现了一个45个基因特征的标志物,可以评估HF患者的风险^[26]。然而上述标志物部分是单个分子,部分是几十个基因的特征,相比人类将近3万个基因的数量,标志物涉及到的基因数量还是特别少,这就会导致容易产生过拟合,使得标志物的扩展应用性不强。因此非常有必要寻找基于转录组的扩展性好的标志物。

使用全转录组的特征而不是仅仅局限于差异基因可以产生更稳定的HF标志物。传统的观点认为蛋白质功能的改变是由于突变导致了氨基酸序列发生了变化,然而有人发现原子含量,尤其是氧原子也与蛋白质功能有关。大气中氧含量的增加促进细胞交流与细胞器的形成,跨膜蛋白长度与蛋白质氧含量有关,而蛋白质氧含量又受到大气氧含量的约束^[27]。与无氧呼吸比较,有氧呼吸能够有效的获取能量,发生更多的代谢反应^[28-29]。对于HF患者,其特征就是组织缺氧,细胞氧化应激增加^[30]。因此,我们猜想在这样的缺氧情况下,转录组会产生系统性的缺氧反应,不同氧含量的蛋白质其表达量会有不同的变化,这种不同的变化可以用来寻找HF标志物。为了证实此猜想,本课题组从GEO(Gene Expression Omnibus, GEO)下载了两套HF数据(GSE57345和GSE21610),然后开发了一个基于全转录组的预测HF的标志物^[25,31],发现基因的表达水平与氧原子含量(oxygen atom content, OAC)的Spearman's相关系数能够很好的区分HF。

1 数据与方法

1.1 蛋白质、mRNA序列与基因表达数据 本研究从Ensembl数据库下载了蛋白质序列和mRNA序列^[32]。从GEO数据库下载了两套HF数据,筛选标准为:组织来源为心肌组织活检并且有正常心肌组

织,最终筛选获得GSE57345(Affymetrix Human Gene 1.1 ST Array)和GSE21610(Affymetrix Human Genome U133 Plus 2.0 Array)^[33]。在GSE57345中,心脏组织样本是MAGNet联盟从患者取得。本文的数据分析采用R语言进行统计分析(version 3.6.2, <https://www.r-project.org/>)。

1.2 计算蛋白质氧原子含量以及OAC与基因表达的Spearman相关性 蛋白质氧原子含量定义为:蛋白质序列中氧原子的数量/蛋白质总原子数。在计算相关性之前,先根据转录本最优密码子的比例(codon optimality ratio of protein, CORP)进行排序,然后分成10等份。CORP定义为:mRNA中最优密码子数量/总的密码子数量。最优密码子定义为:密码子的最佳稳定性相关系数得分(correlation coefficient, CSC),而正值得分的密码子就是最优密码子^[34]。首先计算所有的蛋白质编码转录本的表达量与其相应蛋白质OAC的Spearman相关系数,然后针对上述10份的每一份再计算一次。因此,对于每个样本,会得到11个相关系数,1个总的和10个10等份的。

1.3 构建分类器 使用randomForest R包构建随机森林模型,模型参数为:树的数量为800,其他参数为默认值^[35]。在构建分类器这里,将两套数据的相关系数范围使用公式(1)缩放至[-1,1],y'表示放缩后的值,y表示放缩前的值,min表示原数据值中最小值,max表示原数据中最大值。10个10等份的相关系数作为输入特征。然后通过特征重要性筛选特征。使用准确率、阳性预测值(positive predictive value, PPV)和阴性预测值(negative predictive value, NPV)作为分类器性能评估参数。使用5折交叉验证、测试集以及独立验证集测试模型准确性。准确率=(真阳性数量+真阴性数量)/总量;PPV=真阳性数量/(真阳性数量+假阳性数量);NPV=真阴性数量/(真阴性数量+假阴性数量)。

$$y' = -1 + (1 - (-1)) \times \frac{y - \min}{\max - \min} \quad (1)$$

2 结果

2.1 数据统计学特征 数据的人口统计学资料和基本临床信息见表1。GSE57345包含313个样本,包括177个HF样本和136个正常样本(non-HF controls, NF)。HF样本中有男性144例和女性33例,正常样本有男性73例和女性63例。两组的平均年龄分别为55.4和49.4。GSE21610有38个样本,30(28例男性和2例女性)和HF样本和8个正常样本(6例男性和2例女性),两组平均年龄分别为51.3和29.0。

2.2 HF患者OAC与基因表达的相关性 首先对所有编码蛋白的转录本计算了OAC与基因表达的相关性。发现在GSE57345数据集中, HF样本的相关系数远高于正常样本(图1A, $P=2.25 \times 10^{-11}$)。然后

再使用相关系数作为参数做受试者工作曲线(receiver operating characteristic curve, ROC),曲线下面积(the area under curve, AUC)为 0.772,如图 1B 所示。数据集 GSE21610 的结果类似($P=0.04$),AUC 为 0.729,见图 1C 和 1D。数据集 GSE21610 没有数据集 GSE57345 显著可能是由于其样本量的原因(38 VS 313)。虽然两个数据集相关系数差别比较大,但是 HF 样本都高于正常样本。

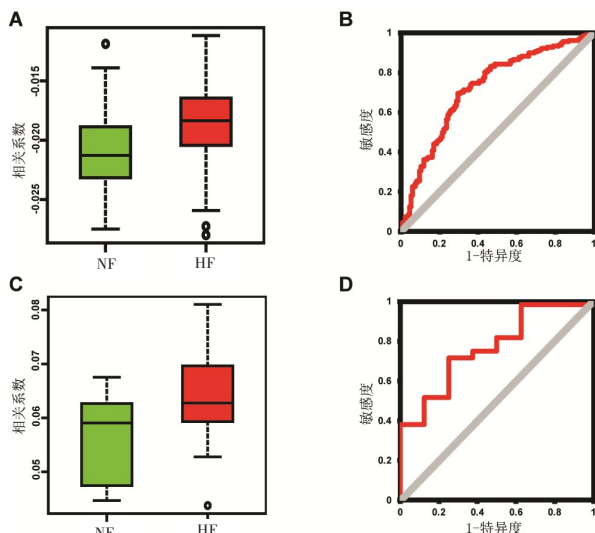


图 1 两个数据集中所有转录本的相关系数

A: 数据集 GSE57345 中相关系数在 HF 和 NF 中的分布; B: GSE57345 中相关系数区分 HF 与 NF 的 ROC 曲线; C: 数据集 GSE21610 中相关系数在 HF 和 NF 中的分布; D: GSE21610 中相关系数区分 HF 与 NF 的 ROC 曲线

2.3 基于选择特征的随机森林分类器 转录本最优密码子的比例与转录本的翻译速度有关系^[34]。因此,我们根据转录本的 COPR 将所有转录本等分为 10 组,每组分别计算表达水平与氧含量的 Spearman 相关系数。在数据集 GSE57345 中,第 3、5、6、7、8、9、10 部分有显著的区别能力,AUC 大于 0.6,表 2。尤其是第 5 部分和第 8 部分,AUC 大于 0.8。第 4 部分的 AUC 也将近 0.6(0.599)。在所有 10 部分中,HF 样本的相关系数都大于正常样本。

为了构建分类器,本研究将 GSE57345 随机分为训练集和测试集两部分,训练集包含 250 个样本(142 HF 和 108 NF),测试集包含 63 个样本(35 HF 和 28NF)。为了评估分类器的性能,把数据集 GSE21610 作为完全独立的一个测试集,仅仅在构建好分类器后才使用。为了保持数据的通用性,再将两套数据的相关系数范围都缩放至[-1,1]。然后使用训练集进行模型构建。10 部分的相关系数作为特征输入随机森林模型,根据特征重要性得分,选择了前三个重要性得分最高的特征(图 2A),分别为第 3、5、8 部分。最终使用这三个特征作为输入特征构建随机森林分类器。首先,使用 5 折交叉验证测试分类器性能。5 折交叉验证的 AUC 为 0.851(图 2B),其准确率为 83.6%,PPV 为 85.7%,NPV 为 81.5%,结果虽然不错,但是也可能是由于过拟合导致,必须通过测试集进行验证。使用训练集所有样

表 1 数据集基本信息

	GSE57345		GSE21610	
HF(是/否)	177(HF)	136(NF)	30(HF)	8(NF)
性别(男/女)	144 /33	73/63	28/2	6/2
年龄	55.42±11.61	49.36±15.00	51.27±12.77	29.00±17.40
HF 病因				
DCM	82		21	
ISCH	95		9	

表 2 总体和 10 部分相关系数的区分能力

	GSE57345		GSE21610					
	HF	NF	AUC	P	HF	NF	AUC	P
A	0.126±0.367	-0.168±0.372	0.722	2.252e-11	0.111±0.424	-0.326±0.469	0.729	0.04222
F1	0.00881±0.291	0.110±0.299	0.586	2.86e-3	0.165±0.389	-0.0530±0.545	0.642	3.15e-1
F2	0.207±0.264	0.190±0.261	0.538	5.72e-1	0.130±0.388	-0.201±0.485	0.688	1.06e-1
F3	0.0851±0.331	-0.0509±0.281	0.638	1.08e-4	-0.201±0.518	-0.174±0.579	0.517	9.04e-1
F4	-0.163±0.274	-0.243±0.371	0.599	3.83e-2	-0.032±0.530	-0.259±0.628	0.617	3.72e-1
F5	0.204±0.298	-0.187±0.335	0.814	9.173e-23	0.164±0.346	-0.313±0.326	0.875	3.66e-3
F6	0.222±0.322	-0.0146±0.326	0.703	6.01e-10	0.0659±0.365	-0.260±0.332	0.738	3.27e-2
F7	0.0640±0.356	-0.193±0.386	0.698	4.90e-9	0.132±0.410	-0.133±0.619	0.613	2.81e-1
F8	0.261±0.307	-0.161±0.289	0.844	6.02e-29	0.0891±0.389	-0.364±0.289	0.875	2.56e-3
F9	-0.0488±0.322	-0.155±0.326	0.678	7.52e-8	0.0970±0.448	-0.319±0.476	0.742	4.88e-2
F10	0.301±0.306	0.160±0.277	0.656	2.55e-5	0.174±0.419	-0.0641±0.560	0.588	2.91e-1

“A”所有转录本.F1 到 F10 表示 1 到 10 部分

本构建模型,然后使用GSE57345的测试集部分数据进行测试,其AUC为0.793(图2C),准确率、PPV和NPV分别为77.8%、80.0%和75.0%。最终,在完全一个独立的数据集GSE21610上测试了我们的分类器(图2D),AUC高达0.908。准确率、PPV和NPV分别为86.8%、86.7%和87.5%。总体上,本研究构建的分类器能够很好地区分HF和NF(表3)。

在独立的测试集GSE21610中,HF和NF的年龄不匹配。为了排除年龄的影响,我们挑选了10个年龄在11.6至46.4(NF样本年龄的 $\bar{x}\pm s$)的HF样本来匹配NF样本。这18个样本(10个HF和8个NF)的准确率、PPV和NPV分别为83.3%、80.0%和87.5%。这说明我们的分类器在独立测试集上的效果不受年龄影响,也并非正负样本差引起。

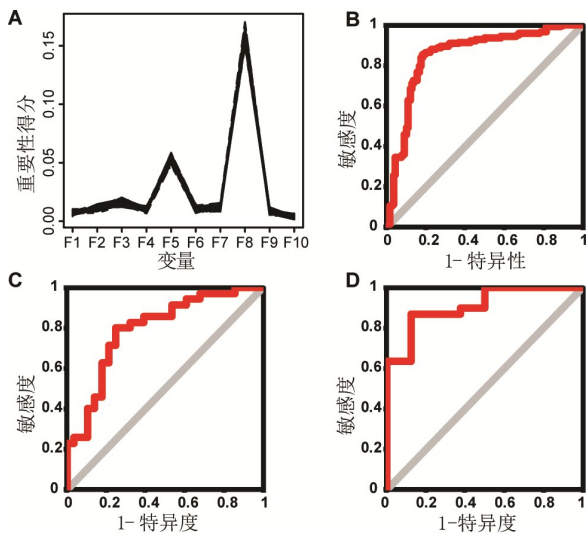


图2 随机森林模型性能评估

A:特征重要性得分;B:模型在训练集5折交叉验证ROC曲线;C:模型在测试集的ROC曲线;D:模型在独立测试集GSE21610中的ROC曲线

表3 分类器的性能评估

	5折交叉验证	测试集	独立测试集
曲线下面积	0.851	0.793	0.908
准确率	83.6% (209/250)	77.8%(49/63)	86.8% (33/38)
阳性预测值	85.8% (121/142)	80.0% (28/35)	86.7% (26/30)
阴性预测值	81.5% (88/108)	75.0% (21/28)	87.5% (7/8)

3 讨论

在本研究中,本研究发现转录本表达量与转录本对应蛋白质的氧原子含量的相关性能很好地区分HF与NF。相关性在HF组中更高,这说明在缺氧环境下,转录组受到了很大影响,系统更倾向于表达氧原子含量更高的蛋白质,可能是为了补偿缺氧或者是缓解氧化应激压力。通过交叉验证、测

试集以及独立测试集评估,发现利用表达量-氧原子含量相关性构建的分类器展现了非常稳定的预测性能。本研究开发的基于全转录组的策略能够避免少量表达波动带来的影响。

然而本研究仍然存在缺陷。首先,受限于数据库中HF患者心脏组织表达谱数据的限制,仅仅评估了其区分HF和NF的性能,没有评估其与HF患者预后的关系。其次,使用的数据集,其样本来源是心肌组织(作为临床预测标志物,更容易取材或者非侵袭的取材更方便合适,比如血液、尿液等)。现实中,常用的很多HF标志物确实来自血液或者是身体检查。尽管如此,也有基于心肌组织的标志物。Heidecker用心肌活检组织找到了HF的转录组标志物^[26]。Morgun使用心肌活检组织转录组构建了预测是否发生心脏移植排斥反应的预测器,提高了心脏移植排斥诊断率^[36]。实际上,心肌活检组织是诊断心脏移植排斥反应的金标准。最后,本研究开发的标志物是基于整个转录组的,不是基于个别基因或者少数基因的,无法对其功能做阐释。我们发现转录效率低的转本部分(F1、F2、F3)对HF鉴别无作用,而转录效率高的部分对HF有鉴别作用,这可能与人类进化过程中氧含量的变化有关,也可能与人体反应机制有关。在缺血情况下,心肌组织氧含量高的转录本转录更多,可能是反应性的存储更多的氧而导致的。随着技术的发展,心肌组织样本提取会变得越来越容易,侵袭性越来越小,基于心肌组织的标志物会越来越多。由于HF是临床死亡最多原因之一,多变量的标志物将会更好的预测HF,改善其预后^[37]。因此,如果有合适的数据,我们非常乐意在新的数据中检验我们预测器预测HF及HF预后的能力。

总的来说,本研究构建了一个基于全转录组的HF标志物,能够准确的区分HF与NF。随着全转录组技术的成熟,价格越来越低,本研究提出的策略能够很容易的实施,未来将会有更多的转录组标志物甚至多组学整合策略的标志物出现。

【参考文献】

- [1] Fordjour PA, Wang L, Gao H, et al. Targeting BNIP3 in inflammation-mediated heart failure: a novel concept in heart failure therapy[J]. Heart Fail Rev, 2016,21(5):489-497.
- [2] Selman L, Harding R, Beynon T, et al. Modelling services to meet the palliative care needs of chronic heart failure patients and their families: current practice in the UK[J]. Palliat Med, 2007 ,21(5):385-390.
- [3] Blinderman CD, Homel P, Billings JA, et al. Symptom distress and quality of life in patients with advanced congestive heart failure[J]. J Pain Symptom Manage,2008,35(6):594-603.
- [4] Chen J, Normand SL, Wang Y, et al. National and regional trends in heart failure hospitalization and mortality rates for Medicare beneficiaries, 1998-2008[J]. JAMA,2011,306(15):1669-1678.

- [5] Mozaffarian D, Benjamin EJ, Go AS, et al. Heart disease and stroke statistics-2016 update: a report from the American Heart Association[J]. *Circulation*,2016, 133(4): e38-e360.
- [6] Adler ED, Goldfinger JZ, Kalman J, et al. Palliative care in the treatment of advanced heart failure[J]. *Circulation*, 2009, 120(25): 2597-2606.
- [7] Gadoud A, Jenkins SM, Hogg KJ. Palliative care for people with heart failure: summary of current evidence and future direction [J]. *Palliat Med*, 2013, 27(9): 822-828.
- [8] Truby LK, Rogers JG. Advanced Heart Failure: Epidemiology, Diagnosis, and Therapeutic Approaches[J]. *JACC Heart Fail*, 2020, 8(7): 523-536.
- [9] Mosterd A, Hoes AW. Clinical epidemiology of heart failure[J]. *Heart*,2007, 93(9): 1137-1146.
- [10] Zarrinkoub R, Wettermark B, Wandell P, et al. The epidemiology of heart failure, based on data for 2.1 million inhabitants in Sweden[J]. *Eur J Heart Fail* , 2013, 15(9): 995-1002.
- [11] Avery CL, Loehr LR, Baggett C, et al. The population burden of heart failure attributable to modifiable risk factors: the ARIC (Atherosclerosis Risk in Communities) study[J]. *J Am Coll Cardiol*, 2012, 60(17): 1640-1646.
- [12] Go AS, Mozaffarian D, Roger VL, et al. Executive summary: heart disease and stroke statistics--2014 update: a report from the American Heart Association[J]. *Circulation*, 2014, 129(3): 399-410.
- [13] Onoue Y, Izumiya Y, Hanatani S, et al. A simple sarcopenia screening test predicts future adverse events in patients with heart failure[J]. *Int J Cardiol*, 2016, 215: 301-306.
- [14] Soyama Y, Mano T, Goda A, et al. Prognostic value of diastolic wall strain in patients with chronic heart failure with reduced ejection fraction [J]. *Heart vessels*, 2017 ,32(1):68-75.
- [15] Huang J, Tan L, Shen R, et al. Decreased peripheral mitochondrial DNA copy number is associated with the risk of heart failure and long-term outcomes[J]. *Medicine*, 2016, 95(15): e3323.
- [16] Lotierzo M, Dupuy AM, Kalmanovich E, et al. sST2 as a value-added biomarker in heart failure[J]. *Clin Chim Acta*, 2020, 501: 120-130.
- [17] Devaux Y. Cardiomyocyte-specific cell-free DNA as a heart failure biomarker?[J] *Can J Cardiol*, 2020, 36(6): 807-808.
- [18] Nagai T, Nishimura K, Honma T, et al. Prognostic significance of endogenous erythropoietin in long-term outcome of patients with acute decompensated heart failure[J]. *Eur J Heart Fail*, 2016, 18(7):803-813.
- [19] Liu S, Iskandar R, Chen W, et al. Soluble glycoprotein 130 and heat shock protein 27 as novel candidate biomarkers of chronic heart failure with preserved ejection fraction[J]. *Heart Lung Circ*, 2016,25(10):1000-1006.
- [20] Gaggin HK, Januzzi JL, Jr. Biomarkers and diagnostics in heart failure[J]. *Biochimica et biophysica acta*,2013, 1832(12): 2442-2450.
- [21] Wettersten N, Maisel AS. Biomarker developments in heart failure: 2016 and beyond[J]. *Curr Opin Cardiol*, 2019, 34(2): 218-224.
- [22] Hall PA, Reis-Filho JS, Tomlinson IP, et al. An introduction to genes, genomes and disease[J]. *J Pathol*,2010, 220(2): 109-113.
- [23] Heidecker B, Hare JM. The use of transcriptomic biomarkers for personalized medicine[J]. *Heart Fail Rev*,2007, 12(1): 1-11.
- [24] Maciejak A, Kiliszek M, Michalak M, et al. Gene expression profiling reveals potential prognostic biomarkers associated with the progression of heart failure[J]. *Genome Med* ,2015, 7(1): 26.
- [25] Liu Y, Morley M, Brandimarto J, et al. RNA-Seq identifies novel myocardial gene expression signatures of heart failure[J]. *Genomics*, 2015, 105(2): 83-89.
- [26] Heidecker B, Kasper EK, Wittstein IS, et al. Transcriptomic biomarkers for individual risk assessment in new-onset heart failure [J]. *Circulation* ,2008, 118(3): 238-246.
- [27] Acquisti C, Kleffe J, Collins S. Oxygen content of transmembrane proteins over macroevolutionary time scales[J]. *Nature*, 2007, 445(7123): 47-52.
- [28] Raymond J, Segre D. The effect of oxygen on biochemical networks and the evolution of complex life[J]. *Science*, 2006, 311(5768): 1764-1767.
- [29] Falkowski PG. Evolution. Tracing oxygen's imprint on earth's metabolic evolution[J]. *Science*, 2006, 311(5768): 1724-1725.
- [30] Kunutsor SK, Laukkanen JA, Bluemke DA, et al. Baseline and long-term gamma-glutamyltransferase, heart failure and cardiac arrhythmias in middle-aged Finnish men: Prospective study and pooled analysis of published evidence[J]. *Eur J Prev Cardiol*, 2016, 23(13):1354-1362.
- [31] Schwientek P, Ellinghaus P, Steppan S, et al. Global gene expression analysis in nonfailing and failing myocardium pre- and post-pulsatile and nonpulsatile ventricular assist device support[J]. *Physiol Genomics* ,2010, 42(3): 397-405.
- [32] Cunningham F, Amode MR, Barrell D, et al. Ensembl 2015[J]. *Nucleic Acids Res*,2015,43(Database issue):D662-D669.
- [33] Barrett T, Wilhite SE, Ledoux P, et al. NCBI GEO: archive for functional genomics data sets--update[J]. *Nucleic Acids Res*, 2013, 41(Database issue): D991-D995.
- [34] Presnyak V, Alhusaini N, Chen YH, et al. Codon optimality is a major determinant of mRNA stability[J]. *Cell*, 2015, 160(6): 1111-1124.
- [35] Liaw A, Wiener M. Classification and regression by randomForest[J]. *R news* ,2002, 2(3): 18-22.
- [36] Morgun A, Shulzhenko N, Perez-Diez A, et al. Molecular profiling improves diagnoses of rejection and infection in transplanted organs[J]. *Circ Res* ,2006, 98(12): e74-e83.
- [37] Toth GG, Vanderheyden M, Bartunek J. Novel device-based interventional strategies for advanced heart failure[J]. *Postepy Kardiol Interwencyjne*, 2016, 12(1): 13-16.