

文章编号: 1000-0615(2008)05-684-06

## 用焦磷酸测序技术检测凡纳滨对虾组织蛋白酶 基因单核苷酸多态性

曾地刚, 陈晓汉, 彭 敏, 李咏梅, 杨春玲, 马 宁,  
蒋伟明, 黎 铭

(广西水产研究所, 广西 南宁 530021)

**摘要:** 焦磷酸测序是一种新型的 DNA 测序技术, 近年来开始应用于单核苷酸多态性(SNPs)检测。组织蛋白酶基因(Cathepsin-L, CTSL)在对虾的蜕皮周期中起重要的作用。本研究利用焦磷酸测序技术, 对 96 个凡纳滨对虾的 CTSL 基因序列(GenBank 登录号: AY366355)的 C681G SNP 位点进行检测分型。结果发现 C/C、C/G 和 G/G 基因型, 频率分别是 0.81、0.16 和 0.03, 频率分布符合 Hardy-Weinberg 平衡。统计分析结果表明在该群体中 C681G SNP 和体重关系不显著( $P > 0.05$ )。研究结果显示焦磷酸测序技术是一种高效的 SNP 检测技术, 可以在对虾的遗传研究中发挥重要作用。

**关键词:** 凡纳滨对虾; 焦磷酸测序; 组织蛋白酶基因; 单核苷酸多态性

**中图分类号:** S 917

**文献标识码:** A

单核苷酸多态性(single nucleotide polymorphisms, SNPs)是指基因组 DNA 中某一特定核苷酸位置上发生转换、颠换、插入或缺失等变化, 而且任何一种等位基因在群体中的频率不小于 1%<sup>[1]</sup>。

单核苷酸多态性的快速检测方法是开展基因变异与遗传效应相关分析研究的重要基础<sup>[2]</sup>。水产动物的遗传分析容易受到环境因子的影响, 往往需要较大的样本数量以减少分析的误差, 因此有必要采用自动化、高通量检测 SNP 的方法。目前, 检测已知 SNP 最常用的方法是 PCR-RFLP<sup>[3]</sup>, 即 PCR 扩增产物的限制性片段长度多态检测方法。然而, PCR-RFLP 需要酶切和电泳等繁琐过程, 不符合自动化、高通量检测的要求, 致使 PCR-RFLP 的实际应用具有很大的局限性<sup>[4]</sup>。焦磷酸测序技术(pyrosequencing)是于

1987 年发展起来的一种新型的短片段 DNA 测序技术<sup>[5]</sup>, 是在 DNA 聚合酶、ATP 硫酸化酶、荧光素酶和三磷酸腺苷双磷酸酶 4 种酶的协同作用下, 将引物上每一个 dNTP 聚合与一次荧光信号释放偶联起来, 通过检测荧光的释放和强度, 达到实时测定 DNA 序列的目的。操作中不需要电泳、样品标记和染色, 具有高度的可重复性、并行性和自动化特点<sup>[6]</sup>。与 Sanger 测序法相比, 焦磷酸测序技术有其特定的优势, 已成为小片段 DNA 分析的重要手段, 目前已开始用于 SNP 的检测<sup>[7]</sup>。

组织蛋白酶在对虾的蜕皮周期中起重要的作用, 因此组织蛋白酶基因(CTSL)被作为对虾生长相关的候选基因。凡纳滨对虾(*Litopenaeus vannamei*)的 CTSL 基因已被克隆, 然而 CTSL 基因在对虾体内的作用机制仍不清楚, CTSL 基

收稿日期: 2007-08-12

资助项目: 国家科技支撑计划项目(2006BAD01A13); 广西十百千人才专项基金(2007113)

作者简介: 曾地刚(1971-), 男, 广西南宁人, 硕士, 助理研究员, 主要从事水生动物分子遗传育种研究。Tel: 0771-5307517, E-mail: zengdigang@126.com

通讯作者: 陈晓汉, Tel: 0771-5316577

因变异与对虾生长性能的相关性仍有待探明<sup>[8]</sup>。而对上述问题的研究有可能筛选到凡纳滨对虾生长性状显著相关的遗传标记。Glenn 等<sup>[9]</sup>利用 PCR-RFLP 方法检测了凡纳滨对虾 CTSL 基因的 G178C 和 C681G 单核苷酸多态性位点,结果发现 C681G 可能和生长性状相关,虽然实验数据分析结果没有达到统计显著水平,但是可能是因为样品量太少(30 个样品)所至。

为了进一步研究凡纳滨对虾 CTSL 基因多态性和生长性状的关系,本研究选择一个凡纳滨对虾群体作为试验材料,采用焦磷酸测序技术对 CTSL 基因序列(GenBank 登录号:AY366355)中的 C681G 单核苷酸多态性位点进行检测,并作统计分析。

## 1 材料与方法

### 1.1 实验动物的采集

实验凡纳滨对虾来自广西水产研究所南美白对虾良种场,是从美国夏威夷引进的 SPF 凡纳滨对虾种虾,经过多代选育的后代群体,采集的方法是在同一个池塘中拉网捕捉对虾,然后随机采集。采集后立即用电子天平称量体重,然后样品在  $-80\text{ }^{\circ}\text{C}$  保存。

### 1.2 基因组 DNA 的提取

基因组 DNA 提取参照 Aljanabi 等<sup>[10]</sup>的方法略加改动:(1)取 50 mg 肌肉组织放入 1.5 mL 离心管中,加入 400  $\mu\text{L}$  DNA 提取缓冲液( $0.4\text{ mol}\cdot\text{L}^{-1}$  NaCl, $10\text{ mmol}\cdot\text{L}^{-1}$  Tris HCl-Cl(pH 8.0), $2\text{ mmol}\cdot\text{L}^{-1}$  EDTA(pH 8.0),1% SDS,20  $\mu\text{g}\cdot\text{mL}^{-1}$  胰 RNA 酶),用剪刀剪碎。(2)加入 8  $\mu\text{L}$   $20\text{ mg}\cdot\text{mL}^{-1}$  蛋白酶 K(终浓度  $400\text{ }\mu\text{g}\cdot\text{mL}^{-1}$ ),充分混匀后,放入  $55\text{ }^{\circ}\text{C}$  水浴锅中温浴 3 h。(3)加 300  $\mu\text{L}$   $6\text{ mol}\cdot\text{L}^{-1}$  NaCl,在旋涡混合器上高速混合 30 s, $10\text{ }000\text{ g}$  离心 30 min,移上清到另一离心管中。(4)加入等体积异丙醇,混匀,置  $-20\text{ }^{\circ}\text{C}$  1 h,然后  $10\text{ }000\times\text{g}$  离心 15 min,弃上清。(5)用 70%乙醇洗涤沉淀,干燥后将其溶于适量 TE 中, $-20\text{ }^{\circ}\text{C}$  备用。基因组 DNA 用紫外分光光度计和琼脂糖凝胶电泳 EB 染色检查浓度和纯度。

### 1.3 PCR 引物

PCR 引物根据 CTSL 基因序列(GenBank 登录号:AY366355)中的 C681G 的两端序列,用 Pyrosequencing-AB 软件自己设计。

PCR-F: CTCCTGCTGGGCTTTCTCC

PCR-R: AGGGACCCCGTCTGCAAC

PCR-F 的 5'端用生物素(biotin)标记。

### 1.4 PCR 条件

先  $94\text{ }^{\circ}\text{C}$  变性 3 min;然后  $94\text{ }^{\circ}\text{C}$  30 s, $58\text{ }^{\circ}\text{C}$  30 s, $72\text{ }^{\circ}\text{C}$  30 s 45 个循环;最后  $72\text{ }^{\circ}\text{C}$  延长 3 min。

### 1.5 单链 DNA 分离纯化

(1) 在 PSQ 96 板中预先加入 45  $\mu\text{L}$  含有  $0.3\text{ }\mu\text{mol}\cdot\text{L}^{-1}$  测序引物的复性缓冲液;使用 Vertex 混匀琼脂糖珠;(2)将需要使用的琼脂糖珠总量(每样本 4  $\mu\text{L}$ )转移到一个 Eppendorf 管中;在琼脂糖珠中加入结合缓冲液,使得平均每个样品约有 40  $\mu\text{L}$  的体积,将混合物混匀;将以上混合物加入 PCR 产物(40  $\mu\text{L}$  反应体积)中,每样本 50  $\mu\text{L}$ ;(3)将 PCR 产物在常温下混匀 5 min,使得琼脂糖珠与生物素结合;(4)在真空制备工作站中,四个样品板中依次加入 180 mL 高纯水、70%乙醇、洗涤缓冲液和变性缓冲液;(5)打开真空制备工作站的泵,将 vacuum prep tool 在高纯水中清洗 30 s;(6)然后将 vacuum prep tool 移到 PCR 板中,抓取琼脂糖珠;(7)将 vacuum prep tool 放入 70%乙醇中 5 s;然后移到变性缓冲液中 5 s;再移到洗涤缓冲液中清洗 5~10 s;关掉泵;(8)将 vacuum prep tool 放入含有测序引物的板中,摇动,释放琼脂糖珠;(9)使用高纯水清洗 vacuum prep tool;(10)将放有样品的 PSQ 96 板加热到  $80\text{ }^{\circ}\text{C}$  2 min,再冷却到室温,即可进行焦磷酸测序反应。

### 1.6 焦磷酸测序

利用 PyroMark ID System (PSQ 96 MA) 和 SNP 试剂盒,按照仪器和试剂盒说明书进行焦磷酸测序。测序由基因科技(上海)有限公司完成。

## 2 结果和分析

焦磷酸测序结果发现实验凡纳滨对虾群体中 CTSL 有 3 种基因型,分别为 C/C,C/G 和 G/G,分别如图 1 至图 3 所示。测序图中,下边所标字母第一个是 E,代表酶,应该是没有峰;第二个是 S,代表底物,有一个高度不等小峰;第三个是 T,是程序特地设置的一个阴性对照峰,应该是没有峰的;然后是 C 和 G 这两个是可能出现峰的,后面的 C 是为了检查前面 C 处是否反应完全,也应该是没有峰;再后面的 A,G,C,T,G 就是后面的

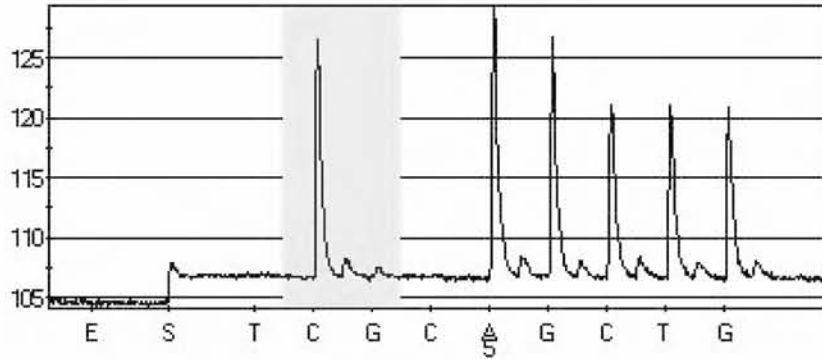


图1 CTSL 基因焦磷酸测序结果(基因型为 C/C)

Fig. 1 Pyrosequencing of CTSL(showed the result of C/C genotype)

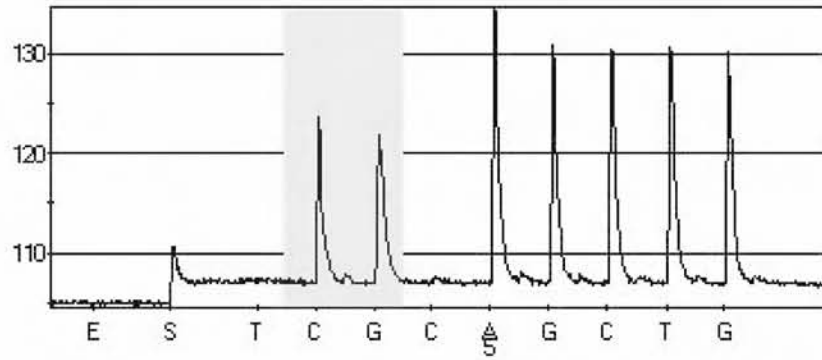


图2 CTSL 基因焦磷酸测序结果(基因型为 C/G)

Fig. 2 Pyrosequencing of CTSL(showed the result of C/G genotype)

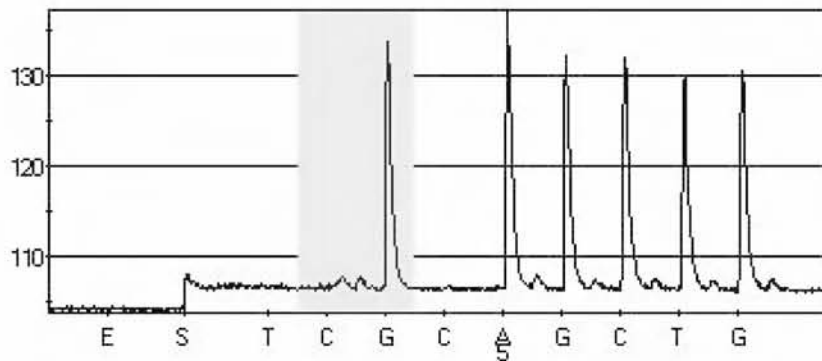


图3 CTSL 基因焦磷酸测序结果(基因型为 G/G)

Fig. 3 Pyrosequencing of CTSL(showed the result of G/G genotype)

序列,峰高近似相等,表示序列为 AGCTG,和 GenBank 收录的序列相同,说明测序是准确的。

本研究一共进行 96 个样品的焦磷酸测序,这套引物测的是反义链,统计结果基因型 C/C 的样

品 78 个,频率 0.81;基因型 C/G 的样品 15 个,频率 0.16;基因型 G/G 的样品 3 个,频率 0.03,卡方检验结果  $X^2 = 3.76$ ,  $P = 0.0524$ ,频率分布符合 Hardy-Weinberg 平衡( $P > 0.05$ )。

统计分析采用单标记均值差检验的方差分析法<sup>[11]</sup>检验同一标记座位上不同基因型间数量性状均值的差异,若差异显著,则表明被检标记与 QTL 连锁。用 SPSS 统计软件(version 12.0)的 GLM 过程进行方差分析,检验不同 SNP 基因型间凡纳滨对虾体重均值差异的显著性,结果 C/C 样品的平均体重为(2.34 ± 1.45) g, C/G 样品的平均体重为(2.18 ± 1.41) g, G/G 样品的平均体重为(3.19 ± 0.59) g, 方差分析结果差异不显著( $P=0.5786$ )。

### 3 讨论

**3.1 凡纳滨对虾 CTSL 基因 SNP 和生长性状的相关性** 凡纳滨对虾亦称南美白对虾,以其壳薄体肥、肉质鲜嫩、生长迅速、群体增长均匀、抗病力强等优点,而逐渐成为南方主要养殖虾种,深受国内外市场的青睐。近年来不少学者对凡纳滨对虾生物学特性和养殖技术进行了研究,但是凡纳滨对虾生产性能提高的选育研究进展缓慢<sup>[12-14]</sup>。目前凡纳滨对虾的选育方法主要还是传统的表型性状选育法,存在选育周期长、选育效率低和受环境因素影响大等缺点。虽然国内外也有一些学者采用 RAPD、AFLP 和微卫星等分子标记技术开展了分子标记选育的研究,但是由于目前可用的凡纳滨对虾 AFLP 或微卫星标记数量少,建立参考家系及作图群体难度大,缺乏标记遗传图谱和物理图谱,因此分子标记选育的研究进展不够快。为此,我们采用候选基因法,以 CTSL 基因作为凡纳滨对虾生长相关的候选基因,研究基因的多态性和生长性状的关系,以期找到生长相关的遗传标记。候选基因法一个重要的优点是适用于任何可以取得表型数据的群体,不一定要使用费用昂贵的参考家系,因而具有相当的灵活性<sup>[15]</sup>。

本研究检测了凡纳滨对虾 CTSL 基因序列中的一个 SNP(C681G),统计分析和对虾体重的相关性。C681G 位于 CTSL 基因第 3 内含子中,既不处于蛋白质编码区也不处于基因调控区,因此它对个体的表现型是无影响的,连锁分析时不需要考虑 SNP 本身的效应。相关性分析结果 C681G 和体重的关系并不显著。这和 Glenn 等<sup>[9]</sup>的研究结果一致。Glenn 等<sup>[9]</sup>认为 C681G 可能和生长性状相关,相关分析没有达到显著水平的原因可能是因为样品量太少(30 个样品)导

致的误差。本研究采用的样品虽然相对比较多,但是由于该实验群体中等位基因 G 的频率很低,因此基因型 C/G 和 G/G 的样品数量也比较少,同样存在实验误差大的问题。今后需要采用更多的实验群体和样品数量进行研究。

### 3.2 焦磷酸测序技术检测 SNP 的优点

作为新一代分子标记,SNP 在高密度遗传和物理图谱的构建与整合、功能基因的鉴定与关联分析、群体遗传学与生物进化研究、作物遗传改良与分子标记辅助育种等理论和应用研究方面有广阔的应用前景<sup>[16]</sup>。因此,许多学者提出了许多 SNP 检测方法,

目前用于已知 SNP 检测的方法主要有 PCR 扩增产物限制性片段长度多态、双向等位特异 PCR、TaqMan 探针技术、能量转移标记的等位特异 PCR、焦磷酸测序、基质辅助激光解吸电离飞行时间质谱等<sup>[17-19]</sup>。其中焦磷酸测序技术由于自动化、高通量、检测时间短、成本低和结果准确而倍受关注。由于焦磷酸测序技术可以进行很精确的定量分析,因此可以把样品的基因组 DNA 进行等量混合,进行一次焦磷酸测序,就能得到该 SNP 等位基因在这些样本中的频率。Nordfors 等<sup>[20]</sup>已经做过高达 1126 样本的 PCR 产物的混合来确定 SNP 的频率。Lavebratt 等<sup>[21-22]</sup>把基因组 DNA 等量混合成 DNA 池来检测等位基因的频率。Wasson 等<sup>[23]</sup>也利用焦磷酸测序技术来确定 SNP 频率。这种做法极大地减少了所需的实验次数和实验消耗。当然,由于样品的 DNA 难以做到很精确的等量混合,所以只适合于样品的初步筛选检测,精确的统计还需要单个样品的分别检测;另一方面,这种做法只能检测等位基因的频率,无法进行基因型的检测。

本研究只检测了 CTSL 基因的一个 SNP 位点,实际上 CTSL 基因可能含有多个 SNP 位点。SNP 位点并不是独立遗传的,而是在染色体上倾向于以一个整体遗传给后代。成组遗传的 SNP 位点在一代又一代的遗传中绝少发生重组。于是,这样的一组 SNP 位点类型也就是单体型(haplotype)<sup>[24]</sup>。由于单体型包含着多个 SNP 的遗传信息,许多研究表明,在与复杂性状的相关分析中,采用单体型比单个 SNP 具有更好地统计分析效果。而用焦磷酸测序技术检测单体型是十分方便和有效的<sup>[25]</sup>。

#### 4 小结

单核苷酸多态性(SNP)是第三代分子标记, 由于其在基因组中分布的高密度性, 被认为在遗传分析方面有着广阔的应用前景。利用遗传标记与突变基因足够靠近时与其连锁不平衡的遗传现象, 通过二者间相关性分析, 使研究动植物复杂遗传特征中的主要遗传因素成为可能。由于 SNP 的数量十分庞大, 因此要求检测方法有效、准确、价廉、DNA 样本用量小、并能够实现自动化。焦磷酸测序技术是一种实时定量的小片段 DNA 测序技术, 不依赖电泳、荧光标记、分子杂交, 适宜已知 SNP 的检测。现在已经有专门的商品化的仪器系统具备同时对大量样品(96个)进行测序分析的能力, 因此, 可以说焦磷酸测序技术为大通量、低成本、适时、快速、直观地进行 SNP 研究和临床检验提供了非常理想的技术平台。

#### 参考文献:

- [1] Wang D G, Fan J B, Siao C. Large-scale Identification, mapping, and genotyping of single nucleotide polymorphisms in the human genome[J]. *Science*, 1998, 280:1077-1082.
- [2] Nishida N, Tanabe T, Takasu M, *et al.* Further development of multiplex single nucleotide polymorphism typing method, the DigiTag2 assay [J]. *Anal Biochem*, 2007, 364(1):78-85.
- [3] Chessa S, Chiatti F, Ceriotti G, *et al.* Development of a single nucleotide polymorphism genotyping microarray platform for the identification of bovine milk protein genetic polymorphisms [J]. *J Dairy Sci*, 2007, 90(1):451-464.
- [4] Alifrangis M, Enosse S, Pearce R, *et al.* A simple, high-throughput method to detect Plasmodium falciparum single nucleotide polymorphisms in the dihydrofolate reductase, dihydropteroate synthase, and P. falciparum chloroquine resistance transporter genes using polymerase chain reaction- and enzyme-linked immunosorbent assay-based technology [J]. *Am J Trop Med Hyg*, 2005, 72(2):155-162.
- [5] Nyren P. The history of pyrosequencing [J]. *Methods Mol Biol*, 2006, 373:1-14.
- [6] Gharizadeh B, Akhras M, Nourizad N, *et al.* Methodological improvements of pyrosequencing technology [J]. *J Biotechnol*, 2006, 124 (3): 504-511.
- [7] Wang H, Elbein S C. Detection of allelic imbalance in gene expression using pyrosequencing [J]. *Methods Mol Biol*, 2006, 373:157-176.
- [8] Le Boulay C, van Wormhoudt A, Sellos D. Cloning and expression of cathepsin L-like proteinases in the hepatopancreas of the shrimp *Penaeus vannamei* during the intermolt cycle [J]. *J Comp Physiol*, 1996, 166:310-318.
- [9] Glenn K L, Grapes L, Suwanasopee T, *et al.* SNP analysis of AMY2 and CTSL genes in *Litopenaeus vannamei* and *Penaeus monodon* shrimp [J]. *Anim Genet*, 2005, 36(3):235-236.
- [10] Aljanabi S M, Martinez I. Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques [J]. *Nucleic Acids Res*, 1997, 25 (22):4692-4693.
- [11] 吴为人, 李维明. 基于性状标记回归的 QTL 区间测验方法 [J]. *遗传*, 2001, 23(2):143-146.
- [12] 王鸿霞, 吴长功, 相建海. 凡纳滨对虾繁殖中不同亲本对子代遗传贡献率的差异 [J]. *动物学报*, 2006, 52 (1):175-181.
- [13] 王兴强, 曹梅, 马旻, 等. 盐度对凡纳滨对虾存活、生长和能量收支的影响 [J]. *海洋水产研究*, 2006, 27(1):8-13.
- [14] 袁路, 蔡生力. 温度、盐度对凡纳滨对虾精荚再生和精子质量的影响 [J]. *水产学报*, 2006, 30 (1): 63-68.
- [15] 高会江, 郝立国, 罗静, 等. 数量性状两个候选基因检测效率分析 [J]. *东北农业大学学报*, 2006, 37 (3):325-332.
- [16] Botstein D, Risch N. Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease [J]. *Nat Genet*, 2003, 33:228-237.
- [17] 汪维鹏, 倪坤仪, 周国华. 单核苷酸多态性检测方法的研究进展 [J]. *遗传*, 2006, 28(1):117-126.
- [18] King C R, Scott-Horton T. Pyrosequencing (R): a simple method for accurate genotyping [J]. *Methods Mol Biol*, 2006, 373:39-56.
- [19] Wang H, Elbein S C. Detection of allelic imbalance in gene expression using pyrosequencing [J]. *Methods Mol Biol*, 2006, 373:157-176.
- [20] Nordfors L, Jansson M, Sandberg G, *et al.* Large-scale genotyping of single nucleotide polymorphisms by Pyrosequencing trade mark and validation against the 5' nuclease (Taqman (R)) assay [J]. *Hum Mutat*,

- 2002,19(4):395-401.
- [21] Lavebratt C, Sengul S. Single nucleotide polymorphism (SNP) allele frequency estimation in DNA pools using Pyrosequencing[J]. Nat Protoc, 2006,1(6):2573-2582.
- [22] Lavebratt C, Sengul S, Jansson M, *et al.* Pyrosequencing-based SNP allele frequency estimation in DNA pools[J]. Hum Mutat, 2004 Jan;23(1):92-97.
- [23] Wasson J, Skolnick G, Love Gregory L, *et al.* Assessing allele frequencies of single nucleotide polymorphisms in DNA pools by pyrosequencing technology [J]. Biotechniques, 2002, 32 (5): 1144-1150.
- [24] Klein R D. The pain protective haplotype; introducing the modern genetic test[J]. Clin Chem, 2007,53(6):1007-1009.
- [25] Lotsch J, Belfer I, Kirchof A, *et al.* Reliable Screening for a pain-protective haplotype in the GTP Cyclohydrolase 1 Gene (GCH1) through the use of 3 or fewer single nucleotide polymorphisms[J]. Clin Chem,2007,53(6):1010-1015.

### Detection of CTSL gene single nucleotide polymorphisms in *Litopenaeus vannamei* shrimp by pyrosequencing

ZENG Di-gang, CHEN Xiao-han, PENG Min, LI Yong-mei, YANG Chun-ling,  
MA Ning, JIANG Wei-ming, LI Ming  
(Guangxi Institute of Fisheries, Nanning 530021, China)

**Abstract:** Pyrosequencing is a new DNA sequencing method, it was used to analyze single nucleotide polymorphisms recently years. Cathepsin-L gene is a lysosomal protease gene that is involved in the intermolt cycle of shrimp. In this study, genotyping of cathepsin-L gene (GenBank accession no. AY366355) SNP (C681G) in 96 *Litopenaeus vannamei* shrimps was performed using pyrosequencing. The C/C, C/G and G/G genotype were found in this population, and the frequency were 0.81, 0.16 and 0.03, respectively. The frequency distribution was consistent with the law of Hardy-Weinberg. The statistic analyses showed that C681G SNP had no significantly association BW in this population ( $P>0.05$ ). The results of this study showed that pyrosequencing is a versatile technique that could improve the efficiency of SNP analysis for shrimp genomic research.

**Key words:** *Litopenaeus vannamei*; pyrosequencing; cathepsin-L gene; single nucleotide polymorphisms (SNP)