**生信分析报告**

**项目标题： 预测结合转录子 ;**

**单 号： BSJF240124 ;**

**分析人员： 黄礼闯 ;**

**分析类型： 生信协助 ;**

**委 托 人： 蒋镥 ;**

**受 托 人： 杭州铂赛生物科技有限公司 .**

# 1 分析流程

使用 hTFtarget 和 JASPAR 网站分析可以与 BMP9 (GDF2) 启动子结合的转录因子（X）

# 2 材料和方法

## 2.1 数据分析平台

在 Linux pop-os x86\_64 (6.9.3-76060903-generic) 上，使用 R version 4.4.2 (2024-10-31) (<https://www.r-project.org/>) 对数据统计分析与整合分析。

## 2.2 TFBS 转录因子结合位点数据 (Dataset: BMP9)

以 RSelenium 抓取 Transcription Factor Target Gene Database 数据库 (<https://tfbsdb.systemsbiology.net/>) (2016, **IF:9**, Q1, Cell systems)1 基因与转录因子结合数据。对**基因集** (STAT1, 来自于Venn 交集[Section: BMP9]) 进行结合位点可视化。以 R 包 JASPAR2022 (0.99.8) 和 TFBSTools (1.44.0) 绘制 STAT1 Motif sequence logo。以 R 包 Gviz (1.51.0) 绘制 ENSEMBL 数据库对应基因的转录子 (transcript)。

## 2.3 hTFtarget 转录因子数据 (Dataset: BMP9)

以 RSelenium 抓取 hTFtarget 数据库 (<http://bioinfo.life.hust.edu.cn/hTFtarget#!/>) 基因与转录因子结合数据。

# 3 分析结果

## 3.1 TFBS 转录因子结合位点数据 (BMP9)

从 Transcription Factor Target Gene Database 数据库获取与 GDF2 结合的转录因子数据。

**Tab.** **1** BMP9 Transcription Factor binding sites

| Target | TF symbol | Motif | Source | Strand |
| --- | --- | --- | --- | --- |
| GDF2 | FOXB1 | FOXB1 forkhead DB... | SELEX | + |
| GDF2 | FOXB1 | FOXB1 forkhead DB... | SELEX | - |
| GDF2 | HOXD12 | HOXD12 homeodomai... | SELEX | + |
| GDF2 | HOXC10 | Hoxc10 homeodomai... | SELEX | + |
| GDF2 | MEF2D | MEF2D MADS DBD di... | SELEX | + |
| ... | ... | ... | ... | ... |

**(File path: Figure+Table/3.1\_TFBS\_转录因子结合位点数据\_(BMP9)/BMP9-Transcription-Factor-binding-sites.csv)**

Tab. **[1](#BMP9-Transcription-Factor-binding-sites)** 为转录因子与基因启动子结合位点数据。

## 3.2 hTFtarget 转录因子数据 (BMP9)

从 hTFtarget 数据库获取与 GDF2 结合的转录因子数据。

**Tab.** **2** BMP9 transcription factor data

| Target symbol | Target ensembl | Browse peaks | No. of datasets | TF |
| --- | --- | --- | --- | --- |
| GDF2 | ENSG00000263761 | Details | 3 | SPI1 |
| GDF2 | ENSG00000263761 | Details | 1 | CTCF |
| GDF2 | ENSG00000263761 | Details | 35 | CTCF |
| GDF2 | ENSG00000263761 | Details | 2 | EP300 |
| GDF2 | ENSG00000263761 | Details | 2 | ERG |
| ... | ... | ... | ... | ... |

**(File path: Figure+Table/3.2\_hTFtarget\_转录因子数据\_(BMP9)/BMP9-transcription-factor-data.xlsx)**

Tab. **[2](#BMP9-transcription-factor-data)** 从 hTFtarget 数据库获取的 GDF2 结合的转录因子附表。

## 3.3 JASPAR

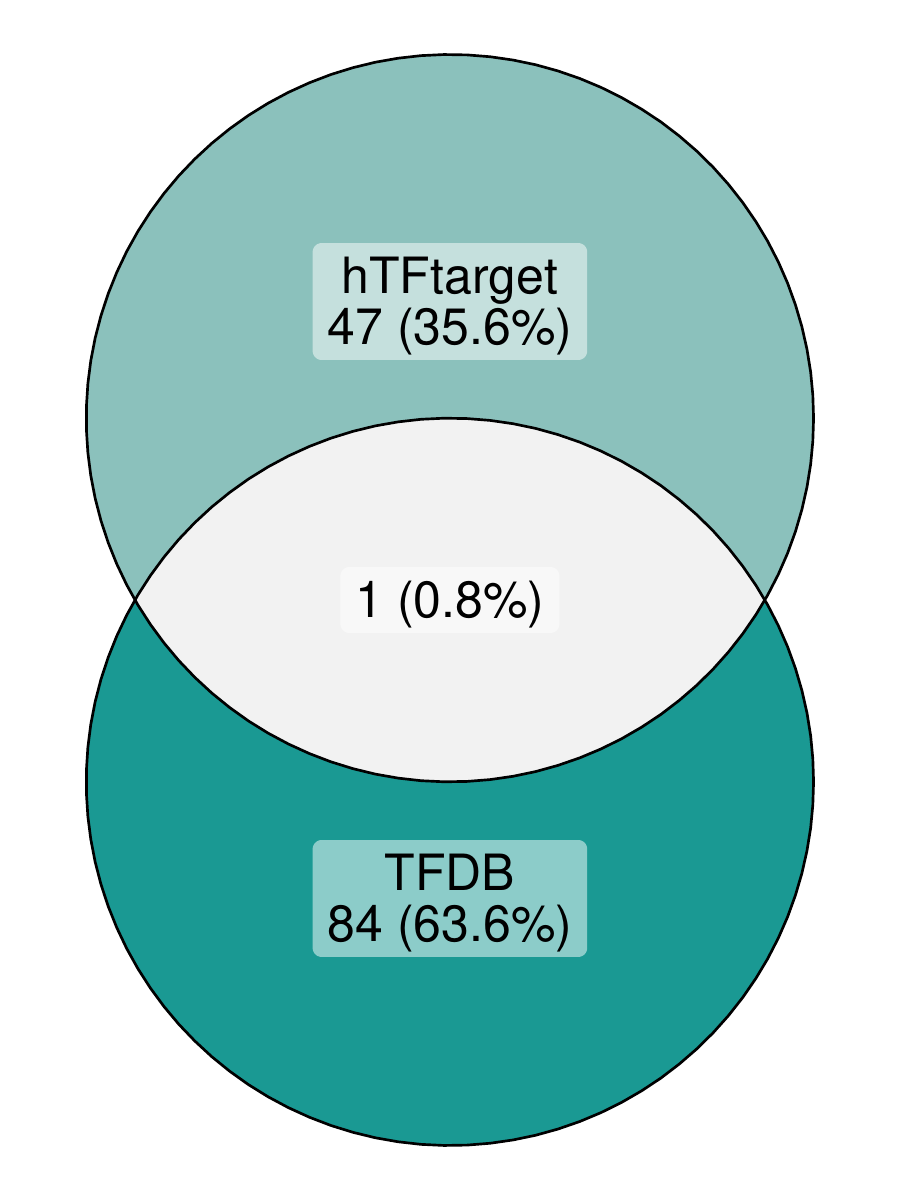
无数据 (BMP9 或 GDF2)。

<https://jaspar.elixir.no/search?q=BMP9&collection=all&tax_group=all&tax_id=all&type=all&class=all&family=all&version=all>

## 3.4 汇总: TFDB + hTFtarget (BMP9)

数据集为：

* **基因集** (GDF2, 来自于TFBS 转录因子结合位点数据[Section: BMP9])
* **基因集** (GDF2, 来自于hTFtarget 转录因子数据[Section: BMP9])



**Fig.** **1** BMP9 Intersection of TFDB with hTFtarget

**(File path: Figure+Table/3.4\_汇总:\_TFDB\_+\_hTFtarget\_(BMP9)/BMP9-Intersection-of-TFDB-with-hTFtarget.pdf)**

* All\_intersection: STAT1

**(See: Figure+Table/3.4\_汇总:\_TFDB\_+\_hTFtarget\_(BMP9)/BMP9-Intersection-of-TFDB-with-hTFtarget-content)**

Fig. **[1](#BMP9-Intersection-of-TFDB-with-hTFtarget)** 为 TFDB, hTFtarget 各自交集。

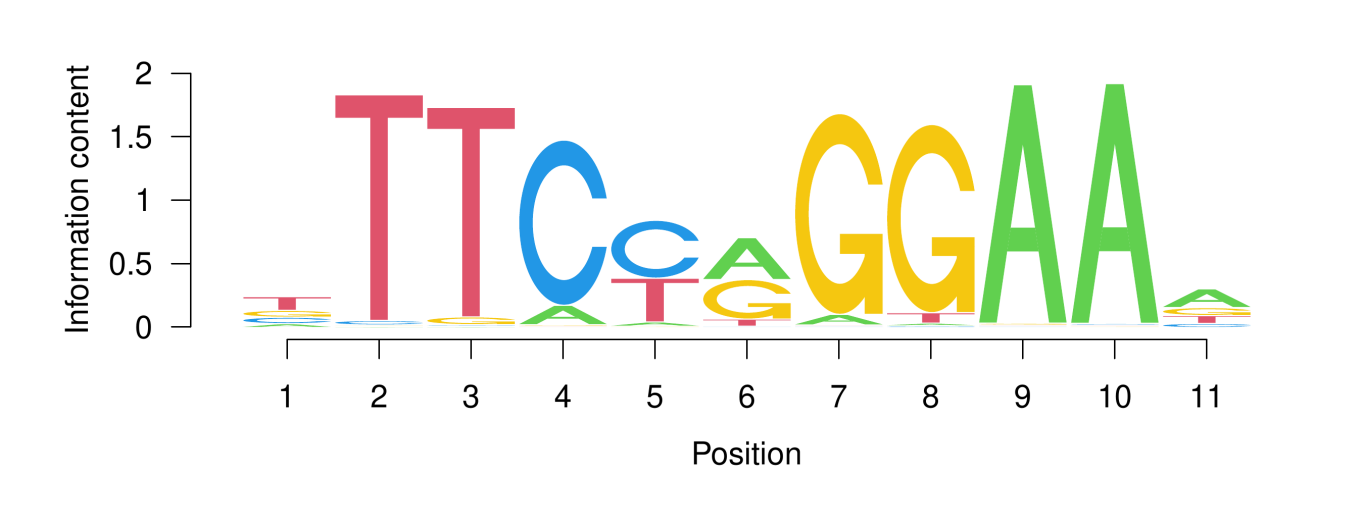
## 3.5 TFBS 结合位点可视化 (BMP9)

**Tab.** **3** BMP9 filtered TF binding site

| Target | Target entrez | TF symbol | Motif | Source |
| --- | --- | --- | --- | --- |
| GDF2 | 2658 | STAT1 | V STAT1 05 M01260 | TRANSFAC |

**(File path: Figure+Table/3.5\_TFBS\_结合位点可视化\_(BMP9)/BMP9-filtered-TF-binding-site.csv)**

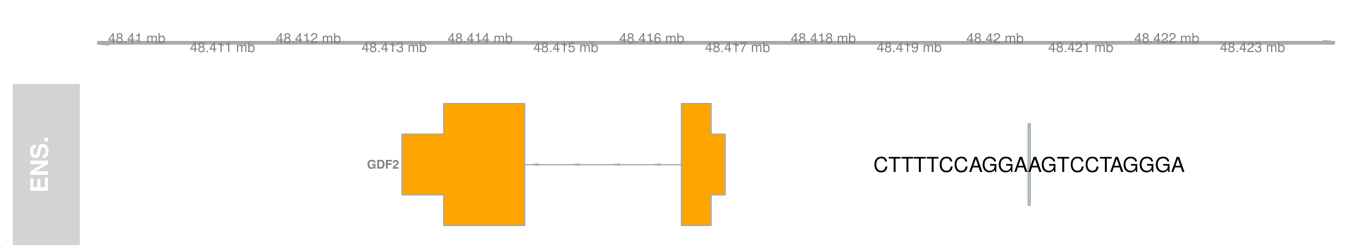
Tab. **[3](#BMP9-filtered-TF-binding-site)** 为转录因子与基因启动子结合位点数据。



**Fig.** **2** BMP9 STAT1 motif sequence logo

**(File path: Figure+Table/3.5\_TFBS\_结合位点可视化\_(BMP9)/BMP9-STAT1-motif-sequence-logo.pdf)**

Fig. **[2](#BMP9-STAT1-motif-sequence-logo)** 为 STAT1 motif sequence logo。字母高度通常代表信息量（以bits为单位）或频率。信息量高意味着该位置对结合很重要，变异较少；频率则直接显示不同碱基的出现比例。字母排列顺序：高度高的字母在上方，表示该碱基在该位置更常见。总高度（Y轴）：可能代表该位置的信息量，信息量越高，表示该位置越保守，对结合越关键。



**Fig.** **3** BMP9 STAT1 GDF2 transcript factor binding Illustrate

**(File path: Figure+Table/3.5\_TFBS\_结合位点可视化\_(BMP9)/BMP9-STAT1-GDF2-transcript-factor-binding-Illustrate.pdf)**

Fig. **[3](#BMP9-STAT1-GDF2-transcript-factor-binding-Illustrate)** 为转录因子结合基因启动子示意 (STAT1\_GDF2) (Transcript 源于 BioMart 获取 ENSEMBL 数据库对应基因的注释区域；转录因子结合位点源于 Transcription Factor Target Gene Database)。

# 4 总结

使用了两个数据集 (TFBS, hTFtarget) (JASPAR 无相关数据) ，取交集 (Fig. **[1](#BMP9-Intersection-of-TFDB-with-hTFtarget)** ) 得到 STAT1 (Fig. **[2](#BMP9-STAT1-motif-sequence-logo)** )。 其结合位点见 Fig. **[3](#BMP9-STAT1-GDF2-transcript-factor-binding-Illustrate)**

# Reference

1. Plaisier, C. L. *et al.* Causal mechanistic regulatory network for glioblastoma deciphered using systems genetics network analysis. *Cell systems* **3**, 172–186 (2016).