

lncRNA 芯片服务

长链非编码 RNA (Long noncoding RNAs, lncRNAs) 是一类长度超过 200nt, 无或少有蛋白编码能力的 RNA 转录本, 在大部分真核生物基因组多由 RNA 聚合酶 II 转录。lncRNA 的表达具有组织特异性, 既参与表观遗传、可变剪接、入核转运等过程, 也能以细胞微结构原件、小 RNA 前体等发挥功能, 且其转录和功能失调可能参与疾病的发生。因此, 对 lncRNA 生物学功能及其与疾病关系的深入研究可为临床提供新的思路和应对策略。

一 芯片规格

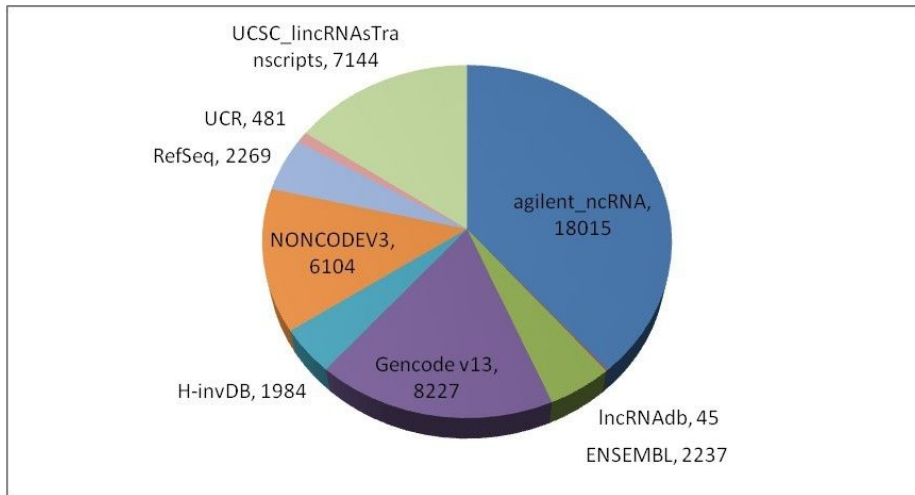
| 芯片名称 | 物种 | 格式 | 非冗余探针统计 |
|-------------------------------|----|--------|---------------------------|
| OE_Biotech Human lncRNA 芯片 | 人 | 4*180K | lncRNA 46506条+mRNA 30656条 |

二 芯片特点

该款芯片覆盖了众多数据库最新 lncRNA, 是目前涵盖 lncRNA 序列最全面的一款产品, 同时也包含美国国家生物技术信息中心 RefSeq 基因数据库中最新 mRNA, 可在一张芯片上同时对 lncRNA 和 mRNA 进行检测, 挖掘二者之间的关联。

该款芯片利用 Agilent SurePrint 专利技术合成 60mer 的长寡核苷酸探针, 具有超过五个数量级的检测动态范围和极高的灵敏度, 可同时检测样本中的低表达和高表达转录本。对于每个 lncRNA 探针至少重复 3 次, 保证了实验的检测可靠度。

三 lncRNA 来源



1) agilent_ncRNA

<https://earray.chem.agilent.com/earray/>

我们沿用安捷伦 SurePrint G3 Human Gene Expression 8x60K v2 版中的非编码 RNA 18015条，该款芯片中的非编码 RNA 来自博德研究所。

数据收集时间：2012年7月

2) lncRNAdb

<http://lncrnadb.com/>

该数据库的 lncRNA 来自发表的文献，我们的芯片收录了该数据库45条 lncRNA 序列。

数据收集时间：2012年6月

3) Gencode V13

<http://www.gencodegenes.org/>

该数据库收录了包括 UTR 数据在内的人类转录本数据，还加入了计算机和试验输入及验证数据。我们的芯片收录了该数据库8227条 lncRNA 序列。

数据收集时间：2012年8月

4) H-invDB

<http://h-invitational.jp/hinv/>

该数据库是由日本生物信息研究中心(JBIRC)建立并由它和日本 DNA 数据库共同维护的关于人类基因的数据库, 我们的芯片收录了该数据库1984条 lncRNA 序列。

数据收集时间: 2012年5月

5) NONCODE V3

<http://noncode.org/NONCODERv3/guide.htm>

该数据库是中国科学院计算技术研究所生物信息学研究组和中国科学院生物物理研究所生物信息学实验室共同开发和维护的一个提供给科学研究人员分析非编码 RNA 基因的综合数据平台。我们的芯片收录了6104条 lncRNA 序列。

数据收集时间: 2012年6月

6) RefSeq

<http://www.ncbi.nlm.nih.gov/projects/RefSeq/>

参考序列数据库(RefSeq)是NCBI提供的包含多物种序列数据信息及相关资料的数据库, 用于医学、基因功能和基因功能比较研究。RefSeq 数据库中所有的数据是一个非冗余的、提供参考标准的数据, 我们的芯片收录了该数据库2269条 lncRNA 序列。

数据收集时间: 2012年8月

7) Ultra-conserved region encoding lncRNAs (UCR)

<http://users.soe.ucsc.edu/~jill/ultra.html>

人类基因中极度保守长链非编码区域命名为超保守区域, 这些区域常位于肿瘤相关的基因区域或脆性位点。这些区域编码的 RNA 称为超保守区域转录本(T-UCRs)。我们的芯片收录了该数据库481条 lncRNA 序列。

数据收集时间: 2012年1月

8) UCSC_lincRNAsTranscripts

<http://genome.ucsc.edu>

该数据库是由加州大学圣克鲁兹分校 (UCSC) 创立和维护的, 该站点包含有多个物种的基因组草图, 并提供一系列的网页分析工具。我们的芯片收录了该数据库7144条 lncRNA 序列。

数据收集时间: 2012年7月

9) Ensembl

<http://www.ensembl.org>

该数据库是由欧洲生物信息研究所 (EBI) 和桑格学院研究所 (Sanger Institute) 共同合作开发的数据库项目, 我们的芯片包含了该数据库2237条 lncRNA 序列。

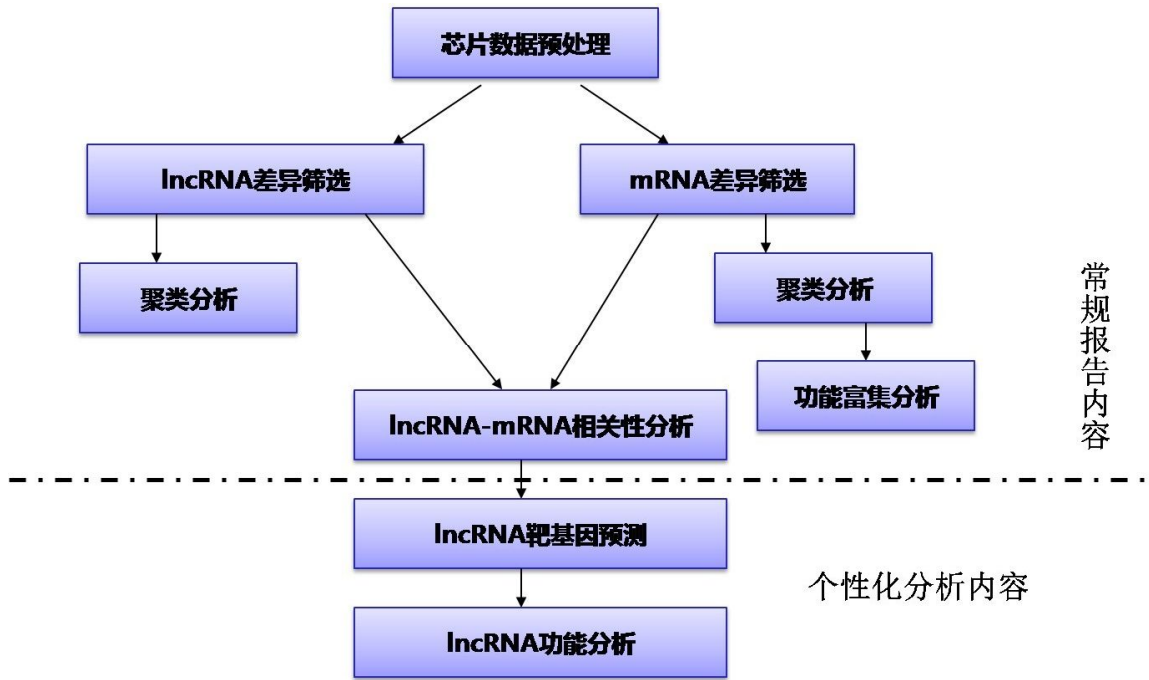
数据收集时间: 2012年8月

lncRNA 总数: 46506

四 lncRNA 研究策略

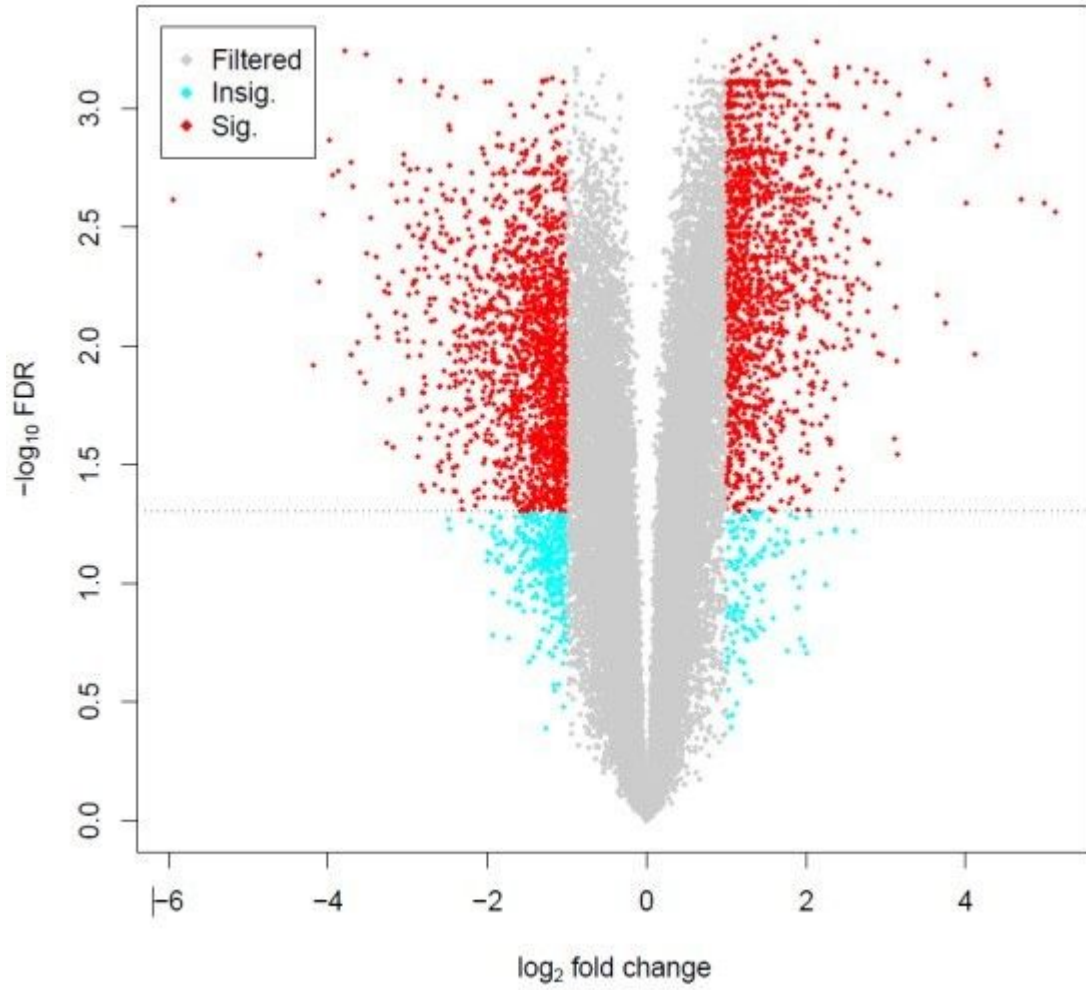


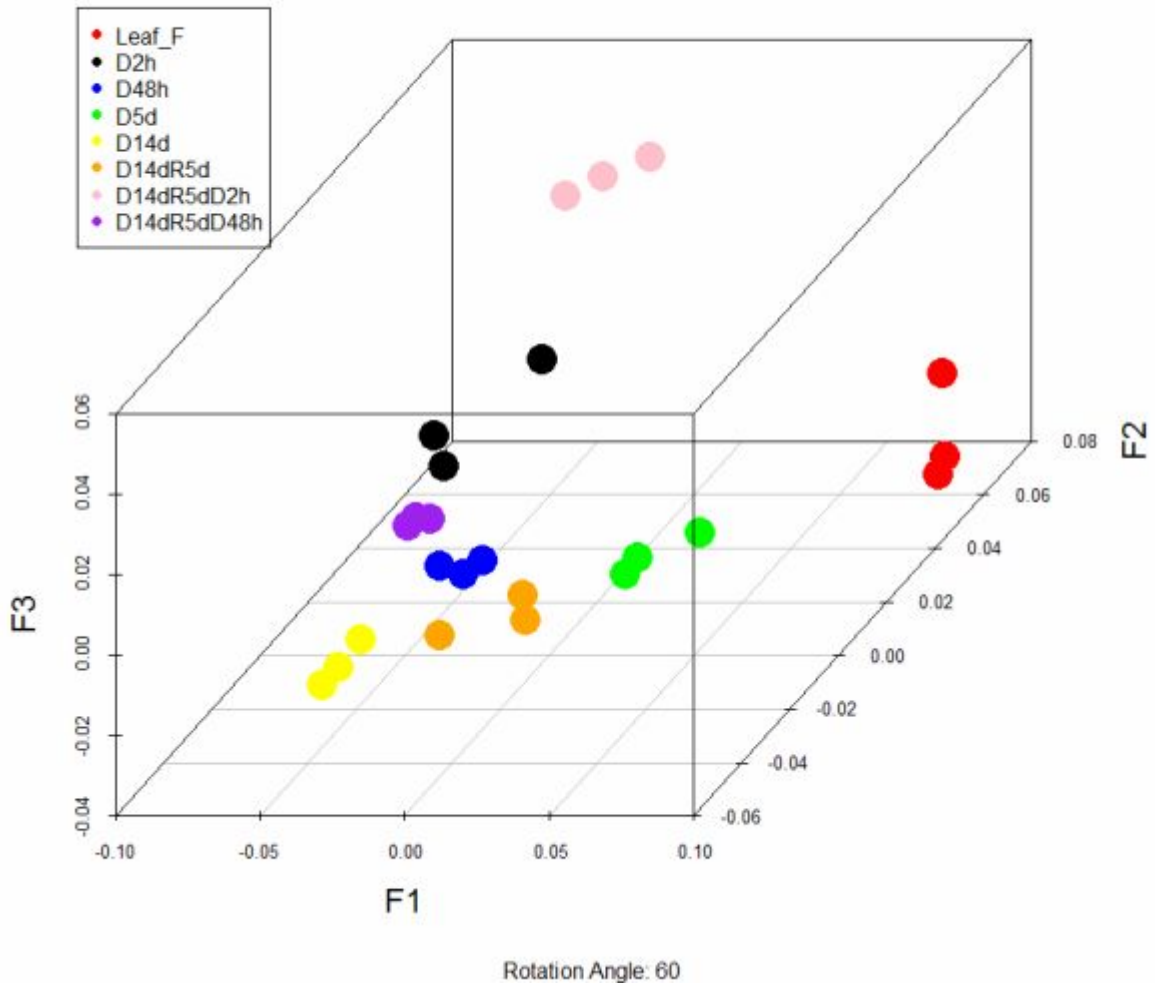
五 lncRNA 芯片数据分析策略



1) 芯片数据预处理：对实验数据质量评估，预处理及均一化处理。

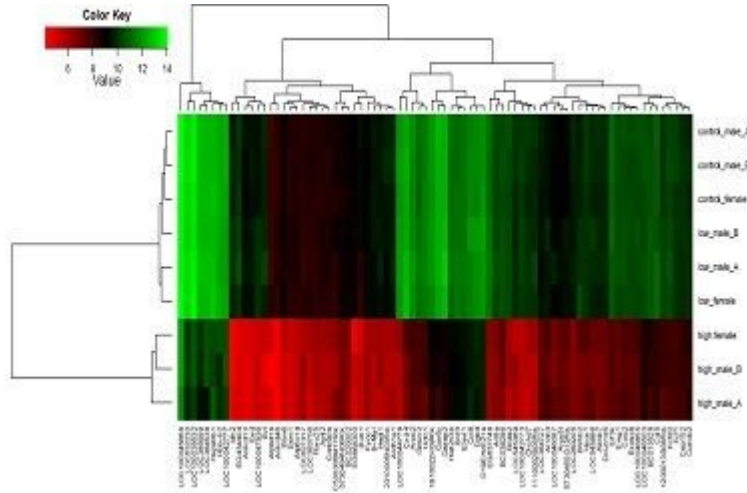
2) 差异 lncRNA 和 mRNA 的筛选：根据客户提供样本量的大小与分布或实验目的，应用倍数法、多重假设检验等手段，对两条件或多条件下的表达差异的 lncRNA 和 mRNA 分别进行计算和筛选。





| lncRNA | 信号值 | | | | | | 差异分析 | |
|---------|-------|-------|-------|-------|-------|-------|------------------|--------|
| NAME | C1 | C2 | Cn | N1 | N2 | Nn | log2 fold change | FDR |
| Probe_1 | 5.946 | 6.115 | | 7.269 | 7.342 | | -1.35 | 0.0037 |
| Probe_2 | 4.525 | 4.046 | | 5.036 | 5.285 | | -1.09 | 0.0111 |

3) 聚类分析: 针对芯片结果进行样本及差异表达 lncRNA 和 mRNA 的聚类, 寻找属于同一表达趋势的基因或样本。

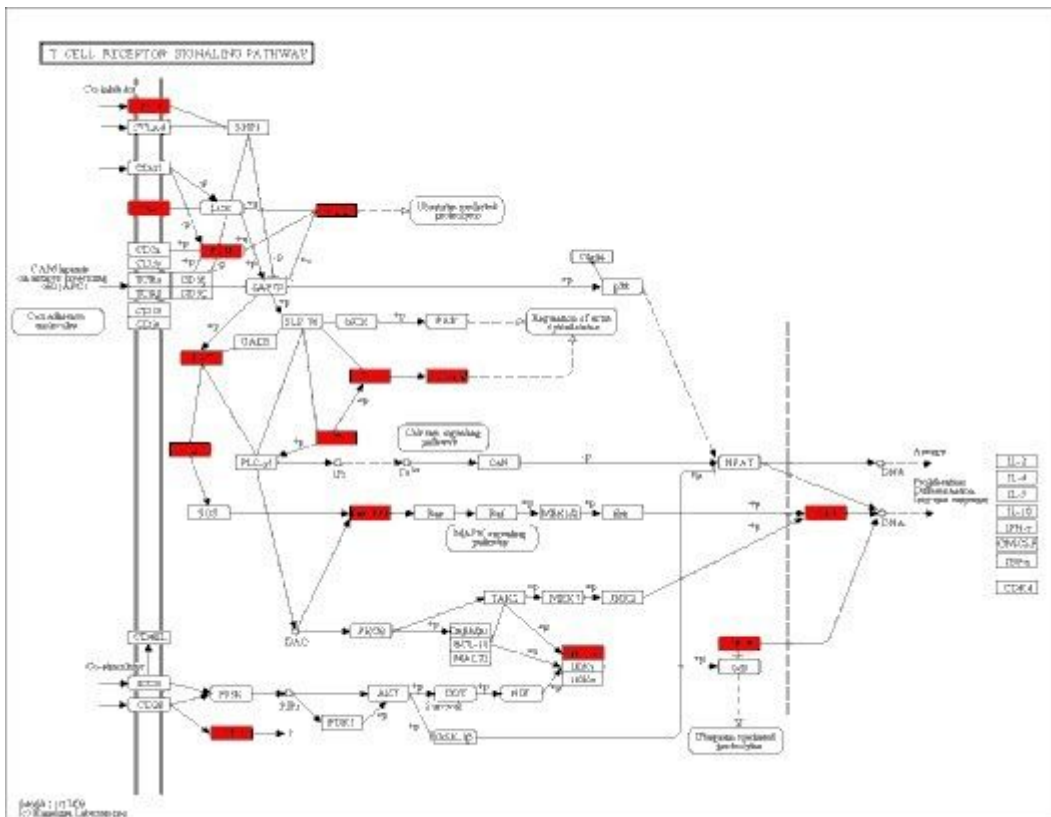
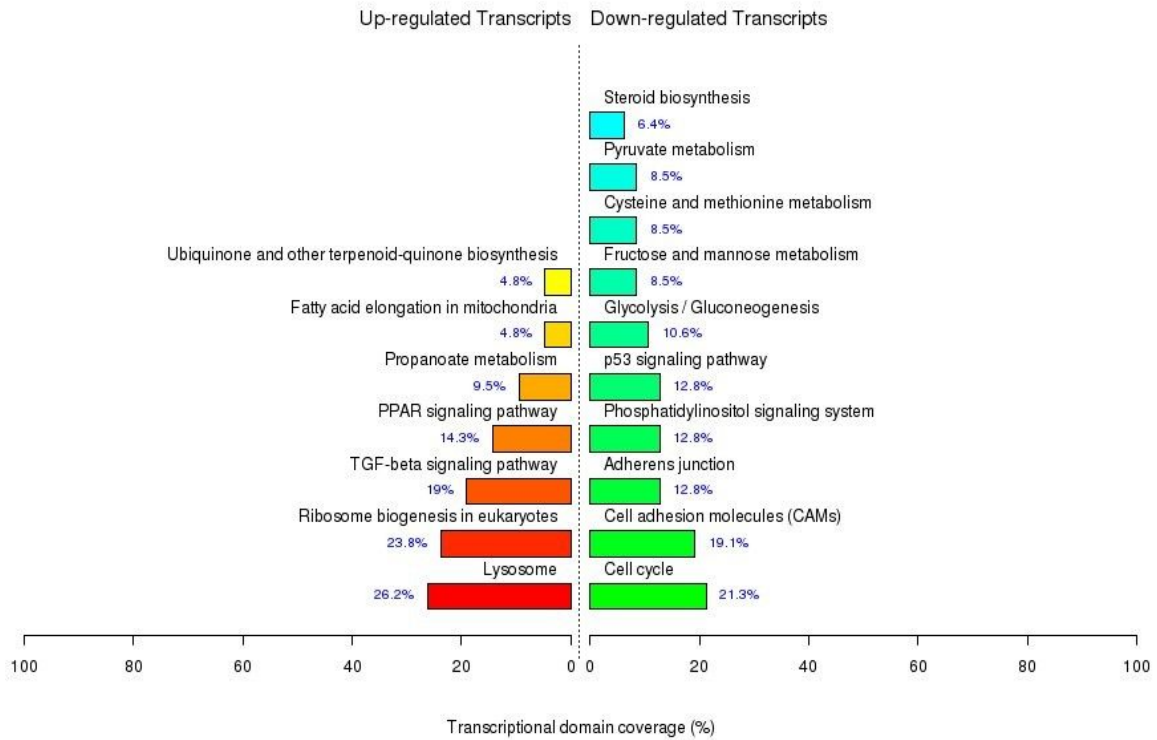


4) 功能富集分析: 差异基因, 应用数据库进行功能富集分析, 挖掘具有统计学意义的差异表达基因的功能类别。显著性 P 值越小, 则它随机聚集差异表达基因的概率越小, 其功能相关性的非随机性就越小, 该功能模块有较大的可能与疾病 (或药物作用) 相关。

Terms

| Rank | ID | Term | List Hits | List Total | Population Hits | Population Total | P-value | Genes ID's | Genes Names |
|------|------------|-----------------------------|-----------|------------|-----------------|------------------|----------|------------------|--|
| 1 | GO:0055085 | transmembrane transport | 10 | 86 | 670 | 14200 | 7.13e-03 | 118980 140803 | sideroflexin 2 |
| 2 | GO:0007049 | cell cycle | 6 | 86 | 403 | 14200 | 3.54e-02 | 1104 23176 | regulator of chromosome condensation 1 |
| 3 | GO:0007608 | sensory perception of smell | 6 | 86 | 434 | 14200 | 4.79e-02 | 1813 219981 | dopamine receptor D2 |
| 4 | GO:0051301 | cell division | 5 | 86 | 262 | 14200 | 2.14e-02 | 1104 311 | regulator of chromosome condensation 1 |
| 5 | GO:0034329 | cell junction assembly | 3 | 86 | 89 | 14200 | 1.68e-02 | 10326 253559 | signal-regulatory protein beta 1 |

KEGG



5) lncRNA-mRNA 相关性分析: 根据设定的阈值筛选 lncRNA 与 mRNA 关系对, 构建 lncRNA-mRNA 共表达网络。

| lncRNA-mRNA对 | | | | | | | | | | | |
|--------------|-----------|-----------|----------------------|--------|------------|-----------|-----------|-----------|------------|----------|-----------|
| lncRNA | pvalue | 相关性系数 | GenomicCoordinates | mRNA | PrimaryAcc | RefSeqAcc | GenbankAc | UniGeneID | EntrezGene | GeneSymb | GeneNa |
| probe_1 | 0.0011257 | 0.9218195 | chr7:7306930-7306989 | mRNA_1 | NM_002960 | NM_002960 | NM_002960 | Hs.557609 | 6274 | S100A3 | S100 ca |
| probe_1 | 0.0018777 | 0.9069247 | chr7:7306930-7306989 | mRNA_2 | NM_003258 | NM_003258 | NM_003258 | Hs.515122 | 7083 | TK1 | thymidin |
| probe_1 | 0.0060354 | 0.8609858 | chr7:7306930-7306989 | mRNA_3 | NM_004256 | NM_004256 | NM_004256 | Hs.225941 | 9390 | SLC22A13 | solute c |
| probe_1 | 0.006037 | 0.8609725 | chr7:7306930-7306989 | mRNA_4 | NM_000187 | NM_000187 | NM_000187 | Hs.368254 | 3081 | HGD | homoger |
| probe_1 | 0.0046665 | 0.8728033 | chr7:7306930-7306989 | mRNA_5 | NM_002205 | NM_002205 | NM_002205 | Hs.505654 | 3678 | ITGA5 | integrin, |
| probe_1 | 0.0085575 | 0.8430877 | chr7:7306930-7306989 | mRNA_6 | NM_001079 | NM_001079 | NM_001079 | Hs.234569 | 7535 | ZAP70 | zeta-cha |
| probe_1 | 0.0092032 | 0.8390654 | chr7:7306930-7306989 | mRNA_7 | NM_033102 | NM_033102 | NM_033102 | Hs.278695 | 85414 | SLC45A3 | solute c |

6) lncRNA 靶基因预测: 基于 lncRNA 与 mRNA 表达相关性以及 lncRNA 与 mRNA 基因组位置近邻关系, 得到 lncRNA 的潜在靶标基因。Cis 作用是 lncRNA 的一种靶基因作用方式。我们利用基因组注释和基因组浏览器鉴定 lncRNA 的可能的靶基因。一般在启动子区域同向转录的靶基因一般是促进表达作用, 反向则一般为抑制作用。trans 作用是 lncRNA 的另一种靶基因作用方式, 其原理基于核酸碱基配对。【个性化分析内容】

7) lncRNA 功能分析: 对于已知功能的 lncRNA, 我们对 lncRNA 进行功能标注。对于未知功能的 lncRNA, 我们预测其功能, 并进行功能富集分析。【个性化分析内容】

| lincRNA | Term | Target_Gene_Count | PVal |
|---------|---|-------------------|------|
| PROBE_1 | hsa04510:Focal adhesion | 33 | 9.05 |
| | hsa04060:Cytokine-cytokine receptor interaction | 35 | 0.01 |
| | hsa00240:Pyrimidine metabolism | 16 | 0.02 |
| | hsa04270:Vascular smooth muscle contraction | 18 | 0.02 |
| | hsa05214:Glioma | 12 | 0.02 |
| | hsa04710:Circadian rhythm | 5 | 0.02 |
| | hsa05414:Dilated cardiomyopathy | 15 | 0.03 |

六 lncRNA 芯片结果验证

上海欧易为您提供 lncRNA 实时定量 PCR 技术服务, 您只需要提供保存完好的样本, 我们可以针对您提供的样本进行实时定量 PCR 实验, 最终提供给您完整规范的实验报告, 为您节省大量的时间和精力。