

# 灵长动物 8 脑区转录组进化分析

Jingwen Yang

2018-10-15

## 目录

### 1 材料与方法

#### 1.1 数据介绍

在本文中我们选取了四个灵长动物 8 个脑区的转录组数据。四个灵长动物分别是 human (*Homo sapiens*) , chimpanzee (*Pan troglodytes*) , gorilla (*Gorilla gorilla*) 和 gibbon (*Nomascus leucogenys*) 。8 脑区的分别为 DPFC (dorsolateral prefrontal cortex 背外侧前额叶皮层) , VPFC (ventrolateral prefrontal cortex 腹外侧前额叶皮层), PMC (premotor cortex 前运动皮层), V1C( primary visual cortex 初级视觉皮层), ACC (anterior cingulate cortex 前扣带皮层), STR (striatum 纹状体), HIP (hippocampus 海马体), CB (cerebellum 小脑)。其中 DPFC, VPFC, PMC, V1C, ACC 属于 neocortical areas (新皮质区域), STR, HIP 属于 subcortical areas (皮质下区域)。CB 是小脑。除去 gibbon 中的 8 个组织只有 1 个生物学重复外, 其余物种中的 8 个组织均有 2~6 个生物学重复。我们使用 RPKM 值作为基因的表达水平度量。四个灵长动物中共得到 27991 个 one-to-one orthologous genes。

#### 1.2 看家基因 (House keeping gene) 数据集

我们选取了文章中所给出的看家基因列表。该数据集给出了一共 3804 个 human 看家基因列表。我们将这 3804 个基因集与 27991 个灵长

动物 one-to-one orthologous genes 取交集，共得到 3405 个灵长动物看家基因。

### 1.3 神经系统发育相关基因 (Nervous System development gene) 数据集

DAVID 数据库中与 “Nervous System development” 相关的 GO Term ID 为 GO:0007399，该 GO Term 包含个子 Term，共包含 265 个基因。我们将这 265 个基因集与 27991 个灵长动物 one-to-one orthologous genes 取交集，共得到 240 个灵长动物神经系统发育相关基因。

## 2 结果与讨论

### 2.1 灵长动物不同脑区进化关系

我们分析了 4 个灵长动物 8 个脑区的转录组数据，共得到了 27991 个 one-to-one orthologous genes。我们使用 sOU 距离计算了表达距离矩阵，并以 gibbon 的 CB 组织作为外类群，构建了这些组织的表达特征树，表达特征树的拓扑结构具有较高的 Bootstrap 值支持，如图一所示。表达特征树的所包含的信息可总结如下：

1. 8 个脑区中的 7 个大脑区域在 4 个灵长动物中分别聚集在一起。而且针对每个物种而言，5 个新皮质区域 (DPFC, VPFC, PMC, V1C, ACC) 倾向聚集在一起，2 个皮质下区域 (STR, HIP) 倾向聚集在一起。
2. 8 个脑区中的小脑作为在形态学上与大脑相对独立的组织展现出了其独有的进化特征。就小脑这个组织而言，四个灵长动物的小脑组织在进化树中与大脑各个区域分开，单独聚在一起。

灵长动物 8 个脑区表达特征树的拓扑结构可反映两部分的内容：

1. 我们所构建的表达特征树与这些脑区的形态学相似度比较一致。在转录组分析中，我们用于计算表达距离的 sOU 方法可帮助我们探索比较转录组分析的进化关系。
2. 不同物种不同组织之间的表达距离由两部分构成，分别是进化距离 (evolutionary distance) 和发育距离 (development distance)。当不同