

哺乳动物 *Gpx* 基因家族进化选择和功能分歧研究*

杨红文^{1,2}, 金丹¹

(1. 华中农业大学 生命科学与技术学院, 武汉 430070; 2. 福建漳州师范学院 生物系, 福建 漳州 363000)

摘要: 研究哺乳动物谷胱甘肽过氧化物酶基因家族(*Gpx*)的亲缘关系和进化选择压力性质。从NCBI等网站获取哺乳动物*Gpx*基因编码区核酸序列和GPX蛋白氨基酸序列,并用MEGA、Clustal构建亲缘关系树,用SignalP和TargetP分析信号肽和亚细胞定位信息,用K estimator和PAML分析其*Gpx*基因家族进化先后顺序和选择压力;用Diverge分析GPX蛋白家族功能分歧。结果表明,GPX蛋白家族进化中形成两个亚类,其一包括GPX1和GPX2,其二包括GPX3、GPX5和GPX6,各GPX蛋白成员均有高度保守的GPX蛋白家族基序和高度保守的硒代半胱氨酸或半胱氨酸残基,氨基端信号肽分析显示,GPX1蛋白为胞浆线粒体型,其余家族成员均为胞外分泌型。哺乳动物*Gpx1*基因进化上受到严格的负选择作用,*Gpx*基因家族成员保守区也受到负选择作用,但用枝位点模型却检出了正选择作用,GPX蛋白家族内存在轻微的功能分歧。哺乳动物GPX蛋白家族的进化上的负选择作用表明其在哺乳动物清除自由基抗氧化方面的关键作用。

关键词: *Gpx* 基因; 哺乳动物; 结构域; 进化分析; 功能分歧

中图分类号: Q951.3

文献标志码: A

文章编号: 1672-6693(2012)02-0030-06

动物细胞在氧化磷酸化产能代谢中产生了大量高反应活性的超氧自由基和过氧化物,如不及时清除会攻击细胞中膜结构、酶和DNA,加重动脉粥样硬化及诱发恶性肿瘤。机体清除自由基和过氧化物的系统包括维生素E、维生素C、谷胱甘肽GSH等非酶还原性成分及超氧化物歧化酶(SOD)、过氧化氢酶(CAT)、谷胱甘肽过氧化物酶(GPX)等酶防御系统,其中因部分家族成员含稀有的硒代半胱氨酸(Sec)而著称的GPX蛋白家族在清除细胞内生物膜中的磷脂氢过氧化物和过氧化氢方面起主要作用^[1]。从小鼠体内分别鉴定出8个GPX蛋白家族成员,其它哺乳动物中也陆续发现GPX蛋白成员的存在。依效应部位及氨基酸序列不同,GPX蛋白家族大致分为4类:胞浆型(cGPX)、胃肠型(giGPX)、血浆型(pGPX)和磷脂氢谷胱甘肽过氧化物酶(phGPX),除磷脂氢型为单体外,其它家族成员均含4个亚基,不同类型的家族成员在氨基酸序列和催化机制上也各自保守^[2]。

GPX蛋白家族广泛存在于从微生物、植物、无脊椎生物、海洋生物直到高等哺乳动物的全生命界,

在植物的胁迫反应和动物的清除自由基抗氧化方面发挥关键作用,随着大量生物*Gpx*基因序列的破译,对其生物信息学的分析成为可能,寄生性吸虫和拟南芥中*Gpx*基因生物信息学已陆续开展^[3-4],哺乳动物的相关研究则还未见报道。本文尝试对常见哺乳动物*Gpx*基因家族的结构域、进化先后顺序、进化选择压力和蛋白质功能分歧加以分析,以期对哺乳动物GPX蛋白家族的深入研究有所裨益。

1 材料和方法

1.1 基因序列获取

在NCBI网站通过关键词“*Gpx*”搜索各数据库中的人(*Homo sapiens*)、小鼠(*Mus musculus*)等哺乳动物的*Gpx*基因家族成员,并通过获得的Gene ID的网页获取各基因的CDS和氨基酸序列,同一基因有差异剪接体的选取其序列较完整者。由于大多数哺乳动物*Gpx*基因家族成员未完全破译,主要选取了人和小鼠的*Gpx1*~*Gpx8*成员及搜索到的常见哺乳动物的*Gpx1*基因,同时选用青蟹(*Scylla paramamosain*)*Gpx1*基因、人和小鼠的*Klf1*基因作为构建亲缘

* 收稿日期 2011-10-29 修回日期 2011-11-24 网络出版时间 2012-03-14 19:27:00

资助项目:福建省青年创新基金(No. 2008F3111)福建省教学质量与教学改革工程项目(闽教高[2011]69号)

作者简介:杨红文,男,讲师,博士,研究方向为动物功能基因组学。

网络出版地址: http://www.cnki.net/kcms/detail/50.1165.N.20120314.1927.201202.30_006.html

关系树的外枝。选用的各基因 mRNA 收录号如下, 人(Hu)*Gpx* 基因家族 :*Gpx1*(NM_000581. 2)、*Gpx2*(NM_002083. 2)、*Gpx3*(NM_002084. 3)、*Gpx4*(NM_002085. 3)、*Gpx5*(NM_001509. 2)、*Gpx6*(NM_182701. 1)、*Gpx7*(NM_015696. 4)、*Gpx8*(NM_001008397. 2); 小鼠(Mu)*Gpx* 基因家族 :*Gpx1*(NM_008160. 6)、*Gpx2*(NM_030677. 2)、*Gpx3*(NM_001083929. 1)、*Gpx4*(NM_001037741. 2)、*Gpx5*(NM_010343. 2)、*Gpx6*(NM_145451. 3)、*Gpx7*(NM_024198. 3)、*Gpx8*(NM_027127. 2); 其它哺乳动物 *Gpx1* 基因 : 大鼠(Rt , *Rattus norvegicus*)(NM_030826. 3)、狗(Do , *Canis lupus familiaris*)(NM_001115119. 1)、黑猩猩(Chp , *Pan troglodytes*)(NM_001077512. 2)、马(Hr , *Equus caballus*)(NM_001166479. 1)、恒河猴(Mc , *Macaca mulatta*)(NM_001159298)、牛(Ox , *Bos tarurs*)(NM_174076. 3)、狨猴(Cj , *Callithrix jacchus*)(AY966403. 1)、兔(Rb , *Oryctolagus cuniculus*)(NM_001085444. 1)、猪(Pg , *Sus scrofa*)(NM_214201. 1) ; 亲缘树外围枝基因 : 青蟹(Cb)*Gpx1*(JN565286. 1)、人(Hu)*Klfl1*(NM_006563. 3)、小鼠(Mu)*Klfl1*(NM_010635. 2)。

1.2 亲缘树构建及氨基酸序列分析

哺乳动物 GPX 蛋白氨基酸序列经 Clustal X 2.1 比对后用 MEGA 5.0 中的 NJ(Neighbor-joining)法构建亲缘关系树, 亲缘树的可靠性用 MEGA 5.0 中的陷阱法(Bootstrap, 重复数 $n = 1\ 000$)来验证。GPX 蛋白成员信号肽的有无用在线软件 SignalP 4.0 分析, 亚细胞定位用在线软件 TargetP 1.1 分析。

1.3 进化顺序和选择压力分析

各 *Gpx* 基因核酸序列经 Clustal X 2.1 比对并经 Bioedit 去除起始密码子、终止密码子和内部特征性的硒代半胱氨酸密码子(同终止密码子, 应去除)后用 K estimator 6.0 计算两两比较的同义突变率 $K_s^{[5]}$, 该值用来表示基因的分化距离, 进一步确定 *Gpx1* 基因所表现的各哺乳动物分化先后顺序和 *Gpx* 家族成员分化先后顺序(以人为例)。 *Gpx1* 直系同源基因和人 *Gpx* 基因家族成员进化上的选择压力用 PAML 4.4 分析^[6], 其中的枝模型 M0 和 M3 的差异显著程度可以确定各基因是否以同一速率进化, 位点模型的 M1 和 M2 之间以及 M7 和 M8 之间的差异程度用来确定各基因进化上是否存在正选择作用, 其中 M1 和 M2 的比较结果更为可靠, 并用软件中贝叶斯模型(BEB)确定正选择位点, 而枝-位点模型

MA 和 MA1 的差异程度则用来进一步确定前景枝相比较于背景枝后正选择的存在及正选择位点的确认。

1.4 功能分歧

Diverge 2.0 采用 MEGA 5.0 构建的亲缘关系树来分析哺乳动物 GPX1 蛋白及人 GPX 蛋白家族成员氨基酸序列间的 I 型功能分歧, I 型功能分歧往往导致蛋白质功能的改变, 功能分歧系数 θ 值在 0 ~ 1 波动, θ 值越大则选取的不同亚类间的功能分歧就越大, 并用似然比检验来确定功能分歧的显著程度^[7]。但 θ 值小并不代表没有功能分歧位点, 可能是极少数功能分歧显著的位点被整体湮没了。

2 结果和分析

2.1 亲缘关系与结构域分析

进化亲缘关系树分析表明, 人、黑猩猩、狨猴和恒河猴 GPX1 蛋白进化距离最为接近(图 1A); 其次依次是狗和马、牛和猪、兔, 和灵长类亲缘关系最远的是啮齿类的小鼠和大鼠。这与哺乳动物的进化分类较为接近, 且陷阱法 Bootstrap 值多在 60% 以上, 可信度较高。图 1B 和图 1C 树形结构高度一致且 Bootstrap 值多在 70% 以上, 表明以人和小鼠为例的 GPX 蛋白家族成员进化关系树较为可信; 其中 GPX3、GPX5 和 GPX6 构成第一亚类, GPX1 和 GPX2 亲缘关系也高度接近, 而 GPX4、GPX7 和 GPX8 与前两类亲缘关系则较远, 且其内部亲缘关系也较为松散。这与 Margis 等^[8]对动物 GPX 蛋白家族的进化分类高度一致。

人 GPX 蛋白家族成员氨基酸序列比对结果与亲缘关系树基本一致(图 2), 人 GPX 蛋白家族所有成员均有 GPX 蛋白的 3 个特征保守序列, 如图 2A、B、C 框所示。其中 A 框(NVA SXU/CG, X 指非保守可变氨基酸)是 GPX 蛋白的催化活性区, 当中含硒代半胱氨酸 U 的 GPX1、GPX2、GPX3、GPX4 和 GPX6 为硒-依赖谷胱甘肽过氧化物酶(sGPX), 取代以半胱氨酸的 GPX5、GPX7 和 GPX8 为硒-非依赖类型(siGPX); B 框(FPCNQFG)为 GPX 蛋白家族的标志性基序, C 框中 GPX4 所特有的序列 KWNFTKLF 表明其属于磷脂氢谷胱甘肽过氧化物酶(LHGPX); 此外, B 框和 C 框中还有另外两个标志性保守氨基酸 Q 和 W(图 2 箭头所指)^[4], GPX1、GPX2、GPX3、GPX5、GPX6 等蛋白邻近的亲缘关系也可由它们共有的保守序列佐证(图 2 中 I 框和 II 框)。GPX 蛋

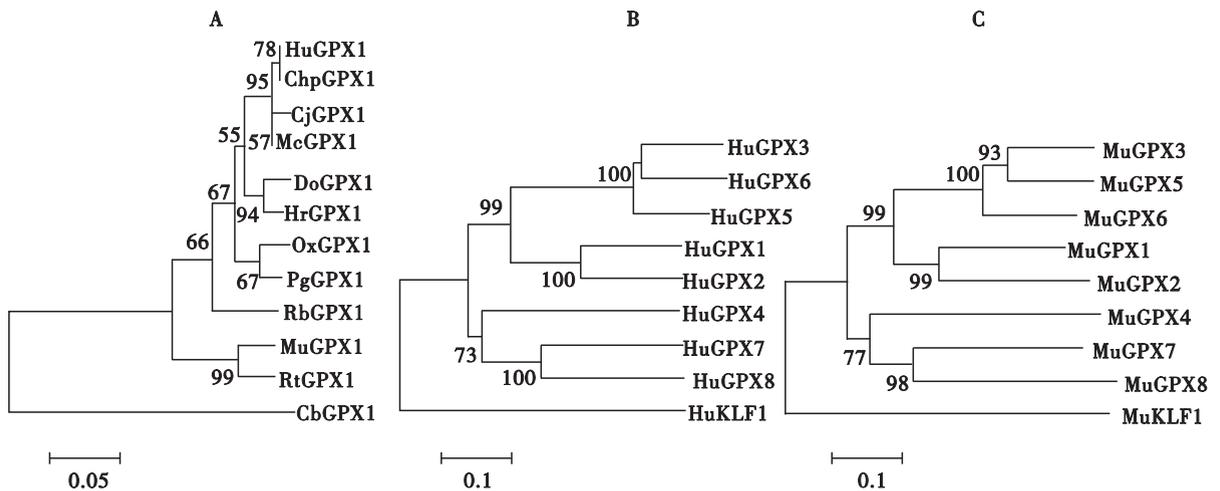
白大多分泌到细胞外或进入线粒体发挥清除自由基抗氧化的作用。SignalP 4.0 分析表明,人 GPX3、GPX4、GPX5、GPX6、GPX7 等蛋白的氨基端均有信号肽序列;进一步的 TargetP 1.1 分析表明,它们均为胞外分泌型。GPX1 蛋白定位为线粒体型,GPX8 蛋白虽无明显的信号肽序列,但亚细胞定位分析表明其可能也是分泌到细胞外发挥作用。

2.2 进化分析

哺乳动物同源基因同义突变和中性突变往往以相近速率进化^[9]。K estimator 6.0 计算的直系同源基因间或基因家族成员间突变未饱和 CDS 的 K_s 值可代表同源基因大致的分化时间;而从直系同源基因两两比较的 K_s 值大小能推导出各哺乳动物大致的分化时间,同样也可由基因家族成员间两两比较的 K_s 值大小推出各成员大致的分化时间和先后顺序。从哺乳动物 *Gpx1* 基因的 K_s 值可以看出,4 种灵长类动物即人、黑猩猩、恒河猴和狨猴亲缘关系非

常接近($K_s < 0.2$);牛和猪、牛和马、狗和马、大鼠和小鼠亲缘关系也都分别较为接近($K_s < 0.3$)。这与哺乳动物实际分类基本一致。以黑猩猩为例,几种哺乳动物分化的先后顺序大致为:啮齿类和牛、猪几乎同时分别分化出去,其次狗和马分离出去,然后是兔分化出去,最后是狨猴、恒河猴、人依次分化出去(表 1)。

表 2 中有效 K_s 值表明,人 *Gpx1* 和 *Gpx2* 基因($K_s = 0.987\ 90$)、*Gpx3*、*Gpx5* 和 *Gpx6* 基因($K_s < 0.8$)分别构成 *Gpx* 基因家族进化亲缘关系密切的两个亚类,与以氨基酸序列为基础的亲缘关系树的结果基本一致,以 *Gpx1* 基因为例,人 *Gpx* 基因家族分化的先后顺序为:*Gpx8* 先分化出去,其次是 *Gpx4*、*Gpx7*,然后 *Gpx3*、*Gpx5* 和 *Gpx6* 分离,最后是 *Gpx1* 和 *Gpx2* 的分离。表 2 中的 K_s 值大多较大,提示 *Gpx* 基因家族是一个哺乳动物分化前就已经形成的古老的基因家族。



注:A 为哺乳动物 GPX1 蛋白亲缘关系树,B、C 分别为人、小鼠的 GPX 蛋白家族亲缘关系树。

图 1 哺乳动物 GPX1 和 GPX 蛋白家族亲缘关系树

Fig. 1 Phylogenetic trees of mammalian GPX1 and GPX protein family

PAML 中的同一模型 M0 和离散模型 M3 比较分析表明(表 3),哺乳动物 *Gpx1* 基因和人 *Gpx* 基因家族成员各自群内都以不同速率进化($p < 0.01$);在位点模型中,各氨基酸位点的错义突变率(dN 或 Ka)和同义突变率(dS 或 Ks)比值 ω 在小于 1、等于 1 及大于 1 时分别代表该氨基酸位点经历了负选择(净化选择)、中性选择和正选择作用,表 3 表明哺乳动物 *Gpx1* 基因和人 *Gpx* 基因家族成员各自整体上都未经历正选择作用($p = 1$)。

位点模型从整体上分析正选择的存在,可能会湮没个别前景枝中存在的正选择作用,而枝-位点模型的 MA 和 MA1 比对则可以弥补,表 4 表明,哺乳动物 *Gpx1* 基因各前景枝仍未检出正选择,但兔、大鼠、小鼠前景枝却发现将近显著水平的正选择位点 143A(后验概率 $p = 0.883$);人 *Gpx* 基因家族中,4 个前景枝均检出正选择作用,和中性模型差异达到显著或极显著水平,并分别检出 3、1、5 和 1 个正选择位点(表 4)。

图 2 人 GPX 蛋白家族氨基酸序列多重比对

Fig. 2 Multiple amino acid alignment of human GPX protein family

表 1 哺乳动物 *Gpx1* 基因进化 *Ks* 值

Tab. 1 The *Ks* values of mammalian *Gpx1* genes

	人	恒河猴	狨猴	牛	猪	狗	马	兔	小鼠	大鼠
黑猩猩	0.010 2	0.093 9	0.188 5	0.472 7	0.532 8	0.383 5	0.377 8	0.320 2	0.424 0	0.525 7
人		0.094 3	0.189 5	0.493 2	0.555 3	0.384 4	0.378 6	0.321 0	0.421 1	0.521 6
恒河猴			0.191 4	0.481 4	0.510 2	0.408 9	0.329 6	0.321 1	0.464 9	0.523 1
狨猴				0.541 3	0.562 4	0.456 3	0.423 4	0.377 1	0.495 5	0.545 7
牛					0.274 8	0.392 5	0.285 9	0.378 6	0.593 0	0.639 9
猪						0.448 9	0.334 7	0.433 9	0.581 1	0.591 8
狗							0.266 4	0.327 5	0.559 2	0.654 3
马								0.317 2	0.456 8	0.508 1
兔									0.457 0	0.482 5
小鼠										0.145 2

表 2 人 *Gpx* 基因家族分化顺序相关的 *Ks* 值

Tab. 2 The *Ks* values of human *Gpx* gene family

	<i>Gpx2</i>	<i>Gpx3</i>	<i>Gpx5</i>	<i>Gpx6</i>	<i>Gpx4</i>	<i>Gpx7</i>	<i>Gpx8</i>
<i>Gpx1</i>	0.987 90	1.652 16	0.828 11(NA)	1.729 12	0.6258 3(NA)	1.261 33	0.976 78(NA)
<i>Gpx2</i>		1.363 32	0.778 68(NA)	1.358 64	1.547 95	1.567 48	0.766 87(NA)
<i>Gpx3</i>			0.750 90	0.623 08	0.746 48(NA)	1.616 62	0.777 09(NA)
<i>Gpx5</i>				0.477 12	0.796 69(NA)	0.784 81(NA)	0.945 80(NA)
<i>Gpx6</i>					0.822 42(NA)	1.854 22	0.797 29(NA)
<i>Gpx4</i>						0.761 93(NA)	0.847 35(NA)
<i>Gpx7</i>							1.992 10

注 :NA 表示非常用模型计算的无效数据。

表 3 *Gpx* 基因家族位点模型分析结果

Tab. 3 Results of likelihood ratio test of *Gpx* gene family for site model

模型	自由度(<i>df</i>)	哺乳动物 <i>Gpx1</i>		人 <i>Gpx</i> 基因家族	
		似然比检验值(<i>LRT</i>)	相似概率(<i>p</i>)	似然比检验值(<i>LRT</i>)	相似概率(<i>p</i>)
M3 vs M0	4	68.720 7	<0.01	101.785 3	<0.01
M2 vs M1	2	0	1	0	1
M8 vs M7	2	2.669 4	>0.05	0.192 9	>0.05

表 4 *Gpx* 基因家族枝位点模型(MA vs MA1)似然比检验结果(*df* = 1)

Tab. 4 Results of likelihood ratio test of *Gpx* gene family for branch-site model(MA vs MA1 *df* = 1)

背景枝	前景枝	似然比检验值(<i>LRT</i>)	相似概率(<i>p</i>)	正选择位点及概率
哺乳动物 <i>Gpx1</i> 基因	<i>Gpx1</i> (人、黑猩猩、恒河猴和狨猴)	0	1	
	<i>Gpx1</i> (牛、猪、狗和马)	0	1	
	<i>Gpx1</i> (兔、大鼠和小鼠)	3.422 9	>0.05	143A(0.883)
人 <i>Gpx</i> 基因家族	<i>Gpx1</i>	8.631 0	<0.01	34Q(0.827) ,121K(0.819) ,153I(0.821)
	<i>Gpx2</i>	4.247 2	<0.05	148H(0.870)
	<i>Gpx3</i> 、 <i>Gpx5</i> 和 <i>Gpx6</i>	9.508 9	<0.01	60V* ,88E* ,109P* ,113K* ,147K**
	<i>Gpx4</i> 、 <i>Gpx7</i> 和 <i>Gpx8</i>	4.509 9	<0.05	154F*

注：* 表示 $p > 0.95$; ** 表示 $p > 0.99$ 。

2.3 蛋白质功能分歧

Diverge 2.0 以 MEGA 的亲缘关系树为基础计算不同亚类的蛋白质之间功能分歧程度,分别计算氨基酸位点间的 I 型功能分歧系数 θ_1 ,并求出所有参比氨基酸功能分歧系数的中位数代表亲缘关系树亚类间的蛋白质功能分歧程度。表 5 表明,哺乳动物直系同源基因 *Gpx1* 之间不存在功能分歧,提示

Gpx1 基因进化上的高度保守性;人 *Gpx* 基因家族中 *Gpx3*、*Gpx5* 和 *Gpx6* 这一亚类与 *Gpx4*、*Gpx7* 和 *Gpx8* 组成的亚类间存在显著的功能分歧($p < 0.05$)。I 型功能分歧系数 θ_1 (0.296 034)较小可能是中位数湮没了少数显著功能分歧位点,当后验概率阈值设为 0.70 时,检出 120、179 两个功能分歧氨基酸位点。

表 5 人 *Gpx* 基因家族和哺乳动物 *Gpx1* 的 I 型功能分歧

Tab. 5 Type I functional divergence of human *Gpx* family and mammalian *Gpx1*

类别	I 型功能分歧系数(θ_1)	似然比检验值(<i>LRT</i>)	差异显著检验(<i>p</i>)	位点后验概率阈值(Q_k)	功能分歧位点
<i>Gpx3</i> 、 <i>Gpx5</i> 和 <i>Gpx6</i> vs <i>Gpx4</i> 、 <i>Gpx7</i> 和 <i>Gpx8</i>	0.296 034	5.850 049	<0.05	0.70	120 ,179
<i>Gpx1</i> (牛和猪)vs <i>Gpx1</i> (兔、大鼠和小鼠)vs <i>Gpx1</i> (狗和马)	0	0	1		

3 讨论

Gpx 基因家族成员的亲缘关系和进化顺序表明, *Gpx8*、*Gpx4* 和 *Gpx7* 较早分化出来,而后两个近缘亚类即 *Gpx3*、*Gpx5*、和 *Gpx6* 组成的亚类与 *Gpx1* 和 *Gpx2* 组成的亚类依次分化出来,这与 Topppo 等^[10]研究的 *Gpx1* 基因是该家族最新成员的结果一致。作为 *Gpx* 基因家族新成员, *Gpx1* 基因广泛存在于从微生物、植物、无脊椎动物直到高等哺乳动物的全生命界。这表明,其它的家族老成员可能也存在于各生物界, *Gpx* 基因家族可能在物种分化前就已完成家族扩张和形成,是一个非常古老的关键基因家族,在

清除生物体内各种产能代谢产生的有害自由基和过氧化物方面发挥不可替代的主要作用,这从其家族进化上严格的负选择作用也可以看出。GPX 蛋白质家族成员依是否含硒代半胱氨酸可分为硒-依赖谷胱甘肽过氧化物酶(sGPX)和硒-非依赖型两类(siGPX),其中硒依赖型通过无机硒原子从还原性谷胱甘肽获得电子并还原过氧化物或清除各类自由基,同时谷胱甘肽由还原型被氧化为氧化型,再由体内谷胱甘肽还原酶系统催化还原并循环利用;硒-非依赖性 GPX 蛋白又称谷胱甘肽转硫酶(GST),主要通过转硫作用催化内外源的亲电子集团和还原性谷胱甘肽的巯基偶联以增加其疏水性,进而容易透过

细胞膜并分解排出体外,减轻对酶蛋白和 DNA 的氧化损伤²¹。

参考文献:

- [1] 王咏梅. 自由基与谷胱甘肽过氧化物酶[J]. 解放军药学报, 2005, 21(5) : 369-371.
- [2] 马森. 谷胱甘肽过氧化物酶和谷胱甘肽转硫酶研究进展[J]. 动物医学进展, 2008, 29(10) : 53-56.
- [3] 章莹, 蔡国斌, 何立, 等. 寄生性吸虫谷胱甘肽过氧化物酶生物信息学分析[J]. 中国人兽共患病学报, 2010, 26(4) : 293-298.
- [4] 周立敬, 周宜君, 高飞, 等. 拟南芥谷胱甘肽过氧化物酶的生物信息学分析[J]. 中央民族大学学报:自然科学版, 2010, 19(2) : 11-17.
- [5] Comeron J M. K-Estimator : calculation of the number of nucleotide substitutions per site and the confidence intervals

- [J]. Bioinformatics, 1999, 15(9) : 763-764.
- [6] Yang Z. PAML 4 : phylogenetic analysis by maximum likelihood[J]. Mol Biol Evol, 2007, 24(8) : 1586-1591.
- [7] Gu X. Statistical methods for testing functional divergence after gene duplication[J]. Mol Biol Evol, 1999, 16(12) : 1664-1674.
- [8] Margis R, Dunand C, Teixeira F K, et al. Glutathione peroxidase family-an evolutionary overview[J]. FEBS J, 2008, 275(15) : 3959-3970.
- [9] Kumar S, Subramanian S. Mutation rates in mammalian genomes[J]. Proc Natl Acad Sci USA, 2002, 99(2) : 803-808.
- [10] Toppo S, Vanin S, Bosello V, et al. Evolutionary and structural insights into the multifaceted glutathione peroxidase (Gpx) superfamily[J]. Antioxid Redox Signal, 2008, 10(9) : 1501-1514.

Animal Sciences

Evolutionary Selection and Functional Divergence of *Gpx* Gene Family in Mammals

YANG Hong-wen^{1, 2}, JIN Dan¹

(1. College of Life Science and Technology, Huazhong Agricultural University, Wuhan 430070 ;

2. Biology Department, Zhangzhou Normal University, Zhangzhou Fujian 363000, China)

Abstract : This paper studies the phylogenetic relationship and evolutionary selective pressure of *Gpx* gene family in mammals. And we collect CDS and protein sequence of *Gpx* genes from NCBI and construct phylogenetic tree with MEGA and Clustal, identify signal peptide and subcellular localization with SignalP and TargetP respectively, conduct evolutionary analysis with K estimator and PAML, analyze functional divergence with diverge. There are two subclasses, GPX1, GPX2 and GPX3, GPX 5, GPX 6 in the protein family; all members share highly conserved token amino acid sequence of GPX family, as well as the same highly conserved amino acid Sec or Cys. Analysis of the signal peptides in the N-termini indicates that GPX1 protein was mitochondrion-localized while all the other members were guided outside cells. Both human *Gpx* gene family and mammalian *Gpx1* genes have endured negative selection, however, positive selection was examined in human *Gpx* gene family when using the branch-site model in PAML and there are slight functional divergence in human GPX protein family. The high conservation and evolutionary purified selection of mammalian *Gpx* gene family imply their essential roles in scavenging free radicals.

Key words : *Gpx* gene ; mammals ; structural domain ; evolutionary selection ; functional divergence

(责任编辑 方 兴)