

# VASH2 序列分析

2024-03-27

LiChuang Huang



@ 立效研究院

# Contents

<b>1</b>	<b>摘要</b>	<b>1</b>
1.1	需求 . . . . .	1
1.2	结果 . . . . .	1
<b>2</b>	<b>前言</b>	<b>1</b>
<b>3</b>	<b>材料和方法</b>	<b>1</b>
3.1	材料 . . . . .	1
3.2	方法 . . . . .	1
<b>4</b>	<b>分析结果</b>	<b>1</b>
<b>5</b>	<b>结论</b>	<b>1</b>
<b>6</b>	<b>附：分析流程</b>	<b>1</b>
6.1	获取蛋白序列 . . . . .	1
6.2	预测位点 . . . . .	2
6.2.1	MusiteDeep . . . . .	2
6.2.2	GPS-MSP . . . . .	3
6.2.3	PRmePred . . . . .	3
	<b>Reference</b>	<b>4</b>

# List of Figures

1	PTM score . . . . .	3
---	---------------------	---

# List of Tables

1	Prediction PTM of Methylarginine . . . . .	2
2	High score prediction PTM of Methylarginine . . . . .	3
3	PRmePred results . . . . .	4

# 1 摘要

## 1.1 需求

为了鉴定 VASH2 被 PRMT5 甲基化的精氨酸残基 (methylarginine), 使用甲基化预测工具包括 PRmePred 和 GPS-MSP 分析 VASH2 的蛋白序列

protein post-translational modification site (PTM)

## 1.2 结果

VASH2 的序列见 6.1。

共尝试使用了以下三种工具预测位点。

- MusiteDeep <https://www.musite.net> [https://github.com/duolinwang/MusiteDeep\\_web](https://github.com/duolinwang/MusiteDeep_web)
- GPS-MSP <https://msp.biocuckoo.org/>
- PRmePred <http://bioinfo.icgeb.res.in/PRmePRed/>

更推荐 MusiteDeep 的预测结果<sup>1</sup> (发表于 Nucleic Acids Research), 其结果已整理, 见 Fig. 1, 和 Tab. 1。

# 2 前言

# 3 材料和方法

## 3.1 材料

## 3.2 方法

Mainly used method:

- Python tool **MusiteDeep** was used for protein post-translational modification site prediction and visualization<sup>1</sup>.
- R version 4.3.2 (2023-10-31); Other R packages (eg., **dplyr** and **ggplot2**) used for statistic analysis or data visualization.

# 4 分析结果

# 5 结论

# 6 附：分析流程

## 6.1 获取蛋白序列

‘VASH2 protein fasta’ 数据已提供。

(对应文件为 `fasta/Seq.fasta`)

## 6.2 预测位点

使用以下工具：

### 6.2.1 MusiteDeep

Table 1 (下方表格) 为表格 Prediction PTM of Methylarginine 概览。

(对应文件为 `Figure+Table/Prediction-PTM-of-Methylarginine.csv`)

注：表格共有 29 行 5 列，以下预览的表格可能省略部分数据；表格含有 1 个唯一 ‘Sequence\_name’。

Table 1: Prediction PTM of Methylarginine

Sequence_name	PTM_type	Position	Residue	PTM_score
VASH2	Methylarginine	309	R	0.7
VASH2	Methylarginine	102	R	0.43
VASH2	Methylarginine	177	R	0.258
VASH2	Methylarginine	65	R	0.215
VASH2	Methylarginine	307	R	0.198
VASH2	Methylarginine	203	R	0.151
VASH2	Methylarginine	10	R	0.146
VASH2	Methylarginine	212	R	0.137
VASH2	Methylarginine	85	R	0.132
VASH2	Methylarginine	211	R	0.113
VASH2	Methylarginine	134	R	0.107
VASH2	Methylarginine	21	R	0.081
VASH2	Methylarginine	316	R	0.072
VASH2	Methylarginine	324	R	0.071
VASH2	Methylarginine	23	R	0.069
...	...	...	...	...

Table 2 (下方表格) 为表格 High score prediction PTM of Methylarginine 概览。

(对应文件为 `Figure+Table/High-score-prediction-PTM-of-Methylarginine.csv`)

注：表格共有 1 行 5 列，以下预览的表格可能省略部分数据；表格含有 1 个唯一 ‘Sequence\_name’。

Table 2: High score prediction PTM of Methylarginine

Sequence_name	PTM_type	Position	Residue	PTM_score
VASH2	Methylarginine	309	R	0.7

Figure 1 (下方图) 为图 PTM score 概览。

(对应文件为 Figure+Table/PTM-score.pdf)

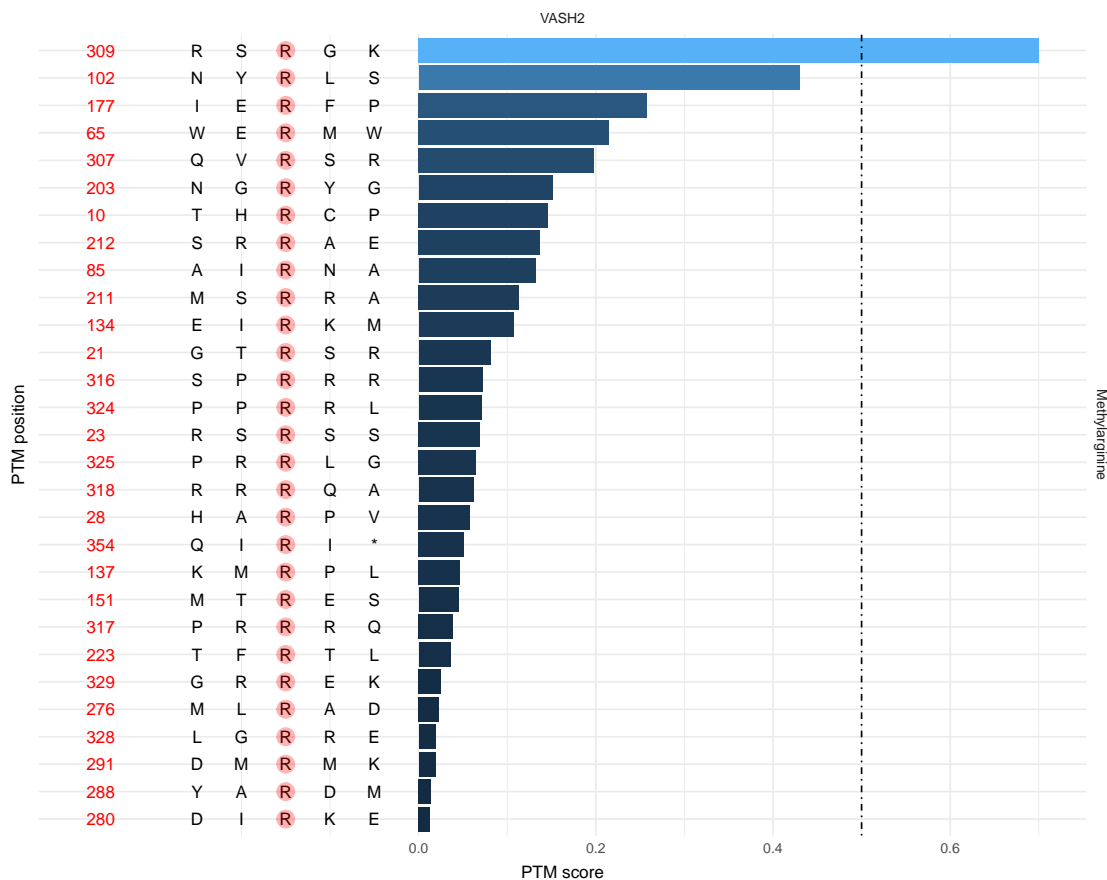


Figure 1: PTM score

### 6.2.2 GPS-MSP

该网络服务器好像并没有提供甲基化的精氨酸残基 (methylarginine) 位点预测工具 (没有 type: R.Me)。

<https://msp.biocuckoo.org/online.php>

### 6.2.3 PRmePred

<http://bioinfo.icgeb.res.in/PRmePred/>

Table 3 (下方表格) 为表格 PRmePred results 概览。

(对应文件为 Figure+Table/PRmePred-results.csv)

注：表格共有 19 行 4 列，以下预览的表格可能省略部分数据；表格含有 1 个唯一 ‘SeqId’。

Table 3: PRmePred results

SeqId	R site	Peptides	Prediction Score
VASH2	10	MTGSAADTHRCPPHKGAKG	0.753787
VASH2	21	PHPKGAKGTRSRSSHARPV	0.939423
VASH2	23	PKGAKGTRSRSSHARPVSL	0.837776
VASH2	28	GTRSRSSHARPVSLATSGG	0.959969
VASH2	85	KGEMVGAIRNA AFLAKPS	0.651501
VASH2	134	HTGTQFFEIRKMRPLSGLM	0.532284
VASH2	151	LMETAKEMTRESLPIKCLE	0.602389
VASH2	203	VVLGIYCNGRYGSLGMSRR	0.884941
VASH2	211	GRYGSLGMSRRAELMDKPL	0.815102
VASH2	212	RYGSLGMSRRAELMDKPLT	0.7399
VASH2	307	ASAHPTQVR SRGKSLSPR	0.963406
VASH2	309	AHSPTQVR SRGKSLSPRRR	0.98479
VASH2	316	RSRGKSLSPRRRQASPPRR	0.819054
VASH2	317	SRGKSLSPRRRQASPPRRL	0.921694
VASH2	318	RGKSLSPRRRQASPPRRLG	0.96432
...	...	...	...

## Reference

1. Wang, D. *et al.* MusiteDeep: A deep-learning based webserver for protein post-translational modification site prediction and visualization. *Nucleic Acids Research* **48**, W140–W146 (2020).