基于硒蛋白序列的蛋白结构和功能研究

Sequence-Based Structural and Functional Study of Selenoproteins

报告人：贾国赓

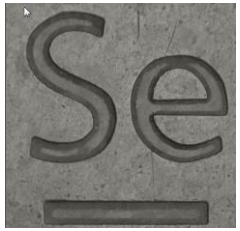
组长：刘东阳

组员：夏冰心，赵锐驰，郑文娟，张锡臣，郑浩楠

2020-01-04

硒元素的历史

Jöns Jacob Berzelius



初次发现

1817

毒性

1930s

必需微量元素

1950s

Chem. Soc. Rev. 2013, *42*, 8870
https://en.wikipedia.org/wiki/Selenium

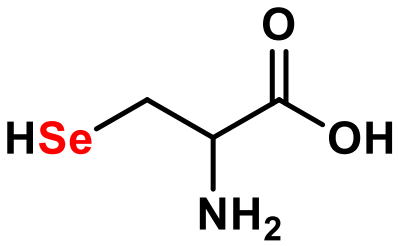
保健品

富硒茶

富硒米

富硒水果

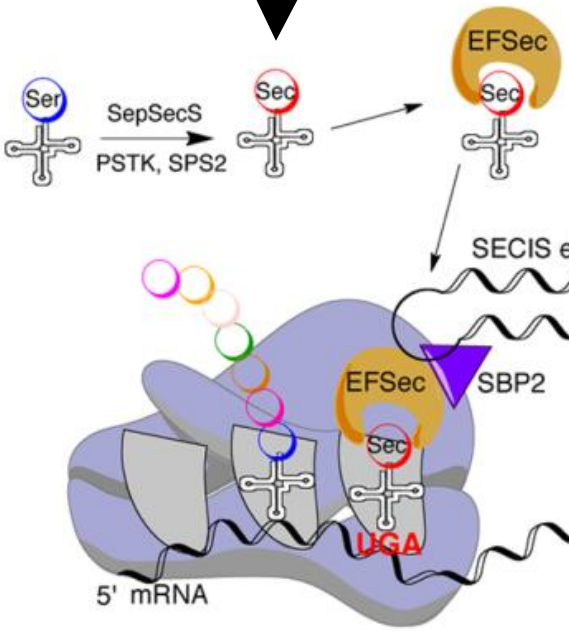
硒蛋白的生物功能



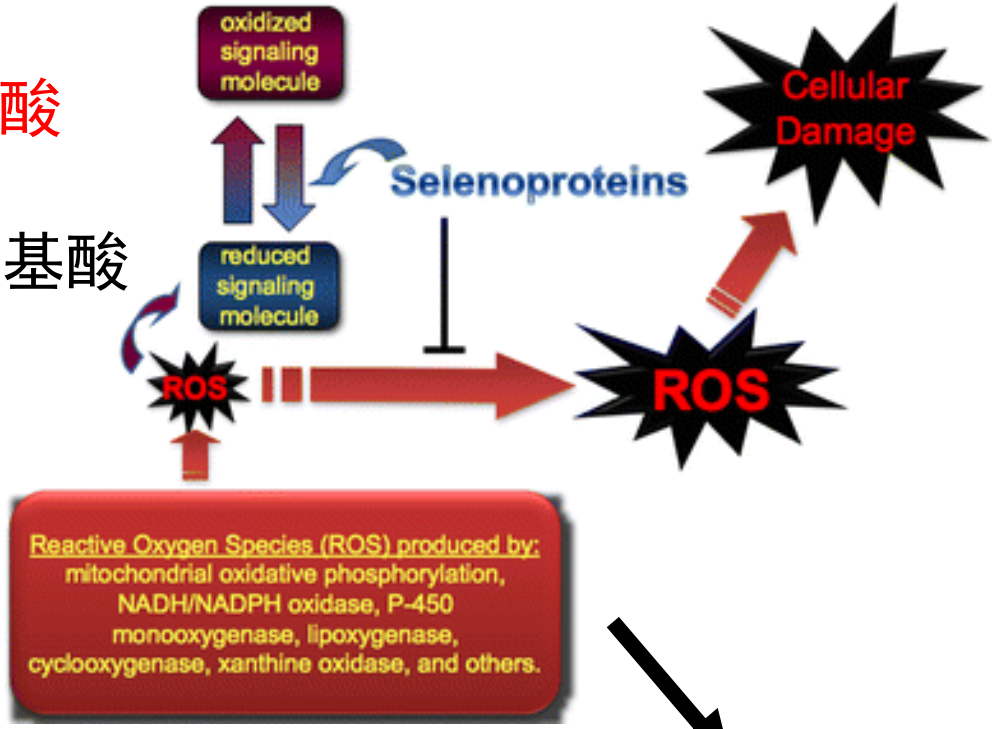
硒代半胱氨酸

Sec，U

第21种天然氨基酸



ACS Chem. Biol. 2016, *11*, 821-841

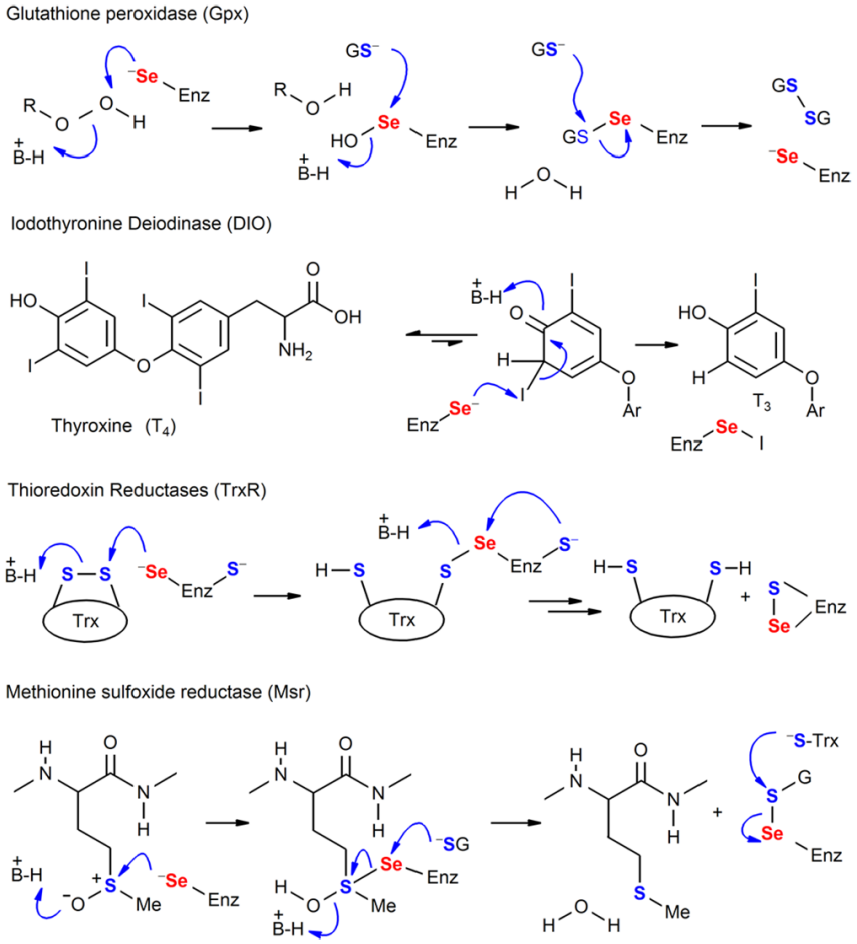


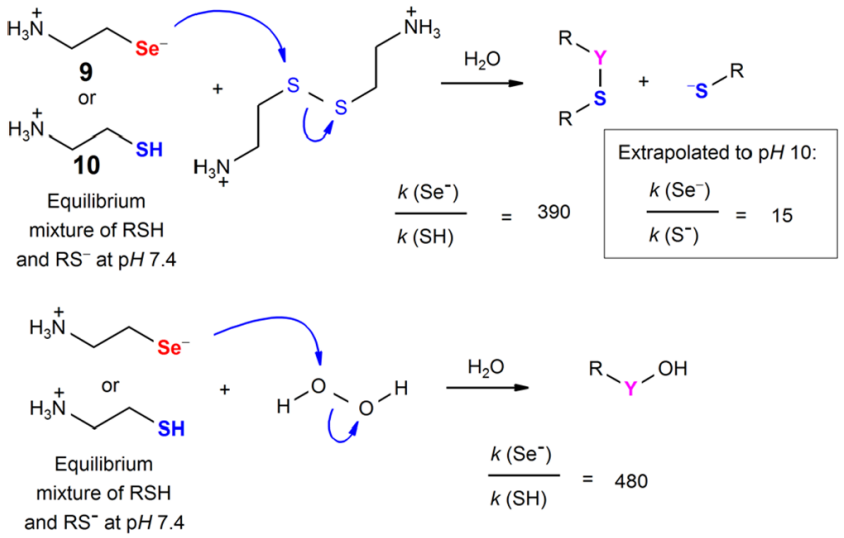
1. 癌症
2. 克山病
3. 大骨节病
4. 帕金森病
5. 糖尿病

Physiol Rev. 2014, *94*, 739-777

Antioxid. Redox Signaling 2007, *9*, 775-806

硒蛋白生物功能的化学基础



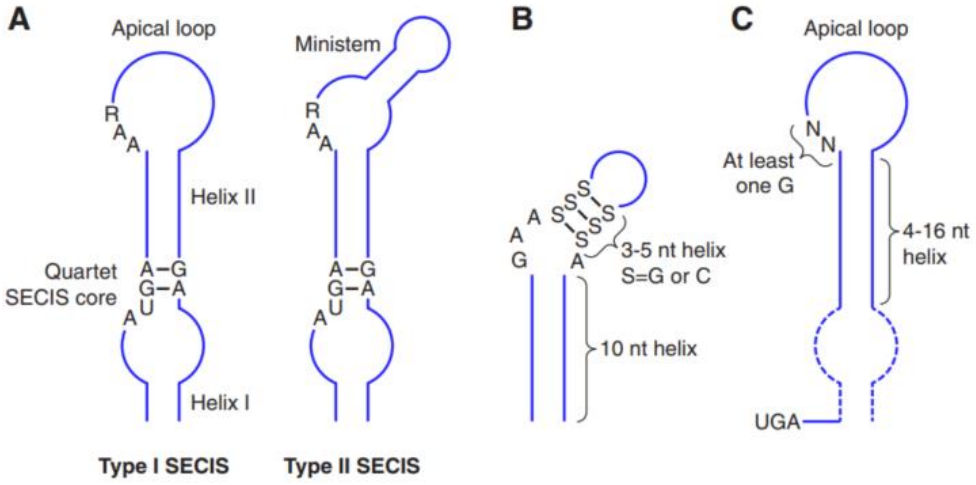


相对于半胱氨酸，硒代半胱氨酸还原性与酸性更强，在生理pH下更易电离使其亲核性较强，因此还原反应速率更高

ACS Chem. Biol. 2016, *11*, 821-841

基于生物信息学的硒蛋白鉴定方法

★ UGA密码子
★ 硒代半胱氨酸插入元件(SECIS)



SECIS候选

↓

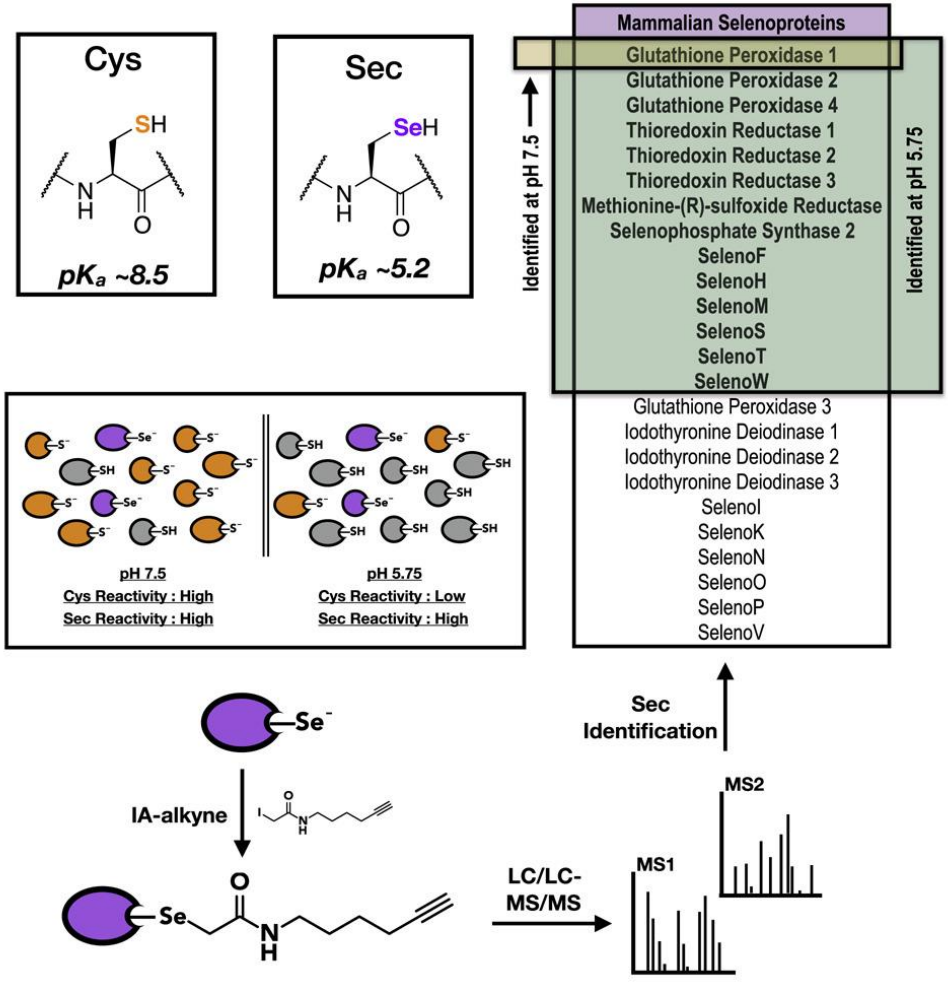
上游UGA密码子

↓

^{75}Se放射性代谢标记

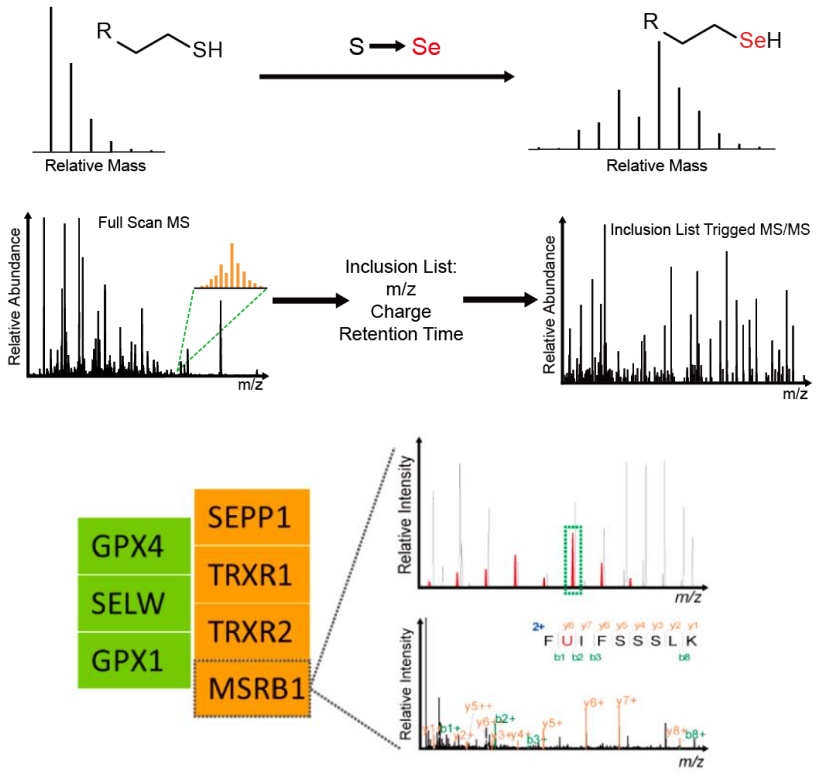
Science 2003, *300*, 1439-1443.
Physiol Rev. 2014, *94*, 739-777

检测硒蛋白的蛋白质组学方法



Cell Chem. Biol. 2018, *25*, 1157-1167

SESTAR: Selenium-Encoded isotopic Signature TARgeted profiling



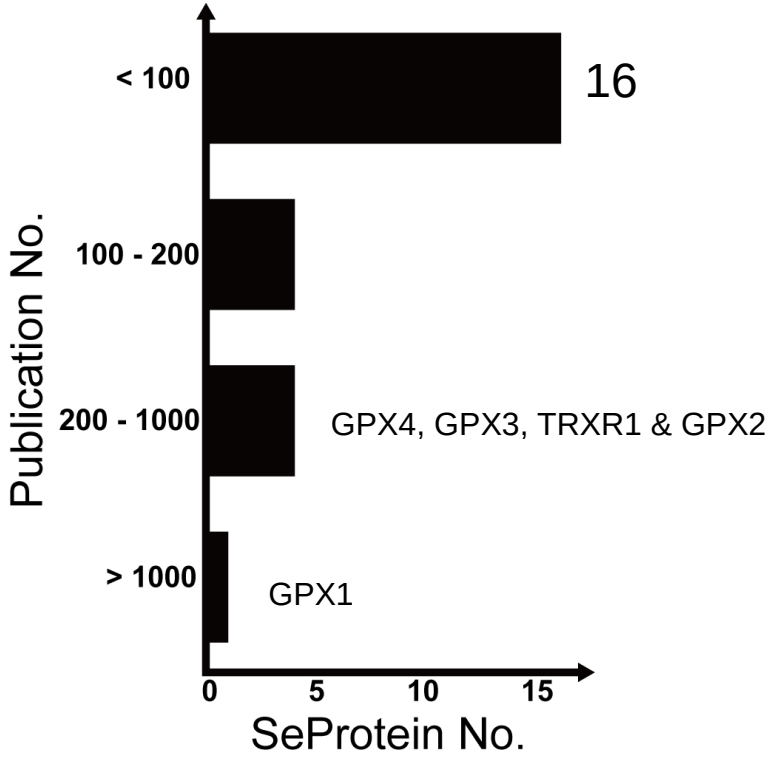
ACS Cent. Sci. 2018, *4*, 960-970

有限的硒蛋白研究

Selenoprotein	Abbreviation	Sec location (protein size)	Function
Glutathione peroxidase 1	GPx1, GPX1	47 (201)	Cytosolic glutathione peroxidase
Glutathione peroxidase 2	GPx2, GPX2	40 (190)	Gastrointestinal glutathione peroxidase
Glutathione peroxidase 3	GPx3, GPX3	73 (226)	Plasma glutathione peroxidase
Glutathione peroxidase 4	GPx4, GPX4	73 (197)	Phospholipid hydroperoxide glutathione peroxidase
Glutathione peroxidase 6	GPx6, GPX6	73 (221)	Olfactory glutathione peroxidase
Iodothyronine deodinase 1	DI1, D1, DIO1	126 (249)	Thyroid hormone-activating iodothyronine deodinase
Iodothyronine deodinase 2	DI2, D2, DIO2	133, 266 (273)	Tissue-specific thyroid hormone-activating iodothyronine deodinase
Iodothyronine deodinase 3	DI3, D3, DIO3	144 (278)	Tissue-specific thyroid hormone-deactivating iodothyronine deodinase
Thioredoxin reductase 1	TR1, TrxR1, TXNRD1	498 (499)	Reduction of cytosolic thioredoxin
Thioredoxin/glutathione reductase	TGR, TR2, TrxR3, TXNRD3	655 (656)	Testis-specific thioredoxin reductase
Thioredoxin reductase 3	TR3, TrxR2, TXNRD2	522 (523)	Reduction of mitochondrial thioredoxin and glutaredoxin
Methionine-R-sulfoxide reductase	MsrB1, SelR, SelX, MSRB1	95 (116)	Reduction of oxidized methionine residues
Selenophosphate synthetase 2	SPS2, SEPHS2	60 (448)	Involved in synthesis of selenoproteins
Selenoprotein W	SelW, SEPW1	13 (87)	Unknown
Selenoprotein T	SelT	36 (182)	Unknown
Selenoprotein H	SelH	38 (116)	Unknown
Selenoprotein V	SelV	273 (346)	Unknown
Selenoprotein I	SelI, SEPI, EPT1	387 (397)	Unknown
15 kDa selenoprotein	Sep15	93 (162)	Putative role in quality control of protein folding in the ER
Selenoprotein M	SelM, SEPM	48 (145)	Unknown
Selenoprotein K	SelK	92 (94)	Putative role in ER-associated degradation
Selenoprotein S	SelS, SEPS1, VIMP	188 (189)	Putative role in ER-associated degradation
Selenoprotein O	SelO	667 (669)	Unknown
Selenoprotein N	SelN, SepN, SEPN1,	428 (556)	Putative role during muscle development
Selenoprotein P	SelP, SEPP1	59, 300, 318, 330, 345, 352, 367, 369, 376, 378 (381)	Se transport

Physiol Rev. 2014, *94*, 739-777

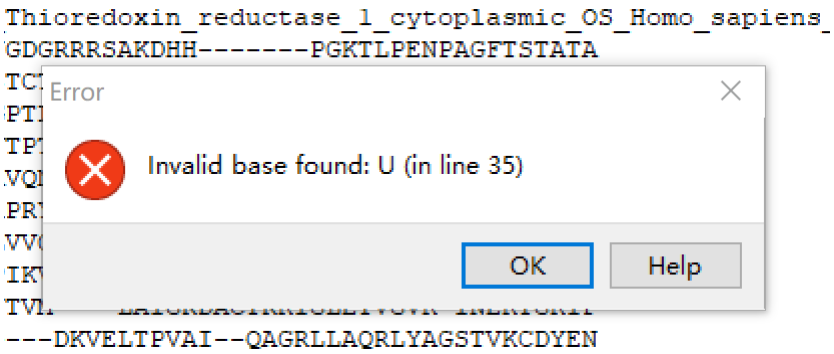
Pubmed搜索标题与摘要内的蛋白质名称得到的结果数目



硒蛋白含量极少→研究难度大

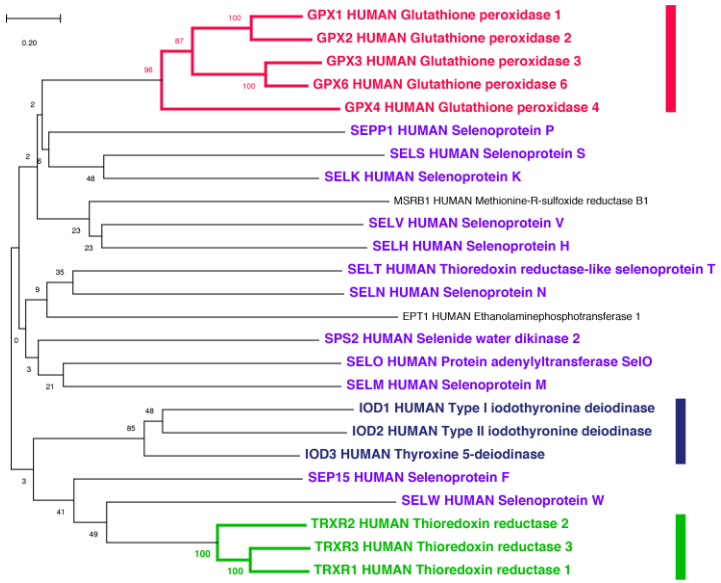
尝试使用生信方法从序列分析功能，重点在于如何处理常见工具不支持的硒代半胱氨酸

人类硒蛋白的系统发生树分析

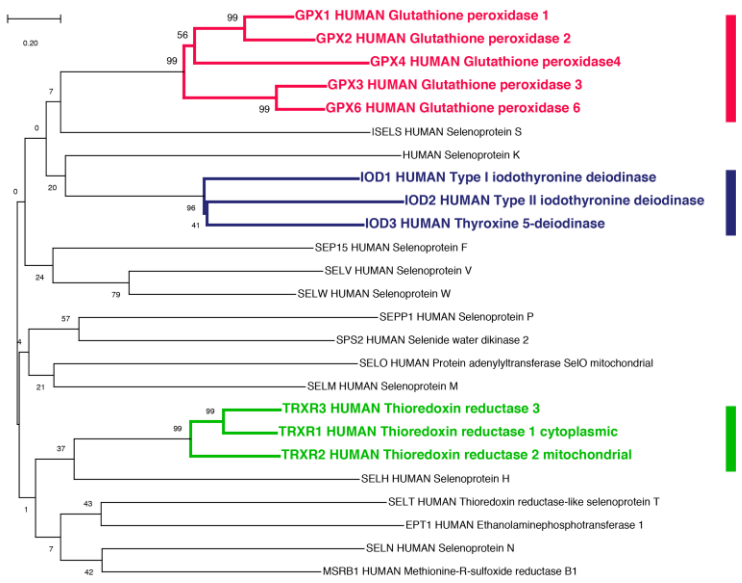


MEGA7对硒蛋白建树时会报错，提示不支持U字符，以前解决方法为将U替换为C。新版MEGAX使用muscle进行比对会将U替换成“?”，可以顺利建树。

U替换C建树结果



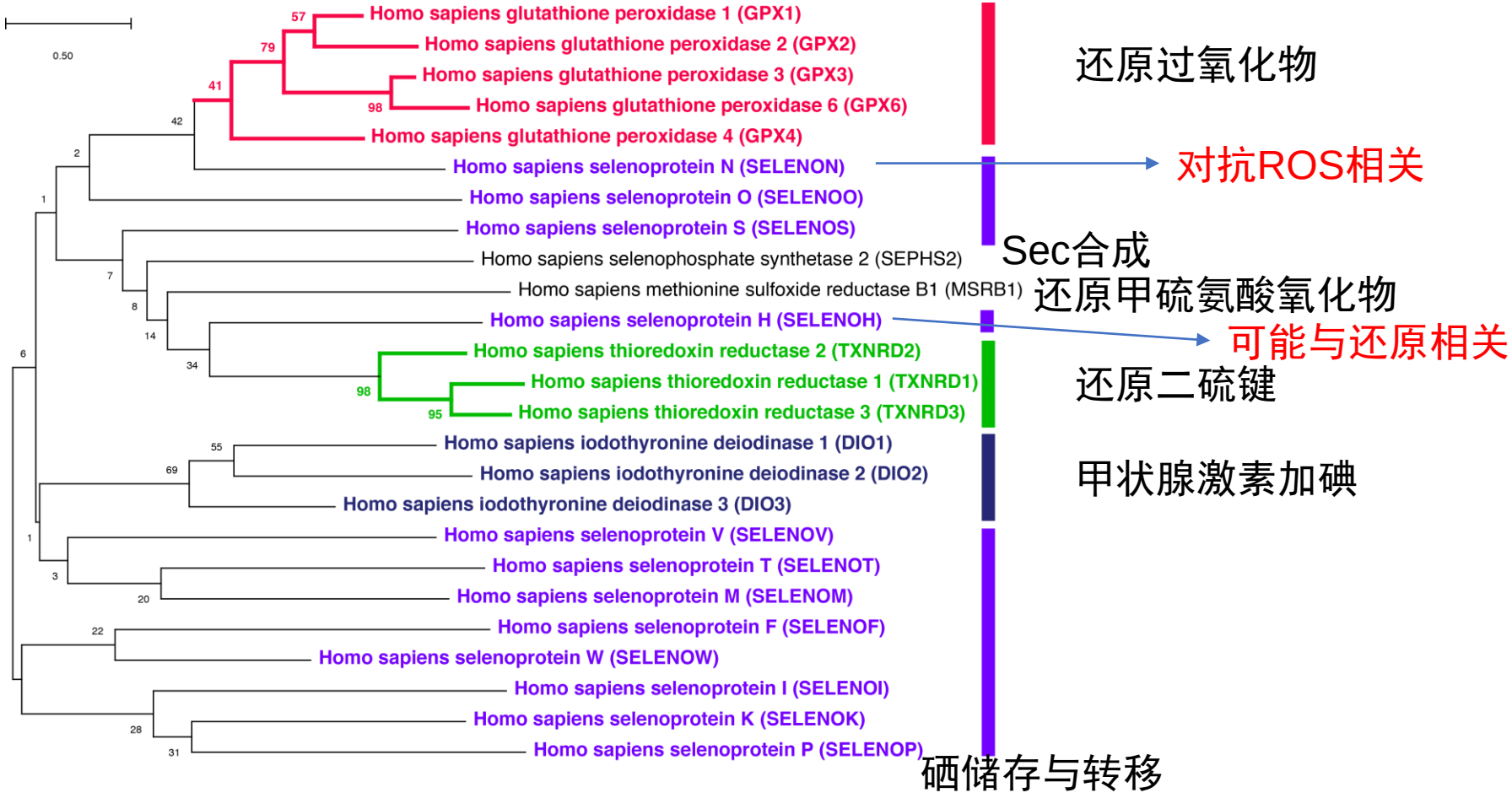
无替换(U自动变为?)建树结果



但无论哪种方法都使序列丧失了硒代半胱氨酸这一关键位点，因此尝试用基因建树

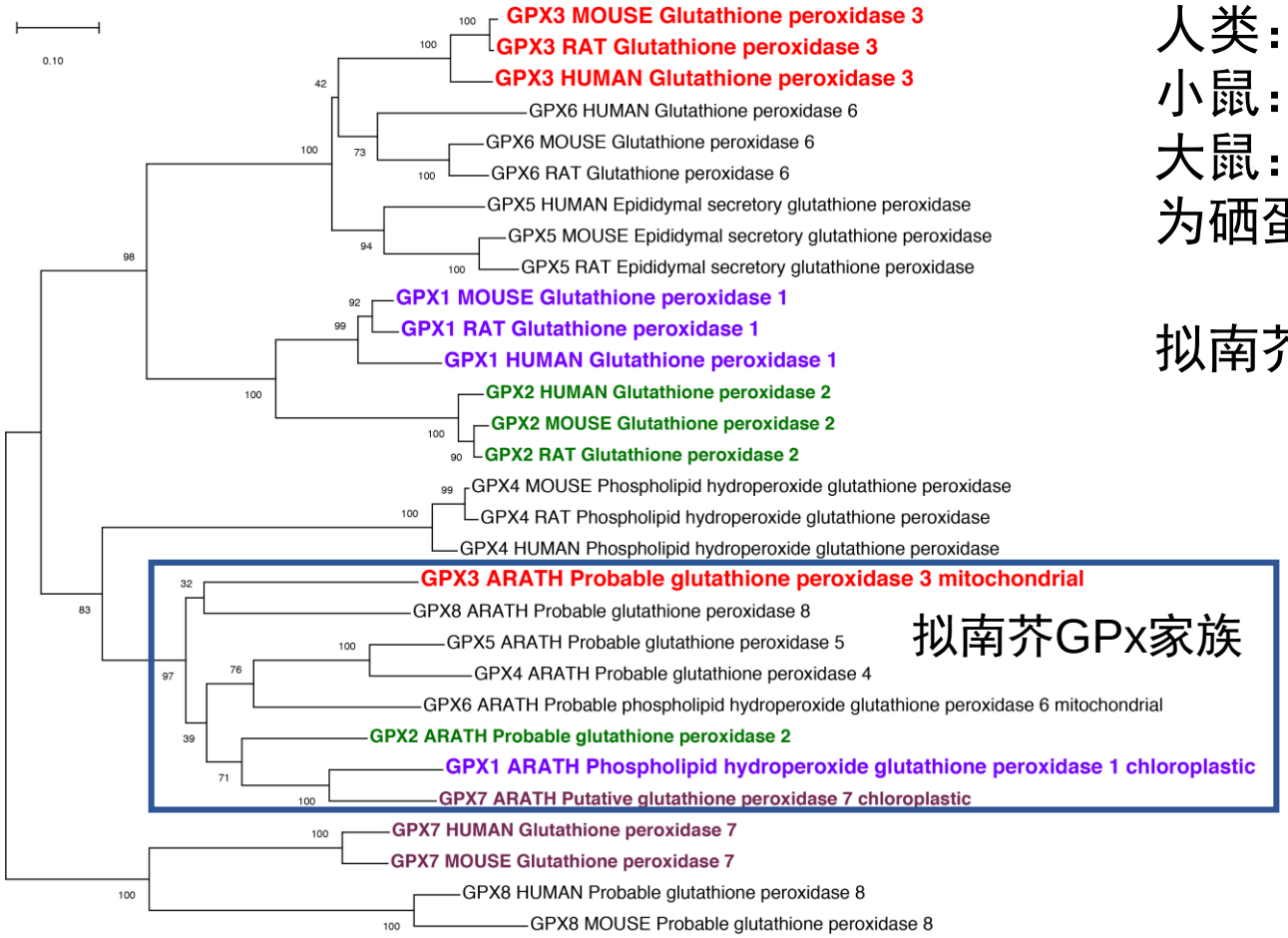
人类硒蛋白的系统发生树分析

从NCBI数据库上下载25个人类硒蛋白对应的蛋白编码序列后构建进化树（邻接法，自举500次）



大部分未知功能的硒蛋白聚到了一起，很难由此推测其功能

不同物种GPx家族蛋白的系统发生树分析



人类：GPx1，2，3，4，6
小鼠：GPx1，2，3，4
大鼠：GPx1，2，3，4
为硒蛋白

拟南芥无硒蛋白

硒蛋白未完全聚为一支

拟南芥的GPx家族蛋白与人类、小鼠、大鼠差别很大。
可能源于是否含硒的差别，或是动物与植物的差别。
需要找另几种含硒的植物加入进行进一步分析。

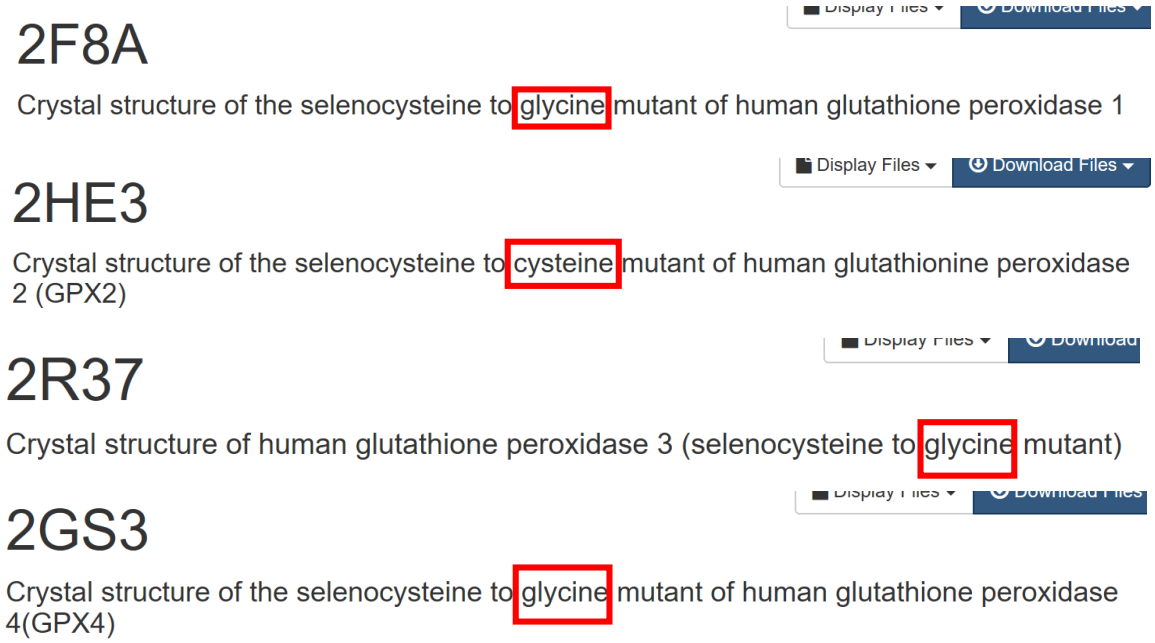
有限的硒蛋白结构数据

Entry	Entry name		Protein names	Gene names
Q86VQ6	TRXR3_HUMAN		**Thioredoxin reductase 3**	**TXNRD3** TGR, TRXR3
Q16881	TRXR1_HUMAN		**Thioredoxin reductase 1, cytoplasmi...**	**TXNRD1** GRIM12, KDRF
Q9NNW7	TRXR2_HUMAN		**Thioredoxin reductase 2, mitochondr...**	**TXNRD2** KIAA1652, TRXR2
Q9BQE4	SELS_HUMAN		**Selenoprotein S**	**SELENOS** SELS, VIMP, AD-015, SBBI8
Q9Y6D0	SELK_HUMAN		**Selenoprotein K**	**SELENOK** SELK, HSPC030, HSPC297
P07203	GPX1_HUMAN		**Glutathione peroxidase 1**	**GPX1**
P22352	GPX3_HUMAN		**Glutathione peroxidase 3**	**GPX3** GPXP
P36969	GPX4_HUMAN		**Phospholipid hydroperoxide glutathi...**	**GPX4**
P18283	GPX2_HUMAN		**Glutathione peroxidase 2**	**GPX2**
Q9NZV6	MSRB1_HUMAN		**Methionine-R-sulfoxide reductase B1**	**MSRB1** SEPX1, HSPC270

硒蛋白有结构数据的很少，尝试用生信方法通过序列预测其它硒蛋白结构，补充对硒蛋白的认识

GPx家族中的硒蛋白仅有GPx6没有结构，适合用来做同源结构预测

硒代半胱氨酸在蛋白结构数据中的缺失

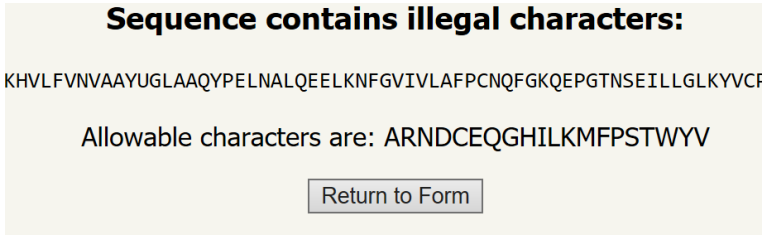


PDB中硒蛋白结构数据中均不含硒代半胱氨酸，多数采用将硒代半胱氨酸突变成Gly、Cys、Ser的方式测定

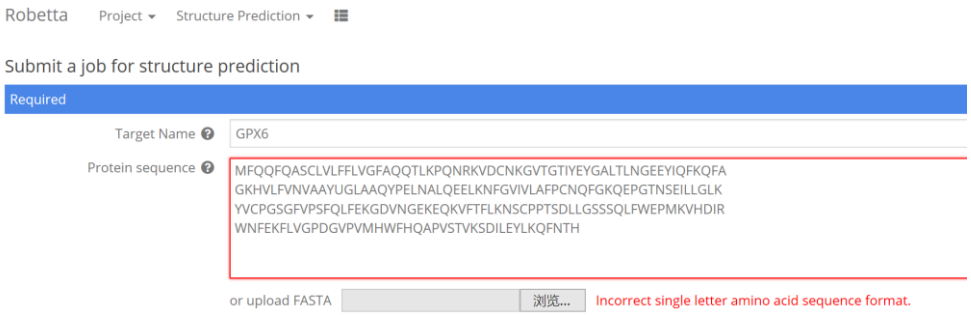
可能源于硒代半胱氨酸还原性太强，结晶后即被氧化，或因为硒代半胱氨酸独特的插入方式不便于构建质粒进行表达以得到足量蛋白进行提纯

蛋白质建模工具不支持硒代半胱氨酸

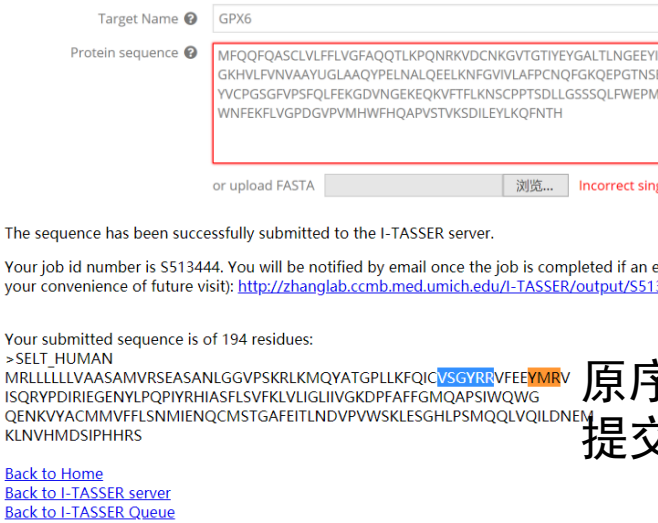
Phyre2报错



Rosetta(Robetta公共服务器)报错



I-TASSER直接删除序列中的U

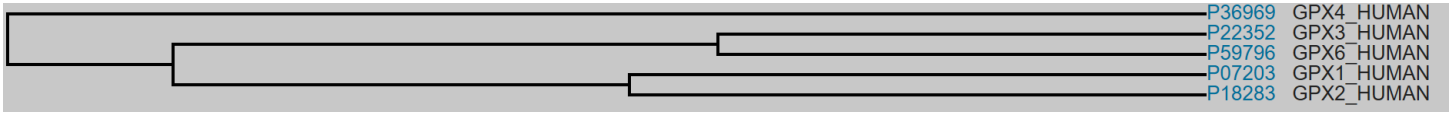


原序列：…VSUG…
提交后：…VSG…

解决方法：仿照结构测定，将U替换为G、C、S，预测突变体结构

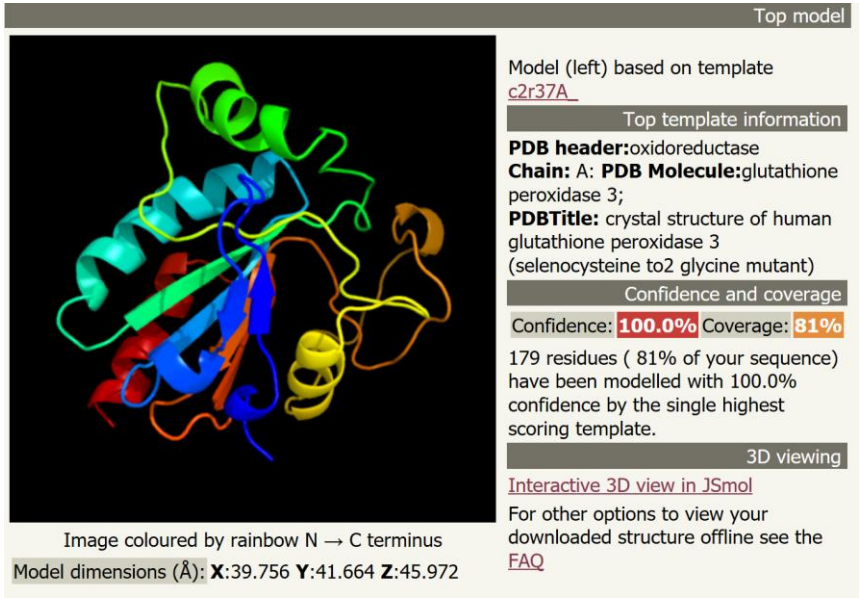
根本解决方法：Rosetta支持添加非常规氨基酸，可自行通过分子建模、量化计算等方法向库中添加Sec，从而预测结构

GPx6突变体结构预测结果



GPx6与GPx3序列相似度最高，因此选GPx3作为Rosetta结构预测模板

Phyre2



Rosetta

GPx6：绿色，GPx3：蓝色

GPx6突变体结构预测结果比较

RMSD = 0.395

Rosetta：绿色，包含1-221 所有氨基酸

Phyre2：红色，仅包含38-217位氨基酸

预测出结构的部分基本一致，但Rosetta能预测出更完整的结构。而Phyre2则无需事先指定模板。

可能可以用Phyre2找到模板，再用Rosetta构建完整结构。如果在Rosetta中加入了硒代半胱氨酸的数据可做到真正的结构预测

致谢与小组分工

感谢罗老师一学期以来的教导

感谢大组成员的帮助与支持

感谢在座各位的参与

刘东阳G21A：结构预测
夏冰心G21B：结构预测
赵锐驰G21C：结构预测
贾国赓G21D：背景介绍及结构预测
郑文娟G22A：系统发生树
张锡臣G22B：系统发生树
郑浩楠G22C：系统发生树

硒与人体健康及疾病的关系

★ 硒在人体中的功能

甲状腺激素代谢

抗氧化防御系统与氧化代谢

免疫

★ 硒的两面性

缺乏：克山病、大骨节病等

过多：毒性

★ 与硒相关的疾病

癌症

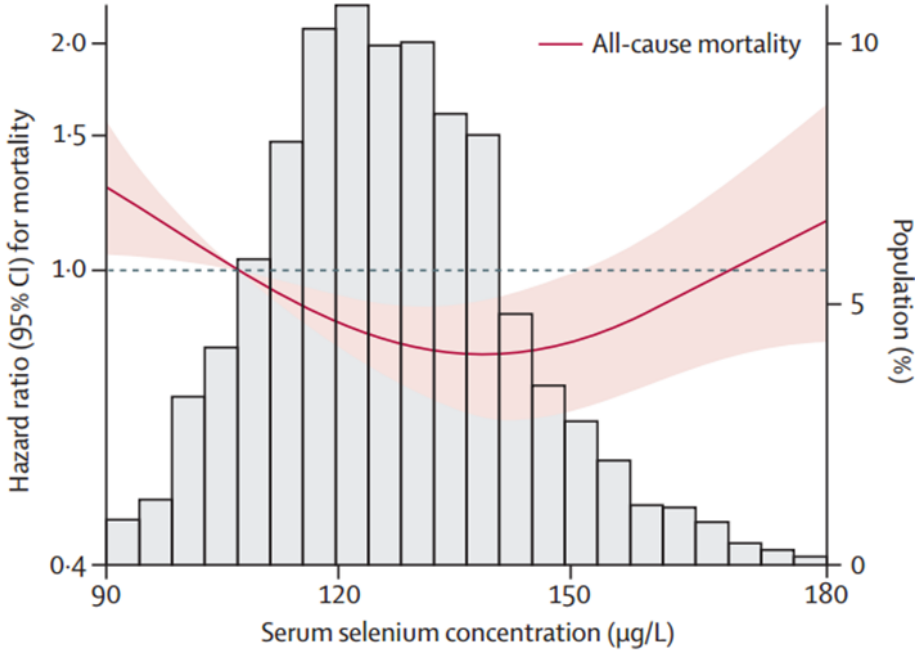
糖尿病

不育

炎症

神经退行性疾病

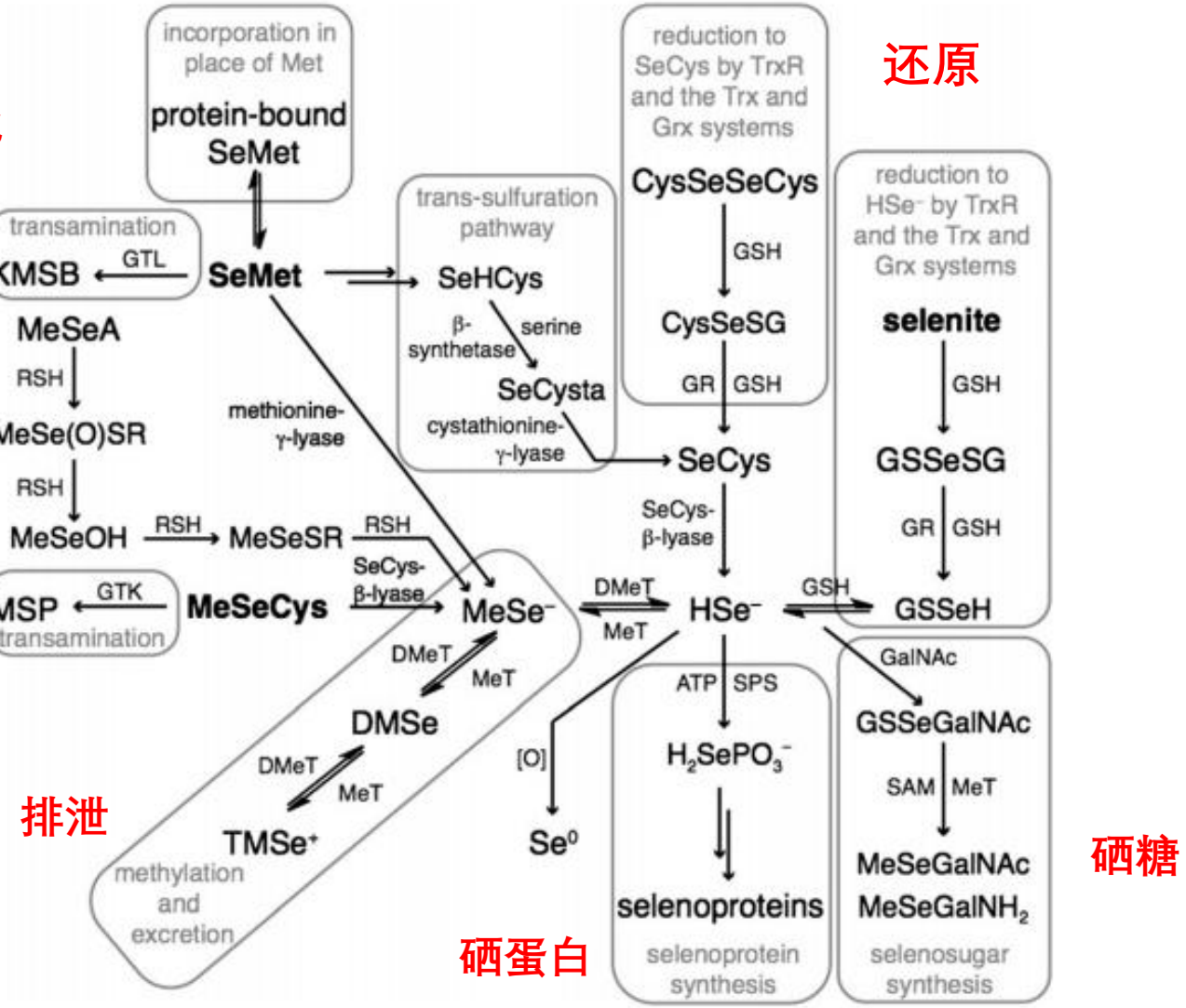
心血管疾病



Lancet 2012, *379*, 1256

Antioxid. Redox Signaling 2011, *14*, 1338

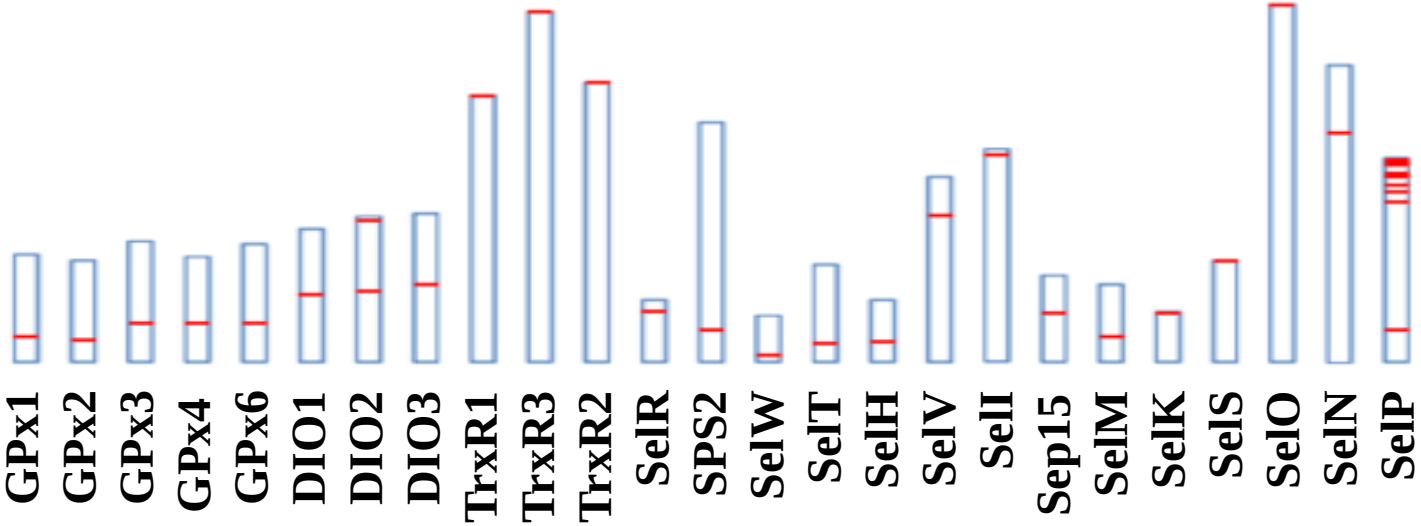
硒的代谢通路



Chem. Soc. Rev., 2013, *42*, 8870

人类硒蛋白质组

抗氧化	还原	甲状腺激素代谢	Sec合成	硒的储存与转运	蛋白质折叠（推测）	未知
GPx1 GPx2 GPx3 GPx4 GPx6 SelK SelR SelW	TrxR1 TrxR2 TrxR3	DIO1 DIO2 DIO3	SPS2	SelP	Sep15 SelM SelS	SelH SelI SelO SelT SelV



Physiol Rev. 2014, *94*, 739-777
Antioxid Redox Signal. 2007, *9*, 775-806