

巨蚶属(*Crassostrea*)四种牡蛎精子 Bindin 蛋白海藻糖凝集素结构域(Fucose binding lectin)多样性研究*

吴琪^{1,2} 李莉¹ 张国范¹

(1. 中国科学院海洋研究所 青岛 266071; 2. 中国科学院研究生院 北京 100049)

摘要 提取巨蚶属中熊本牡蛎、葡萄牙牡蛎和日本巨牡蛎的基因组 DNA, 应用 PCR 技术扩增了三种牡蛎的海藻糖凝集素结构域(Fucose binding lectin, F-lectin), 结合 GenBank 中长牡蛎海藻糖凝集素结构域的序列利用邻接法对得到的序列进行了系统进化分析, 并利用非同义替换率和同义替换率之间的比值分析了巨蚶属四种牡蛎海藻糖凝集素结构域上可能的受正选择的位点。结果表明, 四种巨蚶属牡蛎的海藻糖凝集素结构域序列中长牡蛎与葡萄牙牡蛎的 F-lectin 比长牡蛎与熊本牡蛎的关系更近, 葡萄牙牡蛎三种单倍型(an1, an2, an3)定位在不同的支上, an1 与长牡蛎海藻糖凝集素结构域的关系更近, an2 与 an3 与熊本牡蛎之间的关系更近。在巨蚶属中, 海藻糖凝集素结构域的 3D 模型中, 发现五个可能的受正向选择的位点(40, 66, 67, 123, 125), 并且这五个位点都围绕在三个重要的识别氨基酸(H37, R64, R70)周围, 推测这五个正向选择位点也许与牡蛎种特异的受精识别有关。

关键词 巨蚶属, 结合素蛋白, 海藻糖凝集素结构域, 多样性, 正选择位点

中图分类号 Q18

对于那些在开放的水体环境中行体外受精繁殖的水生生物来说, 配子识别蛋白在精卵识别的过程中发挥着重要作用, 在完成受精的同时, 还必须保障受精的种特异性, 以维持物种的独特性(Swanson *et al*, 2002)。牡蛎是世界性广泛分布的重要的海产无脊椎动物, 繁殖季节到来的时候, 雌、雄配子被排入海水中, 行体外受精。长牡蛎(*Crassostrea gigas*)是我国海水贝类养殖中重要的经济贝类(阙华勇等, 2003), 在我国北部沿海大面积养殖(毛玉泽等, 2005; 王绍宗等, 2010; 滕爽爽等, 2010), 其精子有大量的顶体颗粒, 当顶体反应发生时, 精子顶体发生胞吐作用, 释放其中的顶体颗粒, 其中主要是结合素蛋白 Bindin (Moy *et al*, 2008a, b), 能帮助精子牢固附着在卵黄膜上。海藻糖凝集素结构域(Fucose binding lectin, F-lectin)是 Bindin 蛋白发挥功能的重要结构域。本文

对获得两个熊本牡蛎(*Crassostrea sikamea*)、三个葡萄牙牡蛎(*C. angulata*)以及一个日本巨牡蛎(*C. nippona*)的 Bindin 蛋白 F-lectin 的单倍型序列进行了研究, 以期更深入地理解不同种牡蛎配子识别蛋白之间的进化关系, 进一步认识复杂的细胞识别过程。

1 材料与方法

1.1 材料

熊本牡蛎(*Crassostrea sikamea*)取自江苏省南通市, 葡萄牙牡蛎(*C. angulata*)和日本巨牡蛎(*C. nippona*)分别取自福建省莆田市和浙江舟山地区。经过严格的形态鉴定和 CO I 分子鉴定。样品经过活体解剖取闭壳肌组织于 70%乙醇中固定, 取性腺组织于显微镜下进行性别鉴定后, 保存于 Trizol (Invitrogen) 中, 于实验室-80 冰箱中保存、待用。

* 国家自然科学基金重点项目, 40730845 号, 国家重点基础研究发展计划(973 项目)资助, 2010CB126401 号; 国家高技术研究发展计划(863 计划)资助, 2010AA1019004 号; 国家公益性行业(农业)科研专项, 3-53 号。吴琪, 博士研究生, E-mail: qi_wu@126.com, q909501@163.com

通讯作者: 张国范, 博士, 研究员, E-mail: gfzhang@ms.qdio.ac.cn

收稿日期: 2010-04-13, 收修改稿日期: 2010-06-10

1.2 方法

1.2.1 基因组 DNA 提取 基因组 DNA 提取采用经典的酚-氯仿法, 每只个体取 50mg 闭壳肌组织, 置入 700 μ l DNA 提取缓冲液中(10mmol/L EDTA, 50mmol/L Tris-HCl pH 8.0, NaCl 0.7mol/L, 2-巯基乙醇 0.1% *V/V*)中剪碎, 加入 100 μ g/ml 的蛋白酶 K, 于 60 $^{\circ}$ C 水浴中消化 2.5h, 最后用 10 μ g RNase 水浴 30min。先后用等体积的酚、酚 氯仿 异戊醇(25:24:1)提取 2 次。加 0.2 体积的 10mol/L 乙酸钠和 2 倍体积预冷的无水乙醇沉淀, 经 70%乙醇洗涤两遍后干燥, 50 μ l 双蒸水溶解, 置于-20 $^{\circ}$ C 冰箱中保存。DNA 纯度、浓度及分子量用(Beckman DU-600, USA)紫外分光光度计及琼脂糖电泳检测。

1.2.2 目的片段扩增 根据 GenBank 中长牡蛎 F-lectin 序列设计正向引物: 5'TAGTAAGCCTCGTTC TCTCT3'和反向引物: 5'AACGCAGATAAAAGGCAA-GAA3'。扩增反应在 My Cycler PCR 仪(BIO-RAD)上进行, 反应条件为 94 $^{\circ}$ C 变性 40s, 60 $^{\circ}$ C 退火 50s, 72 $^{\circ}$ C 延伸 50min, 30 个循环, 最终 72 $^{\circ}$ C 延伸 10min。产物经 E.B 染色后经 1.0%琼脂糖凝胶电泳检测。PCR 产物用 TaKaRa DNA 纯化试剂盒纯化, 连接载体 pMD18-T (Takara), 感受态细胞为 *E. coli* (Top10), 阳性转化子经 PCR 筛选后送深圳华大基因测序, 全部采用双向测序。

1.2.3 数据分析 使用 BioEdit 7.0 (Hall, 1998)序列拼接, 序列同源性在 NCBI 的 BLAST 上搜索。使用 ClustalX 1.83 (Thompson *et al*, 1997)进行多重比对, 使用 MEGA4.0 (<http://www.megasoftware.net>)建立系统发生树并检测可能受正选择的位点, 使用 DNAsp 4.10 (Rozas *et al*, 2003)进行单倍型分析。三维结构预测使用 SWISS-MODEL (<http://swissmodel.expasy.org/workspace>)。

2 结果

2.1 序列分析

从两个雄性熊本牡蛎中扩增得到两个 F-lectin 单倍型序列, 并从六个雄性葡萄牙牡蛎中扩增得到了三个单倍型序列, 从一个日本巨牡蛎中得到一个 F-lectin 序列, 经过 NCBI 比对后确认与长牡蛎的五个单倍型序列是同源序列(表 1)。将这 11 条核酸序列对应的 133 个氨基酸组成的多肽序列进行比对(图 1)。三个识别位点(H37, R64, R70)和四个半胱氨酸位点(C27, C36, C88, C105)都非常保守。

表 1 四个种的 F-lectin 单倍型序列在 NCBI 上获取的 GenBank 号

Tab.1 The GenBank number of F-lectin haplotypes from four species

单倍型号	GenBank 号	单倍型号	GenBank 号
s1	GQ131801	g1	EF219425*
s2	GQ131802	g2	EF219426*
an1	GQ408895	g3	EF219427*
an2	GQ408896	g4	EF219428*
an3	GU064556	g5	EF219429*
ni	GU954544		

注: s、g、ni、an 分别表示熊本牡蛎、长牡蛎、日本巨牡蛎、葡萄牙牡蛎; *表示序列来自 GenBank 库

2.2 系统进化分析

利用 MEGA 分析软件, 用邻接法 NJ (Neighbor-Joining)构建 F-lectin 序列的系统发生树见图 2。这个树的结构分为两个大支, 一支包含有全部的长牡蛎序列和一个葡萄牙牡蛎的单倍型序 an1; 另一支包含熊本牡蛎的两个单倍型序列、葡萄牙牡蛎的单倍型序列 an2 和 an3 和日本巨牡蛎的一个 F-lectin 序列 ni。

2.3 检测正选择位点

同义替换率(D_s)和非同义替换率(D_n)常用来检测基因所受到的选择压力。当 $D_n/D_s > 1$ 时, 表示该变异在进化过程中受到正向选择, 间接表明该变异对物种生存有利; 当 $D_n/D_s < 1$ 时, 表示该基因进化中受纯化选择, 进化中较保守(Yang *et al*, 2002)。通常配子识别蛋白都在适应性进化中快速产生分歧, 多数情况下如此快速产生分歧是由正选择造成的, 所以检测了可能受正选择压力的位点。经过 χ^2 检验($2\Delta l = 20.562$, $P < 0.01$)共发现 5 个受正选择压力的位点(95%), 其中三个位于 F-lectin 外显子 1 中(40, 66 and 67; mean $D_n/D_s = 6.848$), 两个位于 F-lectin 外显子 2 中(123 and 125; mean $D_n/D_s = 6.346$)。见图 3。

3 讨论

3.1 巨蛎属 F-lectin 进化分析

F-lectin 是一种能够特异性结合岩藻糖的、古老的结构域, 在原核生物和真核生物中都有发现。在一些鱼类的血清中发现很多含有 F-lectin 结构域的蛋白质都与免疫识别有关(Odom *et al*, 2006), 在 *Tachypleus tridentatu* (Salto *et al*, 1997)和 *Anguilla japonica* (Honda *et al*, 2000)中被认为是免疫识别分子。对于巨蛎属的牡蛎来说, F-lectin 是受精时主要的精子和卵子识别

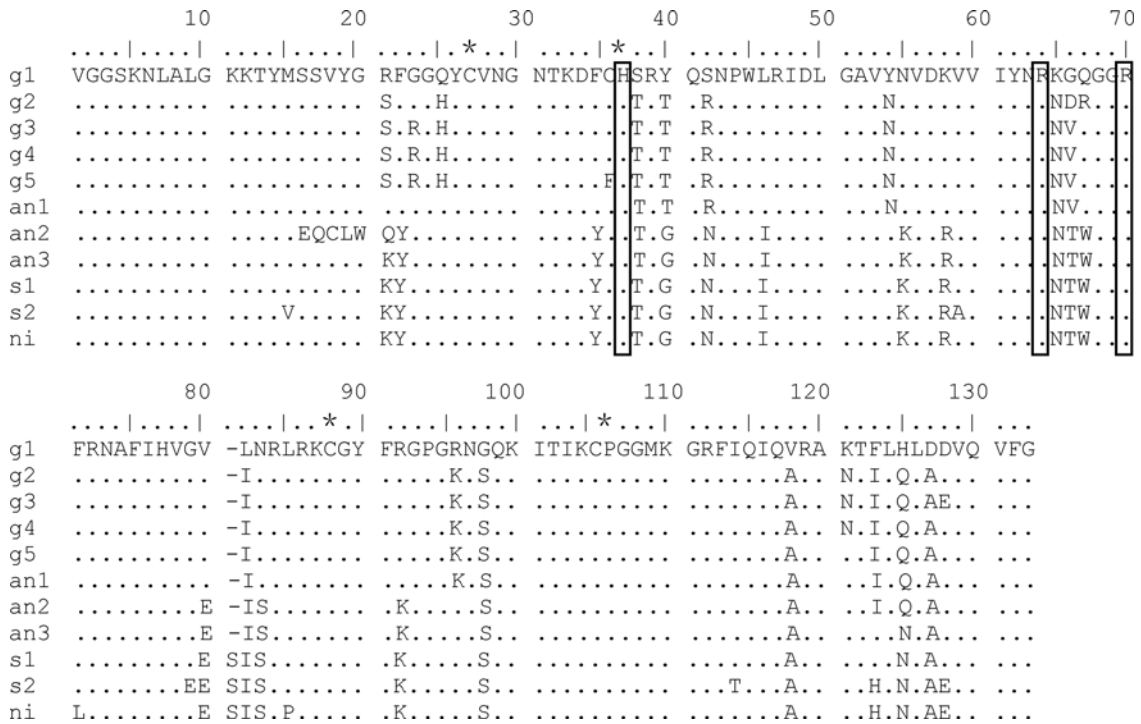


图 1 来自四种牡蛎的 11 个 F-lectin 单倍型的氨基酸序列的比对结果

Fig.1 Alignment of amino acid sequences of eleven F-lectin repeat haplotypes of four species

注: 相同的序列用点表示, 黑框的位置是三个保守的氨基酸位点, 星号下面的是四个保守的氨基酸位点; s、g、ni、an 分别表示熊本牡蛎、长牡蛎、日本巨牡蛎、葡萄牙牡蛎

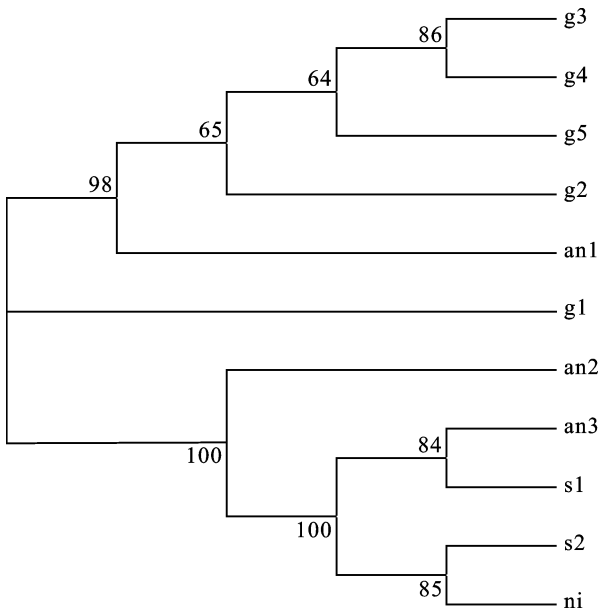


图 2 用 MEGA 3.1 软件的邻接法对四个种的 F-lectin 序列建树(bootstrap 值为 1000)

Fig.2 Phylogenetic tree of eleven haplotypes based on Neighbor-joining method with 1000 bootstrap trials by MEGA 3.1 program

注: 节点上的数值表示 bootstrap 值; s、g、ni、an 分别表示熊本牡蛎、长牡蛎、日本巨牡蛎、葡萄牙牡蛎

的功能域。通过扩增得到 F-lectin 序列可以考察其在种间的变化。虽然得到的数据有限, 但从已有的数据来看, F-lectin 在葡萄牙牡蛎种内、熊本牡蛎种内、长牡蛎种内无论是核苷酸序列还是氨基酸序列都表现出多样性, 并在种间也表现出多样性。

四种巨牡蛎属的牡蛎 F-lectin 序列建立的 NJ 树显示长牡蛎与葡萄牙牡蛎的 F-lectin 比长牡蛎与熊本牡蛎的关系更近, 这与 mtCOI 基因建立的三种牡蛎之间的亲缘关系是一致的(Reece *et al*, 2008)。更加令人深思的是这个树中 an2 和 an3 的位置。葡萄牙牡蛎三种单倍型定位在不同的支上, an1 与长牡蛎 F-lectin 的关系更近, an2 与 an3 与熊本牡蛎之间的关系更近。长牡蛎和葡萄牙牡蛎亲缘关系很近, 尽管二者之间能够进行双向无障碍的杂交, 但它们在生理水平上(Haure *et al*, 2003)、遗传水平上(Huvet *et al*, 2000; Leitão *et al*, 2007)已经被证明是两个独立的物种(Batista *et al*, 2008)。熊本牡蛎也是长牡蛎的近缘种, 但二者之间的杂交仅仅是单向的, 熊本牡蛎的卵子能够与长牡蛎的精子受精, 但是熊本牡蛎的精子却不能与长牡蛎的卵子受精(Banks *et al*, 1994)。作者认

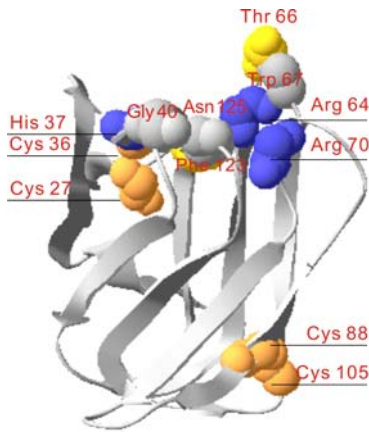


图3 F-lectin 重复结构域 3D 模型

Fig.3 3D modeling of threaded oyster bindin F-lectin repeat structure
注: 五个受正向选择的位点用灰色球(40G, 67W, 123F)和黄色球(66T, 125N)标出, 三个保守的氨基酸位点(37H, 64R, 70R)用蓝色球表示, 四个保守的半胱氨酸位点(C27, C36, C88, C105)用桔黄色球表示

精过程中建立了强有力的生殖隔离屏障(Metz *et al.*, 1996), Bindin 在这两个种间表现出强烈的序列多样性。在鲍鱼中, 对两个配子识别蛋白的全序列分析显示了这两种蛋白都在正向的达尔文选择压力下不断产生分歧, 向各自的方向进化(Vacquier *et al.*, 1997)。在巨蛎属中, F-lectin 的 3D 模型中, 这五个受正向选择的位点都位于 F-lectin 的表面, 并且围绕在三个重要的识别氨基酸(H37, R64 and R70)周围(图 3)。作者推测这五个正向选择位点也许与牡蛎种特异的受精识别有关。

Bindin 与其卵子上的受体通过受精识别过程不断适应、协同进化。配子识别蛋白在种内快速产生分歧最终导致配子之间的不兼容性, 最终引起新的物种形成(Springer *et al.*, 2008)。对配子识别系统的深入研究, 能够帮助人们更深入地了解复杂的细胞识别过程, 以及在这个过程中蛋白质分子之间的进化, 更深入理解在漫长的生物进化过程中生殖隔离和新种形成的历史过程。

参 考 文 献

王绍宗, 李 莉, 亓海刚等, 2010. 长牡蛎(*Crassostrea gigas*)17 个 EST-SNP 标记的开发. 海洋与湖沼, 41(2): 274—281
毛玉泽, 周 毅, 杨红生等, 2005. 长牡蛎(*Crassostrea gigas*)代谢率的季节变化及其与夏季死亡关系的探讨. 海洋与

为葡萄牙牡蛎 an2 和 an3 这两个单倍型也可以看作是 F-lectin 演化的一种过渡状态。

3.2 F-lectin 的正选择位点

在产卵、产精于外部环境的体外受精的物种中, 配子识别蛋白通常都经历着快速的分歧和进化, 有的甚至比线粒体 DNA 进化速度还要快(Metz *et al.*, 1998)。在两种近缘的夏威夷海胆中线粒体和核基因组上的微小变化就在受

- 湖沼, 36(5): 445—451
阙华勇, 张国范, 刘 晓等, 2003. 雄性四倍体与雌性二倍体杂交培育全三倍体长牡蛎(*Crassostrea gigas*)的研究. 海洋与湖沼, 34(6): 656—662
滕爽爽, 李 琪, 李金蓉, 2010. 长牡蛎(*Crassostrea gigas*)与熊本牡蛎(*C. sikamea*)杂交的受精细胞学观察及子一代的生长比较. 海洋与湖沼, 41(6): 914—922
Banks M A, McGoldrick D J, Borgeson W *et al.*, 1994. Gametic incompatibility and genetic divergence of Pacific and *Kumamoto oysters*, *Crassostrea gigas* and *C. sikamea*. *Marine Biology*, 121: 127—135
Batista F M, Ben-Hamadou R, Fonseca V G *et al.*, 2008. Comparative study of shell shape and muscle scar pigmentation in the closely related cupped oysters *Crassostrea angulata*, *C. gigas* and their reciprocal hybrids. *Aquatic Living Resources*, 21: 31—38
Hall T A, 1998. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, 41: 95—98
Haure J, Huvet A, Palvadeau H *et al.*, 2003. Feeding and respiratory time activities in the cupped oysters *Crassostrea gigas*, *Crassostrea angulata* and their hybrids. *Aquaculture*, 218: 539—551
Honda S, Kashlwal M, Mlyamoto K *et al.*, 2000. Multiplicity, structures, and endocrine and exocrine nature of eel fucose-binding lectins. *The Journal of Biological Chemistry*, 275: 33151—33157
Huvet A, Lap è gue S, Magoulas A *et al.*, 2000. Mitochondrial and nuclear DNA phylogeography of *Crassostrea angulata*, the Portuguese oyster endangered in Europe. *Conservation Genetics*, 1: 251—262
Leitão A, Chaves R, Santos S *et al.*, 2007. Interspecific hybridization in oysters: Restriction enzyme digestion chromosome banding confirms *Crassostrea angulata* × *Crassostrea gigas* F₁ hybrids. *Journal of Experimental Marine Biology and Ecology*, 343: 253—260
Metz E C, Palumbi S R, 1996. Positive selection and sequence rearrangements generate extensive polymorphism in the gamete recognition protein bindin. *Molecular Biology and Evolution*, 13: 397—406
Metz E C, Robles-Sikisaka R, Vacquier V D, 1998. Nonsynonymous substitution in abalone sperm fertilization genes exceeds substitution in introns and mitochondrial DNA. *Proceedings of the National Academy of Sciences USA*, 95: 10676—10681
Moy G W, Springer S A, Adams S L *et al.*, 2008a. Extraordinary intraspecific diversity in oyster sperm bindin. *Proceedings of the National Academy of Sciences USA*, 105: 1993—1998
Moy G W, Vacquier V D, 2008b. Bindin genes of the Pacific oyster *Crassostrea gigas*. *Gene*, 423: 215—220
Odom E W, Vasta G R, 2006. Characterization of a binary tandem domain F-type lectin from striped bass (*Morone saxatilis*).

- The Journal of Biological Chemistry, 281: 1698—1713
- Reece K S, Cordes J F, Stubbs J B *et al*, 2008. Molecular phylogenies help resolve taxonomic confusion with Asian *Crassostrea* oyster species. *Marine Biology*, 153: 709—721
- Rozas J, Sanchez-DelBarrio J C, Messeguer X *et al*, 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, 19: 2496—2597
- Salto T, Hstada M, Iwangaga S, 1997. A newly identified horse-shoe crab lectin with binding specificity to O-antigen of bacterial lipopolysaccharides. *The Journal of Biological Chemistry*, 272: 30703—30708
- Springer S A, Moy G W, Friend D S *et al*, 2008. Oyster sperm bindin is a combinatorial fucose lectin with remarkable intra-species diversity. *The International Journal of Developmental Biology*, 52: 759—768
- Swanson W J, Vacquier V D, 2002. Rapid evolution of reproductive proteins. *Nature Reviews Genetics*, 3: 137—144
- Thompson J D, Gibson T J, Plewniak F *et al*, 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, 25: 4876—4882
- Vacquier V D, Swanson W J, Lee Y H, 1997. Positive Darwinian selection on two homologous fertilization proteins: what is the selective pressure driving their divergence? *Molecular Biology and Evolution*, 44: S15—S22
- Yang Z, Nielsen R, 2002. Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Molecular Biology and Evolution*, 19: 908—917

STUDY ON SPERM BINDIN FUCOSE-BINDING LECTIN OF FOUR SPECIES OF *CRASSOSTREA*

WU Qi^{1,2}, LI Li¹, ZHANG Guo-Fan¹

(1. Institute of Oceanology, Chinese Academy of Sciences, Qingdao, 266071;
2. Graduate School, Chinese Academy of Sciences, Beijing, 100049)

Abstract Gamete recognition proteins (GRPs) play an important role in recognition between sperm and egg. Bindin is a major protein packaged as insoluble granules in sperm acrosomes. Fucose binding lectin (F-lectin) is the main functional domain of bindin. F-lectin repeats of *Crassostrea sikamea*, *C. angulata* and *C. nippona* have been got by polymerase chain reaction (PCR). NJ tree is constructed to analyze the relationship of eleven F-lectin repeats from *C. sikamea*, *C. angulata*, *C. nippona* and *C. gigas*. Five positive selected positions (40, 66, 67, 123 and 125) are identified among eleven haplotypes of four species. They are all located on the F-lectin binding face around the three recognition motif residues (H37, R64 and R70). We assume that these five clustered amino acids are related with species specific recognition.

Key words *Crassostrea*, Bindin, Fucose-binding lectin, Polymorphism, Positive selection