



# 大麻中金属硫蛋白的生物信息学分析

## Bioinformatics analysis of metallothionein in *Cannabis sativa*

4G05成员：  
4G05A 伍叶娜  
4G05B 李阿蕾  
4G05C 王元昌  
4G05D 汤雅萍

汇报人：王元昌



# 目录

CONTENTS

**01. 背景**

**02. 寻找同源序列**

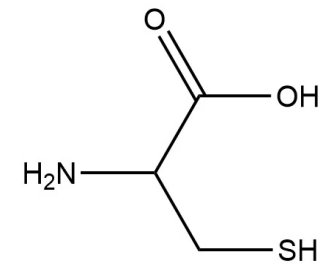
**03. 序列比对**

**04. 引物设计**



# 一、背景

●金属硫蛋白（Metallothionein, MT）是一类普遍存在的、低相对分子质量（6~7 ku）的胞质蛋白，其结构高度保守。含有丰富的半胱氨酸残基（10%~30%）和金属含量极高的蛋白质。通过聚集的巯基簇可以螯合 $\text{Cd}^{2+}$ 、 $\text{Pb}^{2+}$ 、 $\text{Bi}^{3+}$ 、 $\text{Ag}^{2+}$ 、 $\text{Cu}^{2+}$ 、 $\text{Zn}^{2+}$ 和 $\text{Ni}^{2+}$ 等离子，广泛地存在于从微生物到人类各种生物中。

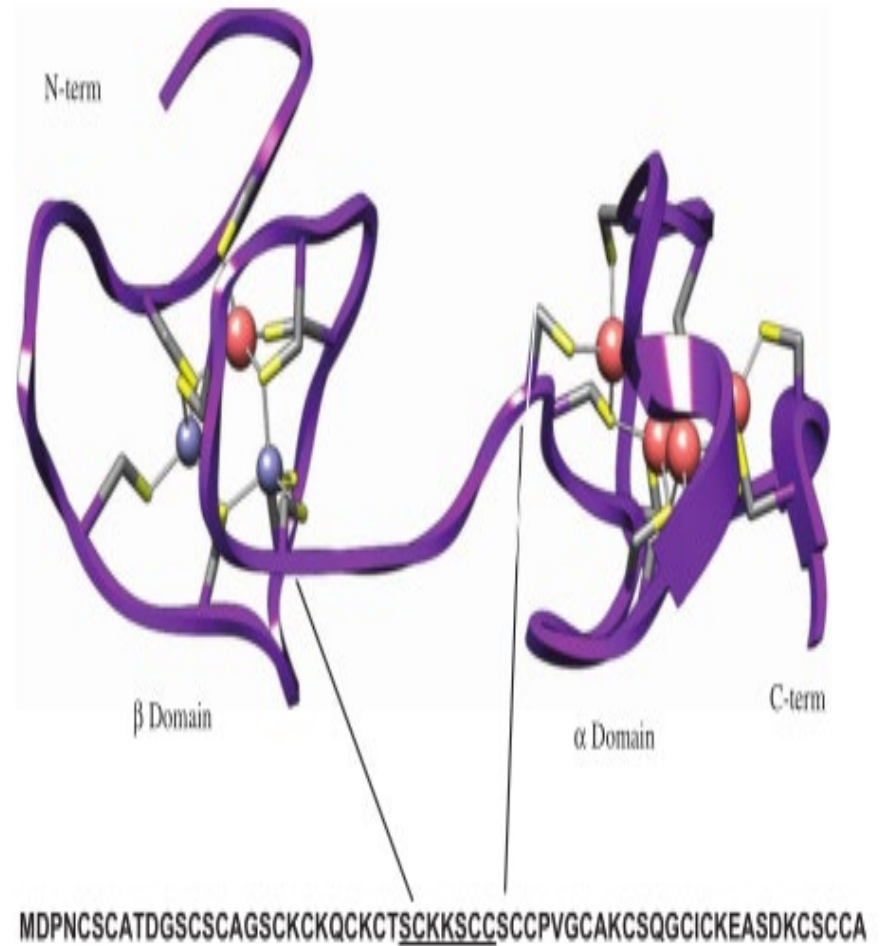


# 植物 MTs 的结构特征

MTs 与金属离子的复合物形成两个结构域： $\alpha$  结构域和 $\beta$  结构域。

MTs 铰链区的存在使 2 个结构域存在较大的柔性和可变性，从而使 MTs 中的金属离子与溶液中的金属离子易于交换，为调节体内金属离子的代谢提供结构基础。

参考文献：Sílvia Atrian, Mercè Capdevila. Metallothionein-protein interactions[J]. *BioMolecular Concepts*, 2013, 4(2).

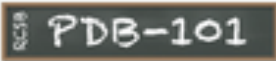







# 搜索网站

**RCSB PDB** PROTEIN DATA BANK  
177426 Biological Macromolecular Structures  
Enabling Breakthroughs in Research and Education

metallothionein

[Advanced Search](#) | [Browse Annotations](#)

Search History Browse Annotations MyPDB

◀ ▶ **QUERY:** Full Text = "metallothionein" **AND** Scientific Name of Source Organism = "Triticum aestivum"

Advanced Search Query Builder ⓘ

Refinements ⓘ

Summary Gallery Compact -- Tabular Report -- ↓ Score

SCIENTIFIC NAME OF SOURCE ORGANISM

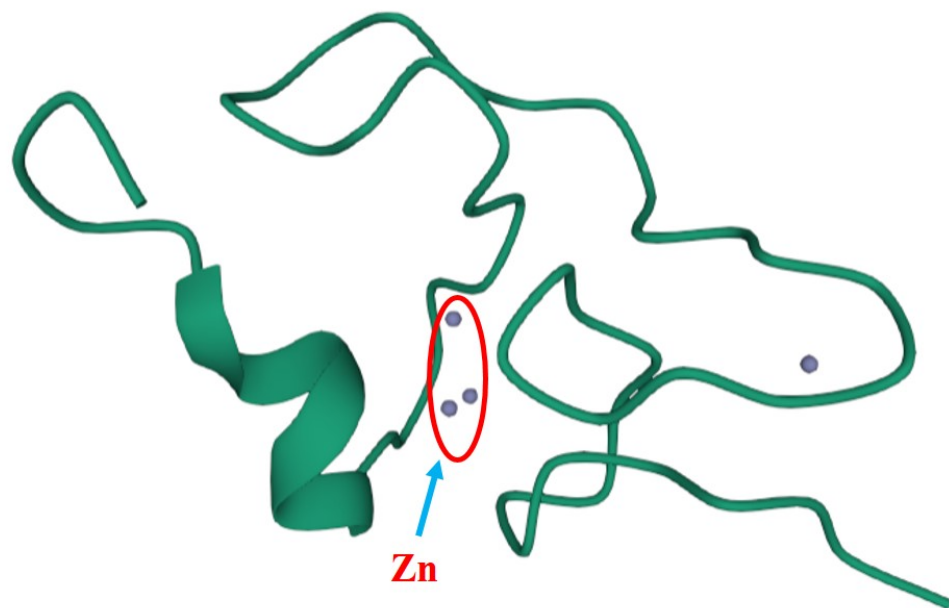
Displaying 1 to 4 of 4 Structures Page 1 of 1

Triticum aestivum (4)

2021/5/12



# 小麦MTs结构



●PDB登录号：  
2KAK



# 植物 MTs 的功能及其应用

- 自由基的清除
- 参与金属离子的供给与运输
- 解除重金属离子毒害与植物修复
- MT 的重要生理功能表现在它是一种目前所知的最有效的自由基清除剂，其清除自由基 ( $\cdot\text{OH}$ ) 的能力约为 SOD 的几千倍，而清除氧自由基 ( $\cdot\text{O}$ ) 的能力约是谷胱甘肽 (GSH) 的 25 倍，而且具有很强的抗氧化活动。研究成果已涉及医学、毒理学、营养学、生物工程、美容护肤等各个领域。

# 何为大麻？

- 大麻（学名：*Cannabis sativa* L.）是桑科大麻属植物，一年生直立**草本**，高1~3米。
- 有雌有雄。雄株叫**梟**（xi）雌株叫**苴**（jū）。
- 其主要有效化学成分为四氢大麻酚（简称**THC**），THC在吸食或口服后有精神和生理的活性作用。
- **温度**：种子在1~3℃能发芽，适宜温度为25~30℃；幼苗能忍耐-3~-5℃低温，幼苗期适宜生长温度10~15℃；快速生长期19~23℃；成熟期18~20℃。
- **光照**：大麻为喜光、短日照作物，晚熟品种对光照反应更为敏感。



营养期

生殖期





# 对照、低浓度和高浓度大麻生长状况



**CK**

**Low  
concentration**

**High  
concentration**

# 大麻为何适合修复重金属污染？

- 不破坏环境, 不会造成二次污染等;
- 生物量大, 重金属耐性强, “不与粮食争好地”。





## 二、寻找同源序列

登录号	拟南芥中 MT 蛋白	大麻中的蛋白名	序列登录号	序列长度	同源相似性	基因覆盖率
Q42377	MT4B_ARATH	EC protein I/II-like	XP_030486179	87 aa	44.19%	96.00%
P93746	MT4A_ARATH	EC protein I/II-like	XP_030486179	87 aa	43.66%	96.00%
		uncharacterized protein LOC115716228	XP_030500840	1268 aa	22.06%	70.00%
P43392	MT1A_ARATH	metallothionein-like protein 2	XP_030504357	76 aa	60.00%	55.00%
P25860	MT2A_ARATH	metallothionein-like protein type 2	XP_030505178	81 aa	58.44%	95.00%
		metallothionein-like protein type 2	XP_030480212	79 aa	60.32%	77.00%
		metallothionein-like protein 2	XP_030504357	76 aa	45.59%	80.00%
Q38804	MT1C_ARATH	metallothionein-like protein 2	XP_030504357	76 aa	60.00%	55.00%
Q38805	MT2B_ARATH	metallothionein-like protein type 2	XP_030505178	81 aa	55.26%	94.00%
		metallothionein-like protein 2	XP_030504357	76 aa	41.38%	68.00%
		metallothionein-like protein type 2	XP_030480212	79 aa	58.06%	75.00%



# 寻找到大麻中的蛋白

蛋白名称↵	登录号↵	Cys 含量↵
EC protein I/II-like↵	XP_030486179↵	19.50%↵
uncharacterized protein LOC115716228↵	XP_030500840↵	3.10%↵
metallothionein-like protein 2↵	XP_030504357↵	15.80%↵
metallothionein-like protein type 2↵	XP_030505178↵	17.30%↵
metallothionein-like protein type 2↵	XP_030480212↵	17.70%↵

# 三、序列比对

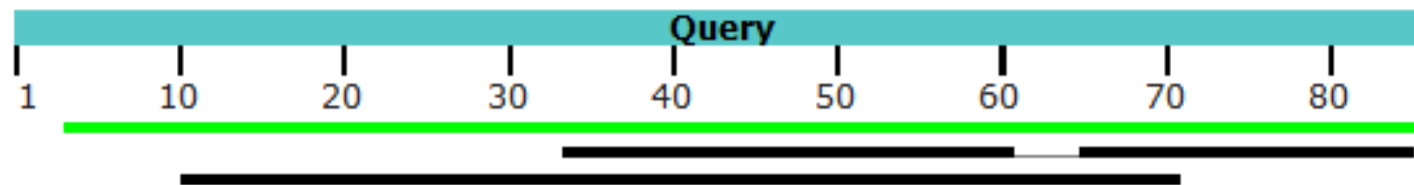
select all 3 sequences selected

[Graphics](#) [Distance tree of results](#) [Multiple alignment](#) New [MSA Viewer](#)

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/>	<a href="#">XP_030486179.1 EC protein I/II-like [Cannabis sativa]</a>		64.2	64.2	96%	2e-20	44.19%	87	Query_46587
<input checked="" type="checkbox"/>	<a href="#">XP_030480212.1 metallothionein-like protein type 2 [Cannabis sativa]</a>		24.8	47.3	55%	2e-04	33.33%	79	Query_46591
<input checked="" type="checkbox"/>	<a href="#">XP_030500840.1 uncharacterized protein LOC115716228 [Cannabis sativa]</a>		27.2	27.2	70%	6e-05	20.59%	1268	Query_46588

Alignment Scores  < 40  40 - 50  50 - 80  80 - 200  >= 200

## Distribution of the top 4 Blast Hits on 3 subject sequences



XP\_030486179.1 EC protein I/II-like [Cannabis sativa]



# 多序列比对

Alignment view Flat query-anchored with letters for identities Line length: 60 Restore defaults

Download

Query range 1: 1 to 60

Query	4	TGKGSASASCNDRCGCPSECPGGESCPHMMS-----EASGGD--QEHNTCECGEHC	53
Query_46587	2	SDTAGNMKVCDDKCGCTVPCPGGLACCTPGE-----KVSSGSSGDDHIKCSCGQHC	53
Query_46588	1028	PPSPQGCDCISGCSDSDKCAVKNGGEIPFNFGAIVE--AKPLVYECGQSC	1078
Query_46591	52	LE-----GSEMSE--VAENGCEGDMC	71

Download

Query range 2: 61 to 94

Query	54	GCNFCICPKTQTQTSAG--FCGEGCTCAICAA	85
Query_46587	54	GCNFCICARSVVSTGVGKAYCICGEGCTCVICTS	87
Query_46588	1079	KCFP--SCNRRVSQHGK	1094
Query_46591	72	KCFPC	78
Query_46591	58	SEVAENG--CICGDMICICICT	77



# Cys含量

Amino acid composition:

CSV format

Ala (A)	4	4.6%
Arg (R)	2	2.3%
Asn (N)	2	2.3%
Asp (D)	5	5.7%
Cys (C)	17	19.5%
Gln (Q)	1	1.1%
Glu (E)	2	2.3%
Gly (G)	13	14.9%
His (H)	2	2.3%
Ile (I)	1	1.1%
Leu (L)	1	1.1%
Lys (K)	6	6.9%
Met (M)	2	2.3%
Phe (F)	0	0.0%
Pro (P)	4	4.6%
Ser (S)	9	10.3%
Thr (T)	8	9.2%
Trp (W)	0	0.0%
Tyr (Y)	1	1.1%
Val (V)	7	8.0%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%

●cys 17个, 19.5%

(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/>	<a href="#">XP_030486179.1 EC protein I/II-like [Cannabis sativa]</a>		64.2	64.2	96%	2e-20	44.19%	87	Query_37089
<input type="checkbox"/>	<a href="#">XP_030480212.1 metallothionein-like protein type 2 [Cannabis sativa]</a>		24.8	47.3	55%	2e-04	33.33%	79	Query_37093
<input type="checkbox"/>	<a href="#">XP_030500840.1 uncharacterized protein LOC115716228 [Cannabis sativa]</a>		27.2	27.2	70%	6e-05	20.59%	1268	Query_37090

Sort by description

↓  
**tBlastN**

select all 2 sequences selected [GenBank](#) [Graphics](#)

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/>	<a href="#">PREDICTED: Cannabis sativa EC protein I/II-like (LOC115702869), mRNA</a>	<a href="#">Cannabis sativa</a>	131	131	100%	5e-43	100.00%	495	<a href="#">XM_030630319.1</a>
<input checked="" type="checkbox"/>	<a href="#">PREDICTED: Cannabis sativa transport and Golgi organization protein 2 homolog (LOC115702868...)</a>	<a href="#">Cannabis sativa</a>	88.6	137	96%	1e-26	100.00%	1471	<a href="#">XM_030630317.1</a>





## 四、引物设计

- 概念：引物设计是一小段单链DNA或RNA，在核酸合成反应时，作为每个多核苷酸链进行延伸的出发点而起作用的多核苷酸链。
- 原则：
  - 1、长度：15-30bp，其有效长度 $[L_n=2(G+C)+(A+T)]$ 一般不大于38，否则PCR的最适延伸温度会超过Taq酶的最佳作用温度(74度)，从而降低产物的特异性。
  - 2、G+C含量：应在40%-60%之间，PCR扩增中的复性温度一般较 $T_m$ 值低等于引物的 $T_m$ 值减去5-10度。引物长度小于20时，其 $T_m$ 恒等于 $4 \times (G+C) + 2 \times (A+T)$ 。
  - 3、碱基分布的随机性：应避免连续出现4个以上的单一碱基。尤其是不应在其3'端出现超过3个的连续G或C，否则会使引物在G+C富集序列区错误引发。



# 引物设计原则

- 4、引物自身：不能含有自身互补序列，否则会形成发夹样二级结构。
- 5、引物之间：两个引物之间不应有多于4个的互补或同源碱基，不然会形成引物二聚体，尤应避免3'端的互补重叠。
- 6、上下游引物的互补性：一个引物的3'末端序列不允许结合到另一个引物的任何位点上。
- 7、3'末端：如果可能的话，每个引物的3'末端碱基应为G或C。
- 8、引物应当超出限制性内切酶识别位点至少3个核苷酸。



# 引物设计

## Primer pair 8

	Sequence (5'→3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	GGAGGCCTGAAAATGAGTGAC	Plus	21	46	66	58.91	52.38	6.00	3.00
Reverse primer	AGTAGGCTTTGCCAACTCCAG	Minus	21	277	257	60.27	52.38	7.00	2.00
Product length	232								



# 参考文献

- [1]W. Braun,M. Vasak,A. H. Robbins,C. D. Stout,G. Wagner,J. H. R. Kagi,K. Wuthrich. Comparison of the NMR Solution Structure and the X-Ray Crystal Structure of Rat Metallothionein-2[J]. Proceedings of the National Academy of Sciences of the United States of America,1992,89(21).
- [2]Jianping Zhang,Yanni Qi,Limin Wang,Lili Wang,Xingchu Yan,Zhao Dang,Wenjuan Li,Wei Zhao,Xinwu Pei,Xuming Li,Min Liu,Meilian Tan,Lei Wang,Yan Long,Jing Wang,Xuwen Zhang,Zhanhai Dang,Hongkun Zheng,Touming Liu. Genomic Comparison and Population Diversity Analysis Provide Insights into the Domestication and Improvement of Flax[J]. iScience,2020,23(4).
- [3]Sílvia Atrian,Mercè Capdevila. Metallothionein-protein interactions[J]. BioMolecular Concepts,2013,4(2).
- [4]Dar S, Shuja R N, Shakoori A R.A synthetic cadmium metallothionein gene (PMCd1syn) of Paramecium species:expression, purification and characteristics of metallothionein protein[J]. Molecular Biology Reports, 2013, 40(2):983-997.
- [5]Leszczyszyn O I, Imam H T, Blindauer C A. Diversity and distribution of plant metallothioneins: a review of structure,properties and functions[J].Metalomics, 2013, 5(9):1146-1169.



# 我们的体会

- 生物信息学使我们在试验开展前对其结果进行简单的预测，试验结果也用生物信息学的方法进行验证，两者互为补充，有机结合。
- 正确并且适当的使用生物信息学能使科研的高效且成功。
- 本学期的生信课程为我们叩开生物信息学的大门，在以后的科研中更应该结合自己课题多学、多练、多思考。



# 特别感谢

- 感谢小组成员的共同努力。
- 感谢罗老师对我们汇报的耐性指导，以后在科研中遇到问题多多向您请教，感谢老师。

邮箱：[wangyuanchang132@163.com](mailto:wangyuanchang132@163.com)