

缢蛏 *HAT* 基因家族的全基因组鉴定及在环境因子、细菌胁迫下的表达分析*

练佳莹¹ 吕丽媛² 姚韩韩¹ 董迎辉^{1, 2①} 林志华^{1, 2①}

(1. 浙江万里学院生物与环境学院 浙江省水产种质资源高效利用技术研究重点实验室 浙江宁波 315100;

2. 浙江万里学院宁海洋生物种业研究院 浙江宁波 315604)

摘要 组蛋白乙酰化酶(histone acetyltransferases, HATs)在机体发育和环境胁迫响应的调控中发挥着重要作用,但在海洋贝类中研究极少。缢蛏(*Sinonovacula constricta*)作为典型的滩涂双壳贝类,经常面临多种极端环境和细菌胁迫的挑战,基于全基因组和转录组数据系统分析了缢蛏 *HAT* 基因家族(*ScHATs*)的系统发育关系,以及它们在不同发育时期、不同环境因子和细菌胁迫下的表达模式。结果共鉴定出 11 个 *ScHATs* 基因,根据序列同源性分为 *GNAT*、*MYST* 和 *p300/CBP* 三类,每个亚家族中几乎所有的 *HAT* 都具有特定的保守结构域和保守基序。基因表达谱和 qRT-PCR 分析显示, *ScHATs* 在幼虫发育前期表达量普遍高于发育后期;在氨氮胁迫下, *ScKAT2B*、*ScKAT8*、*ScKAT9-2*、*Scp300* 在鳃中的表达量显著变化($P<0.05$);在高温胁迫下, *ScHATs* 在鳃和肝胰腺中表现出相似的表达模式,而 *Scp300* 在肝胰腺中的表达量显著降低($P<0.05$);在副溶血弧菌胁迫下, *ScHATs* 的表达均出现显著上调($P<0.05$)。综上结果提示, *ScHATs* 可能在缢蛏的早期发育及抗逆抗病中发挥重要作用,系统研究将进一步理解双壳贝类 *HAT* 基因家族的进化 and 功能奠定理论基础。

关键词 缢蛏; 组蛋白乙酰化酶; 基因家族; 环境胁迫; 细菌感染; 基因表达

中图分类号 S968.3; Q953; Q789 **doi:** 10.11693/hyhz20211000239

在真核生物中,染色质结构的动态调节影响各种生物过程(Ho *et al.*, 2010)。核小体是染色质的基本结构单位,由四个核心组蛋白(H2A, H2B, H3, H4)组成的八聚体和其周围包裹着 147 bp 的 DNA 共同构成(Kouzarides, 2007),每个组蛋白均包含一个结构化的球形结构域和一个从核小体核心延伸出来的非结构化 N 端尾部(Patel *et al.*, 2013),其尾部受到组蛋白乙酰化、甲基化、磷酸化、泛素化等各种翻译后修饰(Kouzarides, 2007)。组蛋白乙酰化是由组蛋白乙酰化酶(histone acetyltransferases, HATs)和组蛋白去乙酰化酶(histone deacetylases, HDACs)共同调控的动态可逆过程(Bannister *et al.*, 2011)。HAT 可将乙酰辅酶 A 的

乙酰基转移至核心组蛋白(主要为 H3 和 H4)N 末端特定赖氨酸的 ϵ -残基上,乙酰基通过中和赖氨酸残基上的正电荷,减弱组蛋白与 DNA 的结合能力,使染色质结构松散,有利于转录因子与 DNA 结合,从而促进基因转录与表达(Struhl, 1998; Eberharter *et al.*, 2002)。根据氨基酸序列和结构相似性, HAT 至少可分为 3 类亚家族: (1) GNAT, 与 GCN5 相关的 N 端乙酰转移酶,包括 KAT1 (Hat1)、KAT2B (PCAF)、KAT9 (Elp3)等; (2) MYST, 包括 KAT5 (Tip60)、KAT6A (MOZ)、KAT7 (Hbo1)、KAT8 (Sas2)等; (3) p300/CREB 结合蛋白(CBP)家族(Hodawadekar *et al.*, 2007; Bannister *et al.*, 2017)。

目前, HAT 已在多种生物中被鉴定,如动物中的

* 浙江省农业新品种选育重大科技专项, 2021C02069-7 号; 宁波市“科技创新 2025”重大专项, 2021Z114 号, 2019B10005 号; 宁波市自然科学基金, 202003N4191 号; 国家海洋水产种质资源库课题。练佳莹, 硕士研究生, E-mail: ljj148369@163.com

① 通信作者: 林志华, 博士生导师, 研究员, E-mail: zhihua9988@126.com; 董迎辉, 博士生导师, 教授, E-mail: dongyinghui118@126.com

收稿日期: 2021-10-10, 收修改稿日期: 2021-11-17

小鼠(Yao *et al.*, 1998)、秀丽隐杆线虫(*Caenorhabditis elegans*) (Goodman *et al.*, 2000)、果蝇(Carre *et al.*, 2005)、弓形虫(Padgett *et al.*, 2018)、马氏珠母贝(*Pinctada martensi*) (杨帅, 2020)等, 植物中的拟南芥(Latrasse *et al.*, 2008)、水稻(Liu *et al.*, 2012)、棉花(Imran *et al.*, 2019)、小麦(Gao *et al.*, 2021)等。研究表明, HAT在植物生长发育(Latrasse *et al.*, 2008)、响应生物和非生物胁迫方面发挥着重要作用(Yin *et al.*, 2019; Gan *et al.*, 2021), 其应激响应很大程度上依赖于翻译后的组蛋白乙酰化(Gao *et al.*, 2021)。但目前关于贝类组蛋白乙酰化修饰鲜有报道。

缢蛏(*Sinonovacula constricta*)是我国重要的海水养殖贝类, 由于其广泛分布于潮间带, 极易受到高温、高氨氮和病原微生物等极端环境和生物压力的影响(郑余琦等, 2017; Dong *et al.*, 2020), 从而导致生长变缓、免疫力降低等问题, 严重时暴发大规模死亡, 造成巨大的经济损失, 因此开展缢蛏抗逆、抗病研究对其养殖业的可持续发展具有重要意义。本研究对缢蛏HAT基因家族进行全基因组鉴定, 并对其保守结构域、保守基序、系统发育关系及其在不同发育时期、环境因子(氨氮和高温胁迫)和副溶血弧菌(*Vibrio parahaemolyticus*)胁迫下的表达特征进行分析, 旨在为深入探索HAT在贝类中的功能及调控机制奠定理论基础。

1 材料与方法

1.1 实验动物及细菌

2021年3月, 从宁波市海洋与渔业科技创新基地取健康缢蛏成贝, 平均壳长(58.43 ± 1.75) mm、壳高(19.93 ± 0.46) mm。缢蛏处理前, 在实验室配制海水中(温度 25 °C, 盐度 20)暂养 3 d。实验用副溶血弧菌为实验室保存菌种, 采用 2216E培养基于 28 °C恒温摇床上培养后测定OD₆₀₀值, 所需终浓度为 1×10^8 CFU/mL。

1.2 数据库资源

缢蛏基因组和转录组数据均为本实验室前期测序获得(Dong *et al.*, 2020)。其余的人(*Homo sapiens*)、白氏文昌鱼(*Branchiostoma belcheri*)、秀丽隐杆线虫、太平洋牡蛎(*Crassostrea gigas*)、虾夷扇贝(*Mizuhopecten yessoensis*)、加州双斑蛸(*Octopus bimaculoides*)和紫色球海胆(*Strongylocentrotus purpuratus*)基因组数据均来源于NCBI数据库(<https://www.ncbi.nlm.nih.gov/>)。

1.3 ScHAT 基因家族的鉴定

为鉴定ScHAT基因家族成员, 先保留与注释高度匹配的序列, 将所获得的序列作为目标序列通过

Blastp方式检索缢蛏基因组数据库, 手动剔除冗余序列, 然后通过SMART、Pfam、Batch-CD Search在线工具结合预测候选序列的保守结构域进行确认, 并根据HAT各家族的相似性进一步分类并命名。使用ProtParam (<https://web.expasy.org/protparam/>)分析HAT蛋白等电点(pI)、分子量(MW)和总疏水性(GRAVY)等理化性质。同时采用相同的方法鉴定出其他物种的HAT基因, 并进行基因数目的分类统计。

1.4 ScHAT 基因家族的系统进化分析

采用MAFFT 7 (Katoh *et al.*, 2013)对缢蛏和其他 7 个物种(人、白氏文昌鱼、秀丽隐杆线虫、太平洋牡蛎、虾夷扇贝、加州双斑蛸和紫色球海胆)的HAT蛋白序列进行多重序列比对。用MrBayes 3.2.6 (Ronquist *et al.*, 2003)进行系统发育分析, 4 条马尔可夫链运行 1.5×10^7 代, 取样频率(sampling frequency)设为 5 000, 样本舍弃比率(burn-in fraction)设为 0.25, 分裂频率平均标准差(average standard deviation of split frequencies)小于 0.01 视为达到收敛。最后用FigTree 1.4.3 对贝叶斯系统树进行可视化。

1.5 ScHAT 基因家族的结构域及 motif 分析

通过在线工具SMART (http://smart.embl-heidelberg.de/smart/set_mode.cgi?GENOMIC=1)、Pfam(<http://pfam.xfam.org/search#tabview=tab1>)预测HAT蛋白的保守结构域。利用MEME (<https://meme-suite.org/meme/tools/meme>)预测HAT蛋白的保守基序(motif), motif数目参数设置为 18, 长度设置为 21~50 aa, 并用TBtools (Chen *et al.*, 2020)进行可视化。

1.6 ScHAT 基因在不同发育时期及环境胁迫下的表达分析

利用已发表的缢蛏不同发育时期、氨氮胁迫(180 mg/L, 72 h)和高温(32 °C, 96 h)胁迫的RNA-seq数据集中检索 11 个ScHAT基因的FPKM (fragments per kilobase per million mapped reads)值, 分析它们在不同发育时期和环境胁迫下的表达模式。其中, 卵子、4 细胞、囊胚、原肠胚、担轮幼虫、D形幼虫、壳顶幼虫和稚贝 8 个时期的SRA登录号分别为: SRR10097413~SRR10097424; 氨氮胁迫和高温胁迫的转录组SRA登录号分别为: SRR9943679~SRR9943690 和 SRR13594753~SRR13594764。利用TBtools绘制基因表达热图。

1.7 ScHAT 基因在副溶血弧菌胁迫下的表达分析

将 150 粒缢蛏成贝随机分成 3 组, 用浓度为 1×10^8 CFU/mL的副溶血弧菌浸泡胁迫。于浸泡后 0、

3、6、12、24、48、72 和 96 h, 每个时间点随机选 6 粒缙蛭, 取其鳃置于液氮速冻后保存在 -80°C 。以 0 h 未胁迫组为对照组。用 Trizol (TaKaRa, Japan) 试剂提取鳃的总 RNA, 并对 RNA 浓度和完整性进行检测。利用 Prime-ScriptTM RT (TaKaRa, Japan) 试剂盒合成 cDNA。用 LightCycler[®] 480II 仪器进行荧光定量反应, 检测 *ScHATs* 在抗菌免疫中的表达规律。实验所用引物如表 1 所示, 利用 Primer 5 软件对 *ScHATs* 的保守结构域进行引物设计, 通过溶解曲线验证引物特异性。以 18S rRNA 为内参基因。qRT-PCR 反应体系 (20 μL) 为: 上下游引物各 1 μL 、cDNA (稀释 100 倍) 8 μL 、SYBR[®] Green Supermix 10 μL 。反应程序为: 95°C 30 s, 40 个循环; 95°C 变性 15 s, 60°C 退火 1 min。每组 3 个生物平行和 3 个技术重复。用 $2^{-\Delta\Delta\text{CT}}$ 法计算 *ScHATs* 的相对表达量, 利用 SPSS 软件统计分析其差异显著性 ($P < 0.05$), GraphPad Prism 8.0.1 绘图。

表 1 qRT-PCR 所需引物及序列
Tab.1 Sequences of the primers used in qRT-PCR

引物名称	序列 (5'~3')
<i>ScKAT1F</i>	CACAAACATTGATGACTTCTCTGC
<i>ScKAT1R</i>	GCTCCATTATCATCATCCACATC
<i>ScKAT2BF</i>	TGAAAACACAGCGGAAGAAGA
<i>ScKAT2BR</i>	ATGGAGGATTGCCAAGAGGA
<i>ScKAT5F</i>	CTAACCGAAGGGTGTAGGCTG
<i>ScKAT5R</i>	CGAGGGTACTCCAACCTTGCTG
<i>ScKAT6AF</i>	AACAAACGGGAGACGGAG
<i>ScKAT6AR</i>	ACACATAGCAGTCATCAGGAAGT
<i>ScKAT6BF</i>	AGAGCAGGGAGTAAAGCGTG
<i>ScKAT6BR</i>	CTGAGATGGTAGCAGCGTAGA
<i>ScKAT7F</i>	GTCTCAAAATGCCCCCTGCC
<i>ScKAT7R</i>	TCATAGCGGGTCTTACACTCG
<i>ScKAT8F</i>	TGGGAAGTATGAGATTGACACCTG
<i>ScKAT8R</i>	CCCATCCACCTCAAACAAAGA
<i>ScKAT9-1F</i>	TGCCAGACTTGCCAAATGTAG
<i>ScKAT9-1R</i>	TCAGGGTGGAAGGGGAATAA
<i>ScKAT9-2F</i>	AGGACCTTGGAACGCAGTGT
<i>ScKAT9-2R</i>	GTAGGCGTAACAGACCAATCAAA
<i>Scp300F</i>	AACCAACAGAGGCACCAAGG
<i>Scp300R</i>	TCCTTTAGATGTGTGGGCTCA
<i>ScCBPF</i>	CCGAGTGTGCCTTCCCAAAT
<i>ScCBPR</i>	ACAAGCCCAGATGTGTGCCC
<i>18SF</i>	TCGGTTCTATTGCGTTGGTTTT
<i>18SR</i>	CAGTTGGCATCGTTTATGGTCA

2 结果

2.1 *ScHAT* 基因家族的鉴定与分类

在缙蛭基因组中共鉴定出 11 个 *ScHAT* 家族成员, 根据其序列同源性可分为 3 类亚家族, GNAT 类 (*ScKAT1*, *ScKAT2B*, *ScKAT9-1*, *ScKAT9-2*)、MYST 类 (*ScKAT5*, *ScKAT6A*, *ScKAT6B*, *ScKAT7*, *ScKAT8*) 和 p300/CBP 类 (*Scp300*, *ScCBP*)。对其家族成员的理化性质分析结果显示, *ScHAT* 家族所有成员均为亲水性蛋白; 除 *Scp300* 为酸性蛋白外, 其余均为碱性蛋白 (表 2)。不同成员外显子数量在 5~26 个之间, 氨基酸长度在 187~2 261 aa 之间, 蛋白分子量在 19.99~92.40 kDa 之间, 等电点 (pI) 在 5.95~9.30, 不稳定系数介于 34.44~62.72 之间。染色体定位发现, 11 个 *ScHATs* 均成功定位在缙蛭的 7 条染色体上。另外, 其他 7 个物种的全基因组鉴定结果如表 3 所示, 太平洋牡蛎和紫色球海胆 *HAT* 基因数目最多为 13 个, 其次人和缙蛭分别为 12 个和 11 个; 其中 MYST 亚家族为最大的一类, 在不同物种中基因数目均为最多; 缙蛭和人 *HAT* 各亚家族基因数最相似。

2.2 *ScHAT* 基因家族的系统进化分析

用 MrBayes 构建了缙蛭与其他物种 *HAT* 基因家族的系统发育树 (图 1)。该进化树共分成 7 个独立的分支, KAT2A/2B 和 KAT9 分支形成独立的一支 (GNAT); KAT5、KAT6A/6B、KAT7 和 KAT8 四个分支形成最大的一支 (MYST)。不同物种 *HAT* 蛋白具有高度同源性, *ScHAT* 各成员均与太平洋牡蛎、虾夷扇贝、加州双斑蛸同源基因形成独立分支, 再依次与紫色球海胆、白氏文昌鱼、秀丽隐杆线虫和人聚类。

2.3 *ScHAT* 基因家族的结构域与 motif 特征分析

基于 SMART 结构域预测, 比较了缙蛭、太平洋牡蛎和人 *HAT* 蛋白的结构域 (表 4)。结果显示, GNAT 亚家族的 KAT2 成员大多含有 PCAF_N 结构域和 Bromo 结构域; KAT9 成员大多包含一个 Radical_SAM 结构域特征; 除 *ScKAT9-1* 之外, 其余 GNAT 亚家族成员均含有 N-Acetyltransf_1 或 C-Acetyltransf_1 结构域。MYST 亚家族成员中, 除 *ScKAT6A* 和 *CgKAT6B-2* 只包含 1 个 PHD-SAM_1 结构域外, 其余成员均含有 zf-MYST 和 MOZ_SAS 结构域; 另外 KAT5 和 KAT8 还有一个 Tudor-knot 结构域 (*HsKAT8* 除外); KAT6 成员存在一个 PHD 结构域特征, 以及 KAT7 成员包含一个 C2HC-zf 结构域。p300/CBP 亚家族成员含有 zf-TAZ、KIX、Bromo 等结构域。

表 2 *ScHAT* 家族基因理化性质
Tab.2 The physicochemical properties of *ScHAT* gene family

基因 ID	基因名	亚家族	氨基酸 数量	染色体号	染色体位置	等电点 /pI	分子量 /kDa	外显子 数量	亲水性	不稳定 系数
evm.model.ctg169.36	<i>ScKAT1</i>	GNAT	379	Hic_asm_8	19899672:19907780	8.65	44.65	9	-0.599	54.55
evm.model.ctg313.2	<i>ScKAT2B</i>		806	Hic_asm_10	42250573:42264982	8.89	92.40	18	-0.534	50.17
evm.model.ctg56.31	<i>ScKAT9-1</i>		390	Hic_asm_0	9015516:9025304	9.3	44.32	13	-0.396	36.33
evm.model.ctg56.32	<i>ScKAT9-2</i>		187	Hic_asm_0	9026664:9031490	7.71	21.23	5	-0.416	34.33
evm.model.ctg367.24	<i>ScKAT5</i>	MYST	473	Hic_asm_1	4261871:4279716	8.29	53.97	14	-0.592	44.51
evm.model.ctg52.16	<i>ScKAT6A</i>		584	Hic_asm_0	53282557:53297791	7.23	66.57	10	-0.785	53.89
evm.model.ctg122.31	<i>ScKAT6B</i>		1 806	Hic_asm_4	73085863:73111539	8.07	19.99	17	-0.846	59.97
evm.model.ctg77.48	<i>ScKAT7</i>		616	Hic_asm_0	32365113:32372689	9.11	70.54	12	-0.872	50.04
evm.model.ctg55.51	<i>ScKAT8</i>	p300/CBP	415	Hic_asm_13	22940213:22957896	7.56	48.76	10	-0.643	40.29
evm.model.ctg75.43	<i>Scp300</i>		670	Hic_asm_6	65986389:66001279	5.95	73.33	10	-0.606	55.83
evm.model.ctg53.44	<i>ScCBP</i>		2 261	Hic_asm_6	62838874:62869874	8.81	25.08	26	-0.767	62.72

表 3 缢蛭与其他物种 *HAT* 家族基因数目的比较
Tab.3 Comparison of gene number of *HAT* family between *S. constricta* and other species

物种	GNAT	MYST	p300/CBP	总计
人(<i>H. sapiens</i>)	5	5	2	12
白氏文昌鱼(<i>B. belcheri</i>)	2	6	2	10
秀丽隐杆线虫(<i>C. elegans</i>)	0	4	1	5
紫色球海胆(<i>S. purpuratus</i>)	3	9	1	13
加州双斑蛸(<i>O. bimaculoides</i>)	2	6	2	10
缢蛭(<i>S. constricta</i>)	4	5	2	11
太平洋牡蛎(<i>C. gigas</i>)	2	6	5	13
虾夷扇贝(<i>M. yessoensis</i>)	2	6	1	9

缢蛭、太平洋牡蛎和人*HAT*家族基因的motif分析结果如图 2 所示。*HAT*不同亚家族之间序列分化明显,共鉴定出 18 个保守的 motif。除 *ScKAT6A* 和 *CgKAT6B-2* 外, MYST亚家族的成员从 5'到 3'方向依次含有 motif 8、motif 5、motif 9、motif 1、motif 2、motif 6, 其中KAT6 成员具有 1 或 2 个 motif 3; 在 GNAT亚家族中, 除KAT1 和KAT9 只含有 1 或 2 个保守motif外, motif 16、motif 7、motif 4、motif 14、motif 15 存在于所有KAT2 成员中; p300/CBP亚家族也具有其独特的motif序列。此外, 同一亚家族成员存在motif 新增或丢失的情况, 如*CgKAT5* 与*ScKAT8* 与其同源基因KAT5和KAT8 分别新增了 1 个 motif 4 和 motif 10; 而*HsKAT2B-2* 与其同源基因KAT2 则丢失了 motif 12。

2.4 *ScHAT* 家族基因在缢蛭早期发育时期的表达模式分析

基于转录组分析*ScHATs*在缢蛭早期发育时期的表达规律, 结果显示, 它们在不同发育时期的表达趋

势呈现多变性(图 3)。GNAT亚家族中, *ScKAT1* 和 *ScKAT9-1* 均在 4 细胞期表达量最高, 而*ScKAT2B*和 *ScKAT9-2* 则在壳顶幼虫时期表达量最高。MYST亚家族中, *ScKAT5* 随着缢蛭的生长发育表达量逐渐升高, 在稚贝期达到最高; 其他 4 个基因呈现出相似的表达模式, 均在发育初期(D形幼虫时期之前)表达量较高, 而在壳顶幼虫和稚贝期表达量明显下降。p300/CBP亚家族的 2 个成员也表现出不同的表达情况, *Scp300* 在稚贝期高表达, 而*ScCBP*在D形幼虫期表达量最高。

2.5 *ScHAT* 家族基因在高氨氮、高温胁迫下的表达特征分析

对*ScHATs*在高氨氮、高温胁迫下的表达模式进行分析(图 4)。在氨氮胁迫下, *ScHATs*在鳃组织中的表达响应高于肝胰腺, 其中*ScKAT9-2*、*ScKAT8* 和*Scp300* 三个基因的表达显著升高($P<0.05$), 而*ScKAT2B*的表达显著降低($P<0.05$)。在高温胁迫后, *ScHATs*在鳃和肝胰腺中表达趋势几乎相同, 其中*Scp300* 在肝胰腺中的表达量显著降低($P<0.05$)。

2.6 *ScHAT* 家族基因在副溶血弧菌感染后的表达特征分析

用qRT-PCR方法, 对 11 个*ScHATs*基因在副溶血弧菌感染后鳃中的表达水平进行检测(图 5)。结果显示, 所有*ScHATs*的相对表达量在副溶血弧菌感染后均出现显著上调($P<0.05$), 并呈现时间依赖性的表达模式。多数基因在 24 h达到较高水平(如*ScKAT6A*和 *ScKAT7*), *ScKAT1* 和*ScKAT6A*在 48 h表达量最高, 其他基因多在 96 h达到最高水平。其中*Scp300* 有所不同, 在感染后 3~12 h表达水平显著下降后, 于 24 h后开始显著上升。

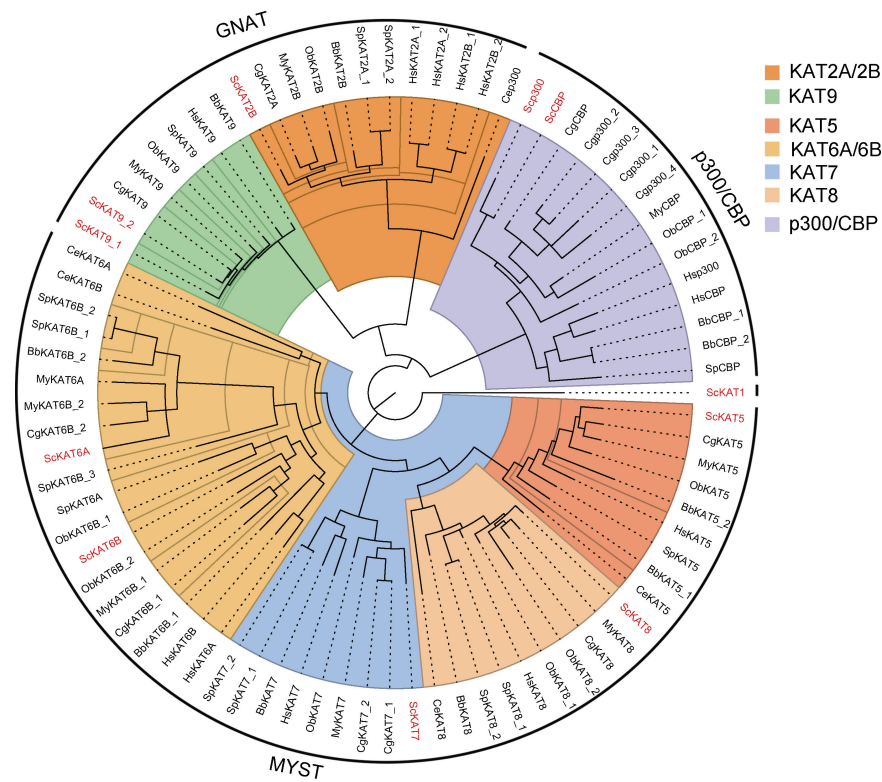


图 1 缢蛏与其他物种 *HAT* 基因家族的系统发育分析

Fig.1 Phylogenetic analysis of *HAT* gene family in *S. constricta* and other species

注: 右上角为 3 类亚家族成员的不同颜色区块, 红色字体代表缢蛏基因。Hs: 人; Bb: 白氏文昌鱼; Ce: 秀丽隐杆线虫; Sc: 缢蛏; Cg: 太平洋牡蛎; My: 虾夷扇贝; Ob: 加州双斑蛸; Sp: 紫色球海胆

表 4 缢蛏、太平洋牡蛎和人 *HAT* 家族成员结构域比较

Tab.4 The domain analysis of *HAT* gene family in *S. constricta*, *C. gigas* and *H. sapiens*

基因 ID	基因名	KAT 分类	基因家族	结构域组成
evm.model.ctg169.36	<i>ScKAT1</i>	KAT1		Hat1_N
NP_066564.2	<i>HsKAT2A-1</i>			PCAF_N-Acetyltransf_1-Bromo
XP_006721881.1	<i>HsKAT2A-2</i>	KAT2A		Acetyltransf_1-Bromo
XP_011432298.2	<i>CgKAT2A</i>			PCAF_N-Acetyltransf_1-Bromo
evm.model.ctg313.2	<i>ScKAT2B</i>			PCAF_N-Acetyltransf_1-Bromo
NP_003875.3	<i>HsKAT2B-1</i>	KAT2B	GNAT	PCAF_N-Acetyltransf_1-Bromo
XP_005265585.1	<i>HsKAT2B-2</i>			PCAF_N-Acetyltransf_1
evm.model.ctg56.31	<i>ScKAT9-1</i>			Radical_SAM-Radical_SAM_C
evm.model.ctg56.32	<i>ScKAT9-2</i>	KAT9		Acetyltransf_1
NP_060561.3	<i>HsKAT9</i>			Radical_SAM-Radical_SAM_C-Acetyltransf_1
XP_011425652.2	<i>CgKAT9</i>			Radical_SAM-Radical_SAM_C-Acetyltransf_1
evm.model.ctg367.24	<i>ScKAT5</i>			Tudor-knot-zf-MYST-MOZ_SAS
XP_006718484.1	<i>HsKAT5</i>	KAT5		Tudor-knot-zf-MYST-MOZ_SAS
XP_034324845.1	<i>CgKAT5</i>			Tudor-knot-zf-MYST-MOZ_SAS
evm.model.ctg52.16	<i>ScKAT6A</i>	KAT6A	MYST	PHD-SAM_1
XP_011542958.1	<i>HsKAT6A</i>			PHD-zf-MYST-MOZ_SAS
evm.model.ctg122.31	<i>ScKAT6B</i>	KAT6B		PHD-zf-C2HC-zf-MYST-MOZ_SAS
XP_016871492.1	<i>HsKAT6B</i>			PHD-zf-MYST-MOZ_SAS

续表

基因 ID	基因名	KAT 分类	基因家族	结构域组成
XP_034318334.1	<i>CgKAT6B-1</i>	KAT7	p300/CBP	Linker_histone-PHD-zf-C2HC-zf-MYST-MOZ_SAS
XP_011442343.2	<i>CgKAT6B-2</i>			PHD-SAM_1
evm.model.ctg77.48	<i>ScKAT7</i>			zf-C2HC-zf-MYST-MOZ_SAS
NP_008998.1	<i>HsKAT7</i>			zf-C2HC-zf-MYST-MOZ_SAS
XP_034330903.1	<i>CgKAT7-1</i>			zf-C2HC-zf-MYST-MOZ_SAS
XP_034329657.1	<i>CgKAT7-2</i>	KAT8	p300/CBP	zf-C2HC-zf-MYST-MOZ_SAS
evm.model.ctg55.51	<i>ScKAT8</i>			Tudor-knot-zf-MYST-MOZ_SAS
XP_011544272.1	<i>HsKAT8</i>			zf-MYST-MOZ_SAS
XP_011434277.2	<i>CgKAT8</i>			Tudor-knot-zf-MYST-MOZ_SAS
evm.model.ctg75.43	<i>Scp300</i>	p300/CBP	p300/CBP	zf-TAZ-KIX
XP_011438001.2	<i>Cgp300-1</i>			zf-TAZ-zf-TAZ-KIX
XP_034303589.1	<i>Cgp300-2</i>			zf-TAZ
XP_034303586.1	<i>Cgp300-3</i>			zf-TAZ-KIX
XP_034303592.1	<i>Cgp300-4</i>			KIX
evm.model.ctg53.44.1	<i>ScCBP</i>			zf-TAZ-KIX-Bromo-DUF902-HAT_KAT11-ZZ-zf-TAZ -Creb_binding
NP_001420.2	<i>Hsp300</i>			zf-TAZ-KIX-Bromo-DUF902-HAT_KAT11-ZZ-zf-TAZ -Creb_binding
XP_016878433.1	<i>HsCBP</i>			zf-TAZ-KIX-Bromo-DUF902-HAT_KAT11-ZZ-zf-TAZ -Creb_binding
XP_034303595.1	<i>CgCBP</i>			zf-TAZ-KIX-Bromo-DUF902-HAT_KAT11-ZZ-zf-TAZ -Creb_binding

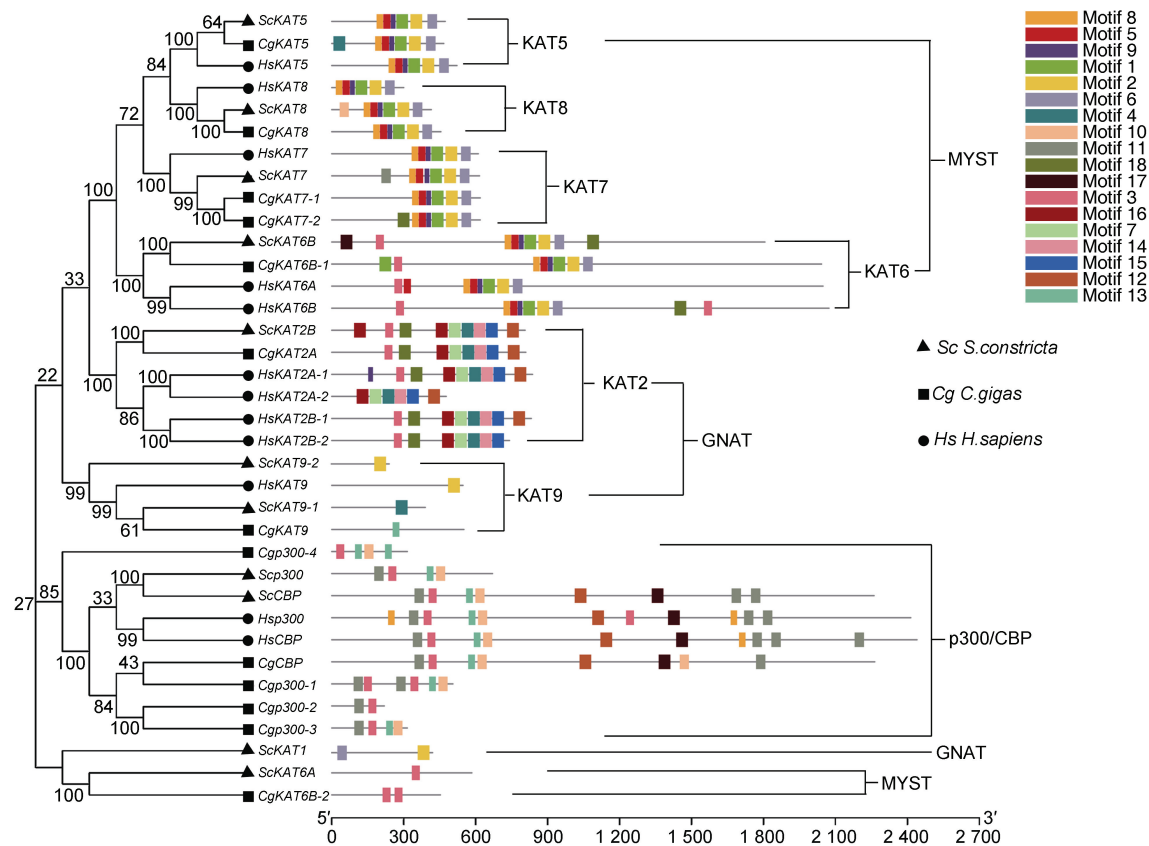


图 2 缢蛏、太平洋牡蛎和人 HAT 基因家族 motif 分析
Fig.2 Motif analysis of HAT gene family in *S. constricta*, *C. gigas* and *H. sapiens*

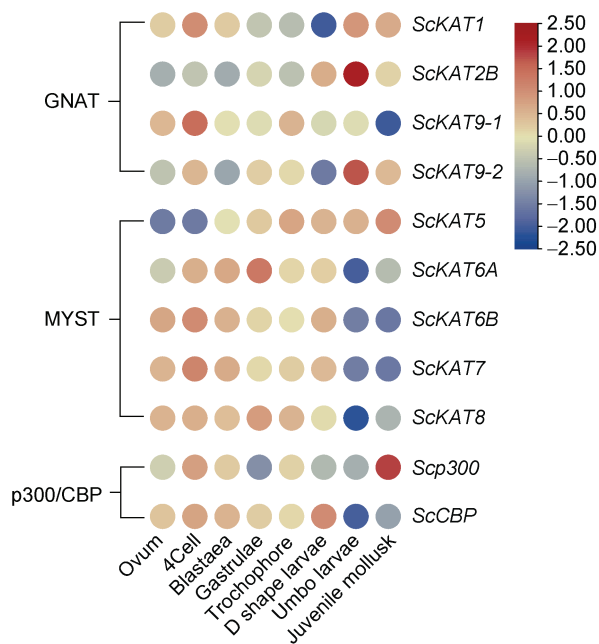


图3 *ScHATs* 在不同发育时期中的表达谱

Fig.3 Expression profiles of *ScHATs* in different developmental stages

注: 8 个发育时期为卵子(Ovum)、4 细胞(4 cell)、囊胚(Blastaea)、原肠胚(Gastrulae)、担轮幼虫(Trochophore)、D 形幼虫(D-shape larvae)、壳顶幼虫(Umbo larvae)、稚贝(Juvenile mollusk)。图例表示 FPKM 对数归一化

3 讨论

由 *HATs* 基因介导的组蛋白乙酰化修饰在转录激活、基因沉默、细胞周期调控、DNA 复制修复、染色体组装等多个细胞过程中发挥至关重要的调控作用(Mai *et al.*, 2009), 而组蛋白乙酰化与基因表达调控

密切相关(Imran *et al.*, 2019)。已有研究发现, *HATs* 在棉花、玉米高等植物生长发育以及对环境胁迫的适应调节中发挥关键作用, 如同属不同种的 3 种棉花 *HAT* 基因数目差异巨大, 且 *HAT* 基因在纤维发育不同时期以及激素、干旱、重金属等不同因子胁迫下的表达差异显著(Imran *et al.*, 2019); 在玉米基因组中鉴定出 6 个 *HAT* 家族基因, 它们在不同组织和生物/非生物胁迫下呈现出明显不同的表达规律(马宇馨等, 2020)。然而, 截至目前, 对动物 *HAT* 家族基因的研究鲜有报道。本研究通过全基因组分析在缢蛏基因组中鉴定出 11 个 *HAT* 家族基因, 根据序列同源性分为 GNAT、MYST 和 p300/CBP 三类亚家族, 这种多个亚家族的存在暗示其在缢蛏长期进化适应中的功能分化; 通过对缢蛏、太平洋牡蛎、人等 8 个物种 *HAT* 亚家族基因数目的比较发现, 每个亚家族在物种间分布不均匀, 其中 MYST 是最大的亚家族, 与以往研究一致(Hodawadekar *et al.*, 2007)。此外, 保守基序和结构域预测结果表明, 在同一系统发育亚群中, 大多数成员具有各自保守的基因结构域, 因而推测 *ScHATs* 与其同源基因具有相似的功能, 在水稻、番茄的研究中也得到相似结论(Liu *et al.*, 2012; 史建磊等, 2020)。

研究发现, *HATs* 在机体发育中发挥关键性作用(Gan *et al.*, 2021)。MYST 亚家族主要作为一种发育调节剂, 参与干细胞的生长和自我修复过程(Voss *et al.*, 2009); GCN5/PCAF 和 p300/CBP 亚家族也被证实在脊椎动物的多个生长发育过程发挥重要的调控功能(Goodman *et al.*, 2000; Koutelou *et al.*, 2021)。GNAT 亚

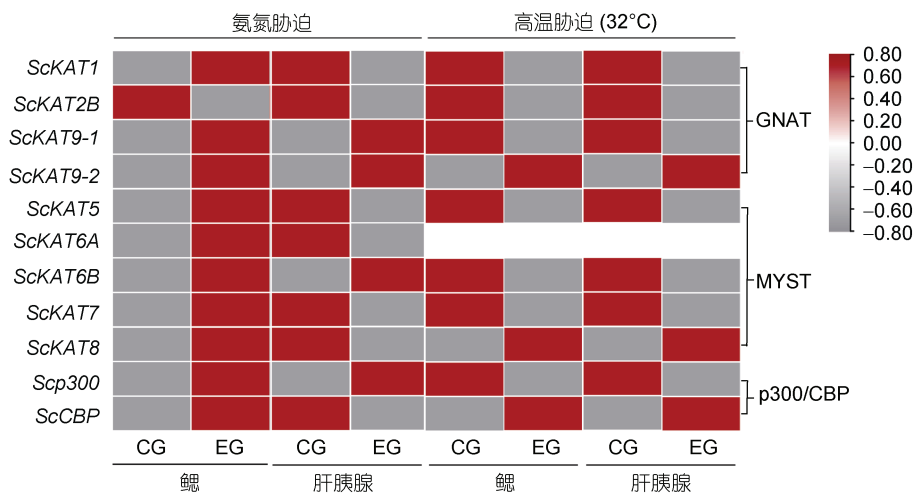


图4 *ScHATs* 在高氨氮、高温胁迫下鳃和肝胰腺的表达模式

Fig.4 Expression profiles of *ScHATs* in gills and hepatopancreas under ammonia nitrogen and thermal stresses
注: CG: 对照组; EG: 实验组。图例表示 FPKM 对数归一化

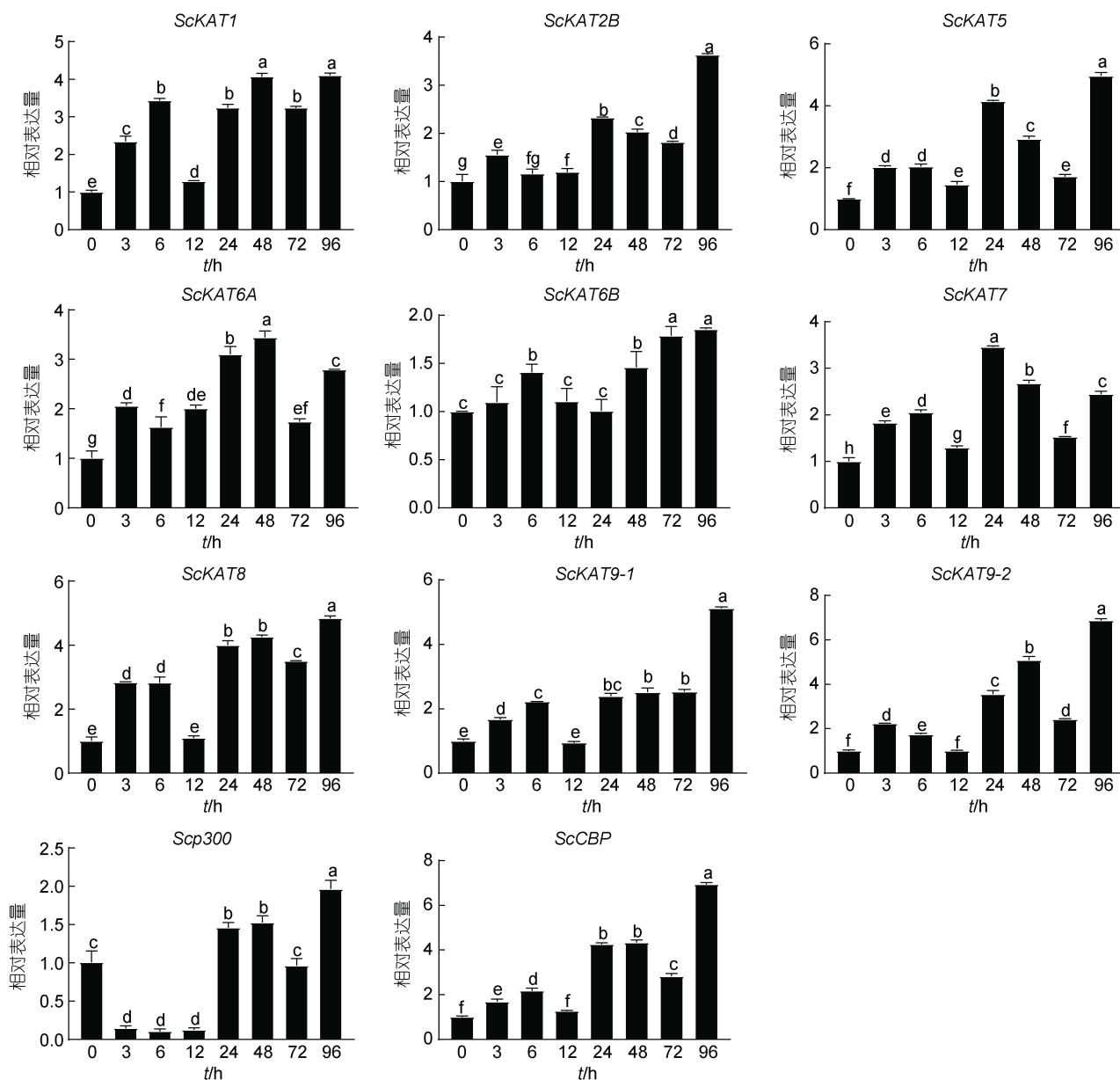


图5 *ScHATs* 在副溶血弧菌感染后鳃中的表达模式

Fig.5 Relative expression of *ScHATs* in gills after *V. parahaemolyticus* infection

注: 以对照组 0 h 的表达量为参照, 用 18S rRNA 归一化计算相对表达量。数据以平均值 $\pm$ S.E. 表示 ($n=3$)。不同字母表示时间点差异显著 ($P < 0.05$)

家族的 KAT2 成员大多具有 Bromo 结构域, 可通过与乙酰化赖氨酸的特异性互作发挥乙酰化酶作用, 进而影响体内基因转录和细胞生长(Zeng *et al*, 2002; Ren *et al*, 2016)。缺失 *GCN5* 等位基因的果蝇突变体出现卵发生、变态以及成虫附肢等异常的现象(Carre *et al*, 2005)。与大多数双壳类一样, 缢蛏幼虫发育从细胞分化迅速的细胞期经卵裂、变态发育到浮游的 D 形幼虫不到 24 h (林笔水等, 1984; 徐小伟等, 2015)。ScHATs 在缢蛏不同发育时期的表达量呈现动态变化规律, 其

在幼虫发育前期表达量普遍高于发育后期(壳顶幼虫和稚贝), 因此推测 *HAT* 基因对缢蛏的正常发育至关重要, 可能通过基因的乙酰化作用改变靶基因的表达水平, 进而参与早期发育的基因调控网络。

已有报道显示, *HAT* 家族基因在响应环境胁迫中发挥重要功能。棉花 *HATs* 的相对表达量在温度、盐度和重金属等胁迫下发生显著变化(Imran *et al*, 2019); 油菜 p300/CBP 亚家族能够提高其应对外界环境变化的能力(Li *et al*, 2014); 拟南芥则通过调节 *GCN5* 基因

的表达响应极端温度变化(Gan *et al*, 2021)。本研究中, 高浓度氨氮胁迫缢蛏后, 鳃中 *ScKAT9-2*、*ScKAT8* 和 *Scp300* 3 个基因的表达量显著升高, 而含有 PCAF 结构域的 *ScKAT2B* 的表达量显著降低($P < 0.05$)。已有研究发现, 在金属镍环境暴露下会导致 p300/CBP 的相关蛋白 PCAF 低表达(Kim *et al*, 2012), 这与本研究结果相似。另外, 在缢蛏鳃和肝胰腺中 *HATs* 的表达模式差异较大, 反映这两个器官在氨氮代谢中的途径和机制有所不同(陈凯锋等, 2020)。在高温胁迫下, *HATs* 在缢蛏鳃和肝胰腺中的表达趋势几乎相同, 这说明不同组织在适应高温响应的调节机制相似, 如通过增加抗氧化酶系统、提高细胞的渗透调节物质等提高耐热性, 在此过程中一些 *ScHATs* 基因表达量的显著变化表明组蛋白修饰的表观遗传在响应高温胁迫中发挥重要调控作用。在对仿刺参(*Apostichopus japonicus*)的高温胁迫下, 也发现组蛋白去乙酰化酶 3 (*HDAC3*)和组蛋白甲基转移酶(histone methyltransferase, *MLL5*)等表观遗传调控基因表达量显著上调(李尚俊等, 2017)。然而, 目前 *HAT* 家族基因调控环境适应的机理尚需开展大量深入研究。

本研究中, 在副溶血弧菌胁迫下所有 *ScHATs* 基因在缢蛏鳃中的表达量均发生显著变化($P < 0.05$)。与对照组相比, 一般呈现三个峰, 即 3~6 h、24~48 h、96 h, 这种表达模式说明随着处理时间的延长, *HATs* 家族基因可能通过调节组蛋白乙酰化水平参与了缢蛏在这些时间点的表达调控, 反映了缢蛏对环境的动态适应过程, 推测 *ScHATs* 的高表达变化可能是缢蛏应对细菌感染时参与调节免疫相关基因转录水平的结果。已有大量研究证明, 许多 *HAT* 家族基因参与免疫及抗病过程, 如敲除稻瘟病菌(*Magnaporthe oryzae*) *HAT* 基因使突变体生长速度和孢子产量显著降低, 进而抑制其侵染能力(Yin *et al*, 2019); *KAT6A* 中的 PHD-zf 结构域存在于一个自身免疫调节因子(AIRE)蛋白中(Musco *et al*, 2008), 通过与生物胁迫耐受性的其他蛋白质伴侣相互作用发挥功能(Waziri *et al*, 2020); 多数 *KAT7* 中包含 C2HC-zf 结构域, 而 C2HC-zf 结构域已在凡纳滨对虾(*Litopenaeus vannamei*)中被证实具有抗菌作用(Yang *et al*, 2021); *KAT9* 中包含的 Radical_SAM 结构域可通过与病毒蛋白结合发挥抗病毒活性, 在病毒感染的先天性免疫反应中起着重要作用(Ghosh *et al*, 2020)。基于上述已有研究, 结合本研究的结构域分析, 不难推测分别包含 PHD-zf、C2HC-zf 和 Radical_SAM 结构域的

ScKAT6A、*ScKAT7*、*ScKAT9-1* 在缢蛏病原侵染的防御过程中发挥重要的抗菌作用, 而不同基因表达量和时间效应的差异也正反映了 *ScHATs* 家族基因不同成员的功能分化及其调控机制差别。

4 结论

本研究在缢蛏基因组中鉴定出 11 个 *ScHAT* 基因, 比较分析了 *HAT* 家族成员的蛋白质理化性质、结构域、保守基序及系统发育关系, 并将其分为 GNAT、MYST、p300/CBP 三类。通过转录组表达谱分析和 qRT-PCR 验证, 探究了缢蛏在不同发育阶段、不同环境因子(高氨氮、高温)胁迫和副溶血弧菌感染条件下 *ScHAT* 家族成员的表达变化, 结果表明 *ScHATs* 基因功能具有多样性, 可能在早期发育及极端环境响应过程中发挥重要作用。该研究将丰富人们对贝类 *HAT* 家族基因系统进化的认识, 也为其功能研究提供重要参考。

参 考 文 献

- 马宇馨, 杜璇玥, 李肖慧, 等, 2020. 玉米组蛋白乙酰转移酶的鉴定与表达规律分析[J]. 河北农业大学学报, 43(5): 20-26.
- 史建磊, 熊自立, 李涛, 等, 2020. 番茄组蛋白乙酰转移酶(HAT)的全基因组鉴定与分析[J]. 江苏农业学报, 36(3): 666-674.
- 李尚俊, 孙国华, 李雪燕, 等, 2017. 高温胁迫下仿刺参表观遗传调控相关基因的表达特征[J]. 中国水产科学, 24(3): 470-476.
- 杨帅, 2020. 组蛋白及其乙酰化修饰在珍珠贝移植免疫中的作用和机制[D]. 湛江: 广东海洋大学: 33-48
- 陈凯锋, 董迎辉, 姚韩韩, 等, 2020. 缢蛏(*Sinonovacula constricta*)氨氮胁迫应答 miR-8245a-5p 靶基因 GOT 验证及其表达特征分析[J]. 海洋与湖沼, 51(2): 388-394.
- 林笔水, 吴天明, 1984. 温度和盐度对缢蛏浮游幼虫发育的影响[J]. 生态学报, 4(4): 385-392.
- 郑余琦, 郑忠明, 秦文娟, 2017. 缢蛏(*Sinonovacula constricta*)生物扰动对养殖废水处理系统中沉积物磷赋存形态垂直分布的影响[J]. 海洋与湖沼, 48(1): 161-170.
- 徐小伟, 张鹏飞, 黄妙琴, 等, 2015. 波纹巴非蛤早期发育的扫描电镜观察[J]. 渔业研究, 37(4): 263-269.
- BANNISTER A J, FALCÃO A M, CASTELO-BRANCO G, 2017. Histone modifications and histone variants in pluripotency and differentiation [M] // GÖNDÖR A. Chromatin Regulation and Dynamics. San Diego, CA: Academic Press: 35-64, doi: 10.1016/b978-0-12-803395-1.00002-2.
- BANNISTER A J, KOUZARIDES T, 2011. Regulation of chromatin by histone modifications [J]. Cell Research, 21(3): 381-395.
- CARRE C, SZYMCAK D, PIDOUX J, *et al*, 2005. The histone H3 acetylase dGcn5 is a key player in *Drosophila*

- melanogaster* metamorphosis [J]. Molecular and Cellular Biology, 25(18): 8228-8238.
- CHEN C J, CHEN H, ZHANG Y, *et al*, 2020. TBtools: An integrative toolkit developed for interactive analyses of big biological data [J]. Molecular Plant, 13(8): 1194-1202.
- DONG Y H, ZENG Q F, REN J F, *et al*, 2020. The chromosome-level genome assembly and comprehensive transcriptomes of the razor clam (*Sinonovacula constricta*) [J]. Frontiers in Genetics, 11: 664.
- EBERHARTER A, BECKER P B, 2002. Histone acetylation: a switch between repressive and permissive chromatin [J]. EMBO Reports, 3(3): 224-229.
- GAN L, WEI Z Z, YANG Z R, *et al*, 2021. Updated mechanisms of GCN5-the monkey king of the plant kingdom in plant development and resistance to abiotic stresses [J]. Cells, 10(5): 979.
- GAO S Q, LI L Z, HAN X L, *et al*, 2021. Genome-wide identification of the histone acetyltransferase gene family in *Triticum aestivum* [J]. BMC Genomics, 22(1): 49.
- GHOSH S, MARSH E N G, 2020. Viperin: An ancient radical SAM enzyme finds its place in modern cellular metabolism and innate immunity [J]. Journal of Biological Chemistry, 295(33): 11513-11528.
- GOODMAN R H, SMOLIK S, 2000. CBP/p300 in cell growth, transformation, and development [J]. Genes & Development, 14(13): 1553-1577.
- HO L, CRABTREE G R, 2010. Chromatin remodeling during development [J]. Nature, 463(7280): 474-484.
- HODAWADEKAR S C, MARMORSTEIN R, 2007. Chemistry of acetyl transfer by histone modifying enzymes: structure, mechanism and implications for effector design [J]. Oncogene, 26(37): 5528-5540.
- IMRAN M, SHAFIQ S, FAROOQ M A, *et al*, 2019. Comparative genome-wide analysis and expression profiling of histone acetyltransferase (HAT) gene family in response to hormonal applications, metal and abiotic stresses in cotton [J]. International Journal of Molecular Sciences, 20(21): 5311.
- KATOH K, STANDLEY D M, 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability [J]. Molecular Biology and Evolution, 30(4): 772-780.
- KIM H L, SEO Y R, 2012. Molecular and genomic approach for understanding the gene-environment interaction between Nrf2 deficiency and carcinogenic nickel-induced DNA damage [J]. Oncology Reports, 28(6): 1959-1967.
- KOUTELOU E, FARRIA A T, DENT S Y R, 2021. Complex functions of *Gcn5* and *Pcaf* in development and disease [J]. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms, 1864(2): 194609.
- KOUZARIDES T, 2007. Chromatin modifications and their function [J]. Cell, 128(4): 693-705.
- LATRASSE D, BENHAMED M, HENRY Y, *et al*, 2008. The MYST histone acetyltransferases are essential for gametophyte development in *Arabidopsis* [J]. BMC Plant Biology, 8: 121.
- LI H, YAN S H, ZHAO L, *et al*, 2014. Histone acetylation associated up-regulation of the cell wall related genes is involved in salt stress induced maize root swelling [J]. BMC Plant Biology, 14: 105.
- LIU X, LUO M, ZHANG W, *et al*, 2012. Histone acetyltransferases in rice (*Oryza sativa* L.): Phylogenetic analysis, subcellular localization and expression [J]. BMC Plant Biology, 12: 145.
- MAI A, ROTILI D, TARANTINO D, *et al*, 2009. Identification of 4-hydroxyquinolines inhibitors of p300/CBP histone acetyltransferases [J]. Bioorganic & Medicinal Chemistry Letters, 19(4): 1132-1135.
- MUSCO G, PETERSON P, 2008. PHD finger of autoimmune regulator: an epigenetic link between the histone modifications and tissue-specific antigen expression in thymus [J]. Epigenetics, 3(6): 310-314.
- PADGETT L R, LENTINI J M, HOLMES M J, *et al*, 2018. Elp3 and RlmN: A tale of two mitochondrial tail-anchored radical SAM enzymes in *Toxoplasma gondii* [J]. PLoS One, 13(1): e0189688.
- PATEL D J, WANG Z X, 2013. Readout of epigenetic modifications [J]. Annual Review of Biochemistry, 82(1): 81-118.
- REN C, ZENG L, ZHOU M M, 2016. Preparation, biochemical analysis, and structure determination of the bromodomain, an acetyl-lysine binding domain [J]. Methods in Enzymology, 573: 321-343.
- RONQUIST F, HUELSENBECK J P, 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models [J]. Bioinformatics, 19(12): 1572-1574.
- STRUHL K, 1998. Histone acetylation and transcriptional regulatory mechanisms [J]. Genes & Development, 12(5): 599-606.
- VOSS A K, THOMAS T, 2009. MYST family histone acetyltransferases take center stage in stem cells and development [J]. BioEssays, 31(10): 1050-1061.
- WAZIRI A, SINGH D K, SHARMA T, *et al*, 2020. Genome-wide analysis of PHD finger gene family and identification of potential miRNA and their PHD finger gene specific targets in *Oryza sativa indica* [J]. Non-coding RNA Research, 5(4): 191-200.
- YANG L W, WANG Z A, ZUO H L, *et al*, 2021. The LARK protein is involved in antiviral and antibacterial responses in shrimp by regulating humoral immunity [J]. Developmental & Comparative Immunology, 114: 103826.
- YAO T P, OH S P, FUCHS M, *et al*, 1998. Gene dosage-dependent embryonic development and proliferation defects in mice lacking the transcriptional integrator p300 [J]. Cell, 93(3): 361-372.
- YIN Z Y, CHEN C, YANG J, *et al*, 2019. Histone acetyltransferase MoHat1 acetylates autophagy-related proteins MoAtg3 and MoAtg9 to orchestrate functional appressorium formation and pathogenicity in *Magnaporthe oryzae* [J]. Autophagy, 15(7): 1234-1257.
- ZENG L, ZHOU M M, 2002. Bromodomain: an acetyl-lysine binding domain [J]. FEBS Letters, 513(1): 124-128.

GENOME-WIDE IDENTIFICATION AND EXPRESSION ANALYSIS OF THE HISTONE ACETYLTRANSFERASE (*HAT*) GENE FAMILY IN *SINONOVACULA CONSTRICTA* UNDER ENVIRONMENTAL AND BACTERIAL STRESSES

LIAN Jia-Ying¹, LYU Li-Yuan², YAO Han-Han¹, DONG Ying-Hui^{1,2}, LIN Zhi-Hua^{1,2}

(1. College of Biological and Environmental Sciences, Zhejiang Wanli University, Zhejiang Key Laboratory of Efficient Utilization Technology of Aquatic Germplasm Resources, Ningbo 315100, China; 2. Ninghai Institute of Mariculture Breeding and Seed Industry, Zhejiang Wanli University, Ningbo 315604, China)

Abstract Histone acetyltransferases (HATs) controls the regulation of body growth and the response to environmental stress. However, little is known about HATs in marine mollusks at present. The razor clam *Sinonovacula constricta* is a typical bivalve living in tidal flats and currently faces extreme environmental challenges. To understand the evolutionary dynamics of *HAT* gene family in shellfish, phylogenetic relationships and gene expression patterns of *HATs* in *S. constricta* (*ScHATs*) at different developmental stages, environmental stresses, and bacterial infection were analyzed and compared. A total of 11 *ScHAT* genes identified from *S. constricta* genome could be divided into three subfamilies (GNAT, MYST, and p300/CBP) in sequence homology. The protein structure prediction showed that *ScHAT* members belonging to each subfamily were conserved in protein domains and motifs. Gene expression analysis revealed that the expression levels of *ScHATs* was generally higher in the early developmental stage than that in the late stage. *ScKAT2B*, *ScKAT8*, *ScKAT9-2*, and *Scp300* were expressed differentially in gills against ammonia-N stress ($P<0.05$). *ScHATs* showed similar expression pattern in gills and hepatopancreas under thermal stress, and *Scp300* were significantly downregulated in hepatopancreas. Additionally, expressions of *ScHATs* increased dramatically after *Vibrio parahaemolyticus* infection ($P<0.05$). Therefore, *ScHATs* is an important class of genes in the development and immune defense of *S. constricta*. This study provides a theoretical foundation for further understanding the evolution and function of *HAT* gene family in mollusks.

Key words *Sinonovacula constricta*; histone acetyltransferase; gene family; environmental stress; bacterial infection; gene expression