

基于 RNA-seq 数据分析长牡蛎(*Crassostrea gigas*) 天然免疫的可变剪接事件*

李若晖^{1, 2, 3} 张琳琳^{1, 2, 3}

(1. 中国科学院海洋研究所实验海洋生物学重点实验室 山东青岛 266071; 2. 青岛海洋科学与技术试点国家实验室海洋生物学与生物技术功能实验室 山东青岛 266237; 3. 中国科学院大学 北京 100049)

摘要 大部分无脊椎动物缺乏基于经典抗体和记忆细胞的适应性免疫, 因而天然免疫在无脊椎动物免疫防御中发挥不可或缺的作用。可变剪接正是产生天然免疫多样性和特异性的的重要途径。牡蛎(*Crassostrea*)是全球范围内的大宗养殖贝类和我国产量最高的海水养殖贝类, 理解其天然免疫系统的多样性和特异性对牡蛎病害防治和养殖业健康可持续发展至关重要。通过对不同病原诱导以及诱导后不同时间的转录组进行生物信息分析, 系统地研究了长牡蛎天然免疫应激反应的可变剪接事件。发现在弧菌诱导下可变剪接事件的总数显著增加, 表明弧菌的感染会诱导长牡蛎可变剪接事件的产生。对病原诱导后可变剪接体的组成类型及表达量显著变化的基因的功能富集分析, 表明免疫系统相关功能被显著富集, 病原感染不同阶段以及不同病原感染后可变剪接体的组成类型和表达量均不一致, 表明长牡蛎免疫系统可通过可变剪接产生免疫响应的特异性和多样性。研究结果为长牡蛎等无脊椎动物天然免疫多样性和特异性提供了典型实例, 也为长牡蛎病害防御提供了理论支撑。

关键词 长牡蛎; 可变剪接; 病原体诱导; 天然免疫

中图分类号 Q789; S944.3; S968.3 doi: 10.11693/hyh20230700148

免疫系统是生物体为了抵御病原体入侵, 在长期进化过程中形成的一套复杂的防御系统, 它可以维持机体的内稳态(蒲新明, 2008)。免疫系统通过一系列免疫应答反应参与免疫调节(王金星等, 2004)。免疫系统可以分为天然免疫系统(Innate immune system)和获得性免疫系统(Adaptive immune system)两种类型, 两种类型对应着不同的作用机制(李颖, 2011)。普遍认为无脊椎动物只存在非特异性的天然免疫, 缺乏获得性免疫(Litman *et al*, 2005)。无脊椎动物应对病原微生物的入侵主要依赖天然免疫系统。为应对各种各样的病原微生物入侵, 基因组需要通过不同的调控产生特定转录组和蛋白质组, 从而完成对病原体入侵做出有效和快速的天然免疫反应。

在漫长的自然进化过程中无脊椎动物进化出一

套功效强大的先天防御体系, 通过细胞代谢的重编程和免疫系统的激活使它们能够有效地应对病原体入侵。许多细胞过程的正常进行在很大程度上取决于 RNA 代谢, 特别是 RNA 的成熟和降解。在真核生物中, 利用不同的剪接位点, 将相同的前体 mRNA 转化成多个信使 RNA (mRNA)的过程, 被称作可变剪接(Alternative splicing, AS) (Kahles *et al*, 2016)。可变剪接通过使基因产生结构和功能不同的 mRNA 和蛋白质, 从而增强转录组可塑性和蛋白质功能多样性(Maniatis *et al*, 2002; Nilsen *et al*, 2010)。可变剪接可以从数量相对较少的基因中产生大量的 mRNA 和蛋白质异构体。可变剪接还可以改变转录本及其编码蛋白的结构, 从而调节蛋白质的功能, 而可变剪接产生的异构体甚至可能具有完全相反的功能(Stamm *et al*,

* 国家重点研发计划, 2022YFD2401400 号; 中国科学院战略性先导科技专项, XDB42000000 号; 国家自然科学基金, 41976088 号。李若晖, 硕士研究生, E-mail: liruohui20@mails.ucas.ac.cn

通信作者: 张琳琳, 博士生导师, 研究员, E-mail: linlinzhang@qdio.ac.cn

收稿日期: 2023-07-14, 收修改稿日期: 2023-08-13

2005)。与哺乳动物中大量关于可变剪接的研究相比,这一机制在无脊椎动物天然免疫中的作用和调控却知之甚少。

近年来,高通量技术和生物信息学工具的快速发展,特别是二代测序技术,如 Illumina、PacBio 和 Oxford Nanopore 直接 RNA 测序,对研究基因表达变化和分析全基因组可变剪接事件做出巨大贡献。目前,利用 RNA-Seq 技术分析了包括人(*Homo sapiens*)、线虫 (*Caenorhabditis elegans*)、果蝇 (*Drosophila melanogaster*)、拟南芥(*Arabidopsis thaliana*)等多个物种基因组的 AS 事件(Pan *et al.*, 2008; Graveley *et al.*, 2011; Ramani *et al.*, 2011; Marquez *et al.*, 2012)。利用生物信息学工具进一步深入挖掘转录组学数据来识别可变剪接事件,将有助于阐明无脊椎动物在生物应激条件下的基因调控机制。

长牡蛎(*Crassostrea gigas*)既是世界上重要的经济贝类,也是一种常见的模式无脊椎动物,已被广泛用于研究(Zhang *et al.*, 2012, 2015)。已有研究表明,AS 在长牡蛎响应高温、高盐和空气暴露等非生物应激中发挥重要作用(Huang *et al.*, 2016),而 AS 在长牡蛎生物应激中的功能研究仍处于空白状态。因此,本研究基于 Illumina 高通量测序的转录组数据(Zhang *et al.*, 2015),分析了长牡蛎在不同病原体和混合弧菌不同感染时间诱导下的可变剪接差异,发现混合弧菌诱导可变剪接事件的增加,分析了在不同病原体和混合弧菌诱导中差异表达的可变剪接基因的功能,从而探究可变剪接对长牡蛎天然免疫多样性和特异性的调控作用。本研究为长牡蛎等无脊椎动物天然免疫多样性和特异性的深入研究提供了理论依据,为后续长牡蛎病害防御提供了数据支持。

1 材料与方法

1.1 数据收集与处理

收集来自 NCBI 的 13 个长牡蛎测序数据,此数据由中国科学院海洋研究所 2013 年上传发布,登录号为 PRJNA194079 和 PRJNA194084。具体来说,测序样品来自:(1) 5 个混合弧菌感染(感染后 0、6、12、24 和 48 h 收集的感染组的鳃);(2) 8 个不同致病菌[4 个革兰氏阳性菌:河口弧菌(*Vibrio aestuarianus*)、鳃弧菌 (*Vibrio anguillarum*)、溶藻弧菌 (*Vibrio alginolyticus*)和 LPS, 1 个革兰氏阴性菌:藤黄微球菌 (*Micrococcus luteus*), PBS 和对照组的鳃]。

为了保证信息分析质量,对原始数据利用 FastQC

质控,使用 Trimmomatic 去除原始数据中带接头的、低质量的以及带 N 碱基的序列,得到 Clean reads 数据。

1.2 可变剪接事件的鉴定

将 Clean reads 与长牡蛎基因组利用软件 STAR 进行比对,比对结果运用 Cufflinks 软件(Trapnell *et al.*, 2012)进行组装,并使用 Cuffcompare 与长牡蛎参考注释进行比较生成 tmap 文件(Trapnell *et al.*, 2010)。使用 ASprofile 中的“extract-as”程序(Florea *et al.*, 2013),利用 gtf、tmap 文件在多个样本中鉴定 12 种可变剪接事件,包括单外显子跳跃(skipped exon, SKIP)、单外显子跳跃(模糊边界) (approximate SKIP, XSKIP)、多外显子跳跃(multi-exon SKIP, MSKIP)、多外显子跳跃(模糊边界) (approximate MSKIP, XMSKIP)、单内含子滞留(intron retention, IR)、单内含子滞留(模糊边界) (approximate IR, XIR)、多内含子滞留(multi-IR, MIR)、多内含子滞留(模糊边界)(approximate MIR, XMIR)、可变 5'或 3'端剪切(alternative exonends, AE)、可变 5'或 3'端剪切(模糊边界) (approximate AE, XAE)、第 1 个外显子可变剪切(alternative 5' first exon, TSS)、最后 1 个外显子可变剪切(alternative 3' last exon, TTS)。

1.3 先天免疫相关基因分析及 GO 功能富集分析

使用 cuffdiff 程序对混合弧菌不同感染时间和不同病原体样本进行不同表达基因分析(Trapnell *et al.*, 2013)。然后使用参数 $|\log_2(\text{fold_change})|>1$ 和 $P < 0.05$ 和 FDR < 0.05 对生成的差异表达基因数据进行过滤。

利用 DAVID 数据库,对显著差异表达的可变剪接基因 ID 进行 GO 分类,获取 GO 编号,并利用 cluster Profiler 对其进行功能富集,以 $P \leq 0.05$ 作为显著富集标准。

1.4 预测的长牡蛎 AS 基因亚型的表达分析

根据 Cuffnorm 利用默认参数对每个样本中鉴定出的异构体进行定量。通过 Cuffdiff 计算混合弧菌感染条件下 FPKM 的转录表达量,然后对所有样本进行归一化处理。

2 结果与分析

2.1 长牡蛎感染混合弧菌后不同时间可变剪接事件的鉴定

为了研究混合弧菌感染下不同时间长牡蛎可变剪接的作用,从 NCBI 数据库下载了 5 个长牡蛎 Illumina 高通量测序平台的转录组测序数据。这些数据来自 5 个混合弧菌不同感染时间的 RNA-seq 读

数, 包括 0、6、12、24 和 48 h, 数据集的总碱基读数为 15.27 Gb。

根据 5 个 NCBI RNA-seq 数据集, 计算了 5 个不同感染时间的 AS 事件的数量(表 1), 其中, 感染 48 h 后有最丰富的 AS 事件(80 566), 而感染 6 h 后的数量最少(74 573)(图 1)。在经混合弧菌感染后, 长牡蛎发生的可变剪接事件类型仍以 TSS、TTS、AE 和 SKIP 为主, 结果表明, 可变剪接在各时间点的数目存在差异, 随着时间的增长, 可变剪接事件也随着增多。

表 1 长牡蛎混合弧菌入侵不同时间点的可变剪接事件
Tab.1 Alternative splicing events of *C. gigas* under *Vibrio* challenge at different time points

组别	0 h	6 h	12 h	24 h	48 h
AE	204	152	165	140	138
XAE	32	22	21	12	20
IR	16	23	16	39	26
XIR	12	11	11	17	10
MIR	1	1	1	2	1
XMIR	2	1	3	0	0
SKIP	184	174	172	146	154
XSKIP	24	38	21	23	17
MSKIP	9	7	10	12	12
XMSKIP	0	0	3	2	1
TSS	38 641	36 931	37 469	39 675	40 022
TTS	38 941	37 213	37 651	39 863	40 165

2.2 长牡蛎感染不同病原体可变剪接事件的鉴定

为了研究不同病原体长牡蛎可变剪接的作用, 从 NCBI 数据库下载了 7 个长牡蛎 Illumina 高通量测序平台的转录组测序数据。这些数据来自不同病原

体入侵的长牡蛎的 RNA-seq 读数, 包括 LPS、河口弧菌(*Vibrio aestuarianus*)、鳃弧菌(*Vibrio anguillarum*)、溶藻弧菌(*Vibrio alginolyticus*)、塔氏弧菌(*Vibrio tubiashii*)、藤黄微球菌(*Micrococcus luteus*)。

根据 7 个 NCBI RNA-seq 数据集, 计算了 7 个不同病原体诱导的 AS 事件的数量(表 2)。结果表明, 不同病原体诱导后长牡蛎的可变剪接的总数目存在差异, 但长牡蛎发生的可变剪接事件类型仍为 TSS、TTS、AE 和 SKIP 四种主要类型(图 2)。

2.3 可变剪接事件差异表达分析

对感染混合弧菌的长牡蛎不同时间的可变剪接基因进行差异表达分析, 分别得到 528、330、302 和 313 个显著差异表达基因。对不同病原体诱导时发生可变剪接的基因进行差异表达分析, 分别得到 528、330、302 和 313 个显著差异表达基因。在这些显著差异基因中, TSS 和 TTS 是主要的可变剪接方式, 这表明它们在响应弧菌的基因表达中可能起着重要的作用。

发生可变剪接事件的差异表达基因有多个基因为天然免疫相关基因家族成员, 混合弧菌不同感染时间发生可变剪接的显著差异表达基因中分别有 38、28、30 和 31 个天然免疫基因(表 3), 不同病原体入侵发生可变剪接的显著差异表达基因中分别有 25、26、25、23、29 和 29 个天然免疫基因(表 4), 这些基因家族包括 *RLR*、*TLR*、*SRCR*、*Clq* 等。但是混合弧菌不同感染时间发生可变剪接的天然免疫相关基因分别有 450、466、482 和 501 个, 占长牡蛎天然免疫相关基因总数的三分之一左右, 说明在病原体感染时, 机体中三分之一的天然免疫基因通过可变剪接发挥作用。

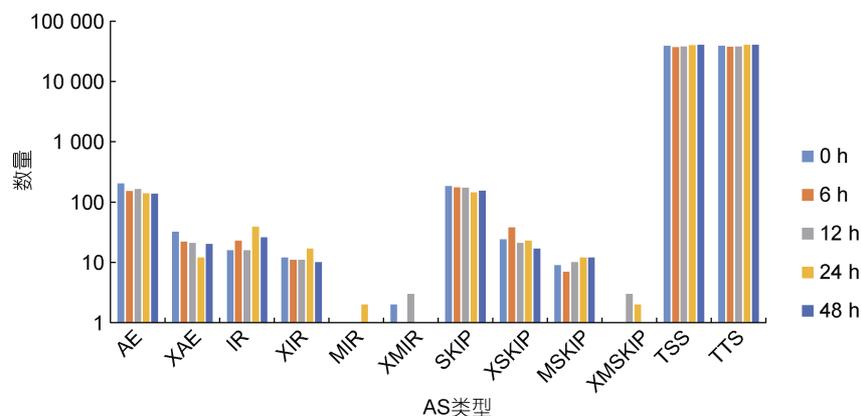


图 1 长牡蛎在混合弧菌胁迫下不同时间点的可变剪接事件

Fig.1 Alternative splicing events of *C. gigas* under *Vibrio* challenge at different time points

表 2 长牡蛎不同病原体入侵下的可变剪接事件
Tab.2 Alternative splicing events of *C. gigas* challenged with different types of bacteria

组别	PBS	LPS	河口弧菌	鳗弧菌	溶藻弧菌	塔式弧菌	藤黄微球菌
AE	152	156	122	150	162	139	175
XAE	28	20	28	30	32	24	26
IR	8	21	19	13	17	15	19
XIR	8	12	8	10	10	11	14
MIR	3	2	2	3	1	2	2
XMIR	1	2	0	1	0	0	0
SKIP	156	138	175	169	175	164	172
XSKIP	31	11	28	24	23	18	25
MSKIP	16	7	7	10	12	16	13
XMSKIP	3	0	1	4	2	1	3
TSS	36 477	35 159	35 767	36 792	35 327	34 901	36 580
TTS	36 716	35 396	36 002	37 050	35 688	35 144	36 874

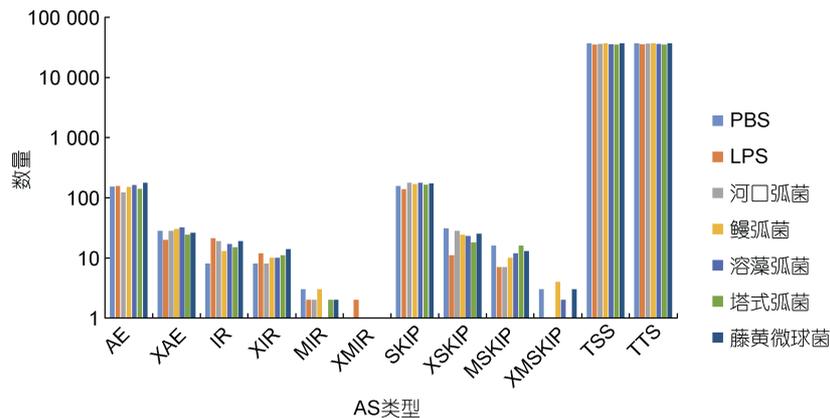


图 2 长牡蛎不同病原体入侵下的可变剪接事件

Fig.2 Alternative splicing events of *C. gigas* challenged with different types of bacteria

2.4 差异可变剪接基因的 GO 富集分析

为了解可变剪接对长牡蛎先天免疫的调控,对弧菌诱导的 5 个不同时间基因组中发生 AS 事件的显著差异表达基因进行了 GO 富集,以 $P < 0.05$ 为显著富集标准,发现得到 143、179、153 和 164 个条目(图 3)。

在 0 h 和 6 h 的差异基因中,生物学过程、细胞构成和分子功能中显著富集的 GO 条目分别有 36、35、72 个。在生物学过程中,发生可变剪接的差异基因主要集中于有机物的分解代谢过程(GO: 1901575)等代谢相关过程;在分子功能中,较多可变剪接基因被富集于催化活性(GO: 0003824)相关 GO 项中;在细胞构成中,发生可变剪接的基因主要集中于膜的固有成分(GO: 0031224)与膜相关条目。

在 0 h 和 12 h 的差异基因中,在生物学过程、细胞构成和分子功能中显著富集的 GO 条目分别有 101、20、58 个。在生物学过程中,发生可变剪接的差

异基因主要集中于 α -氨基酸代谢过程(GO: 1901605)等代谢相关过程;在分子功能中,较多可变剪接基因被富集于辅助因子结合(GO: 0048037)相关 GO 项中;在细胞构成中,发生可变剪接的基因主要集中于膜的固有成分(GO: 0031224)与膜相关条目。

在 0 h 和 24 h 的差异基因中,在生物学过程、细胞构成和分子功能中显著富集的 GO 条目分别有 101、10、72 个。在生物学过程中,发生可变剪接的差异基因主要集中于有机物的分解代谢过程(GO: 1901575)等代谢相关过程;在分子功能中,较多可变剪接基因被富集于催化活性(GO: 0003824)相关 GO 项中;在细胞构成中,发生可变剪接的基因主要集中于膜的固有成分(GO: 0031224)与膜相关条目。

在 0 h 和 48 h 的差异基因中,在生物学过程、细胞构成和分子功能中显著富集的 GO 条目分别有 112、13、39 个。在生物学过程中,发生可变剪接的

表3 弧菌诱导发生可变剪接的差异基因中的天然免疫相关基因

Tab.3 Natural immunity-related genes of differentially expressed alternative splicing genes under *Vibrio* challenge at different time points

基因名称	6 h	12 h	24 h	48 h
<i>a2M</i>	0	1	0	0
<i>C1q</i>	6	5	8	6
<i>C-lectin</i>	8	7	5	5
<i>FREP</i>	3	2	2	2
<i>IKB</i>	1	1	1	1
<i>MBL</i>	1	1	1	1
<i>Galectin</i>	0	0	1	0
<i>MITA_id</i>	2	0	0	1
<i>NLR</i>	3	3	3	3
<i>SRCR</i>	3	2	2	4
<i>TNFR</i>	0	1	1	0
<i>CD36</i>	1	0	0	0
<i>NF-kb</i>	1	0	0	0
<i>SARM</i>	1	0	0	0
<i>TNF</i>	1	0	0	0
<i>LRR-IG</i>	0	0	0	1
<i>SOD</i>	0	0	0	0
<i>TRAF</i>	1	0	0	0
others	6	5	6	7
总计	38	28	30	31

表4 不同病原体入侵发生可变剪接的显著表达差异基因中的天然免疫相关基因

Tab.4 Natural immunity-related genes of differentially expressed alternative splicing genes challenged with different types of bacteria

基因名称	LPS	<i>V. aes</i>	<i>V. ang</i>	<i>V. alg</i>	<i>V. tub</i>	<i>M. lut</i>
<i>C1q</i>	4	5	6	3	4	6
<i>C-lectin</i>	4	5	4	4	5	5
<i>FREP</i>	3	1	1	1	1	2
<i>IKB</i>	1	1	2	1	1	1
<i>MBL</i>	1	1	1	1	1	1
<i>Galectin</i>	0	0	0	0	1	0
<i>orphan</i>	0	0	0	2	0	1
<i>MITA_id</i>	0	1	0	0	1	1
<i>NLR</i>	3	3	3	3	3	3
<i>SRCR</i>	3	2	3	2	4	3
<i>SARM-like</i>	0	1	0	0	1	1
<i>SOD</i>	0	0	0	0	1	0
<i>TRAF</i>	0	0	0	0	1	0
Others	6	6	5	6	5	5
总计	25	26	25	23	29	29

差异基因主要集中于单一有机体代谢过程(GO: 0044710)等代谢相关过程; 在分子功能中, 较多可变剪接基因被富集于催化活性(GO: 0003824)相关 GO 项中; 在细胞构成中, 发生可变剪接的基因主要集中于膜的固有成分(GO: 0031224)与膜相关条目。

由此可见, 这些差异基因主要包括以下主要生物过程: (1) 细胞和代谢过程; (2) 生物过程的调控; (3) 细胞分泌物的产生; (4) 生物黏附过程; (5) 对刺激反应响应先天性免疫系统的启动; (6) 先天性免疫信号转导; (7) 细胞的生长等。其中与免疫相关的比例占据较小, 细胞和代谢过程以及生物调节过程的比例占据较高。这说明在长牡蛎生物应激过程中, 细胞、代谢过程、催化活性和生物调节等途径可能参与机体自身的免疫防御, 与机体的免疫途径共同响应病原体感染。

2.5 差异可变剪接基因的 KEGG 富集分析

对混合弧菌入侵不同时间发生可变剪接的差异表达基因进行 KEGG 富集分析(图 4), 结果表明富集到与免疫相关通路, 在 12 h 富集到细胞凋亡(ko04210)、CD 分子(ko04090)、信号分子与相互作用(ko09133)、细胞生长和死亡(ko09143)等; 在 24 h 富集到 NOD-like 受体信号通路(ko04621)和免疫系统(ko09151)等。

对不同病原体发生可变剪接的差异表达基因进行 KEGG 富集分析(图 5), 结果表明富集到与免疫相关通路, 在 LPS 富集到 PI3K-AKT 通路(ko04151)等; 在 *V. aes* 富集到 NF-kappaB 细胞信号传导通路(ko04064); 在 *V. ang* 富集到 Toll 与 Imd 信号通路(ko04624)、NOD 样受体信号通路(ko04621)、免疫系统(ko09151)和 NF-kappaB 细胞信号传导通路(ko04064); 在 *V. alg* 富集到 NOD 样受体信号通路(ko04621); 在 *V. tub* 富集到免疫系统(ko09151); 在 *M. lut* 富集到 NOD 样受体信号通路(ko04621)、Toll 与 Imd 信号通路(ko04624)。

综合差异表达筛选、GO 富集分析和 KEGG 富集分析发现, 在混合弧菌感染 6 h 时富集到代谢相关通路(图 4a), 推测在感染初期, 机体对能量的需求增加, 迫使体内养分重新分配, 积极提供应激所需的能量物质和氨基酸; 在混合弧菌感染 12 h 时富集到细胞凋亡相关通路(图 4b), 推测在感染中期, 由于混合弧菌对长牡蛎鳃组织的急性感染, 导致鳃细胞出现凋亡, 以至于在感染 6 h、12 h 发生可变剪接事件总个数出现减少; 在混合弧菌感染 24 h 时富集到免疫相关通路、48 h 时富集到异源物质代谢降解相关通路

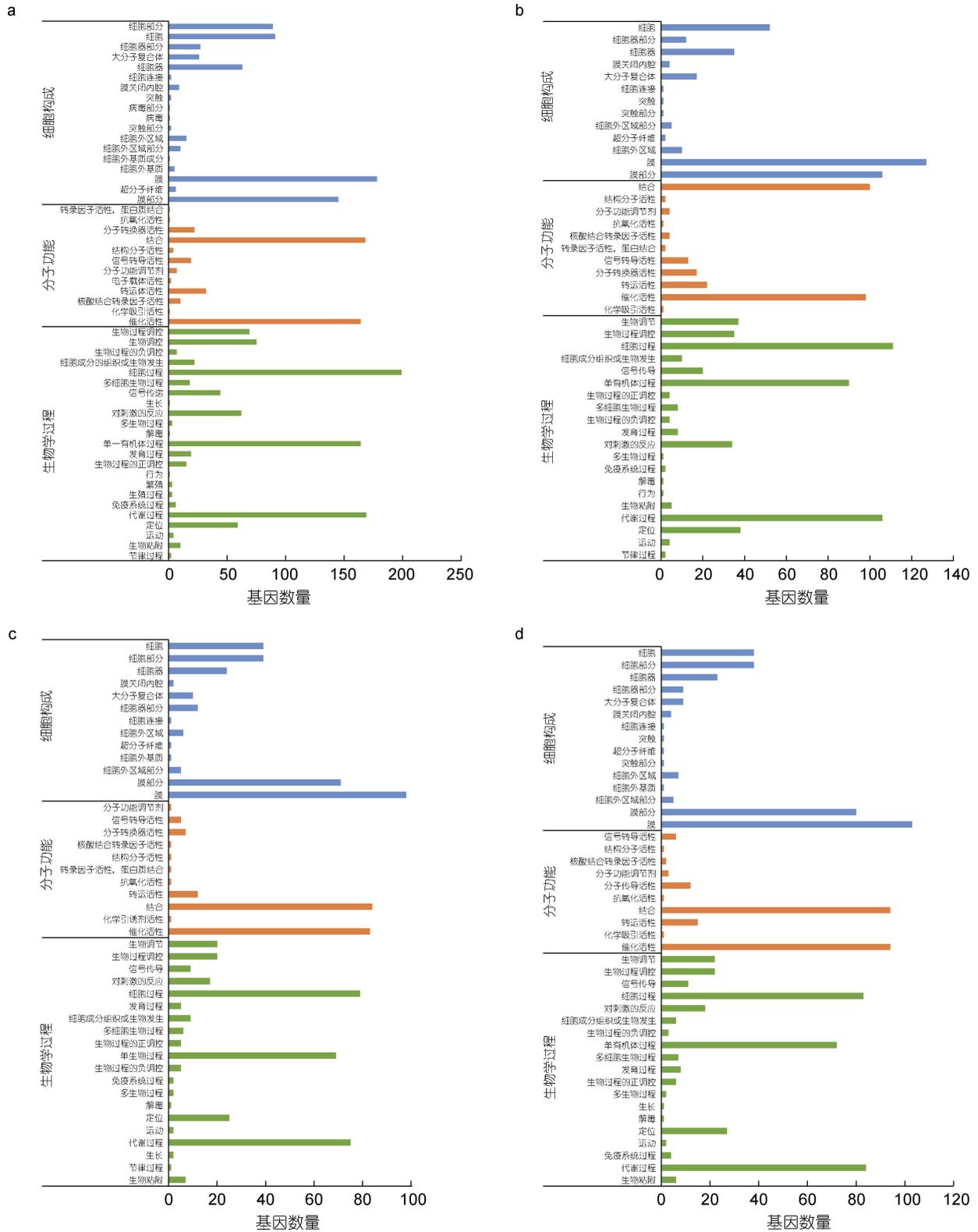


图 3 弧菌诱导下不同时间差异表达可变剪接基因 GO 富集

Fig.3 Statistics of GO annotations for alternative splicing genes with variable expression in response to *Vibrio* challenge at different points
注: a. 6 h 差异可变剪接基因 GO 富集结果; b. 12 h 差异可变剪接基因 GO 富集结果; c. 24 h 差异可变剪接基因 GO 富集结果; d. 48 h 差异可变剪接基因 GO 富集结果

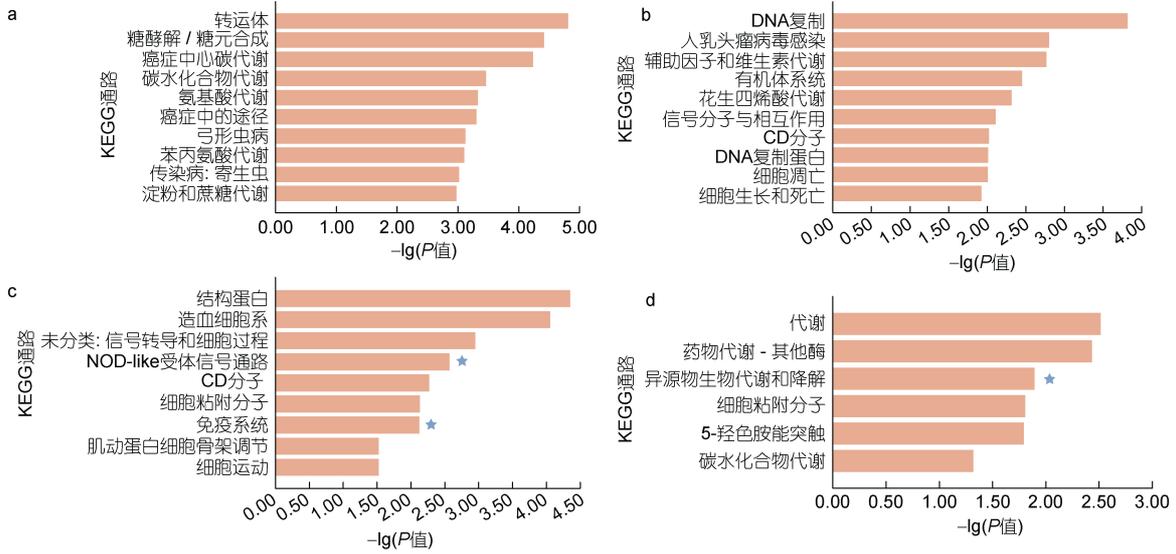


图 4 弧菌诱导下不同时间差异表达可变剪接基因 KEGG 富集

Fig.4 Statistics of KEGG annotations for alternative splicing genes with variable expression in response to *Vibrio* challenge at different points

注: a. 6 h 差异可变剪接基因 KEGG 富集结果; b. 12 h 差异可变剪接基因 KEGG 富集结果; c. 24 h 差异可变剪接基因 KEGG 富集结果; d. 48 h 差异可变剪接基因 KEGG 富集结果; 星号表示免疫相关通路

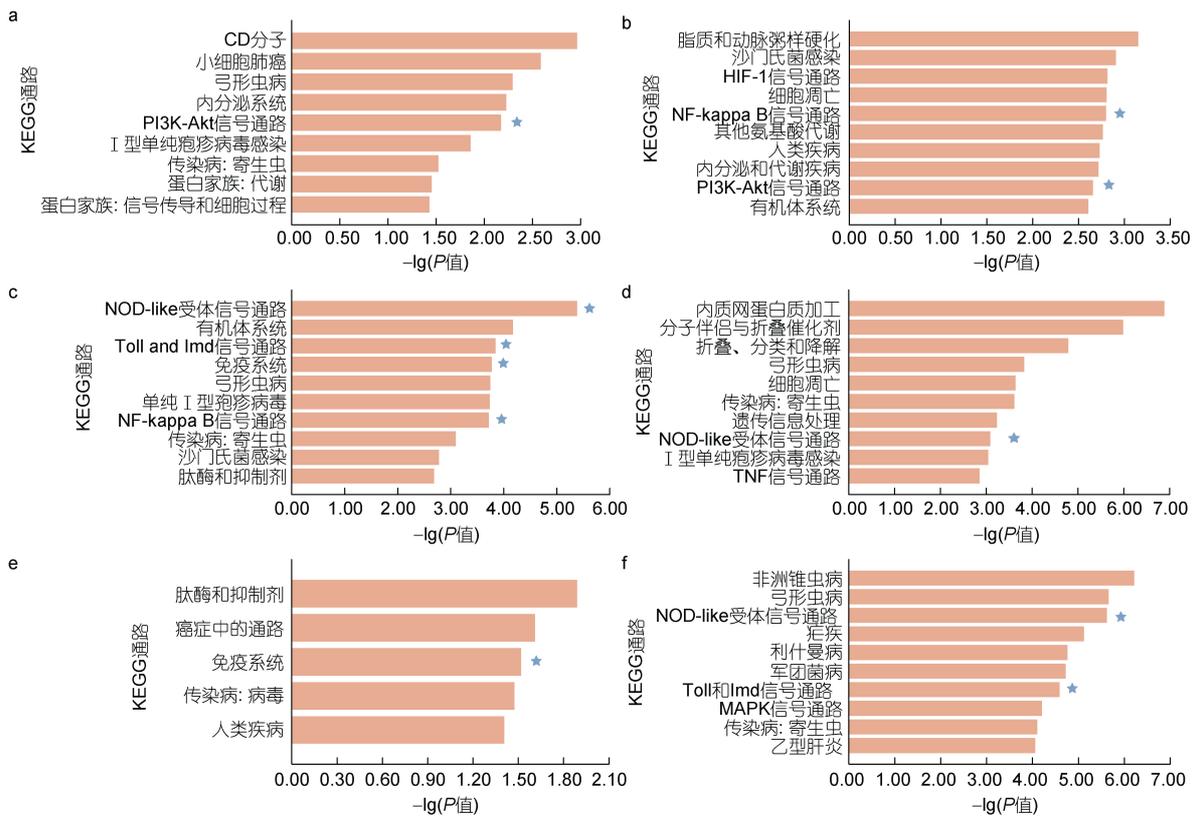


图 5 不同病原体诱导下差异表达可变剪接基因 KEGG 富集

Fig.5 Statistics of KEGG annotations for alternative splicing genes with variable expression in response to various *Vibrio* challenges
注: a. LPS 差异可变剪接基因 KEGG 富集结果; b. *V. aes* 差异可变剪接基因 KEGG 富集结果; c. *V. ang* 差异可变剪接基因 KEGG 富集结果; d. *V. alg* 差异可变剪接基因 KEGG 富集结果; e. *V. tub* 差异可变剪接基因 KEGG 富集结果; f. *M. lut* 差异可变剪接基因 KEGG 富集结果; 星号表示免疫相关通路

(图 4c, 4d), 推测随着感染时间推移, 鳃组织的天然免疫系统发挥作用, 消除了混合弧菌的急性感染, 并且及时清除机体内由于病原体入侵产生的异物, 逐步恢复鳃组织的免疫稳态。

2.6 混合弧菌胁迫下先天免疫基因的可变剪接事件异构体变化

天然免疫在进化上是一个相对保守的系统, 主

要通过 PRRs 识别 PAMPs 进行信号传递。RLRs 是重要的 RNA 传感器分子。对两个 RLR 基因的分析显示, Cg02656 和 Cg26493 基因都拥有六个可变剪接异构体。图 6a 和图 7a 显示了这两个基因的异构体经历了多种可变剪接事件, 表明可变剪接的调节可能非常复杂和多样, 例如, 异构体 Cg02656E 出现了 TSS、IR 和 TTS。另外, 为了探索特定可变剪接异构体与致病

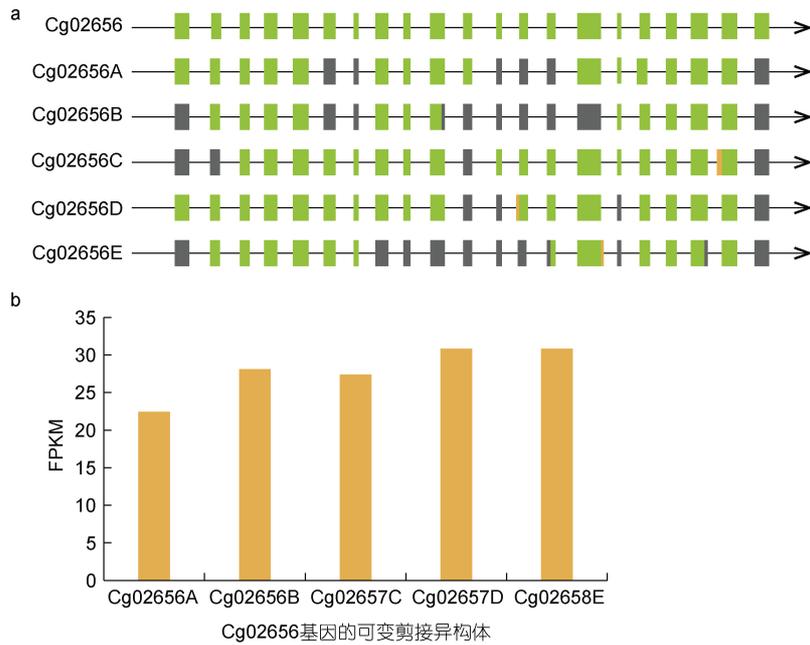


图 6 Cg02656 基因的 AS 异构体(a)及其 FPKM 表达量(b)

Fig.6 The AS isoforms of the Cg02656 gene (a) and their FPKM expressions (b)

注: Cg02656A、B、C、D、E 分别为 0、6、12、24、48 h 所对应的 Cg02656 基因异构体; FPKM: 每千个碱基的转录每百万映射读取的 fragments (Fragments Per Kilobase of exon model per Million mapped fragments), 是基因相对表达量的衡量指标

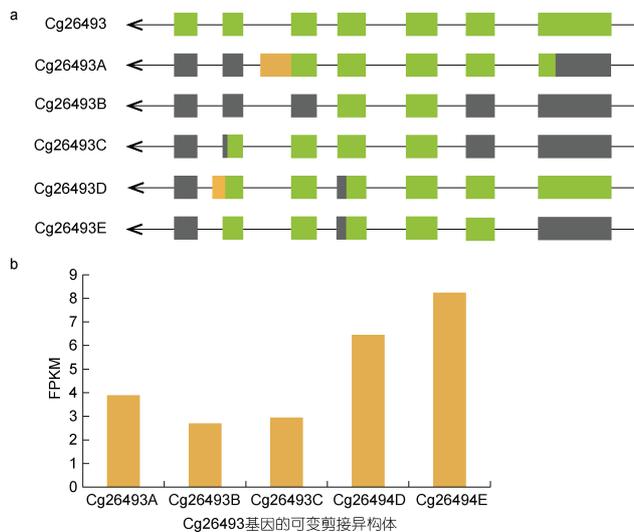


图 7 Cg26493 基因的 AS 异构体(a)及其 FPKM 表达量(b)

Fig.7 The AS isoforms (a) of the Cg26493 gene and their FPKM expressions (b)

注: Cg26493A、B、C、D、E 分别为 0、6、12、24、48 h 所对应的 Cg26493 基因异构体

菌胁迫之间的关系, 计算了不同感染时间可变剪接转录本的表达量(图 6b 和图 7b)。RNA-seq 数据的 FPKM 值表明, Cg02656 和 Cg26493 基因的可变剪接异构体表达量在不同的感染时间存在差异表达。Cg02656 和 Cg26493 基因的表达量均在 48 h 后达到最高值, 其中 Cg02656 基因在 48 h 时的表达量约是感染 0 h 时的 2 倍。

基于这些发现, 我们认为长牡蛎可能具有复杂的可变剪接调节系统, 并且可变剪接可能是长牡蛎产生天然免疫多样性和特异性的重要途径。例如, 这些异构体可能作为 Cg02656 和 Cg26493 编码的完整蛋白的调节因子, 直接参与免疫反应或其他重要的生物学功能。不同感染时间中异构体表达的差异也表明可变剪接调控在不同免疫过程之间存在差异。

3 讨论

机体的免疫是在长期进化过程中形成的, 免疫机制复杂性与各种生物大分子的种类与数量有关, 而在有限长度的遗传信息基础上产生更多的转录本, 无疑能在转录水平上大幅度提高免疫反应的效率, 进而实现对机体最大程度的保护。目前已有大量研究指出许多动物的天然免疫与可变剪接高度相关(Tan *et al.*, 2019; Guo *et al.*, 2022), 这种转录后水平上的调控与免疫应答中多种关键因子的生物合成过程密不可分。此外, 这种调控方式也会造成物种之间代谢通路的特异性。

本研究对感染混合弧菌后长牡蛎的鳃组织中的可变剪接事件进行检测。随着时间的增长, 可变剪接事件的数量也均呈现增加的趋势, 表明混合弧菌的感染会诱导长牡蛎可变剪接事件的产生。关于胁迫反应引起可变剪接事件数量增加的现象在长牡蛎中已有相关报道。Huang 等(2016)的研究结果表明在长牡蛎在高温、高盐、空气暴露胁迫下, 可变剪接事件的数量都有所增加。由此可见, 可变剪接事件的产生可能是长牡蛎响应各种胁迫反应的一种重要调控机制。在本研究中, 12 种类型的可变剪接事件均被鉴定出, 其中可变剪接事件数量最多的是 TSS 和 TTS, 其次是 AE 和 SKIP。

本研究识别了感染混合弧菌前后长牡蛎鳃组织中的发生可变剪接的差异表达基因, 并对这些基因进行了富集分析。其中 GO 的富集分析结果主要包含天然免疫应答、代谢过程。值得注意的是, 长牡蛎感染混合弧菌后 12 h 和 24 h 发生可变剪接的差异表达

基因富集到了免疫系统通路上, 但在 6 h 的样本中并未发现此类富集, 这表明长牡蛎感染后可变剪接事件不会迅速响应。天然免疫应答是非特异性的、机体固有的、反应迅速的、广泛发生的免疫方式, 表明在感染混合弧菌后长牡蛎主要通过天然免疫应答清除入侵病原菌, 对机体进行保护。有机物代谢过程等 GO 词条被认为参与了机体能量代谢过程, 表明免疫应答过程会进行大量能量消耗与物质代谢。在 KEGG 富集分析结果中, 发生可变剪接的差异表达基因被显著富集在细胞凋亡、NOD-like 受体信号通路和免疫系统信号通路上。细胞凋亡是为了维持内环境的稳定, 由基因控制的非被动的细胞死亡, 并不只在机体受到胁迫时才发挥作用, 也是正常细胞代谢的生理现象之一(Elmore, 2007)。细胞凋亡也是免疫系统的一种作用方式。Nagata 等(2017)的研究表明, 许多凋亡调节因子均在免疫系统中发挥着重要作用。NOD-like 受体(NLR)信号通路是生物体的一种重要免疫防御机制。作为天然免疫系统的关键组成部分之一, NLR 信号通路能够识别并响应各种外来病原体、细胞损伤和自身核酸等危险信号, 从而启动免疫防御反应, 维护机体的健康稳定(Van Gorp *et al.*, 2014)。NLR 信号通路主要包括模式识别受体(PRR)、炎症小体、caspase-1 及相关炎性因子等多个组成部分, 这些部分在免疫防御反应中相互协作, 发挥着重要的作用。

本研究识别了不同病原体入侵长牡蛎鳃组织中的发生可变剪接的差异表达基因, 并对这些基因进行了富集分析。在 KEGG 富集分析结果中, 发生可变剪接的差异表达基因被显著富集在 PI3K-AKT 通路、NOD-like 受体信号通路、NF-kappaB 细胞信号传导通路和 Toll 与 Imd 信号通路等信号通路上。核转录因子 κ B (NF- κ B)是转录因子蛋白家族中的一员, 在免疫系统中发挥非常重要的调节作用(Hayden *et al.*, 2006)。NF- κ B 可以调控细胞因子、生长因子以及效应蛋白酶的表达, 从而参与到免疫应答中。在哺乳动物中, Toll 样受体(TLRs)直接识别病原微生物的病原体相关分子模式, 通过 NF- κ B 触发信号反应, 启动先天免疫和适应性免疫反应(Takeda *et al.*, 2003)。Toll 信号途径最先被发现于果蝇, 它主要与机体对真菌和革兰氏阳性细菌的免疫反应有关(Kim *et al.*, 2003; Kaneko *et al.*, 2005)。果蝇 Toll 蛋白能与多种 PMAPs 结合, 从而活化胞内 Tube、Pele 等多种胞内信号分子, 进而调控免疫相关基因的表达。Imd 信号转导途径是一条与机体响应革兰氏阴性细菌的免疫反应相关的

信号转导途径, 其在果蝇体内可调控 220 种基因的表达, 其中就有革兰氏抗阴性细菌的多肽(Romeo *et al*, 2008)。最近的研究发现, Imd 信号通路还可以激活 JNK (c-Jun N-terminal kinase), 实现细胞骨架蛋白的诱导表达。这表明 Imd 信号转导途径不仅在天然免疫应答中, 而且也在组织修复中起着关键作用。

4 结论

本研究基于 RNA-seq 测序数据, 鉴定了长牡蛎病原体诱导过程中发生的可变剪接事件类型及数目, 筛选并分析在不同病原体和混合弧菌诱导中差异表达的可变剪接基因的功能, 从而探究可变剪接对长牡蛎天然免疫多样性和特异性的调控作用。本研究为长牡蛎等无脊椎动物天然免疫多样性和特异性的深入研究提供了理论依据, 为后续长牡蛎病害防御提供了数据支持。

致谢 本文获得中国科学院海洋研究所海洋大数据中心的支持, 谨致谢忱。

参 考 文 献

- 王金星, 赵小凡, 2004. 无脊椎动物先天免疫模式识别受体研究进展[J]. 生物化学与生物物理进展, 31(2): 112-117.
- 李颖, 2011. 细胞抗病毒天然免疫信号转导的调控机制[D]. 武汉: 武汉大学.
- 蒲新明, 2008. 职业环境与机体免疫[J]. 职业与健康, 24(10): 985-987.
- ELMORE S, 2007. Apoptosis: A review of programmed cell death [J]. Toxicologic Pathology, 35(4): 495-516.
- FLOREA L, SONG L, SALZBERG S L, 2013. Thousands of exon skipping events differentiate among splicing patterns in sixteen human tissues [J]. F1000 Research, 2: 188.
- GRAVELEY B R, BROOKS A N, CARLSON J W, *et al*, 2011. The developmental transcriptome of *Drosophila melanogaster* [J]. Nature, 471(7339): 473-479.
- GUO Y N, SUN Z D, ZHANG Y T, *et al*, 2022. Molecular identification and function characterization of four alternative splice variants of trim25 in Japanese flounder (*Paralichthys olivaceus*) [J]. Fish & Shellfish Immunology, 120: 142-154.
- HAYDEN M S, WEST A P, GHOSH S, 2006. NF- κ B and the immune response [J]. Oncogene, 25(51): 6758-6780.
- HUANG B, ZHANG L, TANG X, *et al*, 2016. Genome-wide analysis of alternative splicing provides insights into stress adaptation of the Pacific oyster [J]. Marine Biotechnology, 18: 598-609.
- KAHLES A, ONG C S, ZHONG Y, *et al*, 2016. SplAdder: identification, quantification and testing of alternative splicing events from RNA-Seq data [J]. Bioinformatics, 32(12): 1840-1847.
- KANEKO T, GOLENBOCK D, SILVERMAN N, 2005. Peptidoglycan recognition by the *Drosophila* Imd pathway [J]. Journal of Endotoxin Research, 11(6): 383-389.
- KIM M S, BYUN M, OH B H, 2003. Crystal structure of peptidoglycan recognition protein LB from *Drosophila melanogaster* [J]. Nature Immunology, 4(8): 787-793.
- LITMAN G W, CANNON J P, DISHAW L J, 2005. Reconstructing immune phylogeny: new perspectives [J]. Nature Reviews Immunology, 5(11): 866-879.
- MANIATIS T, TASIC B, 2002. Alternative pre-mRNA splicing and proteome expansion in metazoans [J]. Nature, 418(6894): 236-243.
- MARQUEZ Y, BROWN J W S, SIMPSON C, *et al*, 2012. Transcriptome survey reveals increased complexity of the alternative splicing landscape in *Arabidopsis* [J]. Genome Research, 22(6): 1184-1195.
- NAGATA S, TANAKA M, 2017. Programmed cell death and the immune system [J]. Nature Reviews Immunology, 17(5): 333-340.
- NILSEN T W, GRAVELEY B R, 2010. Expansion of the eukaryotic proteome by alternative splicing [J]. Nature, 463(7280): 457-463.
- PAN Q, SHAI O, LEE L J, *et al*, 2008. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing [J]. Nature Genetics, 40(12): 1413-1415.
- RAMANI A K, CALARCO J A, PAN Q, *et al*, 2011. Genome-wide analysis of alternative splicing in *Caenorhabditis elegans* [J]. Genome Research, 21(2): 342-348.
- ROMEO Y, LEMAITRE B, 2008. *Drosophila* immunity [M] // EWBank J, VIVIER E. Innate Immunity. Totowa, NJ: Humana Press: 379-394.
- STAMM S, BEN-ARI S, RAFALSKA I, *et al*, 2005. Function of alternative splicing [J]. Gene, 344: 1-20.
- TAKEDA K, KAISHO T, AKIRA S, 2003. Toll-like receptors [J]. Annual Review of Immunology, 21(1): 335-376.
- TAN S X, WANG W W, TIAN C X, *et al*, 2019. Post-transcriptional regulation through alternative splicing after infection with *Flavobacterium columnare* in channel catfish (*Ictalurus punctatus*) [J]. Fish & Shellfish Immunology, 91: 188-193.
- TRAPNELL C, HENDRICKSON D G, SAUVAGEAU M, *et al*, 2013. Differential analysis of gene regulation at transcript resolution with RNA-seq [J]. Nature Biotechnology, 31(1): 46-53.
- TRAPNELL C, ROBERTS A, GOFF L, *et al*, 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks [J]. Nature Protocols, 7(3): 562-578.
- TRAPNELL C, WILLIAMS B A, PERTEA G, *et al*, 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation [J]. Nature Biotechnology, 28(5): 511-515.
- VAN GORP H, KUCHMIY A, VAN HAUWERMEIREN F, *et al*, 2014. NOD-like receptors interfacing the immune and reproductive systems [J]. FEBS Journal, 281(20): 4568-4582.

ZHANG G F, FANG X D, GUO X M, *et al.*, 2012. The oyster genome reveals stress adaptation and complexity of shell formation [J]. *Nature*, 490(7418): 49-54.

ZHANG L L, LI L, GUO X M, *et al.*, 2015. Massive expansion and functional divergence of innate immune genes in a protostome [J]. *Scientific Reports*, 5(1): 8693.

ANALYSIS OF ALTERNATIVE SPLICING EVENTS FOR INNATE IMMUNITY OF *CRASSOSTREA GIGAS* INDUCED BASED ON RNA-SEQ

LI Ruo-Hui^{1,2,3}, ZHANG Lin-Lin^{1,2,3}

(1. *Key Laboratory of Experimental Marine Biology, Institute of Oceanology, Chinese Academy of Sciences, Qingdao 266071, China*;
2. *Laboratory for Marine Biology and Biotechnology, Pilot National Laboratory for Marine Science and Technology (Qingdao), Qingdao 266237, China*; 3. *University of Chinese Academy of Sciences, Beijing 100049, China*)

Abstract Most invertebrates lack adaptive immunity based on classical antibodies and memory cells, thus natural immunity plays an indispensable role in invertebrate immune defense. Variable splicing is an important way to generate diversity and specificity of natural immunity. *Crassostrea* is a globally farmed shellfish and the most productive marine shellfish in China. Understanding the diversity and specificity of its natural immune system is essential for the prevention and control of oyster diseases and the sustainable development of oyster aquaculture. The variable splicing events in the natural immune stress response of *Crassostrea gigas* were investigated systematically by bioinformatic analysis of transcriptomes induced by different pathogens at different times after induction. The total number of variable splicing events was found to increase significantly under *Vibrio* induction, indicating that *Vibrio* infection induced the generation of variable splicing events in *C. gigas*. Functional enrichment analysis of genes with significant changes in the composition type and expression of variable spliceosomes after pathogen induction showed that immune system-related functions were significantly enriched, and that the composition type and expression of variable spliceosomes were inconsistent at different stages of pathogen infection as well as after different pathogen infections, suggesting that the immune system of the long oyster can generate specificity and diversity of immune responses through variable splicing. This study provided a typical example of the diversity and specificity of natural immunity in invertebrates such as *C. gigas*, and also a theoretical support for the defense of *C. gigas* diseases.

Key words *Crassostrea gigas*; alternative splicing; pathogen induction; innate immunity