ความถี่จีโนไทป์ของสนิปและระดับการแสดงออกของยีนที่เกี่ยวข้องกับการเติบโตและการเจริญพันธุ์ของกุ้งกุลาดำ *Penaeus monodon*

นาย สายัณห์ ประกอบเพชร

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาเทคโนโลยีชีวภาพ
คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2552

GENOTYPE FREQUENCIES OF SNP AND EXPRESSION LEVELS OF GENES INVOLVED IN GROWTH AND REPRODUCTIVE MATURATION OF THE GIANT TIGER SHRIMP *Penaeus monodon*

Mr. Sayan Prakobpetch

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science Program in Biotechnology
Faculty of Science
Chulalongkorn University
Academic Year 2009

Thesis Title	Genotype frequencies of SNP and expression levels of genes involved in growth and reproductive maturation of the giant tiger shrimp *Penaeus monodon*
By	Mr. Sayan Prakobpetch
Field of Study	Biotechnology
Thesis Advisor	Professor Piamsak Menasveta, Ph.D.
Thesis Co-Advisor	Bavornlak Khamnamtong, Ph.D.

Accepted by the Faculty of Science, Chulalongkorn University in Partial Fulfillment of the Requirements for the Master's Degree

.....Dean of the Faculty of Science

(Professor Supot Hannongbua, Dr.rer.nat.)

THESIS COMMITTEE

.....Chairman

(Associate Professor Chanpen Chanchao, Ph.D.)

.....Thesis Advisor

(Professor Piamsak Menasveta, Ph.D.)

.....Thesis Co-Advisor

(Bavornlak Khamnamtong, Ph.D.)

.....Examiner

(Associate Professor Somkiat Piyatiratitivorakul, Ph.D.)

.....Examiner

(Sirawut Klinbunga, Ph.D.)

.....External Examiner

(Sittiruk Roytrakul, Ph.D.)

สายัณห์ ประกอบแพชร : ความถี่จีโนไทป์ของสนิปและระดับการแสดงออกของยีนที่เกี่ยวข้องกับการเติบโตและการเจริญพันธุ์ของกุ้งกุลาดำ *Penaeus monodon* (GENOTYPE FREQUENCIES OF SNP AND EXPRESSION LEVELS OF GENES INVOLVED IN GROWTH AND REPRODUCTIVE MATURATION OF THE GIANT TIGER SHRIMP *Penaeus monodon*) อ.ที่ปรึกษาวิทยานิพนธ์หลัก: ศ.ดร.เปี่ยมศักดิ์ เมนะเศวต, อ.ที่ปรึกษาวิทยานิพนธ์ร่วม: ดร.บวรลักษณ์ คำน้ำทอง, 150 หน้า.

ปัจจุบันข้อมูลความสัมพันธ์ระหว่างความหลากหลายของจีโนไทป์และฟีโนไทป์ในกุ้งกุลาดำ (*Penaeus monodon*) ยังมีจำกัด การวิเคราะห์สนิปในยีนเป็นวิธีการหนึ่งที่มีประสิทธิภาพสำหรับค้นหายีนที่ถ่ายทอดลักษณะที่มีความสำคัญทางเศรษฐกิจของสิ่งมีชีวิต เช่น ลักษณะการโตเร็ว การเจริญพันธุ์ ทำการค้นหาลำดับนิวคลีโอไทด์ที่บริเวณ 5′ untranslated region (5′UTR) ของยีน *nuclear autoantigenic sperm protein (NASP)* และ *NADP-dependent leukotriene B4 12-hydroxydehydrogenase (LTB4DH)* ซึ่งมีรายงานว่าเป็นยีนที่เกี่ยวข้องกับการเจริญพันธุ์ โดยลำดับนิวคลีโอไทด์บริเวณ 5′UTR ที่หาได้นั้นมีความยาว 640 และ 767 เบส ตามลำดับ จากนั้นออกแบบไพรเมอร์ที่จำเพาะกับบริเวณ 5′UTR ของทั้งสองยีนนี้ (ผลิตภัณฑ์พีซีอาร์มีขนาด 429 และ 592 เบส ตามลำดับ) เพื่อค้นหาสนิปบริเวณดังกล่าวในกุ้งกุลาดำที่มีอายุ 14 เดือน (N = 66) ผลจากการวิเคราะห์โดย PCR-SSCP พบว่ายีน *NASP* มีจีโนไทป์ของ SSCP 1 จีโนไทป์ ส่วนยีน *LTB4DH* พบ 2 จีโนไทป์คือ A และ B (N = 37 และ 29 ตามลำดับ)

นอกจากนี้ ทำการวิเคราะห์สนิปของยีน *dolichyl diphosphooligosaccharyl protein transferase (DDPG), thyroid hormone binding protein (THBP), epidermal growth factor (EGF)* และ *insulin degrading enzyme (IDE)* ซึ่งมีรายงานว่าเป็นยีนที่เกี่ยวข้องกับการเติบโตโดยวิเคราะห์กับตัวอย่างกุ้งอายุ 3 เดือน พบว่ายีน *DDPG* (376 bp, N = 110) และ *IDE* (203 bp, N = 110) มีจีโนไทป์ของ SSCP เพียงจีโนไทป์เดียว ขณะที่ยีน *EGF* (240 bp) พบ 2 จีโนไทป์คือ A และ B (N = 266 และ 74 ตามลำดับ) ส่วนยีน *THBP* (595 bp) นั้นพบทั้งหมด 4 จีโนไทป์ด้วยกันคือ A, B, C และ D (N = 113, 4, 89 และ 134 ตามลำดับ) เมื่อนำตัวแทนของแต่ละจีโนไทป์ไปหาลำดับนิวคลีโอไทด์ พบ สนิป 7 ตำแหน่งและพบ 3 indels ระหว่างจีโนไทป์ A และ B ของยีน *LTB4DH* นอกจากนี้พบสนิป 8 ตำแหน่งระหว่าง 4 จีโนไทป์ของยีน *THBP* คือ $A \longrightarrow T_{80}$, $T \longrightarrow C_{137}$, $G \longrightarrow A_{149}$, $A \longrightarrow G_{306}$, $A \longrightarrow G_{312}$, $T \longrightarrow C_{382}$, $A \longrightarrow G_{422}$ และ $A \longrightarrow G_{494}$ และพบสนิป 2 ตำแหน่งระหว่างจีโนไทป์ A และ B ของยีน *EGF* คือ $A \longrightarrow C_{32}$ และ $C \longrightarrow A_{75}$

เมื่อวิเคราะห์ความสัมพันธ์ระหว่างสนิปจากจีโนไทป์ของ SSCP ของยีน *LTB4DH* กับค่า Gonadosomatic index (GSI) ของกุ้งกุลาดำอายุ 14 เดือน พบว่าไม่มีความแตกต่างอย่างมีนัยสำคัญทางสถิติ (P > 0.05) ในทางตรงกันข้าม พบความแตกต่างอย่างมีนัยสำคัญทางสถิติ (P < 0.05) ระหว่าง SSCP จีโนไทป์ ของยีน *THBP* กับน้ำหนักตัวของกุ้งกุลาดำอายุ 3 เดือน (N = 340) แต่ไม่พบความแตกต่างดังกล่าวในยีน *EGF* (P > 0.05)

เมื่อวิเคราะห์ระดับการแสดงออกของยีน *NASP* และ *LTB4DH* ในรังไข่กุ้งแม่พันธุ์และยีน *IDE*, *EGF* และ *THBP* ในตับกุ้งวัยรุ่นด้วย Quantitative real-time PCR พบว่าไม่มีความแตกต่างของระดับการแสดงออกของยีน *NASP* ในรังไข่กุ้งอายุ 14 เดือนที่มีค่า GSI ต่างกัน (P > 0.05) แต่พบความแตกต่างของระดับการแสดงออกของยีน *LTB4DH* ระหว่างกลุ่มตัวอย่างที่มีค่า GSI ต่างกันโดยพบระดับการแสดงออกของยีนในกุ้งกลุ่มที่ II (ค่า GSI = 0.5 – 1.0%, N = 12) สูงกว่าในกุ้งกลุ่มที่ I (ค่า GSI < 0.5%, N = 6) อย่างมีนัยสำคัญทางสถิติ (P < 0.05) แต่ไม่แตกต่างกับกุ้งกลุ่มที่ III (ค่า GSI > 1.0%, N = 12) อย่างไรก็ตาม ไม่พบความแตกต่างอย่างมีนัยสำคัญทางสถิติ (P > 0.05) ของระดับการแสดงออกของยีน *LTB4DH* ในรังไข่กุ้งที่มี SSCP จีโนไทป์ต่างกัน (จีโนไทป์ A และ B)

นอกจากนี้ ไม่พบความแตกต่างของระดับการแสดงออกของยีน *IDE*, *EGF* และ *THBP* ในตับกุ้งอายุ 3 เดือนระหว่างกลุ่มโตเร็ว (ตัวใหญ่ N = 22) และกลุ่มโตช้า (ตัวเล็ก N = 22) เช่นเดียวกับผลการแสดงออกของยีนเหล่านี้ในตับกุ้งวัยรุ่นที่มี SSCP จีโนไทป์ต่างกัน (P > 0.05)

สาขาวิชา.....เทคโนโลยีชีวภาพ..... ลายมือชื่อนิสิต สายัณห์ ประกอบแพชร

ปีการศึกษา.....2552..... ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก.....

ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์ร่วม บวรลักษณ์ คำน้ำทอง

5072507123 : MAJOR BIOTECHNOLOGY
KEYWORDS: *Penaeus monodons* / SNP / PCR-SSCP / POLYMORPHISM

SAYAN PRAKOBPETCH: GENOTYPE FREQUENCIES OF SNP AND EXPRESSION LEVELS OF GENES INVOLVED IN GROWTH AND REPRODUCTIVE MATURATION OF THE GIANT TIGER SHRIMP *Penaeus monodon* THESIS ADVISOR: PROF. PIAMSAK MENASVETA, Ph.D. THESIS CO-ADVISOR: BAVORNLAK KHAMNAMTONG, Ph.D., 150 pp.

The information on association between genotypic and phenotypic variations in the giant tiger shrimp (*Penaeus monodon*) is limited at present. Analysis of gene-based single nucleotide polymorphism (SNP) is one of the efficient approaches for discovery of genes significantly contributed in commercially important traits of this species. To identify SNP in the 5′ untranslated region (5′UTR) of *nuclear autoantigenic sperm protein* (*NASP*) and *NADP-dependent leukotriene B4 12-hydroxydehydrogenase* (*LTB4DH*), genome walking analysis was carried out and successfully identified their 5′UTR of 640 and 767 bp in length, respectively. Sequence-specific primers were designed and the amplified 429 and 592 bp fragments of respective genes were generated. SNP in *NASP* and *LTB4DH* of 14-month-old shrimp (N = 66) was examined by PCR-SSCP. *NASP* was monomorphic whereas 2 polymorphic SSCP genotypes (A and B found in 37 and 29 individuals, respectively) of *LTB4DH* were observed.

In addition, polymorphism in *dolichyl diphosphooligosaccharyl protein glycotransferase* (*DDPG*, 376 bp), *thyroid hormone binding protein* (*THBP*, 595 bp), *epidermal growth factor* (*EGF*, 240 bp) and *insulin degrading enzyme* (*IDE*, 203 bp) gene segments was also examined in domesticated 3-month-old juveniles. The amplified product of *DDPG and IDE* was monomorphic (N = 110 for each gene). In contrast, 2 SSCP genotypes of *EGF* (A and B found in 266 and 74 individuals, respectively) and 4 SSCP genotypes of *THBP* (A, B, C and D found in 113, 4, 89 and 134 individuals, respectively) were observed. DNA sequencing of different SSCP genotypes of these genes was carried out. Seven SNPs and 3 indels were found between genotypes A and B of *LTB4DH*. In addition, eight SNPs positions ($A \rightarrow T_{80}$, $T \rightarrow C_{137}$, $G \rightarrow A_{149}$, $A \rightarrow G_{306}$, $A \rightarrow G_{312}$, $T \rightarrow C_{382}$, $A \rightarrow G_{422}$ and $A \rightarrow G_{494}$) were found between four genotypes of *THBP* and two SNPs positions ($A \rightarrow C_{32}$ and $C \rightarrow A_{75}$) were found among genotypes A and B of *EGF*.

Association between SNP through SSCP genotypes of *LTB4DH* and the Gonadosomatic index (GSI) of 14-month-old *P. monodon* was not significant ($P > 0.05$). In contrast, a significant difference between frequencies of different SSCP genotypes and the body weight of 3-month-old *P. monodon* (N = 340) was observed in *THBP* ($P < 0.05$) but not in *EGF* ($P > 0.05$).

Quantitative real-time PCR of *NASP* and *LTB4DH* in ovaries of broodstock and *IDE*, *EGF* and *THBP* in hepatopancreas of juvenile *P. monodon* was examined. The expression levels of *NASP* in ovaries of 14-month-old shrimp exhibiting different GSI values were not significantly different ($P > 0.05$). In contrast, *LTB4DH* in ovaries of group II (GSI = 0.5 – 1.0%, N = 12) was significantly greater than that of group I (GSI < 0.5%, N = 6; $P < 0.05$) but not different from that of group III (GSI > 1.0%, N = 12; $P > 0.05$) broodstock. However, the expression levels of ovarian *LTB4DH* in shrimp exhibiting SSCP genotypes A and B were not significantly different ($P > 0.05$).

The expression levels of *IDE*, *EGF* and *THBP* in hepatopancreas of 3-month-old shrimp exhibiting fast (large, N = 22) and slow (small, N = 22) growth rates were not significantly different ($P > 0.05$). Likewise, the expression levels of these genes in hepatopancreas of juveniles exhibiting different SSCP genotypes were also not significantly different ($P > 0.05$).

Field of Study:.....Biotechnology.....Student's Signature.....Sayan Prakobpetch.....

Academic Year:.....2009.....Advisor's Signature.....

Co-Advisor's Signature.....B. Khamnamtong.....

ACKNOWLEDGEMENTS

I would like to express my deepest sense of gratitude to my advisor, Professor Dr. Piamsak Menasveta and my co-advisor, Dr. Bavornlak Khamnamtong and for their great help, guidance, encouragement, valuable suggestion and support throughout my study.

My gratitude is also extended to the thesis committee; Associate Professor Dr. Chanpen Chanchao, Associate Professor Dr. Somkiat Piyatiratitivorakul, Dr. Sirawut Klinbunga and Dr. Sittilak Roytrakul for their recommendations and useful suggestion.

I would like to acknowledge Aquatic Molecular Genetics and Biotechnology Laboratory (AMGB) for providing facilities and National Center for Genetic Engineering and Biotechnology (BIOTEC), Thailand Graduate Institute of Science and Technology (TGIST) for my studentship.

Many thanks are also expressively to all of my friends in the laboratory particularly Miss Sirikarn Prasertluk and Miss Kanchana Sittikhankeaw for their help, suggession and friendship. Thank to all of my friends in Biotechnology Program of RU and CU for friendships. Finally, I would also like to express my deepest sense of gratitude to Assistant Professor Dr. Thidarat Eksittikul for her help and suggestion during a study period.

Finally, I would like to express my deepest gratitude to my parents, my sister and my brother for their love, care, understanding and encouragement extended throughout my study.

CONTENTS

Page

Page

Page

Page

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

Page

Page

Page

Page

Page

LIST OF ABBREVIATIONS

bp	base pair
°C	degree Celcius
dATP	deoxyadenosine triphosphate
dCTP	deoxycytosine triphosphate
dGTP	deoxyguanosine triphosphate
dTTP	deoxythymidine triphosphate
DEPC	dethylpyrocarbonate
DNA	deoxyribonucleic acid
EDTA	ethylene diamine tetraacetic acid
HCl	hydrochloric acid
IPTG	isopropyl-thiogalactoside
Kb	kilobase
M	molar
$MgCl_2$	magnesium chloride
mg	milligram
ml	millilitre
mM	millimolar
NaOH	sodium hydroxide
ng	nanogram
OD	optical density
PCR	Polymerase Chain Reaction
rpm	revolution per minute
RNA	ribonucleic acid
RNase A	ribonuclease A

RT	reverse transcription
SDS	sodium dodecyl sulfate
Tris	tris (hydroxyl methyl) aminomathane
μg	microgram
μl	microlitre
μM	micromolar
UV	ultraviolet
V	voltage

CHAPTER I

INTRODUCTION

1.1 Background information

Farming of the giant tiger shrimp (*Penaeus monodon*) has achieved a considerable economic and social importance in Thailand, constituting a significant source of income and employment. However, the shrimp industry presently faces several problems such as the reduction of high quality wild broodstock and loss of the cultured production due to bacterial and viral infections and large size differences of cultivated *P. monodon* in the same cultured pond. The annual production of *P. monodon* was significantly reduced from approximately 200,000 MT during 1992 - 2002 to about 30,000 MT in 2006 (Limsuwan, 2004; FAO Fishery Statistic, 2009). As a result, the Pacific white shrimp (*Litopenaeus vannamei*) has been introduced to Thailand as an alternative cultured species and initially contributed approximately 20,000 MT of the production in 2002 and dramatically increased to 170,000 and 220,000 MT in 2003 and 2004, respectively (Limsuwan, 2004).

The price of *L. vannamei* is quite low and broodstock used relies almost entirely on genetic improved stocks brought from different sources. In addition, the labor costs in Thailand are higher than other countries (e.g. Vietnam and China) preventing the advantage of competition for the world market. In contrast, the market of premium-sized *P. monodon* is still open because *L. vannamei* is not suitable for that market. Accordingly, *P. monodon* culture is currently promoted for increasing the production of this species in Thailand.

Genetic improvement and other biotechnology applications are crucial to the future development of the shrimp industry (Benzie, 1998; Browdy, 1998). Progress in genetic and biotechnology researches in penaeid shrimp have been slow because a lack of knowledge on fundamental aspects of their biology (Benzie, 1998). A research concerning domestication of *P. monodon* has been carried out by production of high quality pond-reared *P. monodon* broodstock (Withyachumnarnkul *et al*., 1998) but

was recently collapsed by the WSSV infection. The domesticated program of *P. monodon* is start over and it is expected that selective breeding programs of *P. monodon* will provide shrimp having commercially desired phenotypes (e.g. high growth rate and/or disease resistance) and produce *P. monodon* stocks with the ability to induce high quality egg development in domesticated females without the irreversible side-effects caused by a typical eyestalk ablation technique (Lyons and Li, 2002).

Sustainable shrimp industry can be promoted through applications of knowledge from genetic and biotechnological studies including development of genetic markers to assist selective breeding programs of *P. monodon*. The basic knowledge gained can be applied for increasing culture efficiency of *P. monodon*. Researches in both basic (molecular and cellular levels of gonad developmental processes) and applied (microsatellite/SNP genotyping to assist selective breeding of high quality pond-related *P. monodon*) disciplines are required to elevate the culture efficiency and subsequence to sustain the industry of this economically important species.

Appropriate genetic markers can be used to elevate culture and management efficiency of *P. monodon* in Thailand. The long term objectives for domestication and genetic selection of *P. monodon* require comparison of important traits in different stocks/lines of wild and domesticated *P. monodon* using quantitative trait loci (QTL) as marker assisted selection (MAS). Genetic improvement of shrimp can then be carried out more effectively. Isolation and characterization of molecular markers involving important physiological and biological processes of *P. monodon* is thus necessary to provide the means to sustain the production of this species. Integration of the knowledge from various molecular disciplines would provide phenotype-linked molecular markers that can be used to assist the present domestication program of *P. monodon*.

1.2 Objectives of the thesis

The objectives of this thesis are identification of single nucleotide polymorphism (SNP) in growth- and reproductive-related genes of *P. monodon* by single strand conformational polymorphism (SSCP) and DNA sequencing.

Correlations between frequencies of SSCP genotypes and the body weight (for growth-related genes) or gonadosomatic index (GSI; for reproduction-related genes) were examined. In addition, association between SNP by SSCP and expression levels of functionally important genes was also evaluated.

1.3 General introduction

The giant tiger shrimp (*Penaeus monodon* Fabricius, 1798) is one of the world's most economically important cultured crustaceans. Total aquaculture production of *P. monodon* increased gradually from 21,000 tons in 1981 to 200,000 tons in 1988. Later it sharply increased to nearly 500,000 tons in 1993. Since then, the production has been quite variable, ranging from a low of 480,000 tons in 1997 to a high of 676,000 tons in 2001 (FAO Fishery Statistic, 2009). The major producers of *P. monodon* are Thailand, Viet Nam, Indonesia, India, the Philippines, Malaysia and Myanmar. The exportation of the chilled and frozen *P. monodon* from Thailand worldwide since 2002 - 2007 was reported in Figure 1.1 and Table 1.1.

Marine shrimp farms and hatcheries are located along the coastal areas of Thailand where Nakorn Sri Thammarat and Surat Thani in Peninsular Thailand are the major parts of shrimp cultivation. In addition, Chanthaburi (eastern Thailand), Samut Sakhon and Samut Songkhram (central region) also significantly contribute on

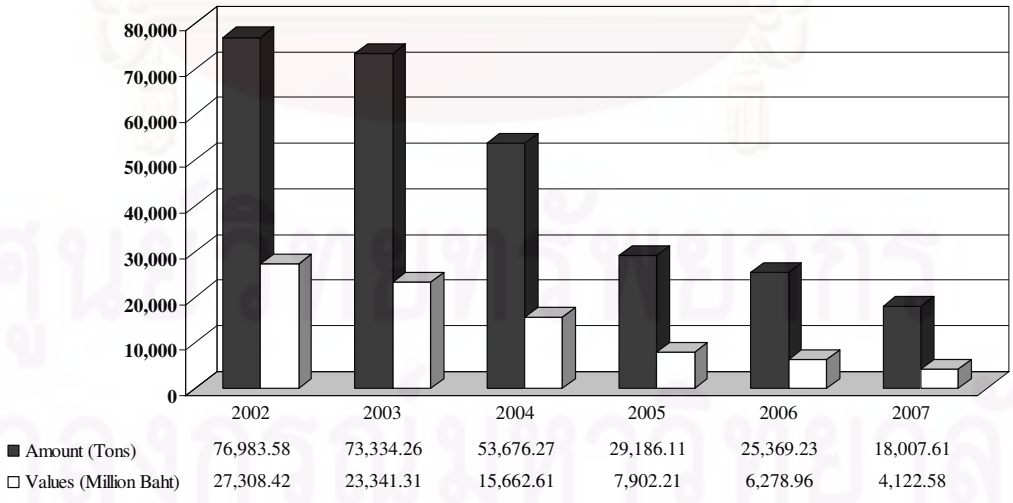


Figure 1.1 The chilled and frozen *P. monodon* exported from Thailand.

Table 1.1 The exportation of the chilled and frozen *P. monodon* from Thailand to worldwide countries between 2002-2007

Chilled and frozen shrimp	Year 2002		Year 2003		Year 2004		Year 2005		Year 2006		Year 2007	
COUNTRY	Amount (T)	Values (MB)	Amount (T)	Values (MB)	Amount (T)	Values (MB)	Amount (T)	Values (MB)	Amount (T)	Values (MB)	Amount (T)	Values (MB)
ALL COUNTRY	**76,983.58**	**27,308.42**	**73,334.26**	**23,341.31**	**53,676.27**	**15,662.61**	**29,186.11**	**7,902.21**	**25,369.23**	**6,278.96**	**18,007.61**	**4,122.58**
U.S.A.	36,609.45	13,690.15	37,700.69	12,156.86	22,623.97	6,534.96	12,146.95	2,992.49	15,248.38	3,599.27	6,822.45	1,498.44
JAPAN	17,131.66	6,476.46	15,237.90	5,380.86	12,725.46	4,304.66	6,662.47	2,260.72	2,523.07	801.97	1,871.61	528.38
KOREA, REPUBLIC OF	3,732.61	1,072.74	5,141.09	1,345.48	5,471.74	1,358.88	3,704.20	864.16	2,738.61	657.02	1,732.33	364.38
CANADA	4,748.05	1,760.06	5,452.53	1,766.59	3,520.70	1,174.95	1,965.59	578.14	2,003.42	528.09	1,725.49	455.28
HONG KONG	1,023.94	259.98	515.66	139.50	707.92	233.90	588.91	169.30	618.16	142.41	1,526.43	352.24
CHINA	942.41	256.76	361.16	73.89	631.25	140.61	663.80	104.85	581.67	71.37	1,520.22	203.63
TAIWAN	2,839.83	757.73	1,886.25	514.63	2,952.10	496.22	1,469.41	381.39	456.40	144.62	633.49	165.04
MALAYSIA	15.16	2.44	44.06	10.22	306.08	97.64	83.38	18.05	7.81	1.62	562.80	136.56
GERMANY, FEDERAL REPUBLIC	258.32	92.50	71.05	23.10	102.37	47.24	11.22	4.18	8.20	2.97	323.49	85.48
AUSTRALIA	2,544.18	818.20	2,668.64	717.87	1,478.74	398.87	295.99	107.53	317.05	107.25	246.39	69.67
VIETNAM	-	-	9.86	3.37	28.95	6.20	265.89	54.66	49.20	4.70	204.37	43.41
U. KINGDOM	242.17	70.76	-	-	284.72	118.55	94.74	33.05	65.88	19.65	198.68	59.61
CYPRUS	64.67	22.77	125.49	44.05	110.67	35.21	96.17	31.16	85.39	27.50	135.64	39.53
RUSSIAN FEDERATION	20.81	10.76	99.84	42.27	97.87	39.49	125.47	46.86	36.72	14.82	98.24	31.12
SINGAPORE	4,139.36	1,181.92	2,366.43	611.73	1,141.54	219.52	379.47	65.82	237.65	37.87	71.66	16.70
U. ARAB EMIRATES	0.01	0.01	14.10	5.29	2.78	1.45	1.66	0.64	4.00	2.08	38.54	5.00
KOREA (NORTH)	1,082.36	343.35	304.23	98.10	77.69	24.05	41.32	14.49	42.73	16.44	31.53	6.39
FRANCE	189.40	74.88	78.27	42.73	136.06	64.92	11.49	5.25	0.27	0.12	30.74	7.43
ITALY	13.20	3.31	10.36	2.49	68.03	9.28	23.46	6.17	12.55	4.11	30.32	4.86
POLAND	30.63	13.15	40.16	14.41	17.32	6.69	2.74	0.98	-	-	24.90	7.94
NETHERLANDS	81.66	26.24	45.44	13.59	3.44	1.29	0.32	0.19	30.18	13.48	23.13	6.91

Table 1.1 (Cont.)

Chilled and frozen shrimp	Year 2002		Year 2003		Year 2004		Year 2005		Year 2006		Year 2007	
COUNTRY	Amount (T)	Values (MB)	Amount (T)	Values (MB)	Amount (T)	Values (MB)	Amount (T)	Values (MB)	Amount (T)	Values (MB)	Amount (T)	Values (MB)
UKRAINE	4.00	1.60	8.95	3.57	21.34	8.10	39.67	14.21	12.38	5.18	21.06	5.53
NEW ZEALAND	315.07	102.24	392.89	120.03	391.39	112.96	174.70	52.47	76.78	19.58	19.47	5.03
MAURITIUS	-	-	-	-	-	-	-	-	-	-	19.16	3.81
BELGIUM	278.17	83.05	80.25	24.51	16.50	5.23	4.57	1.11	113.35	27.72	16.40	4.17
NEW CALEDONIA	-	-	-	-	-	-	-	-	-	-	12.14	1.21
SAUDI ARABIA	-	-	-	-	-	-	-	-	-	-	11.09	1.80
ISARAEL	25.31	9.45	55.32	21.00	12.44	4.35	10.08	3.39	12.01	4.67	8.48	2.56
SWEDEN	-	-	2.61	0.61	0.36	0.21	0.25	0.11	0.50	0.30	6.72	1.04
SWITZERLAND	37.82	14.56	38.91	11.51	138.85	47.47	92.57	30.77	34.01	12.02	5.13	1.18
QATAR	-	-	-	-	-	-	-	-	-	-	5.00	0.57
CHILE	-	-	-	-	-	-	-	-	-	-	4.93	2.30
LEBANON	-	-	0.40	0.24	14.44	8.11	-	-	0.56	0.27	4.90	1.72
CROATIA	12.20	2.57	4.00	0.73	-	-	4.32	1.73	1.50	0.36	2.24	0.97
LAO PEOPLES DEMOCRATIC	10.80	0.55	13.20	0.65	19.60	0.93	18.00	0.85	16.80	0.76	1.51	0.52
INDONESIA	41.00	10.32	39.49	9.10	60.05	17.60	-	-	-	-	1.47	0.20
DENMARK	16.95	5.80	-	-	-	-	-	-	0.02	0.00	0.53	0.06
PHILIPPINES	58.70	8.03	-	-	0.95	0.33	-	-	0.01	0.00	-	-
GUAM	50.93	18.34	69.46	23.44	36.92	8.16	45.84	14.09	6.92	1.95	-	-
SPAIN	44.49	6.84	28.72	9.94	1.28	0.90	34.12	7.04	0.00	0.00	-	-
FRENCH POLYNESIA (TAHITI)	238.44	65.08	287.62	69.61	322.90	87.88	84.74	23.41	-	-	-	-
INDIA	0.09	0.00	34.04	11.80	57.38	18.14	20.48	7.96	-	-	-	-
OTHER COUNTRIES	139.75	45.84	105.19	27.56	92.48	27.63	22.16	4.99	27.05	8.80	14.94	1.91

Source: Department of Fisheries

the country production. The intensive farming system has resulted in consistent production of marine shrimp of Thailand. Thailand has been regarded as one of the leading shrimp producer of cultivated shrimp for over two decades.

Still, *P. monodon* farming relies almost entirely on ocean-caught females for farm seed supply. This open reproductive cycle and reliance on wild stocks of *P. monodon* results in heavy exploitation of female broodstock from wild populations (Klinbunga, 1996). Mass production of larvae using closed-life-cycle, captive populations and selective breeding for improved culture performance, disease resistance, and other desirable traits of this species are required for the sustainable of the shrimp industry.

1.4 Taxonomy of *P. monodon*

Penaeid shrimp belong to the largest phylum in the animal kingdom, the Arthropoda. This group of animals is characterized by the presence of paired appendages and a protective cuticle or exoskeleton that covers the whole animal. The subphylum Crustacea is made up of 42,000, predominantly aquatic species belonging to 10 different classes. Within the class Malacostraca, shrimp, together with crayfish, lobsters and crabs, belong to the order Decapoda (Figure 1.2). Taxonomical recognition of *P. monodon* is illustrated below.

Phylum Arthropoda
Subphylum Crustacea
Class Malacostraca
Order Decapoda
Superfamily Penaeoidea
Family Penaeidae Rafinesque, 1815
Genus *Penaeus* Fabricius, 1798
Subgenus *Penaeus*
Species *Penaeus monodon*

Figure 1.2 Taxonomy of the black tiger shrimp, *Penaeus monodon* Fabricius, 1798 (Brusca and Brusca, 1990).

1.5 Morphology of *P. monodon*

The exterior of penaeid shrimp is distinguished by a cephalothorax with a characteristic hard rostrum, and by segmented abdomen (Figure 1.3). Most organs, such as gills, digestive system and heart, are located in the cephalothorax while the muscles concentrate in the abdomen. Appendages of the cephalothorax vary in appearance and function. In the head region, antennules and antennae perform sensory functions. The mandibles and the two pairs of maxillae form the jaw-like structures that are involved in food uptake. In the thorax region, the maxillipeds are the first three pairs of appendages, modified for food handling, and the remaining five pairs are the walking legs (pereiopods). Five pairs of swimming legs (pleopods) are found on the abdomen.

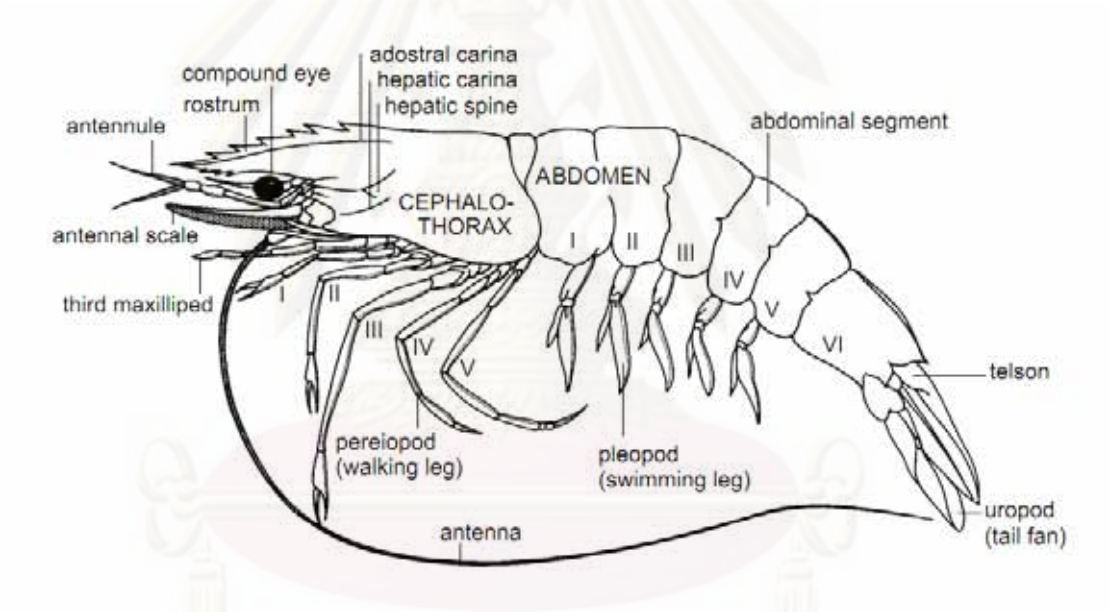


Figure 1.3 Lateral view of the external morphology of *P. monodon* (Primavera, 1990).

The internal morphology of penaeid shrimp is outlined in Figure 1.4. Penaeids and other arthropods have an open circulatory system and, therefore, the blood and the blood cells are called hemolymph and hemocytes, respectively. Crustaceans have a muscular heart that is dorsally located in the cephalothorax. The valved hemolymph vessels leave the heart and branch several times before the hemolymph arrives at the sinuses that are scattered throughout the body, where exchange of substances takes

place. After passing the gills, the hemolymph returns in the heart by means of three wide non-valved openings (Bauchau, 1981).

A large part of the cephalothorax in penaeid shrimp is occupied by the hepatopancreases. This digestive gland consists of diverticula of the intestine. Spaces between these hepatopancreatic tubules are hemolymph sinuses. The main functions of the hepatopancreases are the absorption of nutrients, storage of lipids and production of digestive enzymes (Johnson, 1980). One of the hemolymph vessels that leaves the heart ends in the lymphoid organ, where the hemolymph is filtered. This organ is located ventro-anteriorly to the hepatopancreases. The hemocytes are produced in haematopoietic tissue. This organ is dispersed in the cephalothorax, but mainly present around the stomach and in the onset of the maxillipeds. Lymphoid organ and haematopoietic tissue are not shown in Figure 1.4.

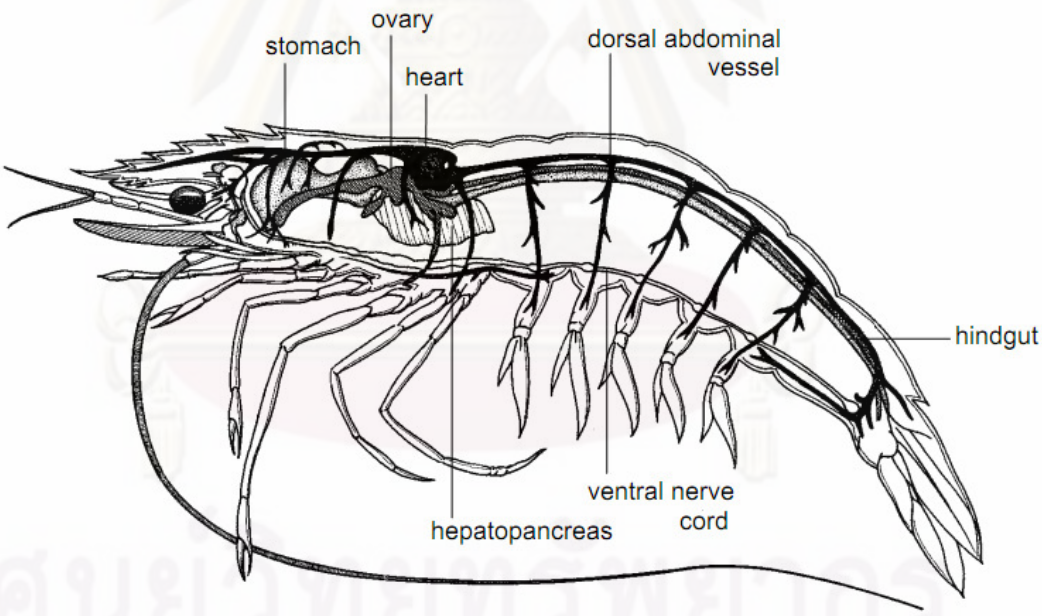


Figure 1.4 Lateral view of the internal anatomy of a female *Penaeus monodon* (Primavera, 1990).

1.6 Molecular biology approaches used in this thesis

1.6.1. Polymerase chain reaction (PCR)

The introduction of the Polymerase Chain Reaction (PCR) by Mullis and Faloona (1987) has opened a new approach for molecular genetic studies. This method is a molecular biology technique for enzymatically replicating DNA without using a living organism, such as *E. coli* or yeast and is a method using specific DNA sequences by the two oligonucleotide primers, 17 - 30 nucleotides in length. Million copies of the target DNA sequence can be synthesized from the low amount of starting DNA template within a few hours.

The PCR reaction components are composed of DNA template, a pair of primers for the target sequence, dNTPs (dATP, dCTP, dGTP and dTTP), buffer and heat-stable DNA polymerase (usually *Taq* polymerase). The amplification reaction consists of three steps; denaturation of double stranded DNA at high temperature, annealing to allow primers to form hybrid molecules at the optimal temperature, and extension of the annealed primers by heat-stable DNA polymerase. The cycle is repeated for 30 - 40 times (Figure 1.5). The amplification product is determined by agarose or polyacrylamide gel electrophoresis.

1.6.2 Genome walking analysis

Genome walking analysis is a method for identifying unknown genomic regions flanking a known DNA sequences. Initially, genomic DNA is separately digested with different blunt-end generating restriction endonucleases (usually, *Hae* III, *Dra* I, *Pvu* II and *Ssp* I). The digested genomic DNA in each tube was then ligated to the adaptor. The ligated product is used as the template for PCR amplification. PCR was carried out with the primer complementary to the adaptor (AP1) and the interesting gene (gene specific primer; GSP). The resulting product is amplified with nested primers (AP2 and nested GSP). The nested PCR products were cloned and characterized (Figure 1.6). This technique allows isolation of the promoter region of interesting genes and 3′ and 5′ untranslated region (UTR) that required further characterization of SNPs at 3′ and 5′ UTR.

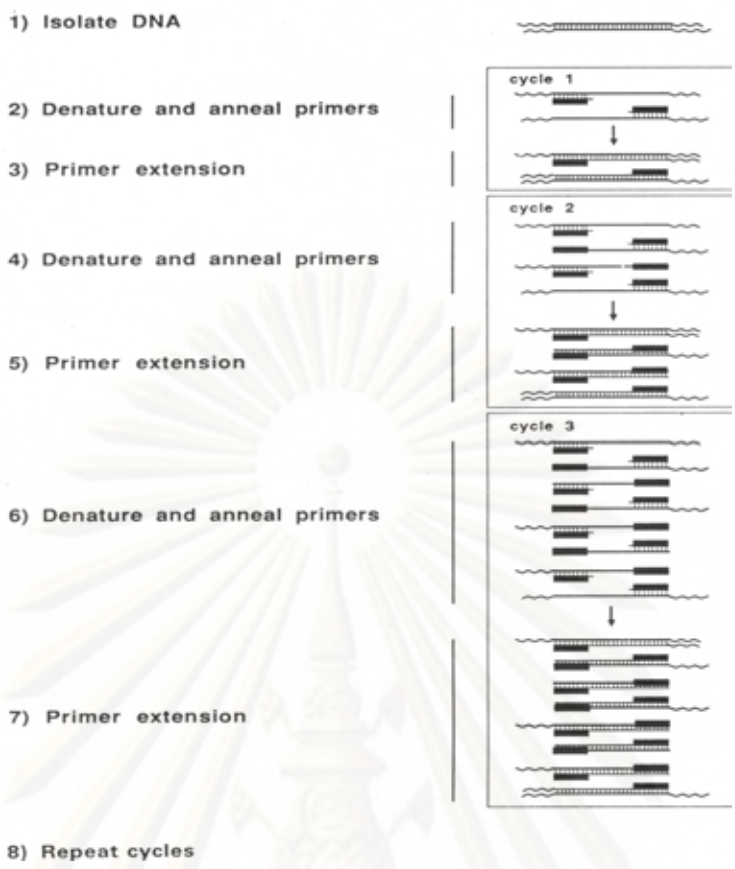


Figure 1.5 General illustration of the Polymerase Chain Reaction (PCR) for amplifying DNA.

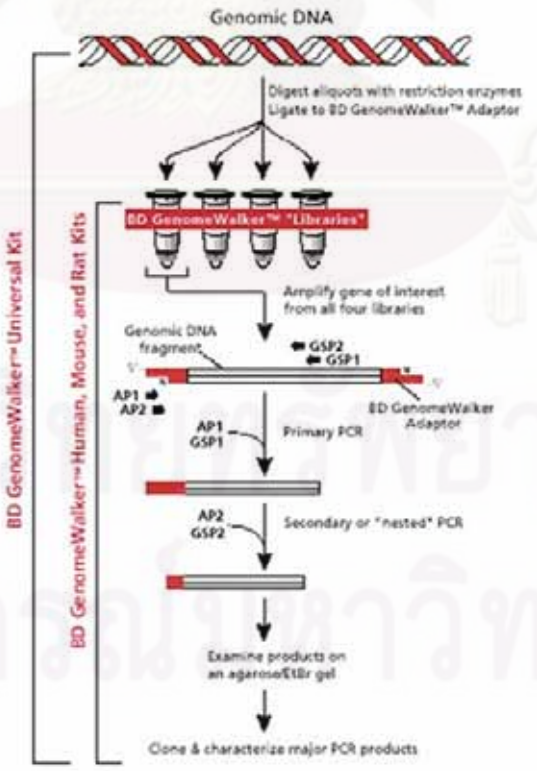


Figure 1.6 A flow chart illustrating the Genome walking analysis protocol.

1.6.3 DNA sequencing

DNA sequencing is the process of determining the exact order of the bases (A, T, C and G) in a piece of DNA. In essence, the DNA is used as a template to generate a set of fragments that differ in length from each other by a single base. The fragments are then separated by size, and the bases at the end are identified, recreating the original sequence of the DNA.

There are two general methods for sequencing of DNA segments: the "chemical cleavage" procedure described by Maxam and Gilbert, 1977 and the "chain termination" procedure was described by Sanger *et al*., 1977. Nevertheless, the latter method is more popular because chemical cleavage procedure requires the use of several hazardous substances. DNA fragments generated from PCR can be directly sequenced or alternatively, those fragments can be cloned and sequenced. This eliminates the need to establish a genome library and searching of a particular gene in the library.

DNA sequencing is the most optimal method for several genetic applications. This technique provides high resolution and facilitating interpretation. DNA fragments generated from PCR can be directly sequenced or alternatively, those fragments can be cloned and sequenced. This eliminates the need to establish a genome library and searching of a particular gene in the library.

The enzymatic sequencing approach has presently been developed to automated method (Figure 1.7). DNA sequences can be detected using a fluorescence-based system following labeling with a fluorescence dye. PCR allow the possibility to isolate homologous DNA sequences from any organism of interest with unprecedented speed. This greatly allows wider application of DNA sequencing analysis for population genetic and systematic studies.

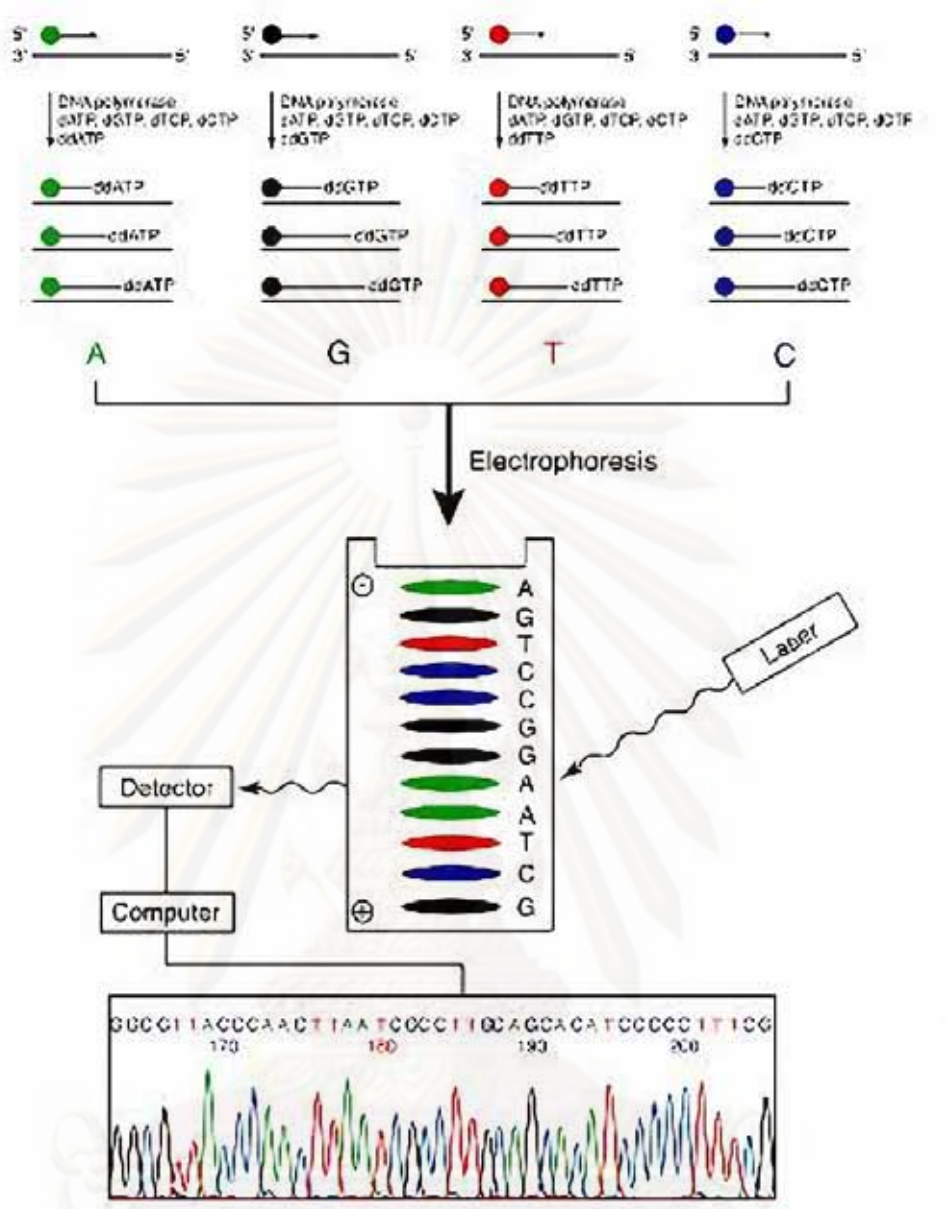


Figure 1.7 A schematic diagram illustrating principles of Automated DNA sequencing.

1.6.4 Single Nucleotide Polymorphism (SNP)

Single nucleotide polymorphism (SNP) is a single base change in a DNA sequence, with a usual alternative of two possible nucleotides at a given position. For such a base position with sequence alternatives in genomic DNA to be considered as an SNP (Figure 1.8), the lowest frequent allele should have a frequency of 1% or

greater. Although in principle, at each position of a sequence stretch, any of the four possible nucleotide bases can be present, SNPs are usually biallelic. One of the reasons for this, is the low frequency of single nucleotide substitutions at the origin of SNPs, estimated to being between 1 x 10^{-9} and 5 x 10^{-9} per nucleotide and per year at neutral positions in mammals. Therefore, the probability of two independent base changes occurring at a single position is very low. Another reason is due to a bias in mutations, leading to the prevalence of two SNP types (transitions).

Mutation mechanisms result either in transitions: purine-purine (A ↔ G) or pyrimidine-pyrimidine (C ↔ T) exchanges, or transversions: purine-pyrimidine and *vice versa* (A ↔ C, A ↔ T, G ↔ C, G ↔ T) exchanges. Theoretically, transversions should be as twice greater than transitions if mutations are random. However, the observed data indicate a clear bias towards transitions. Results obtained from identification of SNPs from EST sequences gave a transition to transversion ratio of 1.7 in human. Such a ration is greater in EST (2.3 - 4.0) and non-coding DNA (138 SNPs with a ratio of 2.36) in chicken (Vignal *et al.,* 2002). One probable explanation for this bias is the high spontaneous rate of deamination of 5-methyl cytosine to thymidine in the CpG dinucleotides, leading to the generation of higher levels of C/T SNPs (or G/A SNPs on the reverse strand). In addition, one base pair indels (insertions or deletions) is also recognized as SNPs even though they are occurred by different mechanisms.

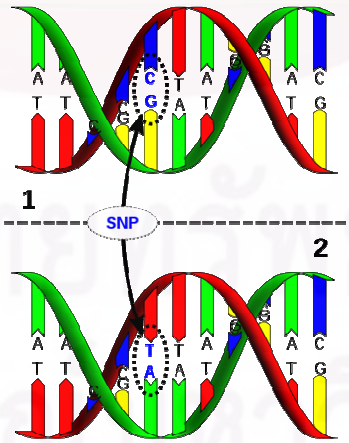


Figure 1.8 General illustration of the Single Nucleotide Polymorphism (SNP). (http://www.snipscreen.com/genetics.php)

1.6.4.1 Classification of SNP

SNP can be divided to noncoding SNP (or type II SNP) which is composed of that found in 5′ or 3′ nontranscribed regions, 5′ or 3′ untranslated regions, introns and intergenic spacers and coding SNP (or type I SNP) which is located in exons and results in non-replacement (synonymous) and replacement (non-synonymous) of amino acids in the polypeptide chains.

1.6.4.2 SNP identification and characterization

Several molecular techniques may be used to identify and subsequently to characterize SNP under investigation. DNA sequencing is the direct method for SNP identification. Nevertheless, several indirect methods are also commonly used for SNP identification (Table 1.2). This include restriction fragment length polymorphism (RFLP), SSCP and Denaturing gradient gel electrophoresis (DGGE), primer extension, oligonucleotide ligation assay (OLA), *Taq*Man exonuclease detection, PCR-allele specific amplification (PASA) and bidirectional PCR-allele specific amplification (Bi-PASA) and mass spectroscopy (MS).

Table 1.2 Comparisons of selected mutation screening methods (Shastry, 2002)

Method	Fragment length (bp)	Advantage	Disadvantage	Efficiency (%)
Single strand conformational polymorphism	~300	No expensive equipment	Small fragments, Temperature variation	80
Heteroduplex analysis	300-600	No expensive equipment	Conditions to be determined	80
Denaturing gradient gel electrophoresis	100-1000	Simple, long and short fragments	Gradient gel required, mutation in GC region may not be detected	100 with GC clamp
Enzymatic mismatch detection	300-1000	Long and short fragments	Identifies all kinds of mutations	100
Base excision sequence scanning	50-1000	Accurate	Expensive instruments	100
RNAase cleavage	1.6 kb	Longer and rapid analysis	Requires special kit	100
Chemical cleavage	1-2 kb	Large fragment	Multi-steps, labor intensive and hazardous chemicals	100
DNA sequencing	500	Rapid and easy, no addition sequencing	Labor intensives	100

1.6.5 Single Strand Conformational Polymorphism (SSCP)

SSCP analysis is one of the most widely used techniques for indirectly detection of mutations and variation of DNA (deletions, insertions and single nucleotide polymorphism, SNP). The amplified PCR product (usually less than 300 bp in length) is denatured and loaded into a low crosslink non-denaturing polyacrylamide gel (with or without glycerol supplementation). The principle of this technique relies on different mobility due to differential folding of the single stranded DNA (Figure 1.9).

Single strand DNA molecules take on secondary and tertiary structures (conformations) due to base pairing between nucleotides within each strand. These conformations depend on the length of the strand, and the location and number of regions of base pairing. They also depend on the primary sequence of the molecule, such that a nucleotide change at a particular position can alter its conformation. Accordingly, molecules differing in their conformations (e.g. due to a single nucleotide change) can be separated.

The major advantage of SSCP is that a large number of individuals may be simultaneously genotyped. Heteroduplexes (products from misannealing of single stranded DNA from different alleles) can occasionally resolve from homoduplexs (products from annealing of single stranded cDNA of the same alleles) give additional information on the presence of variants. Therefore, SSCP is regarded as one of the potential techniques that can be used to detect low polymorphism in various species prior to confirmation of the results by nucleotide sequencing. In addition, SSCP required small PCR amplicons which are relative easy to amplify.

The disadvantages of SSCP are reproducibility of the technique because SSCP patterns are strongly affected by temperature and degree of cross-linking. Additionally, multi-allelic patterns of some nuclear DNA markers may cause the SSCP patterns too complicated for estimation of allele frequencies precisely.

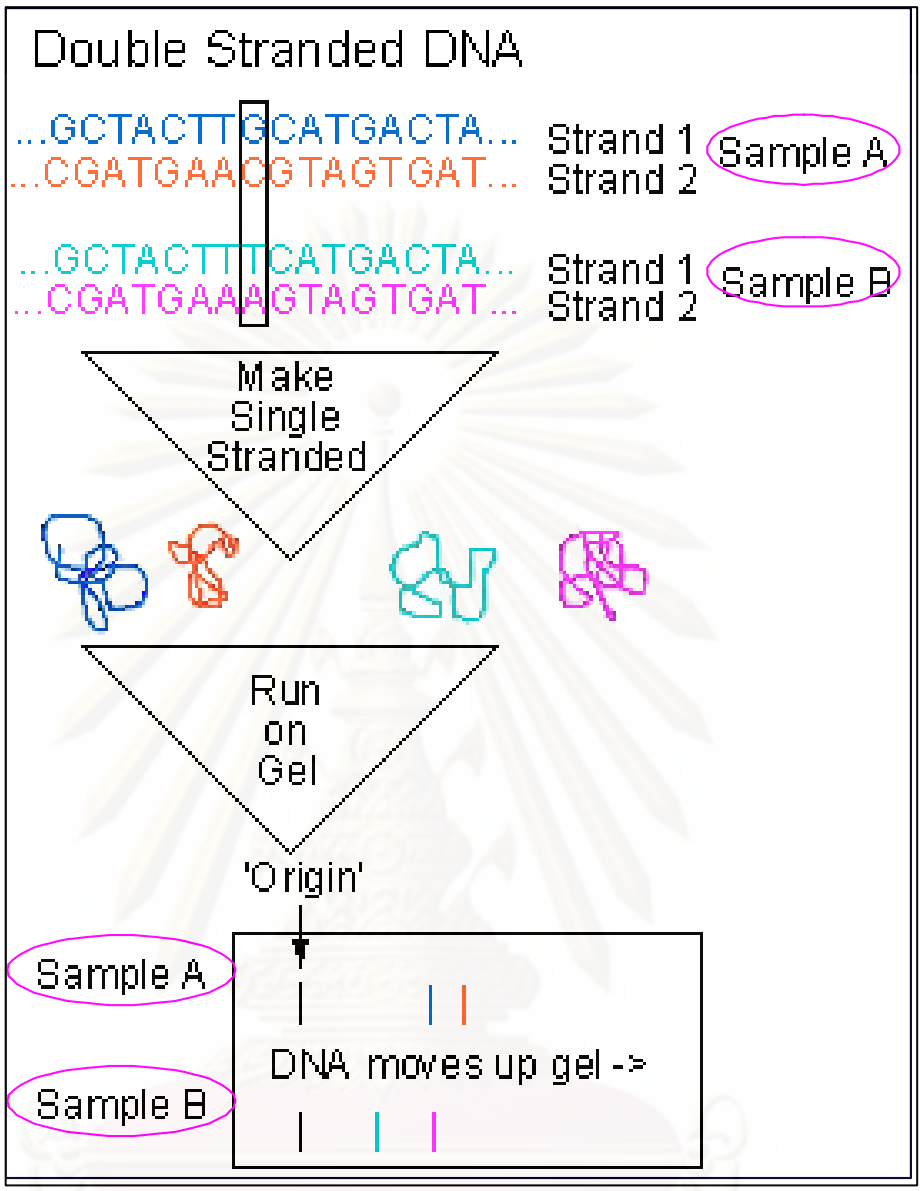


Figure 1.9 A schematic diagram of SSCP analysis.
(http://www.amonline.net.au/evolutionary_biology/images/sscp.gif)

1.6.6 Reverse Transcription-Polymerase Chain Reaction (RT-PCR)

RT-PCR is a comparable method of conventional PCR but the first strand cDNA template rather than genomic DNA was used as the template in the amplification reaction (Figure 1.10). It is a direct method for examination of gene expression of known sequence transcripts in the target species. Alternatively, RT-PCR can also be used to identify homologues of interesting genes by using degenerate primers and/or conserved gene-specific primers from the original species and the first

strand cDNA of the interesting species as the template. The amplified product is futher characterized by cloning and sequencing.

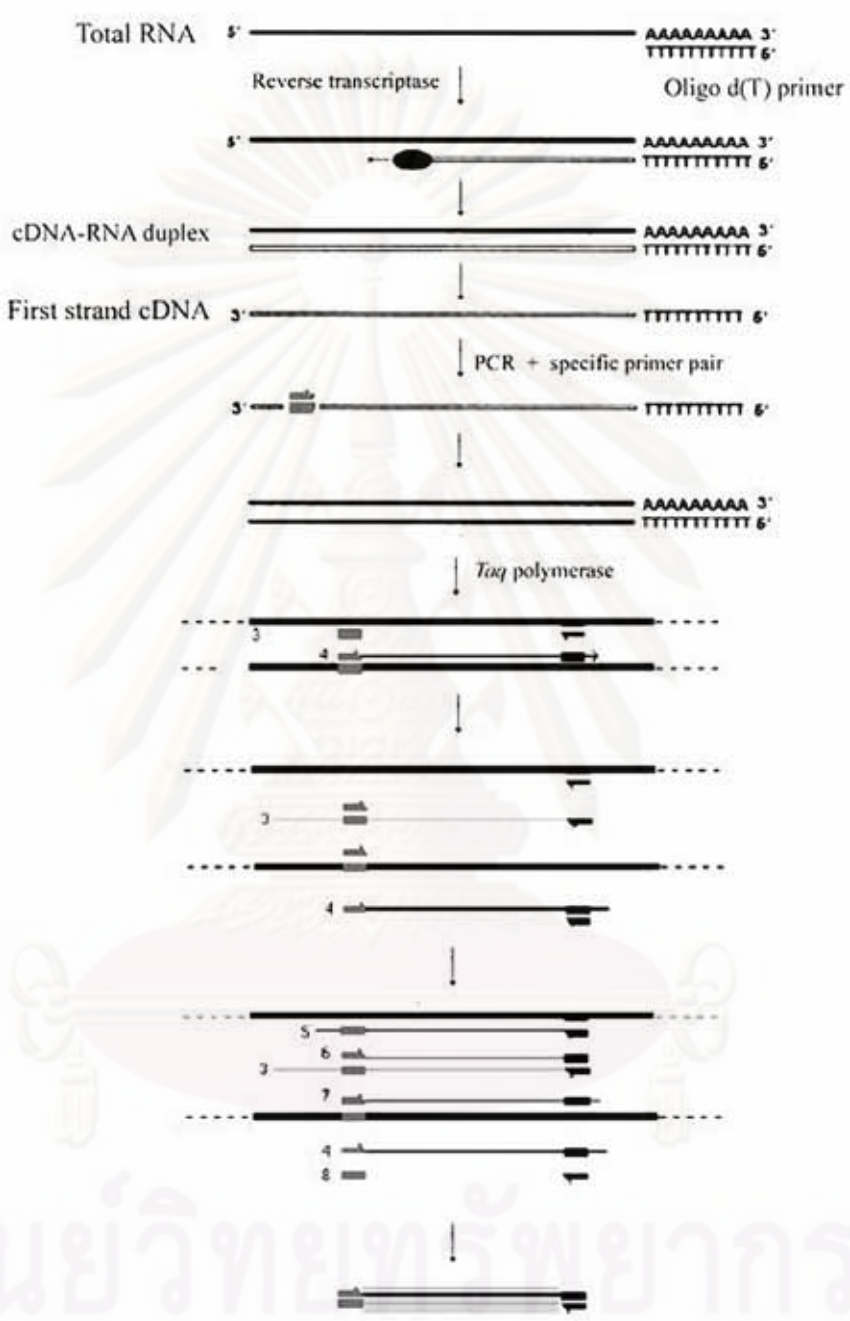


Figure 1.10 Overall concepts of the RT-PCR procedure. During first-strand cDNA synthesis an oligo d(T) primer anneals and extends from sites present within the total RNA. Second strand cDNA synthesis primed by the 18 - 25 base specific primer proceeds during a single round of DNA synthesis catalyzed by *Taq* polymerase**.**

1.6.7 Rapid Amplification of cDNA Ends-Polymerase Chain Reaction (RACE-PCR)

Rapid amplification of complementary DNA (cDNAs) ends (RACE) is a powerful technique for obtaining the ends of cDNAs when only partial sequences are available. Using SMART (**S**witching **M**echanism **A**t 5′ end of **R**NA **T**ranscript) technology, terminal transferase activity of Powerscript Reverse Transcriptase (RT) adds 3 - 5 nucleotides (predominantly dC) to the 3′ end of the first-strand cDNA. This activity is harnessed by the SMART oligonucleotides whose terminal stretch of dG can anneal to the dC-rich cDNA tail and serve as an extended template for reverse transcriptase. A complete cDNA copy of original mRNA is synthesized with the additional SMART sequence at the end.

Essentially, an adaptor with a defined sequence is attached to one end of the cDNA; then the region between the adaptor and the known sequences is amplified by polymerase chain reaction (PCR). Since the initial publication in 1988 (Frohman *et al*., 1988), RACE has greatly facilitated the cloning of new genes. Currently, RACE remains the most effective method of cloning cDNAs ends. It is especially useful in the studies of the temporal and spatial regulation of transcription initiation and differentiation splicing of mRNA. A linker at the 3′ end and an adaptor at the 5′ end are added to the first strand of cDNA during reverse transcription; amplification of virtually any transcript to either end can then make use of this same pool of cDNAs (Figure 1.11). In addition to being simple, the efficiency of 5′-RACE is dramatically increased because the adaptor is added only to the full-length cDNAs.

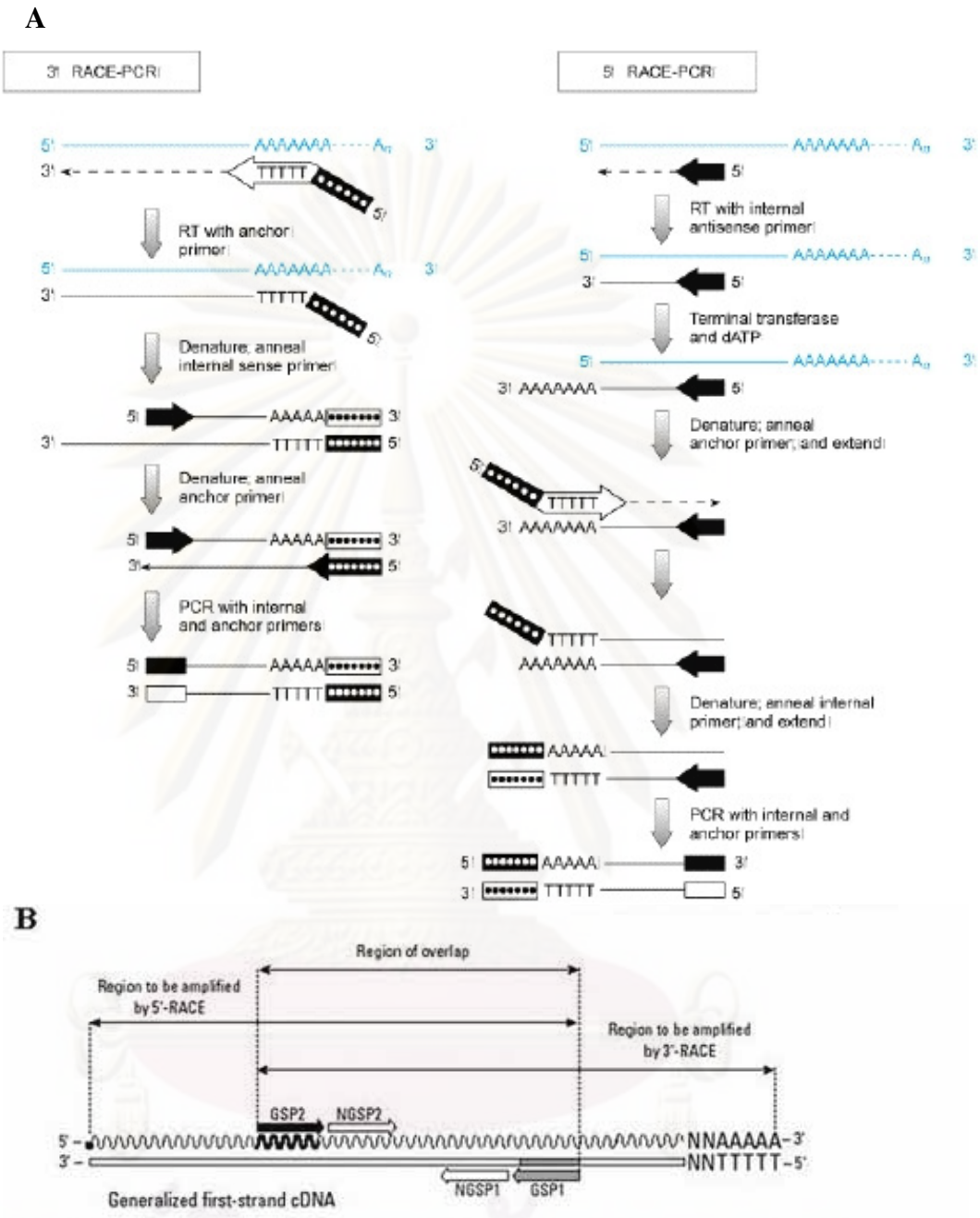


Figure 1.11 Overview of RACE-PCR. A; Mechanism of 5' and 3' RACE cDNA synthesis. First strand synthesis is primed using a modified oligo (dT) primer. After reverse transcriptase reaches the end of the mRNA template, it added several dC residues. The oligonucleotide adaptor anneals to the tail of the cDNA and serves as an extended template for RACE-PCR. B; Relationships of gene-specific primers to the cDNA template. This diagram shows a generalized first strand cDNA template.

1.6.8 Quantitative real-time PCR

Quantitative real time polymerase chain reaction (qPCR) or “kinetic polymerase chain reaction” is a laboratory technique based on the polymerase chain reaction, which is used to amplify and simultaneously quantify a target DNA molecule. It enables both detection and quantification (as absolute number of copies or relative amount when normalized to DNA input or additional normalizing genes) of a specific sequence in a DNA sample (Figure 1.12).

The procedure follows the general principle of polymerase chain reaction. Its key feature is that the amplified DNA is quantified as it accumulates in the reaction in a *real time* manner after each amplification cycle. Two common methods of quantification are: (1) the use of fluorescent dyes that intercalate with double-stranded DNA, and (2) modified DNA oligonucleotide probes that fluoresce when hybridized with a complementary DNA (VanGuilder *et al*., 2008). Typically, the reaction is prepared as usual, with the addition of fluorescent dsDNA dye. The reaction is run in a thermocycler and after each cycle, the levels of fluorescence are measured with a detector; the dye only fluoresces when bound to the dsDNA (i.e., the PCR product). With reference to a standard dilution, the dsDNA concentration in the PCR can be determined.

A DNA-binding dye binds to all double-stranded (ds) DNA in PCR, causing fluorescence of the dye. An increase in DNA product during PCR therefore leads to an increase in fluorescence intensity and is measured at each cycle, thus allowing DNA concentrations to be quantified. However, dsDNA dyes such as SYBR Green bind to all dsDNA PCR products, including nonspecific PCR products (such as "primer dimer"). This can potentially interfere with or prevent accurate quantification of the intended target sequence. Therefore, primers for quantitative real-time PCR should be carefully designed and tested for the high amplification efficiency (usually >95%) without any non-specific amplification product and primer dimer.

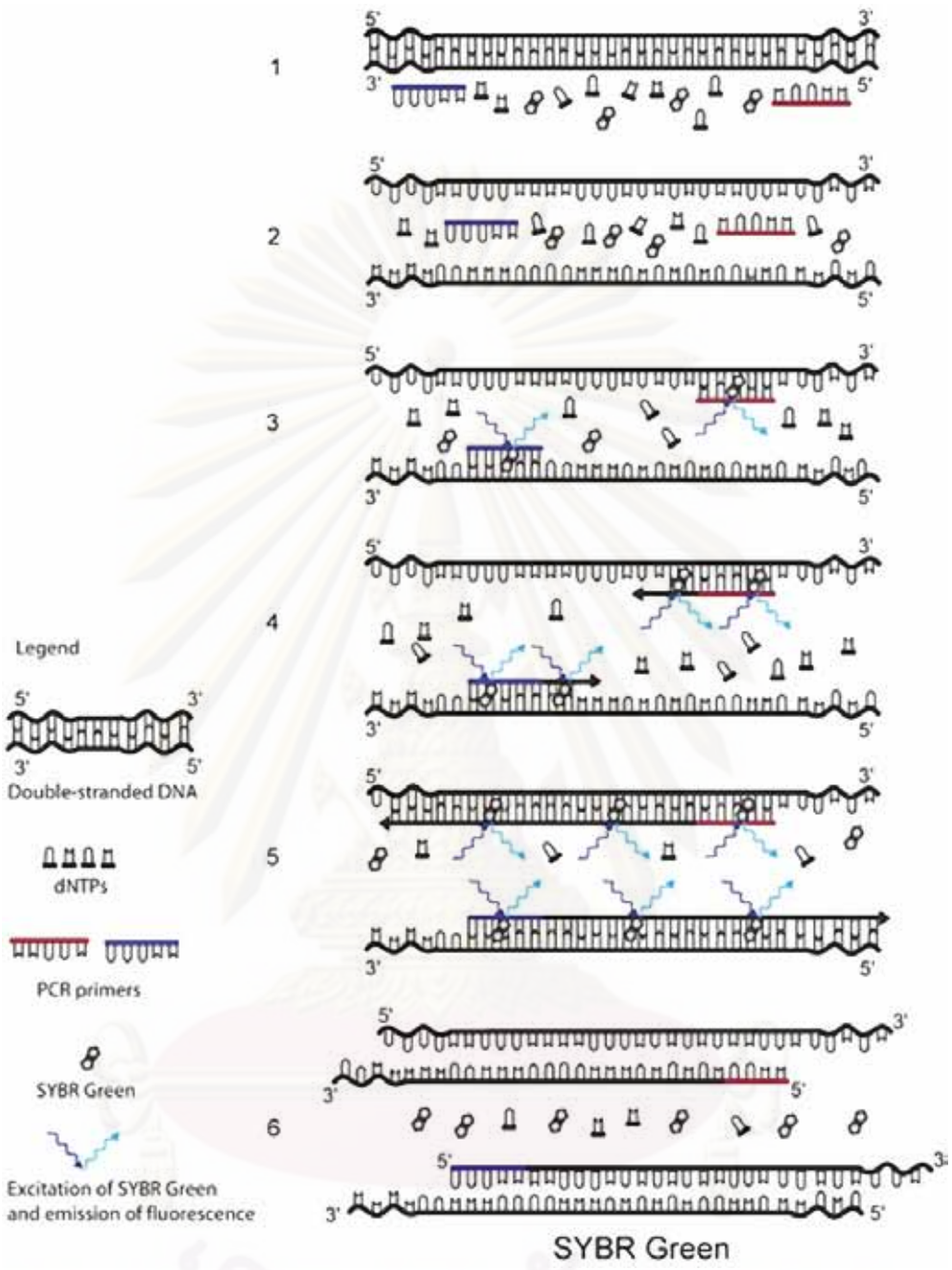


Figure 1.12 The principle of SYBR Green detection in quantitative real-time PCR.

1.7 SNP in functional important genes and its applications

Analysis of gene-based single nucleotide polymorphism (SNP) is one of the efficient approaches for discovery of genes which significantly contribute in production traits of commercially important species (Liu and Cordes, 2004). This opens the possibility to locate major loci responding for quantitative traits of *P. monodon* where this information is not available in this species at present (Tao and Boulding, 2003).

Several molecular techniques may be used to identify and subsequently to characterize SNP under investigation. DNA sequencing is the direct method for SNP identification. Nevertheless, several indirect methods are also commonly used. These include restriction fragment length polymorphism (RFLP), single strand conformational polymorphism (SSCP) and denaturing gradient gel electrophoresis (DGGE), primer extension, oligonucleotide ligation assay (OLA), *Taq*Man exonuclease detection, PCR-allele specific amplification (PASA) bidirectional PCR-allele specific amplification (Bi-PASA) and mass spectrometry (MS).

SNP has been widely used in several applications including evolution and population genetic studies, gene discovery, pharmacology and construction of genetic linkage maps and linkage disequilibrium mapping for identification of quantitative trait loci (QTL). Association between genotype and phenotype (disorders) were found in both non-coding and coding regions. The mutated SNP that influence promoter activity or DNA and pre-mRNA conformation, play a direct or indirect role in phenotypic expression. Studies about correlation between SNP and diseases have been extensively carried out in human (Immervoll *et al*., 2001; Tan *et al*., 2001). Nevertheless, research on association analysis of SNP and gene expression is rather limited and still at the initial stage in aquatic species.

An EST approach (single-pass sequencing of randomly selected clones from cDNA libraries) has been successfully applied and recognized as an effective method for discovery of immune-related genes in *L. vannamei* and *L. sertiferus* (Gross *et al*., 2001), *M. japonicus* (Rojtinnakorn *et al*., 2002), *F. chinenesis* (Shen *et al*., 2004) and *P. monodon* (Supungul *et al*., 2004). This approach has been applied for identification

and characterization of genes involved with reproduction (ovarian and testicular development) and growth of *P. monodon* (Preechaphol *et al*., 2007).

In the channel catfish, He *et al*., (2003) identified putative SNP markers within genes by comparative analysis of ESTs from the blue catfish (*Ictalurus furcatus*) and the channel catfish (*I. punctatus*) deposited in GenBank. Result suggested that SNPs are useful for mapping genes in interspecific resource families (channel catfish x blue catfish). In addition, SNPs that differed within species were found and can be applied to genome scans in channel catfish resource families.

Hayes *et al*., (2007) illustrated an extensive resource of putative SNP markers, derived from EST in the Atlantic salmon (*Salmo salar*). A highly valuable resource for creating a dense genetic map of the salmon genome for aligning genetic and physical maps, and for fine mapping of QTL affecting economic traits in Atlantic salmon aquaculture were found.

Xu *et al*., (2006) characterized two *parvalbumin* (*PVALB1* and *PVALB2*) genes and their association with growth traits in Asian sea bass (*Lates calcarifer*). Pavalbumins are extremely abundant in fish muscle and play an important role in muscle relaxation. The cDNA for *PVALB1* and *PVALB2* were 840 and 667 bp, in length, respectively. Both genes consisted of five exons and fore introns, encoded 109 amino acids. Using real-time polymerase chain reaction, expression of *PVALB1* was detected in all 10 tissues tested, with expression in brain, kidney, muscle and small intestine being 15 to 322-fold higher than in the other tissues. Expression of *PVALB2* was detected only in muscle, brain, and intestine, was up to 10-fold lower than *PVALB1* expression. A $(CT)_{17}$ microsatellite in the 3′UTR region of *PVALB1* and three single nucleotide polymorphism (SNPs) in the third intron of *PVALB2* were identified. The microsatellite in *PVALB1* was significantly associated with body weight and body length at 90 days post-hatch ($P < 0.01$) whereas the SNPs in *PVALB2* were not associated with these traits.

He *et al.,* (2008) studies association between SNP in cytochrome P450-c19a and reproductive traits of the Japanese flounder (*Paralichthys olivaceus*). A SNP in the exon 7 of *CYP19a* gene, SNP2, was significantly associated with 17β-estradiol (E_2) ($P < 0.05$) and gonadosomatic index (GSI) ($P < 0.05$). Individuals with genotype

AB of SNP2 had significantly higher serum E_2 levels ($P < 0.05$) and GSI ($P < 0.05$) than those of genotype AA or BB. In addition, there was significant association between one diplotype based on three SNPs and reproductive trait. The genetic effects for both serum E_2 of diplotype D9 and GSI of diplotype D1 were respectively much higher than those of other diplotypes ($P < 0.05$). Results seem to explain effects of *CYP19a* gene in reproductive endocrinology of the Japanese flounder.

SNP markers in candidate genes can be treated as similar as other biallelic co-dominantly segregated DNA markers. The advantage of coding SNP is that they are located in regions that code for functionally important proteins. Therefore, they are more likely to be near QTL that affect commercially important traits.

Correlations between genotypic and phenotypic variations in shrimp are still not understood. Therefore, effects of SNP on expression levels of functional important genes and phenotypes in *P. monodon* should be carried out. Analysis of gene-based SNP is possibly one of the efficient approaches for discovery of genes which are important in complex traits in *P. monodon*.

Tong *et al*., (2002) developed type I markers for genome mapping and other applications. Primers were designed from ESTs established from the cephalothorax of *P. monodon*. Thirty-four primer pairs successfully generated PCR products from genomic DNA of *P. monodon*. SSCP analysis indicated that approximately 30% of the EST tested were polymorphic in investigated shrimp and exhibited Mendelian segregation patterns. Some ESTs were also cross-species amplified in other shrimp (*P. chinensis*, *P. japonicus* and *P. vannamei*) allowing the possibility to be used for comparative mapping between related species. SSCP analysis of CU89 primer also revealed 2 distinct genotypes in investigated *P. japonicus* individuals from Australian (100% of genotype A, $N = 5$) and the South China Sea (80% and 20% of genotypes B and A, $N = 5$).

Kenway *et al.,* (2006) identifying genetic correlations between growth and survival rates in *P. monodon*. Shrimp were reared in captivity in tanks over three generations with full pedigree information. Weights of animals were measured at six ages between 7 and 54 weeks along with survival in each period. Females were more variable in weight than males after week 16, and variances between each sex were

standardized prior to estimation of heritability and genetic correlations. The phenotypic mean ± standard deviation of weights at week 40 was 35 ± 6 g for males and 44 ± 10 g for females. Heritability with standard errors at 16, 30, 40 and 54 weeks was 0.56 ± 0.04, 0.55 ± 0.07, 0.45 ± 0.11 and 0.53 ± 0.14 respectively. Heritability for family survival, determined from mean survival within each family, was 0.51 ± 0.18, 0.36 ± 0.18 and 0.71 ± 0.17 over periods 4 to 10, 10 to 16 and 16 to 32 weeks respectively. The genetic correlations between weight and survival revealed no significant trend. The results indicated significant concurrent improvements in both growth and survival are possible through selective breeding.

Glenn *et al.,* (2005) studied association analysis of SNP of *alpha-amylase* (*AMY2*) and *cathepsin-L* (*CTSL*) and the body weight in 2 populations of *Litopenaeus vannamei* (LV1 and LV2, $N = 75$ and 30 with the mean BW of 0.35 ± 0.06 and 2.52 ± 0.30 g, respectively) and a mapping population of *P. monodon* ($N = 41$) of investigated shrimp. SNP genotypes were carried out using PCR-RFLP of *AMY2* with *Sca I* and *CTSL* with *Pvu* II. Neither polymorphism of *AMY2* nor *CTSL* were found to be significantly associated with BW of LV1 and LV2 populations ($P > 0.05$).

There are a number of advantages to use SNP of EST as genetic markers. For example; if expressed EST markers are genetically associated with a trait of interest, it is possible that mapped genes directly affect that trait. In addition, EST are derived from coding genes, they are more transferable across pedigree and species boundaries than markers derived from non-coding regions.

PCR-SSCP is a simple and effective method for identification of SNP (Orita *et al*., 1989) before further analyzed by DNA sequencing. SNP by expressed sequence tags (SBE) were recently developed by amplification of 102 previously identified ESTs in ovaries and hemocytes of *P. monodon*. A total of 48 successfully amplified genes were further characterized by SSCP analysis and 44 of which were polymorphic. The full length of *ribophorin I* and *receptor for activated protein kinase C* (*RACK*) were successfully characterized by RACE-PCR. Semi-quantitative RT-PCR of *ribophorin I* and *RACK* were carried out. Significantly different expression levels of *ribophorin I* ($P < 0.05$) but not in *RACK* ($P > 0.05$) were observed in ovaries and testes of *P. monodon* broodstock. Although the preliminary study did not reveal significant association between SNP (SSCP patterns) and expression levels of

ribophorin I and *RACK* in the limited sample size of *P. monodon* broodstock, that approach demonstrates the possibility to further test for association between SNP of candidate genes allied with growth and reproduction (degrees of gene expression by the wild type/mutant alleles and growth rates and reproductive performance) of *P. monodon* (Buaklin, 2005).

As described previously, correlation between genotypic and phenotypic variations in shrimp is still not understood. In this thesis, association between SNP of growth- and reproduction-related genes and phenotypes (body weight for the former and gonadosomatic index, GSI, for the latter) of *P. monodon* were examined. In addition, effects of SNP and expression levels of mRNA of functional important genes on growth and reproductive maturation in *P. monodon* were also studied**.**

CHAPTER II

MATERIALS AND METHODS

2.1 Experimental animals

Broodstock-sized *P. monodon* (14-month-old, $N = 66$) were collected from the Broodstock Multiplication Center, Burapha University, Chanthaburi. For SNP analysis of reproduction-related gene, pleopods of each shrimp are collected and transported back to the laboratory. For gene expression analysis, ovaries were dissected out from each shrimp and immediately placed in liquid N_2. Specimens are kept at -80 °C until used.

Moreover, 342 individuals of domesticated *P. monodon* cultured for 3-month in the same earth pond were also collected. The wet weight of each shrimp is measured ($\overline{X}$ = 13.32 ± 2.42 g). Initially, approximately 16% of shrimp exhibiting from the top ($N = 54$, $\overline{X}$ = 16.78 ± 2.23 g) and the bottom ($N = 56$, $\overline{X}$ = 10.43 ± 1.08 g) according to the body weight were used for polymorphic analysis of various-growth-related genes including *dolichyl diphosphooligosaccharyl protein glycotransferase* (*DDPG*) *thyroid hormone binding protein* (*THBP*), *epidermal growth factor* (*EGF*) and *insulin degrading enzyme* (*IDE*). Subsequently, polymorphism of *THBP* and *EGF* was examined in the entire sample. For SNP analysis, pleopods of each shrimp are collected and transported back to the laboratory. For gene expression analysis, hepatopancreas was dissected out from each shrimp and immediately placed in liquid N_2. Specimens were kept at -80 °C until used.

2.2 Genomic DNA extraction

Genomic DNA was extracted from a piece of pleopod of each *P. monodon* individual using a phenol-chloroform-protenase K method (Klinbunga *et al*., 1999). A piece of pleopod tissue was dissected out from a frozen pleopod and placed in prechilled microcentrifuge tube containing 500 µl of the TEN buffer (200 mM Tris-HCl, 100 mM EDTA and 250 mM NaCl, pH 8.0) and briefly homogenized with a

micropestle. SDS (10%) and RNase A (10 mg/ml) solutions were added to a final concentration of 1.0% (w/v) and 100 µg/ml, respectively. The resulting mixture was then incubated at 37 °C for 1 hour. At the end of the incubation, a protenase K solution (10 mg/ml) was added to the final concentration 100 µg/ml and the sample was further incubated at 55 °C for 3 - 4 hours. After the incubation, an equal volume of buffer-equilibrated phenol was added and gently mixed for 15 minutes. The solution was centrifuged at 12,000 rpm for 10 minutes at room temperature. The upper aqueous phase was transferred to a newly sterile microcentrifuge tube. This extraction process was then repeated once with phenol:chloroform:isoamylalcohol (25:24:1) and twice with chloroform:isoamylalcohol (24:1). The final aqueous phase was transferred into a sterile microcentrifuge tube and mixed with the equal volume of TE buffer (10 mM Tris-HCl, pH 8.0 and 0.1 mM EDTA) and subsequently with one-tenth volume of 3 M sodium acetate, pH 5.2 was added. After that, DNA was precipitated by an addition of two volume of cold absolute ethanol and mixed thoroughly. The mixture was incubated at -80 °C for 30 minutes. The precipitated DNA was recovered by centrifugation at 12,000 rpm for 10 minutes at 4 °C. The DNA pellet was then washed twice with 1 ml of cold 70% ethanol. After centrifugation, the supernatant was removed. The DNA pellet was air-dried and resuspended in 50 µl of TE buffer. The DNA solution was incubated at 37 °C for 1 - 2 hours and stored at 4 °C until further needed.

2.3 Measurement of DNA concentration

2.3.1 Estimation of DNA concentration by spectrophotometry

The concentration of extracted DNA was estimated by measuring the optical density at 260 nanometre (OD_{260}). An OD_{260} of 1.0 corresponds to a concentration of 50 µg/ml double stranded DNA (Sambrook and Russell, 2001). Therefore the concentrations of DNA are estimated in µg/ml by the following equation;

[Nucleic acid] = OD_{260} x dilution factors x nucleotide factor

* nucleotide factor = 50 for DNA and 40 for RNA and 33 for oligonucleotides

The value at OD_{260} allows calculation of total nucleic acids whereas the value at OD_{280} determines the amount of proteins in the DNA and RNA solution. The ratio between OD_{260}/OD_{280} provides an estimate on the purity of extracted DNA and RNA. Pure preparation of DNA and RNA has OD_{260}/OD_{280} ratio of 1.8 and 2.0, respectively. The ratios that much lower than those values indicate contamination of residual proteins or phenol in the extracted DNA and RNA (Sambrook and Russell, 2001).

2.3.2 Estimation of the amount of DNA by gel electrophoresis

DNA concentration can also be estimated on the basis of its direct relationships between the amount of DNA and the level of fluorescence after ethidium bromide staining. DNA was electrophoresed through 0.8 - 1.0% agarose gel prepared in 1X TBE buffer at 100 volts. After electrophoresis, the gel was stained with ethidium bromide. DNA concentration was estimated from the intensity of the fluorescent band by comparing with that of undigested λDNA.

2.4 Agarose gel electrophoresis (Sambrook and Russell, 2001)

Appropriate amount of agarose was weighed out and mixed with 1X TBE buffer (89 mM Tris-HCl, 8.9 mM boric acid and 2.5 mM EDTA, pH 8.0). The gel slurry was heated until complete solubilization in the microwave. The gel solution was left at room temperature to approximately 55 $^{\circ}$C before poured into a gel mould. The comb inserted, the gel was allowed to solidify at room temperature for approximately 45 minutes. When needed, the gel mould was placed in the gel chamber and sufficient 1X TBE buffer was added to cover the gel for approximately 0.5 cm. The comb was carefully withdrawn. One-fourth volume of the gel-loading dye (0.25% bromophenol blue and 25% ficoll, MW 400,000 prepared in sterile deionized H_2O) was added to each sample, mixed and loaded into the well. A 100 bp DNA ladder and/or λ*Hin*d III was used as the standard DNA markers. Electrophoresis was carries out at 100 volts until the tracking dye migrated about three-quartered of the gel. After electrophoresis, the gel was stained with ethidium bromide (0.5 μg/ml) for 5 minutes and destained to remove unbound ethidium bromide by submerged in H_2O for 15 minutes. The DNA fragments were visualized using a UV transilluminator.

2.5 Genome walking analysis

For identification nucleotide sequence of the 5′ untranslated region (5′UTR) of functionally important genes: *NADP-dependent leukotriene B4 12-hydroxydehydrogenase* (*LTB4DH*) and *nuclear autoantigenic sperm protein* (*NASP*) were further characterized by genome walking analysis.

2.5.1 Restriction digestion of genomic DNA

Two and a half micrograms of genomic DNA extracted from a single individual of *P. monodon* were singly digested with 50 units of a blunt-end generating restriction enzyme (*Eco* RV, *Stu* I, *Pvu* II and *Sca* I) in the reaction volume of 100 µl. The reaction was incubated at 37 °C for 4 hours. Five microliters of the digestion product was then electrophoresed on 0.8% agarose gels to determine whether the digestion was complete.

2.5.2 Purification of the restriction digestion DNA

The digested DNA was purified by a phenol-chloroform-proteinase K method (Klinbunga *et al*., 1999). An equal sample volume of buffer-equilibrated phenol was added to the sample. The mixture was vortexed for 5-10 seconds and centrifuged for 5 minutes at room temperature to separate the aqueous and organic phases. The upper layer was collected and mixed with a sample volume of chloroform:isoamylalcohol (24:1). The sample was then centrifuged as in the previous step. The upper layer was collected and mixed with one-tenth sample volume of 3 M sodium acetate, pH 4.5 sequentially with 2.5 folds sample volume of cold absolute ethanol. The mixture was incubated at -80 °C for 30 minutes before DNA pellet was collected by centrifugation at 12,000 rpm for 10 minutes at room temperature. After a brief wash with cold 70% ethanol, the DNA pellet was air-dried and resuspended in 10 µl of TE buffer. The purified products were confirmed by electrophoresis on 0.8 % agarose gels.

2.5.3 Ligation of digested genomic DNA to Genome Walker Adaptors

The ligation reaction was performed in a 10 µl reaction volume containing 2.2 µl of digested purified DNA, 1.8 µl of 50 µM GenomeWalker Adaptor (Clonetech) (GenomeWalker Adaptor : 5′-GTAATACGACTCACTATAGGGCACGCGTGG

TCGACGGCCCGGGCTGGT-3′), 5 µl of 2X ligation buffer and 3 units of T4 DNA ligase. The reaction mixture was incubated at 16 °C overnight and then terminated by incubation at 70 °C for 5 minutes. The product was 10 fold diluted in TE buffer before used as the template for PCR step, is also called a mini-library.

2.5.4 PCR-based genomic DNA walking

The primary PCR of *LTB4DH* and *NASP* of *P. monodon* were carried out in a total volume of 50 µl reactions containing 1 µl of each mini-library, 1X buffer, 200 µM each of dATP, dCTP, dGTP and dTTP, 2 mM $MgCl_2$, 0.2 µM each of Adaptor primer 1 (AP1 : 5′- GTAATACGACTCACTATAGGGC -3′), 0.2 µM of GSP1 (Table 2.1) and 1 unit of Advantage® 2 DNA Polymerase (Clonetech). The amplification reaction was carried out using a amplification condition indicated in Table 2.2. Five microliters of primary PCR product was electrophoretically analyzed by a 1.2% agarose gel. A 100 bp ladder and λ*Hin*d III was included as the DNA marker.

The primary PCR product of *NASP* was 10-fold diluted in H_2O and used as the template for secondary PCR. The secondary PCR was carried out in a total volume of 50 µl reaction containing 1 µl of diluted primary PCR product, 1X buffer, 200 µM each of dATP, dCTP, dGTP and dTTP, 2 mM $MgCl_2$, 0.2 µM of Adaptor primer 2 (AP2 : 5′-ACTATAGGGCACGCGTGG T-3′), 0.2 µM of GSP2 (Table 2.1) and 1 unit of Advantage® 2 DNA Polymerase. The amplified genome walked fragment of each gene was cloned and sequenced.

Table 2.1 Gene specific primer (GSPs) used for genome walking analysis of *LTB4DH* and *NASP* of *P. monodon*

Gene specific primer	Sequence	Tm (˚C)
LTB4DH		
5′ GW1	5′- ACCTCCATCTTCACAAGCCAG -3′	64
NASP		
5′ GW1	5′- CACATCTCTTTCTTGGACGGCG -3′	68
5′ GW2 (nested)	5′- GGCTGGCACTGGTCTCGGCA -3′	68

Table 2.2 The amplification conditions for genome walking analysis of *LTB4DH* and *NASP* of *P. monodon*

Gene	Amplification condition
LTB4DH	
5′ GW1	7 cycles of 94 °C for 25 seconds and 68 °C for 3 minutes
	32 cycles of 94 °C for 25 seconds, 63 °C for 3 minutes
	And the final extension at 63 °C for 7 minutes
NASP	
5′ GW1	7 cycles of 94 °C for 25 seconds and 68 °C for 3 minutes
	32 cycles of 94 °C for 25 seconds, 63 °C for 3 minutes
	And the final extension at 63 °C for 7 minutes
5′ GW2 (nested)	7 cycles of 94 °C for 25 seconds and 68 °C for 3 minutes
	32 cycles of 94 °C for 25 seconds, 63 °C for 3 minutes
	And the final extension at 63 °C for 7 minutes

2.6 Cloning of the PCR products

2.6.1 Elution of DNA from agarose gel

After electrophoresis, desired individual DNA bands were excised from agarose gels (200 - 300 mg) using a sterile scalpel. DNA was extracted from the gel pieces using an illustra^{TM} PCR band purification kit (GE Healthcare). Five hundred microliters of the Capture buffer Type 3 was added to the sample and mixed by vortexing. The mixture was incubated at 55 $^{\circ}$C for 10 - 15 minutes until the gel slice was completely dissolved. During the incubation period, the tube was inverted every 2-3 minutes. An illustra^{TM} MicroSpin column was placed in a collection tube and 800 µl of the sample mixture was applied into the illustra^{TM} MicroSpin column and centrifuged at 8,000 rpm for 30 seconds. The flow-through was discarded. The illustra^{TM} MicroSpin column was placed back in the collection tube. The column was washed by the addition of 500 µl of the ethanol-added Wash buffer type 1 and centrifuged at 8,000 rpm for 30 seconds. After discarding the flow-through, the illustra^{TM} MicroSpin column was centrifuged for 2 minutes at the full speed (14,000 rpm) to dry the column matrix. The dried column was placed in a new microcentrifuge tube and 15 µl of the Elution buffer type 4 or water was added to the center of the column matrix. The column was left at room temperature for 2 minutes before centrifuged for 2 minutes at the full speed to recover the gel-eluted DNA.

2.6.2 Ligation of PCR products to the pGEM®-T Easy Vector

DNA fragment was ligated to the pGEM-T easy vector (Promega) in a 10 µl reactions volume containing 50 ng of DNA insert, 5 µl of 2X Rapid Ligation Buffer (60 mM Tris-HCl, pH 7.8, 20 mM $MgCl_2$, 20 mM DDT, 2 mM ATP and 10% PEG8000), 25 ng of pGEM-T easy vector and 3 units of T4 DNA ligase. The reaction mixture was incubated at 4 $^{\circ}$C overnight before transformed into *E. coli* JM 109.

2.6.3 Transformation of ligation products into *E. coli* JM 109

2.6.3.1 Preparation of competent cell

A single colony of *E. coli* JM 109 was inoculated in 10 ml of LB broth (1% Bacto tryptone, 0.5% Bacto yeast extract and 0.5% NaCl, pH 7.0) with vigorous

shaking at 37 °C overnight. The starting culture was then inoculated into 50 ml of LB broth and continued culture at 37 °C with vigorous shaking to OD_{600} of 0.5 to 0.8. The cells were briefly chilled on ice for 10 minutes and recovered by centrifugation at 2700 g for 10 minutes at 4 °C. The pellets were resuspended in 30 ml of ice-cold $MgCl_2/CaCl_2$ solution (80 mM $MgCl_2$ and 20 mM $CaCl_2$) and centrifuged as above. The cell pellet was resuspended in 2 ml of ice-cold 0.1 M $CaCl_2$ and the suspension was divided into 200 µl aliquots. These competent cells was either used immediately or stored at -80 °C for subsequently used.

2.6.3.2 Transformation

The competent cells were thawed on ice for 5 minutes. Two to four microliters of the ligation mixture were added and gently mixed by pipetting. The mixture was left on ice for 30 minutes. During the incubation period, the ice box was gently moved forward and backward a few times every 5 minutes. The transformation reaction was heatshocked in a 42 °C water bath (without shaking) for exactly 45 seconds. The reaction tube was immediately placed on ice for 2 - 3 minutes. The mixture were removed from the tubes and added to a new tube containing 1 ml of pre-warmed SOC (2% Bacto tryptone, 0.5% Bacto yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$ and 20 mM glucose). The cell suspension was incubated with shaking at 37 °C for 90 minutes. The mixture were centrifuged for 20 seconds at room temperature, and resuspended in 100 µl of the SOC medium and spread onto a selective LB agar plates (containing 50 µg/ml of amplicillin and spread with 20 µl of 25 µg/ml of X-gal and 25 µl of 25 µg/ml of IPTG for approximately 1 hour before used) and further incubated at 37 °C overnight (Sambrook and Russell, 2001). The recombinant clones containing inserted DNA are white whereas those without inserted DNA are blue.

2.6.4 Colony PCR and digestion of the amplified inserts

Colony PCR was performed in a 25 µl reaction mixture containing 10 mM Tris-HCl, pH 8.8, 50 mM $(NH_4)_2SO_4$, 0.1% Triton X-100, 100 mM of each dNTP, 2 mM $MgCl_2$, 0.1 µM each of pUC1 (5′-CCGGCTCGTATGTTGTGTGGA-3′) and pUC2 (5′-GTG GTG CAA GGC GAT TAA GTT GG-3′), 0.5 unit of *Taq* DNA polymerase (Fermentas). A colony was picked by a pipette tip, placed in the culture

tube and served as the template in the reaction. PCR was carried out in a thermocycler consisting of predenaturation at 94 °C for 3 minutes followed by 35 cycles of denaturation at 94 °C for 30 seconds, annealing at 50 °C for 1 minutes and extension at 72 °C for 90 seconds. The final extension was carried out at the same temperature for 7 minutes. The colony PCR products were electrophoresed through a 1.2% agarose gel and visualized after ethidium bromide staining.

The colony PCR products containing the insert were separately digested with *Alu* I and *Rsa* I (Promega) in a 15 µl reaction volume containing 1X buffer (6 mM Tris-HCl, 6 mM $MgCl_2$, 50 mM NaCl and 1 mM DDT, pH 7.5 for *Alu* I and 10 mM Tris-HCl, 10 mM $MgCl_2$, 50 mM NaCl and 1 mM DDT, pH 7.9 for *Rsa* I), 0.1 mg/ml BSA, 2 units of each enzyme and 4 µl of the colony PCR product. The reaction mixture was incubated at 37 °C overnight. The reaction was analyzed by 1.2% agarose gel electrophoresis.

2.6.5 Extraction of recombinant plasmid DNA

Plasmid DNA was isolated using an illustra™ plasmidPrep mini spin kit (GE Healthcare). A discrete white colony was inoculated into a sterile culture tube containing 3 ml of LB broth supplemented with 50 µg/ml of amplicillin and incubated with shaking (200 rpm) at 37 °C overnight. The culture was transferred into a sterile 1.5 ml microcentrifuge tube and centrifuged at 14,000 rpm for 1 minute. The supernatant was discarded. The bacterial pellet was resuspended in 200 µl of the Lysis buffer type 7 containing RNaseA and thoroughly mixed by vortexed. The resuspended cells were lysed by the addition of 200 µl of the Lysis buffer type 8 and mixed gently by inverting the tube for 10 times. The mixture was stood for 2 minutes at room temperature. After that, 400 µl of the Lysis buffer type 9 was added to neutralize the alkaline lysis step and mixed immediately by inverting the tube for 10 times. The mixture was then centrifuged at 14,000 rpm for 15 minutes. The illustra™ plasmid mini column was placed in a collection tube and the clear lysate was applied into the illustra™ plasmid mini column and centrifuged at 8,000 rpm for 30 seconds. The flow-through was discarded. The illustra™ plasmid mini column was placed back in the collection tube. The column was washed by the addition of 400 µl of the Wash buffer type 1 and centrifuged at 8,000 rpm for 30 seconds. After discarding the flow-through, 600 µl of the ethanol-added Wash buffer was added and centrifuged as

above. The illustra^{TM} plasmid mini column was further centrifuged for 2 minutes at the full speed (14,000 rpm) to dry the column matrix. The dried column was placed in a new microcentrifuge tube and 50 µl of the Elution buffer type 4 or water was added at the center of the column matrix. The column was left at room temperature for 2 minutes before centrifuged for 2 minutes at the full speed to recover the purified plasmid DNA. The concentration of extracted plasmid DNA was spectrophotometrically measured.

2.6.6 DNA sequencing

Nucleotide sequences of recombinant plasmids were examined by automated DNA sequencer using M13 forward and/or M13 reverse primer as the sequencing primer by MACROGEN (Korea). Nucleotide sequences were blasted against data in the GenBank (http://www.ncbi.nlm.nih.gov/blast) using Blast*N* (nucleotide similarity against the nr/nt database) and Blast*X* (translated protein similarity against the nr database). Nucleotide sequences of fragments representing different SSCP genotypes of each gene were aligned by Clustal X (Thompson *et al*., 1994).

2.7 Isolation of exon/intron of interesting important genes

The exon/intron structure of the target gene was characterized. Several primer pairs were designed from cDNA sequence of each gene and used for the PCR amplification against genomic DNA as the template (Tables 2.3 and 2.4). The PCR fragment was cloned and sequenced as described above.

2.8 Microsatellite analysis of *P. monodon* juveniles and broodstock used in this study

2.8.1 Specimens and DNA extraction

Forty individuals of 3-month-old juveniles and twenty individuals of 14-month-old broodstock of *P. monodon* from Broodstock Multiplication Center, Burapha University, Chanthaburi. Pleopods were genotyped by microsatellites Genomic DNA was extracted from frozen pleopods of *P. monodon* using a phenol-chloroform-proteinase K method (Klinbunga *et al*., 1999).

Table 2.3 Genes, primer sequence, length and the melting temperature of primers used for exon/intron detection

Gene/ Primer	Sequence	Length (bp)	Tm (°C)
NADP-dependent leukotriene B4 12-hydroxydehydrogenase (*LTB4DH*)			
LTB4DH-2	F : 5′- TGGGAGGCACTCAGTCAAGC -3′	20	64
	R : 5′- TTCCATCCAGCGATCAAACC -3′	20	60
LTB4DH-3	F : 5′- GAGATGGGAGACTTGCCTAA -3′	19	56
	R : 5′- TTTGGTCCACGCTACCTGTT -3′	20	60
DEAD-box protein 52			
Deadbox-3	F : 5′- GAGCGAGCGAAAGAACTGT -3′	19	58
	R : 5′- GGCTTATCTGCCATTGTCC -3′	19	58
Thyroid hormone binding protein (*THBP*)			
RT-THBP	F : 5′- TACAAGCCCGATACTTATGAC -3′	21	60
	R : 5′- AAGGTTGGGAAGGACTGAAT -3′	20	58
Insulin degrading enzyme (*IDE*)			
RT-IDE	F : 5′- GGAGTGGGTGTTTGAAGAGTGT -3′	22	66
	R : 5′- TCTAACTCATCTTCAGCCTACCA -3′	23	66
Epidermal growth factor (*EGF*)			
RT-EGF	F : 5′- GTCACTGCTGAAAATAGCCATCT -3′	23	66
	R : 5′- CGGTTCTGAAACACTGAAATACG -3′	23	66

Table 2.4 The amplification conditions for the exon/intron region of the amplified gene segment

Primer	Amplification condition
LTB4DH-2 LTB4DH-3 Deadbox-3	1 cycle of 94 °C for 3 minutes, 35 cycles of 94 °C for 30 seconds, 55 °C for 45 seconds, 72 °C for 60 seconds and the final extension at 72 °C for 7 minutes
RT-THBP	1 cycle of 94 °C for 3 minutes, 35 cycles of 94 °C for 30 seconds, 55 °C for 45 seconds, 72 °C for 45 seconds and the final extension at 72 °C for 7 minutes
RT-IDE RT-EGF	1 cycle of 94 °C for 3 minutes, 35 cycles of 94 °C for 30 seconds, 53 °C for 45 seconds, 72 °C for 45 seconds and the final extension at 72 °C for 7 minutes

2.8.2 Microsatellite amplification and analysis

Three microsatellite loci of *P. monodon* (*CUPmo02*, *CUPmo13* and *CUPmo15*) were used. PCR conditions were optimized for each microsatellite primer pair as described in Table 2.6. The amplification reaction was carried out in 25 µl reaction mixture containing 10 mM Tris-HCl, pH 8.8, 50 mM KCl, 0.1% Triton X-100, 200 µM each of dATP, dCTP, dGTP and dTTP, 2 mM $MgCl_2$, 0.2 µM of each primer, 25 ng of genomic DNA of *P. monodon* and 1 unit of DyNAzyme™ II DNA Polymerase (Finnzymes). Five microliters of primary PCR product was electrophoretically analyzed by a 1.2% agarose gel and a 100 bp ladder was used as the DNA marker.

2.8.3 Denaturing polyacrylamide gel electrophoresis

2.8.3.1 Preparation of glass plate

The long glass plate was thoroughly wiped with 1 ml of 95% commercial grade ethanol with a piece of the tissue paper in one direction. This process was then repeated twice. Afterwards, the long glass plate was coated with 1 ml of freshly prepared bind silane (4 µl of bind silane, Amersham Biosciences, 991 µl of 95%

ethanol and 5 µl of glacial acetic acid) and left for approximately 10 - 15 minutes. Excess binding solution was removed with a piece of the tissue paper. The long glass plate was further cleaned with 95% ethanol for 3 times.

The short glass plate was treated as described above with the exception that the binding solution was replaced by the Repel silane (2% dimethyldichlorosilane in octamethylcyclotetrasitoxone). The cleaned glass plates were assembled with a pair of 0.4 mm spacers.

2.8.3.2 Preparation of denaturing polyacrylamide gel electrophoresis

The PCR product were electrophoretically separated on 5% of denaturing polyacrylamide gels were prepared by combining 40 ml of the acrylamide solution (19:1 acrylamide:bisacrylamide with 7 M urea in 1X TBE buffer) with 300 µl of freshly prepared 10% ammonium persulphate and 30 µl of TEMED. The acrylamide solution was gently swirled and degassed for 15 minutes. The bottom of assembled glass plates were sealed with the plastic tape. The bottom and both sides of the assembled glass plate were securely clamped. The assembled plate sandwich was hold at a 45 degree angle on the bottom corner. The acrylamide solution was then gently injected into one side of the assembled plates. The filled plate sandwich was left in the horizontal position. The flat edge of the shark-tooth comb was then inserted. The gel was left at room temperature for 1 hour. After that, the polymerized gel was covered by the water-soaked tissue papers and left at room temperature for 4 hours (or overnight) for complete polymerization. When required, the spring clips and the sealing tape were carefully removed. The top of the gel was rinsed with deionized H_2O.

Table 2.5 PCR primers and optimized annealing temperatures for *P. monodon* microsatellite loci

Locus	Sequence	Tm (°C)
CUPmo02	F : 5′-CCAAGATGTCCCCAAGGC-3′	58
	R : 5′-CTGCAATAGGAAAGATCAGAC-3′	60
CUPmo13	F : 5′-GACTTCGTGGTATCAATGACTGT-3′	66
	R : 5′-TGTCAGTTCATGTAGTCTGCTC-3′	64
CUPmo15	F : 5′-CATCACTTGGTCTACAGCA-3′	19
	R : 5′-TAAGGTTACTAATGGGCACT-3′	20

Table 2.6 The amplification conditions for *P. monodon* microsatellite loci

Gene	Amplification condition
CUPmo02	1 cycle of 94 °C for 3 minutes, 35 cycles of 94 °C for 45 seconds, 53 °C for 1 minute, 72 °C for 45 seconds and the final extension at 72 °C for 7 minutes
CUPmo15	1 cycle of 94 °C for 3 minutes, 30 cycles of 94 °C for 45 seconds, 56 °C for 45 seconds, 72 °C for 60 seconds and the final extension at 72 °C for 7 minutes
CUPmo13	1 cycle of 94 °C for 5 minutes, 35 cycles of 94 °C for 30 seconds, 58 °C for 30 seconds, 72 °C for 30 seconds and the final extension at 72 °C for 7 minutes

2.8.3.3 Electrophoresis

The gel sandwich was placed in the vertical sequencing apparatus with the short glass plate inward. The gel sandwich was securely clamped with the integral gel clamps along the sides of the sequencing apparatus. The upper and lower buffer chambers were filled with approximately 300 ml of 1X TBE. The shark-tooth comb was reinserted into the gel until the teeth just touched the surface of the gel. Three

microliters of the polyacrylamide gel loading dye (98% formamide, 200 μl EDTA, 0.25% bromophenol blue and 0.25% xylene cyanol) was loaded into each well. The gel was prerun at 35 W for 20 minutes.

Six microliters of the amplification products were mixed with 3 μl of the loading buffer and heated at 95 °C for 5 minutes before snapped cooled on ice for 3 minutes. The sample was carefully loaded into the well. Electrophoresis was carried out at 35 W for approximately 2 hours (XC moved out from the gel for approximately 30 minutes). Microsatellite band were visualized by silver staining.

2.8.3.4 Silver staining

The gel plates were carefully separated using a plastic wedge. The long glass plate with the gel was placed in a plastic tray containing 2 liters of the fix/stop solution (10% glacial acetic acid) and agitate well for 30 minutes. The gel was briefly soaked with shaking 3 times for 3 minutes with deionized water. The gel was lifted out from the tray between each wash and allowed the washed water draining out of the gel for 5 seconds. The gel was transferred to 0.1% silver nitrate solution (1.5 liters) and incubated with agitation at room temperature for 30 minutes. The gel was soaked in 1.5 liters of deionized water with shaking (10 forward and 10 backward agitation) and immediately placed in the tray containing 1.5 liters of the chilled developing solution. This step is crucial and the time taken to soak the gel in the water and transfer it to chilled developing solution (Sodium carbonate) should be no longer than 5 - 10 seconds. The gel was well agitated until the first bands are visible (usually 1.5 - 2 minutes). The gel was then transferred to another tray containing 1.5 liters of chilled developer and shaken until bands from every lane were observed (usually 2 - 3 minutes). One liter of the fix/stop solution was directly added to the developing solution and continued shaking for 3 minutes. The stained gel was soaked in deionized water twice for 3 minutes each. The gel was placed in the plastic bag and air-dried. The sized of alleles were estimated by reference to 100 and 50 bp marker ladder.

2.9 Identification of single nucleotide polymorphisms (SNP) in 5′UTR and the exon/intron region of interesting important genes using SSCP analysis

SNP in the 5′UTR of reproductively related genes: *NADP-dependent leukotriene B4 12-hydroxydehydrogenase (LTB4DH)* and *nuclear autoantigenic sperm protein (NASP)* and those in the exon/intron region of growth related genes: *thyroid hormone binding protein (THBP), insulin degrading enzyme (IDE), epidermal growth factor (EGF)* and *dolichyl diphosphooligosaccharyl protein glycotransferase (DDPG)* were analyzed by SSCP.

2.9.1 Primer design

Primers were designed from the 5′UTR of *LTB4DH* and *NASP* and the cDNA sequence of *THBP, IDE, EGF* and *DDPG* (Table 2.7) using Primer Premier 5.0 software. The criteria for primer designing were the primer length of 18 - 25 bases, the melting temperature of 55 - 70 °C, and random base distribution of the primers to avoid polypurine and polypyrimidine tracts, less than 5 °C differences of the melting temperature of a primer pair.

2.9.2 PCR

Generally, PCR was carries out in a 25 µl reaction mixture containing 1X Buffer (10 mM Tris-HCl, pH 8.8, 50 mM KCl, 0.1% Triton X-100), 200 µM each of dATP, dCTP, dGTP and dTTP, 2 mM $MgCl_2$, 0.2 µM of each primer, 25 ng of genomic DNA of *P. monodon* and 1 unit of DyNAzyme™ II DNA Polymerase (Finnymes). The amplification condition were initially carries out by predenaturation at 94 °C for 3 minutes followed by 35 cycles of denaturation at 94 °C for 30 seconds, annealing 53 °C for 45 seconds and extension for 45 seconds. The final extension was carries out at 72 °C for 7 minutes. Subsequently, more appropriate amplification conditions were further adjusted as, for example, described in Table 2.8.

Table 2.7 Genes, primer sequence, length and the melting temperature of primers designed from growth and reproductively related genes of *P. monodon*

Gene	Sequence	Length (bp)	Tm (˚C)
NADP-dependent leukotriene B4 12-hydroxydehydrogenase (*LTB4DH*)			
	F : 5′- TTCAGCCTCTACTATGTTT -3′	19	52
	R : 5′- CTGGGATGCTAGGGACTA -3′	18	56
Nuclear autoantigenic sperm protein (*NASP*)			
	F : 5′- TCACAGATACGGTCAACAGCG -3′	21	64
	R : 5′- AACCACTACGGCAAACTCACA -3′	21	62
	R2 : 5′- TCGGAAGCAACAAACGAGGGT -3′	21	64
Thyroid hormone binding protein (*THBP*)			
	F : 5′- TACAAGCCCGATACTTATGAC -3′	21	60
	R : 5′- AAGGTTGGGAAGGACTGAAT -3′	20	58
Insulin degrading enzyme (*IDE*)			
	F : 5′- GGAGTGGGTGTTTGAAGAGTGT -3′	22	66
	R : 5′- TCTAACTCATCTTCAGCCTACCA -3′	23	66
Epidermal growth factor (*EGF*)			
	F : 5′- GTCACTGCTGAAAATAGCCATCT -3′	23	66
	R : 5′- CGGTTCTGAAACACTGAAATACG -3′	23	66
Dolichyl diphosphooligosaccharyl protein glycotransferase (*DDPG*)			
	F : 5′- TTCTGGCAACGGCAAAGTAG -3′	20	60
	R : 5′- ATGGGTCAATGCGAACAAAG -3′	20	58

Table 2.8 Amplification conditions used for PCR-SSCP analysis

Gene	Amplification condition
5′UTR-*LTB4DH*	94 °C for 3 minutes, 1 cycle followed by 35 cycles of
5′UTR-*NASP*	94 °C for 30 seconds, 55 °C for 45 seconds, 72 °C for
THBP	45 second and the final extension at 72 °C for 7 minutes
DDPG	
IDE and EGF	94 °C for 3 minutes, 1 cycle followed by 35 cycles of
	94 °C for 30 seconds, 53 °C for 45 seconds, 72 °C for
	45 second and the final extension at 72 °C for 7 minutes

2.10 Single Strand Conformational Polymorphism (SSCP) analysis

The PCR product of growth- and reproduction-related genes were successfully amplified using the amplification conditions described in Table 2.8 but did not exhibit size polymorphism or allelic variants between individual were further characterized using single strand conformational polymorphism (SSCP) to examine whether the amplification product of the same genes in different shrimp individuals were polymorphic due to the existence of single nucleotide polymorphism (SNP). SSCP analysis allows the identification of polymorphic sites in DNA. Single-stranded DNA is resolved based on both conformation and size.

2.10.1 Preparation of glass plates

The long and short glass plates were treated and prepared as previously described for Microsatellite analysis. The cleaned glass plates were assembled with a pair of 1.5 mm spacer.

2.10.2 Non-denaturing polyacrylamide gel electrophoresis

Different concentrations of low closslink non-denaturing polyacrylamide gels (37.5:1 or 75:1 of acrylamide and bis-acrylamide) were prepared by dilution of a 40% stock solution to required gel concentration and degased *in vacuo*. The acrylamide gel solution (30 - 40 ml) was mixed with 300 µl of 10% ammonium persulphate and 30 µl

of tetramethylethylenediamine (TEMED). The analytical comb was inserted into the prepared gel and allow for polymerization for at least 4 hours or overnight.

2.10.3 PCR and electrophoresis

PCR was carried out against genomic DNA of *P. monodon* using conditions described in Table 2.8. Six microliters of the amplified PCR product were mixed with four volumes of SSCP loading dye (95% formamide, 0.25% bromophenol blue, 0.25% xylene cyanol and 10 mM NaOH), denatured at 95 °C for 5 minutes and immediately cooled on ice for 5 minutes. The denatured products were electrophoretically analyzed in native polyacrylamide gels at 200 - 250 volts for 12 - 16 hours at 4 °C. SSCP band were visualized by silver staining.

2.11 Identification of SNP by direct DNA sequencing of PCR product

The PCR product was amplified from genomic DNA and electrophoretically analyzed in duplication. One was run side-by-side with a 100 bp DNA markers and the other was loaded into the distal well of the gel. After electrophoresis, lanes representing the DNA standard and its proximal SNA sample were cut and stained with ethidium bromide (0.5 µg/ml) for 5 minutes. Position of the DNA marker and the ethidium bromide-stained fragment were used to align the position of the non-stained target DNA fragment.

2.11.1 Direct sequencing of the PCR product

After electrophoresis, the desired DNA fragment was excised from the agarose gel using a sterile scalpel and placed in a pre-weighed microcentrifuge tube. DNA was eluted out from the gel using a illustra™ PCR band purification kit (GE Healthcare).

The gel-eluted PCR products were sequenced for both directions on the automated DNA sequencer (MACROGEN, Korea) using the original forward or reverse primers primers of investigated genes as the sequencing primer. Nucleotide sequences of different SSCP genotypes of each gene were aligned by Clustal X.

2.12 Identification of SNP in cloned PCR product

The PCR product was ligated to the pGEM®-T Easy vector. The reaction mixture was incubated overnight at 4 °C before transformed into *E. coli* JM 109. After that, extracted plasmid DNA using an illustra™ plasmidPrep mini spin kit (GE Healthcare). The recombinant plasmid were sequenced for both direction using the automated DNA sequencer (MACROGEN, Korea). Nucleotide sequences of different SSCP genotypes of each gene were aligned by Clustal X.

2.13 RNA extraction

Total RNA was extracted from hepatopancrease and ovaries of each the shrimp using TRI REAGENT® (Molecular Research Center). A piece of tissue was immediately placed in mortar containing liquid nitrogen and ground to the fine powder. This tissue powder was transferred to a microcentrifuge tube containing 200 µl of TRI REAGENT (1 ml / 50 - 100 mg tissue) and homogenized. Additional 800 µl of TRI REAGENT were added. The homogenate was left at room temperature for 5 minutes before 0.2 ml of chloroform was added. The homogenate was vortexed for at least 15 seconds, left at room temperature for 2 - 15 minutes and centrifuged at 12,000 rpm for 15 minutes at 4 °C. The mixture was separated into the lower phenol-chloroform phase (red), the interphase, and the upper aqueous phase (colorless).

The aqueous phase (inclusively containing RNA) was carefully transferred to a new 1.5 ml microcentrifuge tube. RNA was precipitated by an addition of 0.5 ml of isopropanal and mixed thoroughly. The mixture were left at room temperature for 10 - 15 minutes and centrifuged at 12,000 rpm for 10 minutes at 4 °C, the supernatant was removed. The RNA pellet was washed with 1 ml of 75% ethanol and centrifuged at 12,000 rpm for 10 minutes at 4 °C. The ethanol was removed. The RNA pellet was air-dried for 5 - 10 minutes. RNA was dissolved in DEPC-treated H_2O for immediately used. Alternatively, the RNA pellet was kept under absolute ethanol in -80 °C freezer for long storage.

Concentration of extracted total RNA was measured by spectrophometry as described above. The quality of total RNA was evaluated by electrophoresis of 1 µl of RNA in 1.0% agarose gels as described previously.

2.14 DNase I treatment

Fifty-five micrograms of total RNA were treated with DNase I (0.5 U/1 µg of RNA, usb) at 37 °C for 30 minutes. After the incubation, the sample was gently mixed with the a sample volume of phenol:chloroform:isoamylalcohol (25:24:1) for 10 minutes. The sample was centrifuged at 12,000 rpm for 10 minutes at 4 °C, and the upper aqueous phase was collected. The extraction process was then repeated once with chloroform:isoamylalcohol (24:1) and one with chloroform. The final aqueous phase was mixed with one-tenth final sample volume of 3 M sodium acetate, pH 5.2. After that, RNA was precipitated by adding twice the sample volume of cold absolute ethanol. The mixture was incubated at -80 °C for 30 minutes, and the precipitated RNA was recovered by centrifugation at 12,000 rpm for 10 minutes at 4 °C. The RNA pellet was then washed twice with 1 ml of cold 75% ethanol. Alternatively, the RNA pellet was kept in absolute ethanol at -80 °C until required.

2.15 Synthesis of the first strand cDNA

One and a half micrograms of total RNA from the tissue of *P. monodon* were reverse-transcribed to the first strand cDNA using an Improm- II™ Reverse Transcription System Kit (Promega). Total RNA was combined with 0.5 µg of oligo dT_{12-18} and appropriate DEPC-treated H_2O in final volume of 5 µl. The reaction was incubated at 70 °C for 5 minutes and immediately placed on ice for 5 minutes. Then 5X reaction buffer, $MgCl_2$, dNTP Mix, RNasin were added to final concentrations of 1X, 2.25 mM, 0.5 mM and 20 units, respectively. Finally, 1 µl of Improm- II™ Reverse transcriptase was add and gently mixed pipetting. The reaction mixture was incubated at 25 °C for 5 minutes and at 42 °C for 90 minutes. The reaction mixture was incubated at 70 °C for 15 minutes to terminate the reverse transcriptase activity. Concentration and rough quality of the newly synthesized first strand cDNA was spectrophotometrically examined (OD_{260}/OD_{280}) and electrophoretically analyzed by 1.2% agarose gels.

2.16 RT-PCR

Typically, the amplification reactions were performed in a 25 µl reaction volume containing 10 mM Tris-HCl, pH 8.8, 50 mM KCl and 0.1% Triton X-100, 1.5

- 2.0 mM $MgCl_2$, 100 or 200 µM each of dNTPs, 0.2 µM of each primer, 1 unit of DyNazyme™ DNA Polymerase (Finnzymes) and 2 µl of a 10 fold-dilute first strand cDNA (about 200 ng). RT-PCR was initial performed by predenaturation at 94 °C for 3 minutes followed by 20, 25 and 30 cycles of denaturation at 94 °C for 30 seconds, annealing at 53 °C (or 55 °C) for 45 seconds and extension at 72 °C for 30 seconds. The final extension was carried out at 72 °C for 7 minutes. Five microliters of the amplification products were electrophoretically analyzed though 1 - 2% agarose gel.

Table 2.9 Genes, primer sequence, length and the melting temperature of primers used for RT-PCR

Gene	Sequence	Length (bp)	Tm (°C)
LTB4DH	F : 5′- TGGGAGGCACTCAGTCAAGC -3′	20	64
	R : 5′- TTCCATCCAGCGATCAAACC -3′	20	60
NASP	F : 5′- GCCGTCCAAGAAAGAGATTGATAC -3′	24	70
	R : 5′- CGGCCATAGTAGAAATAAGCATCAC -3′	24	70
THBP	F : 5′- TACAAGCCCGATACTTATGAC -3′	21	60
	R : 5′- AAGGTTGGGAAGGACTGAAT -3′	20	58
IDE	F : 5′- GGAGTGGGTGTTTGAAGAGTGT -3′	22	66
	R : 5′- TCTAACTCATCTTCAGCCTACCA -3′	23	66
EGF	F : 5′- GTCACTGCTGAAAATAGCCATCT -3′	23	66
	R : 5′- CGGTTCTGAAACACTGAAATACG -3′	23	66

Table 2.10 The amplification conditions for RT-PCR

Gene	Amplification condition
LTB4DH	94 °C for 3 minutes, 1 cycle followed by 35 cycles of
NASP	94 °C for 30 seconds, 55 °C for 45 seconds, 72 °C for
THBP	45 second and the final extension at 72 °C for 7 minutes
IDE	94 °C for 3 minutes, 1 cycle followed by 35 cycles of
EGF	94 °C for 30 seconds, 53 °C for 45 seconds, 72 °C for
	45 second and the final extension at 72 °C for 7 minutes

2.17 Tissue distribution analysis by RT-PCR

Gene expression in tissue was studies in ovaries, testes, heart, hemocytes, lymphoid organs, intestine, gill, pleopods, thoracic ganglion, stomach, eyestalk and hepatopancrease of *P. monodon*.

For the target genes, 150 ng of the first strand cDNA from various tissues was used as the template in 25 µl reaction volume containing 10 mM Tris-HCl, pH 8.0, 50 mM KCl and 0.1% Triton X-100, 2 mM $MgCl_2$, 100 mM each of dATP, dCTP, dGTP and dTTP, 0.2 µM of each primer and 1 unit of DyNAzyme^{TM} DNA Polymerase (FINNZYMES). *Elongation factor-1α* (F : 5′- ATGGTTGTCAACTTTGCCCC -3′ and R : 5′- TTGACCTCCTTGATCACACC -3′) were also amplified from the same sample. The reactions were predenaturation at 94 °C for 3 minutes followed by 25 cycles composing of a 94 °C denaturation step for 30 seconds, a 53 °C annealing step for 45 seconds and a 72 °C extension step for 45 seconds. The final extension was carried out at 72 °C for 7 minutes. Five microliters of the amplification product was electrophoretically analyzed though a 1.5% agarose gel.

2.18 Isolation and characterization of the cDNA sequence of *insulin degrading enzyme* (*IDE*) of *P. monod*on using Rapid Amplification of cDNA Ends-Polymerase Chain Reaction (RACE-PCR)

2.18.1 Preparation of the 5′ and 3′ RACE template

Total RNA was extracted from ovaries of *P. monodon* broodstock using TRI Reagent. Messenger (m) RNA was purified using a QuickPrep *micro* Purification Kit (BD Biosciences). The RACE cDNA template was prepared by combining 1 µg of ovarian mRNA with 1 µl of 5′-CDS primer and 1 µl of 10 µM SMART II A oligonucleotide for 5′ RACE-PCR or 1 µg of ovarian mRNA with 1 µl of 3′ CDS primer A for 3′ RACE-PCR (Table 2.11). The components were mixed and spun briefly. The reaction was incubated at 70 °C for 2 minutes and snapcooled on ice for 2 minutes. The reaction tube was spun briefly. After that, 2 µl of 5X First-Strand buffer, 1 µl of 20 mM DDT, 1 µl of dNTP Mix (10 mM each) and 1 µl of PowerScript Reverse Transcriptase were added. The reactions were mixed by gently pipetting and centrifuged briefly to collect the contents at the bottom of the tube. The tube was

incubated at 42 °C for 1.5 hours in a thermocycler. The first strand reaction products were diluted with 125 µl of Tricine-EDTA buffer and heated at 72 °C for 7 minutes. The first strand cDNA tenplate was stored at -20 °C.

2.18.2 Primer designed for RACE-PCR and primer walking

Gene-specific primers (GSPs) for *IDE* were designed from ovaries and hemocytelibraries. The antisense primer (and nested primer) was designed for 5′ RACE-PCR and the sense primer (and nested primer) for the 3′ RACE-PCR (Table 2.12).

For sequencing of genes that showed the full length from the 5′ direction, the product from colony PCR was considered. If the insert was larger than that of the homologues, the 3′ direction was sequenced. Internal primers were designed for primer walking of the inserted cDNA (Table 2.13).

Table 2.11 Primer sequence for the first strand cDNA synthesis and RACE-PCR

Primers	Sequence
SMART II A Oligonucleotide	5′- AAGCAGTGGTATCAACGCAGAGTACGCGGG -3′
3′ RACE CDS Primer A	5′- AAGCAGTGGTATCAACGCAGAGTAC(T)30 N-1 N -3′ (N=A, C, G or T; N-1= A,G or C)
5′ RACE CDS Primer	5′-(T)25 N-1 N-3′ (N=A, C, G or T; N-1= A,G or C)
10X Universal PrimerA Mix (UPM)	Long : 5′-CTAATACGACTCACTATAGGGCAAGCAGTG GTATCAACGCAGAGT-3′ Short : 5′- CTAATACGACTCACTATAGGGC -3′
Nested Universal Primer A (NUP)	5′- AAGCAGTGGTATCAACGCAGAGT -3′

Table 2.12 Gene-Specific Primers (GSPs) and nested GSP used for isolation of the cDNA of *P. monodon IDE*

Gene specific primer	Sequence	Length (bp)	Tm (˚C)
Insulin degrading enzyme (*IDE*)			
5′ RACE1	R : 5′- TCTTTACACTCTTCAAACACCCACTC -3′	26	74
5′ RACE (nested)	R : 5′- CCAACAAGAGAGTTCACCCATCCA -3′	24	72
3′ RACE	F : 5′- GGAGTGGGTGTTTGAAGAGTGT -3′	22	66

Table 2.13 Internal primer used for primer walking sequencing of the full length cDNA of *IDE* in *P. monodon*

Internal primer	Sequence	Length (bp)	Tm (˚C)
3′RACE-Internal	R : 5′- CTACATTCACTGTTGATCCGA -3′	21	60

2.18.3 RACE-PCR

The master mix which is sufficient for 5′ and 3′ RACE-PCR and the control reactions was prepared (Tables 2.14 and 2.15). For each 25 µl amplification reaction, 14.0 µl sterile deionized H_2O, 2.5 µl of 10X Advantage® 2 PCR buffer, 0.5 µl of 10 µM dNTP mix and 0.5 µl of 50X Advantage® 2 DNA Polymerase mix were combined. The reaction was carried out for as described in Table 2.16.

Table 2.14 Compositions for amplification of the 5′end of gene homologues using 5′ RACE-PCR

Compenent	5′RACE-PCR	UPM only (Control)	GSP only (Control)
5′RCAE-ready cDNA template	1.5 µl	1.5 µl	1.5 µl
UPM (10X)	5.0 µl	5.0 µl	-
GSP1 (10 µM)	1.0 µl	-	1.0 µl
GSP2 (10 µM)	-	-	-
H_2O	-	1.0 µl	5.0 µl
Master Mix	17.5 µl	17.5 µl	17.5 µl
Final volume	25 µl	25 µl	25 µl

Table 2.15 Compositions for amplification of the 3′ end of gene homologues using 3′ RACE-PCR

Compenent	3′RACE-PCR	UPM only (Control)	GSP only (Control)
3′RCAE-ready cDNA template	1.5 µl	1.5 µl	1.5 µl
UPM (10X)	5.0 µl	5.0 µl	-
GSP1 (10 µM)	1.0 µl	-	1.0 µl
GSP2 (10 µM)	-	-	-
H_2O	-	1.0 µl	5.0 µl
Master Mix	17.5 µl	17.5 µl	17.5 µl
Final volume	25 µl	25 µl	25 µl

Table 2.16 The amplification conditions for RACE-PCR of various gene homologues of *P. monodon*

Gene	Amplification condition
Insulin degrading enzyme (*IDE*)	
5′ RACE-PCR	25 cycles of 94 °C for 30 seconds, 65 °C for 45 seconds, 72 °C for 90 seconds and the final extension at 72 °C for 7 minutes
5′ RACE-PCR (nested)	20 cycles of 94 °C for 30 seconds, 65 °C for 45 seconds, 72 °C for 90 seconds and the final extension at 72 °C for 7 minutes
3′ RACE-PCR	25 cycles of 94 °C for 30 seconds, 65 °C for 45 seconds, 72 °C for 90 seconds and the final extension at 72 °C for 7 minutes

The primary 5′ and 3′ RACE-PCR products were electrophoretically analyzed through 0.8 - 1.0% agarose gel. If the discrete expected bands were not obtained from the primary amplification, nested PCR was performed using the recipes illustrated in Table 2.16. The primary PCR product was 50-fold diluted. The primary PCR product was performed using 1 - 5 µl of diluted PCR product as a template using the conditions described in Table 2.16.

2.19 Examination of expression levels of interesting genes in hepatopancreases and ovaries of *P. monodon* by quantitative real-time PCR

2.19.1 Primer design

The intron/exon structure of *LTB4DH* and *NASP* was characterized. Several primer pairs were designed from cDNA sequence of each gene and used to PCR against genomic DNA as the template. The PCR fragment was cloned and sequenced. The forward or reverse primer covering intron/exon boundaries or alternatively, a primer pairs sandwiching the large intron was designed (Figure 2.1). A size of the expected PCR product size was approximately 100 - 200 bp or less than 500 bp.

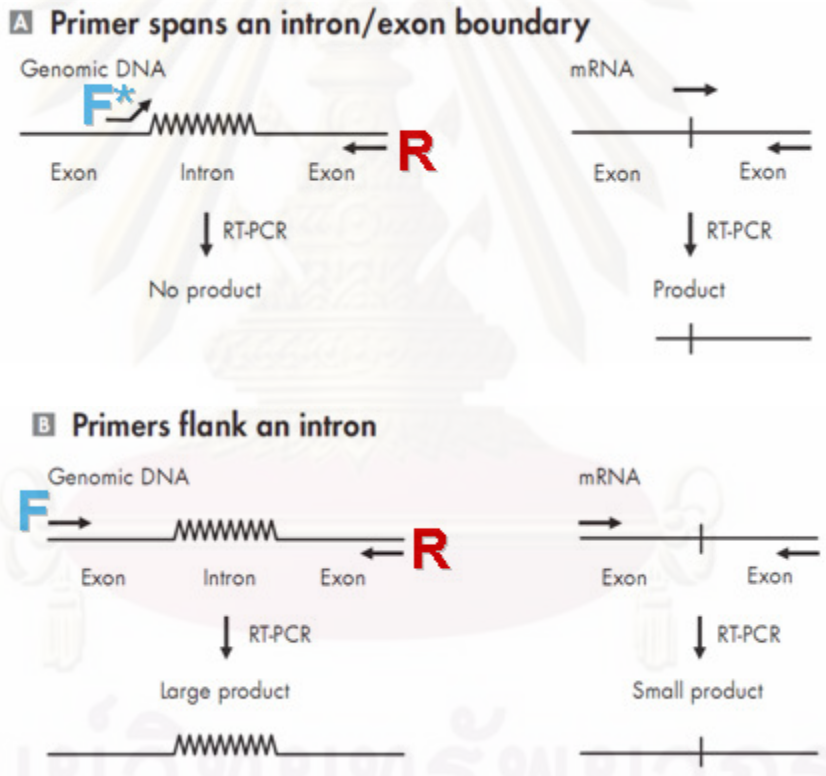


Figure 2.1 Two approaches of primer design to prevent amplification of the residual genomic DNA in the template for quantitative real-time RT-PCR based on primers spanning exon-intron boundaries (A) or primers flanking an intron (B).

For quantitative real-time PCR analysis of *IDE*, *EGF* and *THBP*, primers for SNP detection and RT-PCR were used. These primers did not generate the primer dimmer and any sign for the contamination of genomic DNA in the cDNA template.

For construction of the standard curve of each gene, the DNA seqment covering the target PCR product and *EF-1α* were amplified from primers for quantitative real-time PCR. The PCR product were cloned Plasmid DNA were extracted and used as the template for estimation of the copy number. A 10 fold-serial dilution was prepared corresponding to 10^2 - 10^7 molecules/µl. The copy number of standard DNA molecules can be calculated using the following formula:

X g/µl DNA / [plasmid length in bp x 660] x 6.022×10^{23} = Y molecules/µl

The standard curves (correlation coefficient = 0.995 - 1.000 or efficiency higher than 95%) were drawn for each run. The standard samples were carried out in a 96 well plate and each standard point was run in duplicate.

2.19.2 Quantitative real-time PCR analysis

The target transcripts and the internal control *EF-1α* of the synthesized cDNA were amplified in a reaction volume of 10 µl using 2X LightCycler® 480 SYBR Green I Master (Roche, Germany). The specific primer pairs were used at a final concentration of 0.3 µM. The amplification condition was the same as that described in the Table 2.18. The real-time PCR assay was carried out in a 96 well plate and each sample was run in duplicate using a LightCycler® 480 Instrument II system (Roche).

A ratio of the absolute copy number of the target gene and that of *EF-1α* was calculated. The relative expression level between shrimp possessing different ovarian development (or treatment) were statistically tested using one way analysis of variance (ANOVA) followed by a Duncan's new multiple rang test. Significant comparisons were considered when the *P value* was < 0.05.

Table 2.17 Genes, primer sequence, length and the melting temperature used for quantitative real-time PCR

Gene	Sequence	Length (bp)	Tm (°C)
LTB4DH	F : 5′- GCTGGCTTGTGAAGATGGAGATGTC -3′	25	76
	R : 5′- CCACTCTGGGTTCTTACTCTCAA -3′	23	68
NASP	F : 5′- GCCGTCCAAGAAAGAGATTGATAC -3′	21	64
	R : 5′- CGGCCATAGTAGAAATAAGCATCAC -3′	21	62
THBP	F : 5′- TACAAGCCCGATACTTATGAC -3′	21	60
	R : 5′- AAGGTTGGGAAGGACTGAAT -3′	20	58
IDE	F : 5′- GGAGTGGGTGTTTGAAGAGTGT -3′	22	66
	R : 5′- TCTAACTCATCTTCAGCCTACCA -3′	23	66
EGF	F : 5′- GTCACTGCTGAAAATAGCCATCT -3′	23	66
	R : 5′- CGGTTCTGAAACACTGAAATACG -3′	23	66

Table 2.18 Amplification condition for used quantitative real-time PCR

Gene	Amplification condition
LTB4DH	95 °C for 10 minutes, 1 cycle followed by 40 cycles of 95 °C for 15 seconds, 58 °C for 30 seconds, 72 °C for 30 seconds
NASP	95 °C for 10 minutes, 1 cycle followed by 40 cycles of 95 °C for 30 seconds, 58 °C for 45 seconds, 72 °C for 30 seconds
THBP	95 °C for 10 minutes, 1 cycle followed by 40 cycles of 95 °C for 15 seconds, 55 °C for 30 seconds, 72 °C for 30 seconds
IDE	95 °C for 10 minutes, 1 cycle followed by 40 cycles of 95 °C for 15 seconds, 53 °C for 30 seconds, 72 °C for 30 seconds
EGF	

2.20 Statistic analysis

Correlation between SSCP patterns of *LTB4DH* and GSI values of 14-month-old shrimp and between those of *THBP* and *EGF* and the body weight of 3-month-old juveniles were statistically tested (using independent t-test for *EGF* or one way analysis of variance, ANOVA following by Duncan's multiple range test for others). Significant differences were considered if $P < 0.05$.

The expression levels of *LTB4DH* and *NASP* in ovaries of broodsotck shrimp carrying different GSI values (<0.5% for group I, 0.5 - 1.0% for group II and > 1.0% for group III) were statistically tested using ANOVA and Duncan's multiple range test. In addition, correlation between the expression levels of *LTB4DH* in shrimp carrying different SSCP genotypes (SNP) was also statistically examined as above.

The expression levels of *IDE*, *THBP* and *EGF* in hepatopancreas of juvenile shrimp exhibiting different growth rates according to the body weight tested using the independent t-test. In addition, correlation between the expression levels of *IDE*, *THBP* and *EGF* in shrimp carrying different SSCP genotypes (SNP) was also statistically examined as above.

CHAPTER III

RESULTS

3.1 Genomic DNA extraction

Genomic DNA was extracted from a piece of pleopod of each *P. monodon* using phenol-chloroform-proteinase K method (Klinbunga *et al*., 1996). The quality of extracted genomic DNA was electrophoretically determined using a 0.8% agarose gel. High molecular weight DNA at the similar size as that of undigested λDNA (approximately 50 kb) along with partially degraded DNA was observed (Figure 3.1).

The ratio of OD_{260}/OD_{280} of extracted genomic DNA was 1.8 - 2.0 suggesting that the quality of extracted genomic DNA was acceptable for used in further applications. DNA samples showing the ratio much lower then 1.8 was possibly contaminated with residual proteins or phenol. In contrast, those with the ratio greater than 2.0 may be contaminated with RNA as visualized by the smear at the bottom of gel (as can be seen in Figure 3.1). However, RNA contamination did not affect the PCR amplification. Therefore, the extracted DNA was used for various techniques (e.g. genome walking and PCR-SSCP).

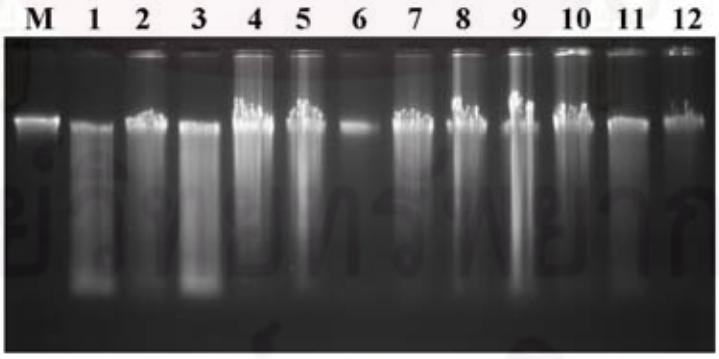


Figure 3.1 A 0.8% ethidium bromide-stained agarose gel showing the quality of genomic DNA (Lanes 1-12) extracted from a pleopod of different individuals of *P. monodon*. Lane M = 120 ng of undigested λDNA.

3.2 Identification of 5′ untranslated region (5′UTR) and introns of functionally important genes

Genome walking analysis was carried out for isolation of the 5′UTR of several functionally important genes including *NADP-dependent leukotriene B4 12-hydroxydehydrogenase (LTB4DH)* and *nuclear autoantigenic sperm protein (NASP)* In addition, the genomic gene segment of *LTB4DH, thyroid hormone binding protein (THBP), DEAD-box protein 52, insulin degrading enzyme (IDE)* and *epidermal growth factor (EGF)* were also identified.

3.2.1 *NADP-dependent leukotriene B4 12-hydroxydehydrogenase (LTB4DH)*

Genome walking analysis of *LTB4DH* was carried out using the *Stu* I, *Pvu* II and *Sca* I mini-libraries as the template. The primary PCR product of 1,400, 900 and 2,000 bp product from the respective mini-libraries was obtained (Figure 3.2). These fragments were cloned and sequenced.

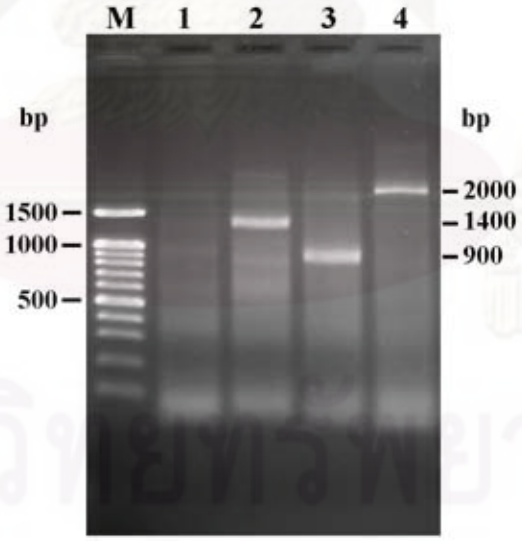


Figure 3.2 The primary PCR product of *NADP-dependent leukotriene B4 12-hydroxydehydrogenase* amplified from the *Eco* RV (Lane 1), *Stu* I (Lane 2), *Pvu* II (Lane 3) and *Sca* I mini-libraries (Lane 4) using reverse gene-specific primers and the adapter primer (AP1). A 100 bp DNA ladder (Lane M) were used as the DNA marker.

After sequence assembly, additional 767 bp of the 5′UTR was deduced from partial nucleotide sequence of cloned 1,400, 900 and 2,000 bp genome walking fragments. Sequences from one end of these fragments were perfectly overlapped with the original full length cDNA sequence of *LTB4DH*. Moreover, primer sequences used for PCR amplification were also found suggesting that part of the 5′UTR of this gene was successfully identified (Figure 3.3).

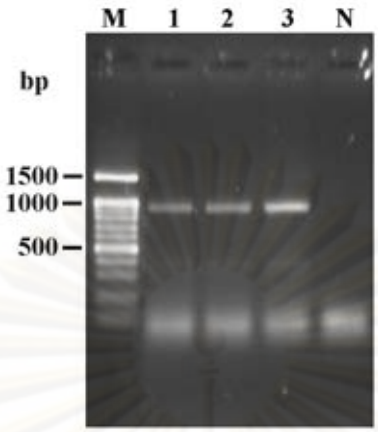
ACGCGTGGTCGACGGCCCGGGCTGGTCTGCTTTTCCTCTCCATCTTTCTTTGCTTCCTCTTCTCTTTGT
ATAATTAGCACCGACCCTTTTACTTTACTTCATCTAATATTTGTTTATTTTCTAAATTAAAAAGTTTGC
TAAAATACTCTCCATCTTCCTTTTGTATTTTCAGCCTCTACTATGTTTCTATTTCCATCCCTTACCACA
CCTATATTTTCAGTTCTAGTTTTTTATGCATTTCCTCAGCTACCTTATATATTGCCTTTTCATTGTCTG
TCATGCCATTTTTTATGACTTCCCAGACCGCTATTTTAACTAACTTCATTGCTTTTCAGTTTGCTTCTT
TACACCTTATTTCATCTTCCTTATTTTTGCTTGTTTCCATTTATTTTTTTGTTTCATAGATTTCTTATT
CCACCACCAAGTCGGTATTTTTTGCTCTTGTATCCTTTGCTGTTATAACCTTTGTCTCCTTCATATTTT
GCCATCCTTCCTCTCTCTTCTGTCTTCTCTTTATATTGTTTCAGAGGATCATCTGGCTTTGTAAATATT
TATAGAAAATACCAGACTCTGTGCTAAATAATTTCTCACTCTACTCTTGACCTGCTTTTTTTCTCCACA
AAGCCTACTCCCTTCCAACATGTCCATTTAAGTCTCCTTTCACTGTCACACACACACACCCCTTTCGGG
GACTTTATTATCTGTAGAGATGCATGTGGCTTTTGTCTTGCAGGAAATCTGTAGTCCCTAGCATCCCAG
AAGCCAAA**ATG**GTGACAGCAAAGGTGTGGAAGCTAGCCAAGCGGCCAGAGGGGATGCCCAAGAAGGAAG
ACTTTGTGTGTGTGGAAGAAGAG*CTGGCTTGTGAAGATGGAGGT*

Figure 3.3 The 5′ untranslated region (5′UTR) of *NADP-dependent leukotriene B4 12-hydroxydehydrogenase*. The putative start codon is boldfaced and underlined. Primers for genome walking analysis is italicized and underlined. Overlapped sequences between the genome walking fragment and full length cDNA of *LTB4DH* is highlighted. A 767 bp fragment was successfully identified.

In addition, a 1001 bp fragment (Figure 3.4A) was amplified by primers LTB4DH-2F/R. Two introns (461 and 276 bp) were found within the amplification product along which exons of 56 (partial), 124 and 84 (partial) bp (Figure 3.4B).

A 738 bp fragment (Figure 3.5A amplified by primer LTB4DH-3F/R was also cloned and sequenced. Two introns (196 and 315 bp) were found within the amplification product along which exons of 52 (partial), 121 and 54 (partial) bp (Figure 3.5B).

A

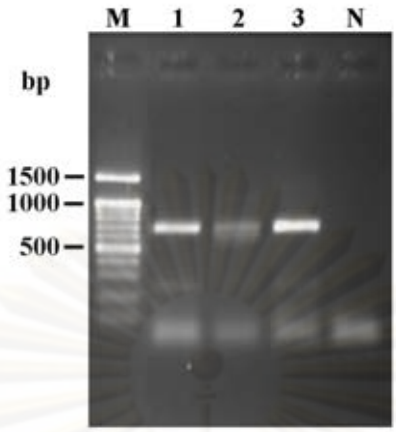


B

TGGGAGGCACTCAGTCAAGCAGCGCCTGATAAGATTAACTGCTTCTTTGAAAATGT*AAGTTTGGACAC*
ATTGTTTTTACCACAGACTTATTTTAATACAATTTAAAGATGAAATGACTCAGAATTTTACTGAATAA
CGTGGGTTATAGAATCTAAAATAGAATTGGTGAATTCTGTAATAGTTTTACATTCACATTTAGAAAAT
GAGAAATTGTGGGTAAAGTATTTAAACTGGACCCATTTATCCAGGTAACCAAGCAACTGTAAGGAGTT
TTGCAGGTATACATAATCCATTAGTCTGAGTCACAAGATTGGTATTTTTCTTGTTTTAATAATATATA
ACAATTTTGCTTCTTGATGTTACAGTTTTGGAGAGTTTTTAATAATGTGGCTATATGAATTTATTTAA
TAGAATTAACAGATATTGAATATACTTAGGTTGATATGGGAGGAAAAAGTAAGTTCTACCCTTGTTTA
*TTACTTGTCTTCTTTATTTTTATTTTTGTTCATTTTAAGGT*TGGAGGACAGTTCACTGCTGAAGCTCT
ACCACACATGGCTGAGAGGGGCAGAATGGCAATTTGTGGTGTGATTTCAGAGTACAATGATGACGACA
AGGATTCTGGTGTGGTGGCAATGACTAGT*AGGTATTTTGTTGTTTCATTGTCAGTGTAATGACTTTTT*
TCCTCCCAGTAGGTTTCTTTTCTGTATATGAATGTCTGTAGAATATATAATTAAAGTTACTTATATGA
TTAGGCCTTTTTAGTAGATAATGTACATTTATTTATAGTGCAGTTTGAAAATTAGAATAGTGGTCTTA
TTTATTTTATTAAATTCTGGCTCCCCTGTGTTGATAGAATCATATGAGGATAATCCTTTGTTATGGTA
*GGGTAAAATGTTTCATCACTCGATGTTCCAGGT*CCGTTGAATCCAGCTACGATACTTTGGAAGCAACT
TGCAGTGGAAGGCTTCATTGTTACAAGGT**GGTTTGATCGCTGGATGGAA**

Figure 3.4 (A) A 1.5% ethidium bromide-stained agarose gel showing the amplification product of *LTB4DH* using primers LTB4DH-2F/R. Lane M = a 100 bp DNA ladder, lanes 1-3 = the amplified product from genomic DNA of *P. monodon* and lane N = negative control (without genomic DNA template). (B) The partial genomic DNA sequence of *LTB4DH* amplified from primer LTB4DH-2F/R. Introns are italicized and underlined. Primer sequences are illustrated in boldface and underlined.

A



B

GAGATGGGAGACTTGCCTAAGAGTCTTGGCTTAGGGATTCTGGGCATGCCGG*GGTATGTCCAGTTATT*
AGTCTTGGGTGGTGTAAAAAAAAGAAAAGTCGAGACTAAGTTAAAGTATAGAGCTCATTTGGATTTGG
AGTTGTTTATATGTTCTTGATTTTATTGTACATGGAAACCATATAATAATATATTTCTTAAAGAGTGT
*TAGTACATTTACAATATATGTATAAAGTTAAATTTTTTATCACA*GGAATACAGCATATTTTGGATTTC
TTGAAATCTGCCGACCGAAAGCAGGAGACACTGTGTTGGTAAATGCTGCTGCTGGAGCAGTTGGAAGT
GCTGTGATTCAGATAGCCAAAATCAAAGG*TTAGAACTAAAACTGTTTCTTTAGAGAAATGCTTAATTA*
AGAAACTGTCTCACACAGTCTGTTGGACACAATTTATGAAAAGATGGATTACAGTGATAGTTGGGTTC
ATCAAAGCTTGTGTTAGAGAATAAGAATCTAATCGCTGGATGTAATAAGCATTGGTTATAGAGTAAAG
CAAATTTAAAACACTCAACCCCATGATAAGAAAAGAAACTTTATGGAAAAAATTTGAATTTATAGTAT
CAGTGTAGCTTATGAATATTTTAGGCTTTATTTGTGAATGCATCATTTTTATATAAAGTTTCCCCCTC
*TAGG*GTGTAAGGTAATTGCCTTTTCGGGATCTGATGA**AACAGGTAGCGTGGACCAAAG**

Figure 3.5 (A) A 1.5% ethidium bromide-stained agarose gel showing the amplification product of *LTB4DH* using primers LTB4DH-3F/R. Lane M = a 100 bp DNA ladder, lanes 1-3 = the amplified product from genomic DNA of *P. monodon* and lane N = negative control (without genomic DNA template). (B) The partial genomic DNA sequence of *LTB4DH* amplified from primers LTB4DH-3F/R. Introns are italicized and underlined. Primer sequences are illustrated in boldface and underlined.

Nucleotide sequences of all characterized fragments were assembled and illustrated by Figure 3.6.

ACGCGTGGTCGACGGCCCGGGCTGGTCTGCTTTTCCTCTCCATCTTTCTTTGCTTCCTCTTCTCTTTGT
ATAATTAGCACCGACCCTTTTACTTTACTTCATCTAATATTTGTTTATTTTCTAAATTAAAAAGTTTGC
TAAAATACTCTCCATCTTCCTTTTGTATTTTCAGCCTCTACTATGTTTCTATTTCCATCCCTTACCACA
CCTATATTTTCAGTTCTAGTTTTTTATGCATTTCCTCAGCTACCTTATATATTGCCTTTTCATTGTCTG
TCATGCCATTTTTTATGACTTCCCAGACCGCTATTTTAACTAACTTCATTGCTTTTCAGTTTGCTTCTT
TACACCTTATTTCATCTTCCTTATTTTTGCTTGTTTCCATTTATTTTTTTGTTTCATAGATTTCTTATT
CCACCACCAAGTCGGTATTTTTTGCTCTTGTATCCTTTGCTGTTATAACCTTTGTCTCCTTCATATTTT
GCCATCCTTCCTCTCTCTTCTGTCTTCTCTTTATATTGTTTCAGAGGATCATCTGGCTTTGTAAATATT
TATAGAAAATACCAGACTCTGTGCTAAATAATTTCTCACTCTACTCTTGACCTGCTTTTTTTCTCCACA
AAGCCTACTCCCTTCCAACATGTCCATTTAAGTCTCCTTTCACTGTCACACACACACACCCCTTTCGGG
GACTTTATTATCTGTAGAGATGCATGTGGCTTTTGTCCTGAGAGAAATCTGTAGTCCCTAGCATCCCAG
AAGCCAAA**ATG**GTGACAGCAAAGGTGTGGAAGCTAGCCAAGCGGCCAGAGGGGATGCCCAAGAAGGAAG
ACTTTGTGTGTGTGGAAGAAGAGCTGGCTTGTGAAGATGGAG***GTAATGTGTATCAGCATATTATATACA***
TAGTCCCCTTTGGTTTGTATATTCCTATATGATATTCAATTACTTTTATCTGTTGATAATTATCATTGG
TTTTTGTCTCTAGTAATTTAAGGCAATACCATTCCCACATAATGAAATACAAGAAAACTGTTATTCTAA
TGTTCTGGCAAATGATTAACTAACCTTTTTCAGTACTGCAATGTCACTAATGATTTACAAATGCTTTTG
CAGAGATGTCATTATTGAAGCTGAATTTTTGAGTGTAGATCCATACATGCGGTACATGATCAAGCAAAT
CCCTCTTGATGTACCTGTCACCGGTACTCAGGTTGCAAAAGTAT***GTGACTTTAACTCTCATTATATTAT***
TTTTGCTTGAAGAGATTATACTATATATGAAAAATTCACTTACCAGAAGGTTAGGCCTAGAATGATTAA
AGAATTTTTCAGATTAAGCAAAAATGCCTAAATTGGTTGAGGAAAAATAAAGGTCTTTTATTCACTTAA
AAAAAAAAAATCCCTCTAAAGGAAATGTCCTACATTATTGTGAATATTGAGCCATGTATTTTTGTTTAA
AAAATGCATAAATATATAGACAAATTAACAATTGCACACTGGAAGTACAAACAAAAAGTTTAGTAAATC
ATTATGTACTTCTTCCTTTAGTTCTCTCCCATGTAAGCATATTTTCTTTGTGATACAAATACTTTGGAG
AGTTATGATGAAGAATGATAAAAAACATCAATTTTGGTTTCAGAGTGATTGAGAGTAAGAACCCAGAGT
GGCCAGTGGGTACGTACGTTGTTTCACCGTCTGGCTGGCGCTCCCACACCCGACTCACTGAGAAAGACT
TCAAGGCTGATAACTTTTTCCAGAAGCCCATGAAGCTGCCAGAGATGGGAGACTTGCCTAAGAGTCTTG
GCTTAGGGATTCTGGGCATGCCGG***GGTATGTCCAGTTATTAGTCTTGGGTGGTGTAAAAAAAAGAAAAG***
TCGAGACTAAGTTAAAGTATAGAGCTCATTTGGATTTGGAGTTGTTTATATGTTCTTGATTTTATTGTA
CATGGAAACCATATAATAATATATTTCTTAAAGAGTGTTAGTACATTTACAATATATGTATAAAGTTAA
ATTTTTTATCACAGGAATACAGCATATTTTGGATTTCTTGAAATCTGCCGACCGAAAGCAGGAGACACT
GTGTTGGTAAATGCTGCTGCTGGAGCAGTTGGAAGTGCTGTGATTCAGATAGCCAAAATCAAAGG***TTAG***
AACTAAAACTGTTTCTTTAGAGAAATGCTTAATTAAGAAACTGTCTCACACAGTCTGTTGGACACAATT
TATGAAAAGATGGATTACAGTGATAGTTGGGTTCATCAAAGCTTGTGTTAGAGAATAAGAATCTAATCG
CTGGATGTAATAAGCATTGGTTATAGAGTAAAGCAAATTTAAAACACTCAACCCCATGATAAGAAAAGA
AACTTTATGGAAAAAATTTGAATTTATAGTATCAGTGTAGCTTATGAATATTTTAGGCTTTATTTGTGA
ATGCATCATTTTTATATAAAGTTTCCCCCTCTAGGGTGTAAGGTAATTGCCTTTTCGGGATCTGATGAA
ACAGGTAGCGTGGACCAAAGACCTTGGGGCAGACCATGCCTTCAACTACAAAACAGCAAACATTGGACA
GTTCACTGCTGGGAGGCACTCAGTCAAGCAGCGCCTGATAAGATTAACTGCTTCTTTGAAAATGT***AAGT***
TTGGACACATTGTTTTTACCACAGACTTATTTTAATACAATTTAAAGATGAAATGACTCAGAATTTTAC
TGAATAACGTGGGTTATAGAATCTAAAATAGAATTGGTGAATTCTGTAATAGTTTTACATTCACATTTA
GAAAATGAGAAATTGTGGGTAAAGTATTTAAACTGGACCCATTTATCCAGGTAACCAAGCAACTGTAAG
GAGTTTTGCAGGTATACATAATCCATTAGTCTGAGTCACAAGATTGGTATTTTTCTTGTTTTAATAATA
TATAACAATTTTGCTTCTTGATGTTACAGTTTTGGAGAGTTTTTAATAATGTGGCTATATGAATTTATT
TAATAGAATTAACAGATATTGAATATACTTAGGTTGATATGGGAGGAAAAAGTAAGTTCTACCCTTGTT
TATTACTTGTCTTCTTTATTTTTATTTTTGTTCATTTTAAGGTTGGAGGACAGTTCACTGCTGAAGCTC
TACCACACATGGCTGAGAGGGGCAGAATGGCAATTTGTGGTGTGATTTCAGAGTACAATGATGACGACA
AGGATTCTGGTGTGGTGGCAATGACTAGT***AGGTATTTTGTTGTTTCATTGTCAGTGTAATGACTTTTTT***
CCTCCCAGTAGGTTTCTTTTCTGTATATGAATGTCTGTAGAATATATAATTAAAGTTACTTATATGATT
AGGCCTTTTTAGTAGATAATGTACATTTATTTATAGTGCAGTTTGAAAATTAGAATAGTGGTCTTATTT
ATTTTATTAAATTCTGGCTCCCCTGTGTTGATAGAATCATATGAGGATAATCCTTTGTTATGGTAGGGT
AAAATGTTTCATCACTCGATGTTCCAGGTCCGTTGAATCCAGCTACGATACTTTGGAAGCAACTTGCAG
TGGAAGGCTTCATTGTTACAAGGTGGTTTGATCGCTGGATGGAAGGAATTGATCAGATGAAGCAGTGGA
TCAGTGAGGGTAAAGTCAAGTACAAGGAAACTGTTGTAAAGGGATTTGACAAAATGCCTGAAGCATTCA
TTGGCCTTTTCCTCGGTGAAAATACAGGCAAGGCCATTGTAGCAGTGTGAGGAACCTTGTGTGGAATGA
AAAGATCAAAACAATCTACTTTATTTTCATATTTTCTTAATTTAATAGGGAACCACAACCTAAAAATAT

AACTTGATAAATTCATCTGTTGAAACTTATTATGGTGGAAAATTGTGCAGCATTGGCAAACATTTTCTT
TGTTTTGGATGTTAAAAATATTTCTATAATAGCTGCATGCACTGGAAAATACTCTTACTTTACAAGTAG
GCCAGTACGTCTATTTCACATCTTCATTTTTGAAGATGCTATTTCATATATATGCAGACCTGTTTATTG
CTTTTCCATTCTGTAATTTATCTATCTTCCATTCTTCTAAATTATCCTTGTTGAGGCATTGCCAACTTC
AAATAAAGTGCAAACCTCGTGCCGAATTCGGCACGAGGGTCAAATACTTTCTTTGTTGTGTTGGCCCTG
GGGTTAGCCCTTGTTGCAGCCAAAGAGGACATCCGAAGCGAACGCAGCGTTACTGCAGACAACCACCCT
TACTCGAAGCTATGTGAGAAACAACCTGACAAATTCATTTGCGGCAACTGCAAGACCTTGATCCAGTGT
GTGAAGGGACAGGCCTTCACTCGCCACTGCATTGAGGACCACTTCTGTTCGGAGAGGCCCCAGTTTGGC
GGTGGTGTCTGCTACCCAAATGAACCTCTGGACTGCACCTGCGTGAAGGCCAACGAGTTCCGAGTGGAC
CCCTACGACTCTCAGAGGTTCTTCTCTTGCAAGGCTGTTGGCTCCACCCCCGAGAACTACAAGTGCCCA
GATGGTATGGTCTTCGATGAAGGCTCGGCACAGTGCCAGACGGCAAGTGGCCTGCCTCCATGCGTGGTG
GCGGGTACTTTTGCCAACCCAAGCAACTGCAGTGAATACTACTCGTGCATTAGCCTGCGCAGTGGATGG
CTGCAGAAGTCGTTCATGTGCACCAATGACATGATGTACAACGAACAAAAGGATGCATGTGAAGATCCG
TGCATATACCAGTTCGTGTGCCAGCAGGAAGGCCGTTACCCTGACCTTCTGAACAAGCAGAATTACTTT
GAATGCTACATGCTTGGTGGCGTGTTGCAGCAGCTACGTTACAGCTGCCCTGAGAGCTACAGGTGGGAT
ATCGTGTCTCCAGGTGTGGGCCAGTGCGTGGAGGACCATGGAGATAAAGATTCTAACTATGCCTTCGGG
CAGTGCGACATCCCTGACAACCTGTGTCCTGGCCCGTGAGAACTTGTTTACTAATAAGGGCAACCCTGT
CGCAACACCACTAAATGGAATAGGGGCCAGACCTCGACTTCTGTCTTAAATAAATGTTGCTTTTTAAAA
AAAAAAAAAAAAAAAAAA

Figure 3.6 The partial genomic DNA sequences of *NADP-dependent leukotriene B4 12-hydroxydehydrogenase (LTB4DH)*. Start and stop codons are illustrated in boldface and underlined. Introns are boldfaced, italicized and underlined. The 5′UTR is highlighted.

3.2.2 *Nuclear autoantigenic sperm protein (NASP)*

Genome walking analysis was also applied to isolate the 5′UTR of *NASP*. The discrete bands were obtained from the secondary PCR using the template from the *Sca* I mini-libraries (Figure 3.7). A fragment of 750 bp was selected, cloned and sequenced. The 5′UTR of *NASP* of 670 bp was obtained (Figure 3.8). Nucleotide sequences of all characterized fragments were assembled and illustrated by Figure 3.9.

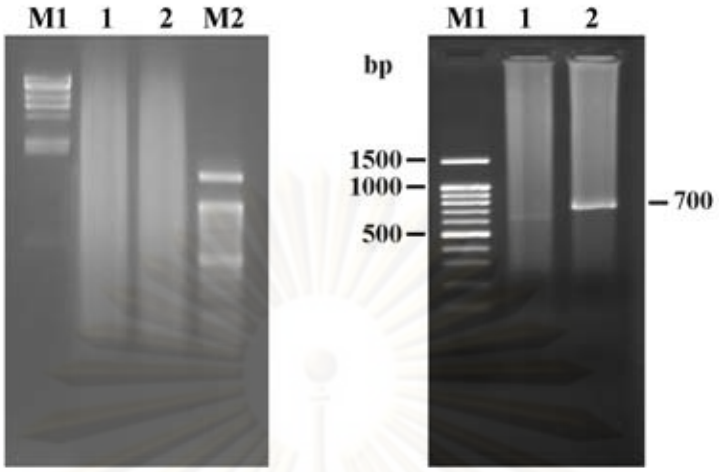


Figure 3.7 The primary (A) and secondary (B) PCR products of *nuclear autoantigenic sperm protein* (*NASP*). The *Stu* I (Lane 1) and *Sca* I mini-libraries (Lane 2) were amplified with the reverse gene-specific primers and the adapter primers (AP1 and AP2). λ-*Hin*d III and 100 bp DNA ladder (Lanes M1 and M2) were used as the DNA markers.

CGACCAGCCCGGTCTGGTACTCAATTTATACCAAAATATCGTAAAAGTATTACATATGGAATAATATTA
TATACAATATATGAAAATAATCGATACATATGTAATTCAGAGTGATTTTCGATACCGTTATACTTTTTG
CCAACAACAATGACCATTCATCACAGGTTTAAATCCCTCGCCAGAGCTCATATGGCGGTCGAAAATAAC
GAATCAATCCTAGTTTTCTATCTTTTCAAATTATAGGATAAATTGGTTTATGAAATATATGAAAAAATT
AAATGTAGCATCTAAAATAAATACAAGTTGCAAAAATAATTGCCACAGACTCCTAATCACAGATACGGT
CAACAGCGAACGTCACACAATAATCACCCCTTACAAACGTAGTGAACCATAAAAAGCAAACTGTGTAAT
AGAAACTAACTTATGAATATACATGGAATCCAAATCTCATCCTAAAAGTAACATAAATGAGAAAAAATA
ACGTCAAAAAATGGATGTAGACACGATGACGTCACAGTGAGGGCGGTGGGGATCGCGCGGGAAAGAAGC
GTCAATGGCGTCCAGTGTCTCGGCGGGTCCACACTCTTTTCCGGCCTTTGTACGCTCCACACGCGCAAC
CCCGTCACCCTCGTTTGTTGCTTCCGAACGTTGCTAATAGTCGACCACG**ATG**TCCGAGTCGCCGGTAAA
AGC*TGCCGAGACCAGTGCCAGCC*

Figure 3.8 The 5′ untranslated region (5′UTR) of *nuclear autoantigenic sperm protein*. The putative start codon is boldfaced and underlined. Primers used for genome walking are italicized and underlined. Overlapped sequences between the genome walking fragment and full length cDNA of *NASP* is highlighted. A 670 bp fragment was successfully identified.

CGACCAGCCCGGTCTGGTACTCAATTTATACCAAAATATCGTAAAAGTATTACATATGGAATAATATTT
ATACAATATATGAAAATAATCGATACATATGTAATTCAGAGTGATTTTCGATACCGTTATACTTTTTGC
CAACAACAATGACCATTCATCACAGGTTTAAATCCCTCGCCAGAGCTCATATGGCGGTCGAAAATAACG
AATCAATCCTAGTTTTCTATCTTTTCAAATTATAGGATAAATTGGTTTATGAAATATATGAAAAAATTA
AATGTAGCATCTAAAATAAATACAAGTTGCAAAAATAATTGCCACAGACTCCTAATCACAGATACGGTC
AACAGCGAACGTCACACAATAATCACCCCTTACAAACGTAGTGAACCATAAAAAGCAAACTGTGTAATA
GAAACTAACTTATGAATATACATGGAATCCAAATCTCATCCTAAAAGTAACATAAATGAGAAAAAATAA
CGTCAAAAAATGGATGTAGACACGATGACGTCACAGTGAGGGCGGTGGGGATCGCGCGGGAAAGAAGCG
TCAATGGCGTCCAGTGTCTCGGCGGGTCCACACTCTTTTCCGGCCTTTGTACGCTCCACACGCGCAACC
CCGTCACCCTCGTTTGTTGCTTCCGAACGTTGCTAATAGTCGACCACG**ATG**TCCGAGTCGCCGGTAAAA
GCTGCCGAGACCAGTGCCAGCCCCAAGAGCTCGCCGTCCAAGAAAGAGAT***GTGAGTTTGCCGTAGTGGT***
TTCATGCCTTTGTGTATTTGGCGGAATAAGGTTCGTCAATTTTTTTAAAAATCTAGAACTATTTTGGTA
GCACCTTCGAAATCTCGCTGCAGTGTAGTTGCAAGCTGCGGTGGAGAAGGTTGGCAACGGAGGCCTCGG
ATCTAATTTGTTTTGAGCTATTAGCGGCTTTTCCATAGACTTCACAAGTCTTTGTGAAAACGAAATTGA
TTTGATTTAGCTTTTGCATAAACCGAGAATTTCTTGCACTTTATGATTGTTAGTAACTTAGGTTATAGT
ACATTCTACGTAATGAGCACAAGTATTACTCTTGCATGTTTTAAATGTCTTACAATTTTATATTACCCA
TTTAGACGGTTAGGAACTTCCCACTATATAACCGTCCATGATACATCTGACTCTCTGCAGTTAGATTTG
CGTTTCCTTGGACTTCAACATCTATGCAGACACCAACTCTTGCTATCATTTTCTGTAGCCTTTGGGATT
CTGGAAAAATGGTAATAAACTTCCATATGTTCTGGCAAGACGGGATCACTTCCTTATGGGAATTGCAAA
TTACAAGCAATTGTATTGCCAGAACACTGTTTAGTCCAGCAATGTCAAAAGTGCAATAAGTTTCAACAG
GGAAGCCTTTGTAGATGGCAGTAAGTTATTCATTTTGCAAGACCATAGCCATTGAACACTTCTTAAGTT
GCACTTGGGTGTCAGGATTTGATAAGTTGATTTTTTTTTTTTTTTTTTTAATGAATTCTTGAACAATTA
CCAAATTCTTTGATTCAACCATTGATATTGATTCAACCATTGATATTGATTCAACCATTGATATTGCAT
TTGACCCATATAAACTCAAGGCTTTGATGATAAAGGGTATAAAGTTGACAAAGTGCATGGCCTCCTCCT
GTTGTGACCATACACTAATACCTTCAAAGAATAGCCCCTGTGATAAGTTAAGGAAATGTGAATACAGGT
ATCAAGTGAGAATGGTATAGATATTTTGAAATCTTAATATATAAAAGTATGGGAATAGTATTGCTTGTC
CCCCCCCCCCCCATGTGGCCTTTTGACTTGAATGTGGCATGTGGTGCCATGACTGCATAACATGCAGATT
TATAACCTAAAAATTTATCTTGATAAAATTATTGAATTAAAATACCCCCCCCCTTTCCAAGGTATTCCT
ATGGTATTCGTGTAGACTTTCCAATTGACATGAATAATCGTAAATCCATTATCACCACTTAAAACATGG
ATGTAAGTACTTTAAGTTCACTTACTGGTTTGGTAATTAATAACTTTATTGTTGCAGTGATACAGCAAC
CCAAGCTTTAAATCACTTTGCTCAGGGCAAGAGACACTTGGTTGTTGGTGACATTTCATCTGCAGTTAA
TTCTTTGCAGGAA***GTTGTCATTCTACGAATTGTATTTACTCCAACCAATGCTTTGGAAGACTCAGTTTC***
AACAGAAATAAAATTAAACCCTAATTTTCATTTTAGAAAGTCCTGTTGTATATTAATCAGTTTCTCGAT
TACTAAAGTTTCACAATTAACTAAGTGATGCTTAAAATTGCAGGAGGCATGTAGACTACTAGCAGAGCA
ATACGGTGAAACTGCTCCCGAGTGTGGTGATGCTTATTTCTACTATGGCCGTGCACTGCTTGAAATGGC
ACGCATGGAGAACGGAGTCTTAGGAAATGCTTTGGATGGAGTTCCCGATGGAGAGGACATGGACAATTC
CCAGGTAGAAAATCCTGAAAAAATGACAGAGGATGAGAAGAACGAGGTAACAGAACAGGTTGGGAAGGC
ATTGGAAGAGAACTTTAAAGATCTTGAGGATGTGTCAAAAAGTAAGTCGGCACAGCAGAATGGAGATGC
AAAGGCAAAGGCAGAAGAGTCTTCAGGTGTTGAGGAGGCTAANATGGATGTAGATTCAGCTGGAGTGTC
AGAATCAAAGGTGAAGATGGAGGCGAGAGGCAAGAGGATAAGTAGAAAGTCGGAAGGGGAGGAAAAGAG
TAAAGAGGAAACCTCGGACACTGATGGCACCACCACTTCCAAAGTAGAGGCTAGCTCAGTAGATAGTGA
AAAGGTAGACAAGGAAATCAAGCCTGAGAAAAAGGAAGTTGTGGATACCAAAGATAGTTCCAAAGAGGA
GGCAGAGGAATCCGAAAAGGTGACGGAGGAGAAGGTTGAGGCTAAGGAGGAAGAAGGGAAAACCACTGA
GAAGGGAGAGGGAGAGAAGGAAAAGGGATCAGGAGACACTAAAGATGAAAAGGGAAAGGAAGATGCCAA
AGTGGAAGAGGAGAATGTAAAAACTGAAGCAAAGGAAGAGAAATGGAAACTGATTCGTTGTGAGAAGGA
AGGCAGTTCCAATGGAGAGAGGAAGAGGAAGCGGATGGCGATGGTGATGGTGATGGCGAAGGCGAGGGA
GAAGATTCTCAAGAAGATTCCCAGGATGAAGGTGAAAAGGA***GTGCCAGCCAAGAAGAAGGCAAGTACTA***
CTTATGCTTGGTATGGTATTTGGAGCCCCCATTATTTTAAAAGCAAGATATTAGGGCAGGTTGGGCACT
TGCTCGAGACAATAATTTTTTTTACTCTTGTCTTCTGCAAGAAACTGAAAACGTGTAGGAAGGAACAATT
TAAATCTTTTGTTGAATTCTTGCCTTAATTGCTTTTCTTAATCAGTCCTGCCTGTATAAACTGATAGGC
AGAGATCCTTTCTCTGTAAACATGGTTTAGGCAATGAGGAATTGAAGGCAACAAGTCTTTAAATTAAAT
TATCTTGAGGCAAAACTTGATTCAGTCCTCTCTCTCCCTCTCTCCATGTTACCAGAGGGTTTTTATACA
AAACAGTGTAAATGTACACTAGTTAATCTCTGCTCAGCAGTAGATGAAACTGAAATATTGCTAAAGGTA
GATTTGAAAAGGTTGGTACTGTTCCAAAGCTCAAATAGATGAATCATAATGTTTAAATAGGGAGTTGAT
ATGGCTAGTCCCACACATCCTGCCAGTATTAGCAATGCAGACCTTGGCTTGCTTGTAGCATCATGATTA
TACTTTACTCTGAATTATAAAGGGCTAAATATCAATATTAGATATTTTGAAGCCTGTGGATCCAATAAC
CTCGAGCAACTAGATCCACCAGAGACCAGGTACCCTAGATGCACAATGGACAAGTGCCCCCCTCCCCCA
AATGGCACCCCTGTCAGGGATGCCAATATCAGGTTGAGGTCACATTAAAGACCTATCTTTATTTCCCTA
CAAAACTCTGCCACAGTATAGGGCAATGTTACCAAATCCCCAATCTGGACAAGACTAAAGATCTAAATG
TTAGAAATTGAGGGCTAAGTTGTTTATAATGCCTTTTCAAAATAAAACTGAAGAGAAATTTTCTGCAGG
GGAAAAGACTGAGGAGGATGAGGAAGAAGTATCAAACTTGCAGCTCTCTTGGGAGATGTTGGAATTGGC

AAAGGTTATCTATCAGAA***GTAAGTAACTTTTTACAGTAGTGTGAAAATGGGTGTTTTAATAACTACAGT***
GGTTGGTTTAATCTTGTAACCTTATTTTTTTCATCTTAGGCAACAGGATGACAACCCAGAGATGGCTAA
GAAAGTTGCCCAAGTGTACCTAAAACTTGGAGAAGTAGGCTTGGAGAGCGAAAATTATTCACAGGGTAT
TGAAGATTTCAAACAGTGTCTGCAAATACAGGAGAAAATTCTTGAGGAAGACAACAGGTGTTTGGCAGA
AACCCATTACCAGCTTGGTGTAGCACACTCCTTCTGCTCGAAATTAACCCTCACTAAAGGGAACAAAAG
CTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTCGGCACG
AGGAGAAGAGCAAAGAAAGAAAGATGCTGCAGAGAGACCTGATCCATTCTACACCGAAGAAGGCGAGAT
TGAAGAATTGAACAAATTGTTACCAGAGATGAAGGAAAAGGTTACAGATATGGAGGAAATGAAGAAAGA
CAGCAAGGACAGACTCCAGAAAGCGGCAAAGGAAGCATTCATGGCAAATGCAATTGGTGGCACCTCCAA
AGCTGGATCTTCATCACAAACTGGATTTGATGCACCTTCAAGTTCTACATCCTCAACCCCCACAGAAAT
AAAGGCTTCCAACATTACTCATCTTGTAAGAAAGAAGCAGAGGAAACCTGAGGATGAGGTTGAGGGAGA
AGAGGTGAAAAAGGCAAAAGGCGAGAATGGGGAACCACATGGAACTGCTAATGGACCACCAATGGCACC
AATGGGCACTCTGAAACCATGGAAACAGAGGAAAAGGATACCCCAACAAATGGGGCAAGCACTGAAGAA
T**TAA**AGGAGAAAGCAGCTGAAGAGATGAAGAAAAAGACGGATATGATCACTGGGAAAACTGAGGCAGCA
TCCTAGATAGTGTATAATGCCTTTAAAACACTGCCAGTTATTGGACTAAGAACTTTTGTATTTTATTGT
AATGGAAGCCATTTTATTACCAGTGAACTGTGTTTTAAGATTTACATAATTTATTAAAAATGAAATGTA
AAAAAAAAAAAAAAAAAA

Figure 3.9 The partial genomic DNA sequences of *nuclear autoantigenic sperm protein (NASP)*. Start and stop codons are illustrated in boldface and underlined. Introns are boldfaced, italicized and underlined. The 5′UTR is highlighted.

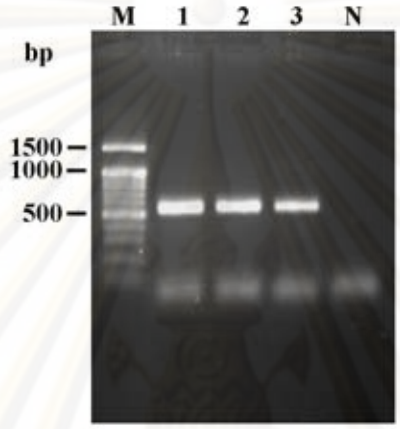
Genome walking analysis of other genes was also carried out. Nucleotide sequence of the amplified product of *dolichyl diphosphooligosaccharyl protein glycosyltransferase (DDPG), thyroid hormone binding protein (THBP), DEAD-box protein 52, endothelial cell growth factor (ECGF), aspartate amino transferase (AST), 3-oxoacid CoA transferase* and *cyclin dependent kinase 7 (Cdk7)* could not be assembled with the original cDNA sequence and did not match any sequence in the GenBank database. They were regarded as non-specific amplification genome walking products.

The exon/intron gene segments of *LTB4DH, DEAD-box protein 52* and *THBP* were further examined along with *insulin degrading enzymes* (*IDE*) and *epidermal growth factor* (*EGF*).

3.2.3 *DEAD-box protein 52*

A 581 bp fragment (Figure 3.10A) was amplified by primer Deadbox-3F/R was cloned and sequenced. One intron (282 bp) was found within the amplification product along with the partial sequences of 2 exons (84 and 215 bp; Figure 3.10B).

A.



B

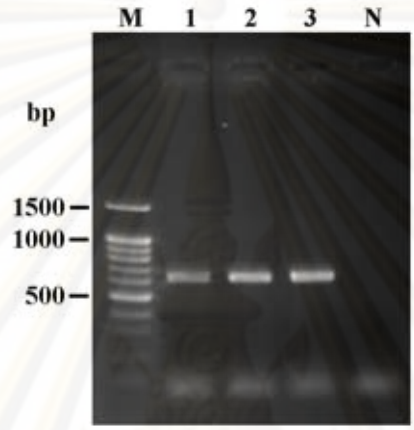
GAGCGAGCGAAAGAACTGTTCAAGGAGCTGATTTACGACAACCTGATGGTGGATGCCATTCATGCAGA
TCGAACACAGCTGCAG*GTTTGTGTTTGCATTTGTGTTAACTAACTGTGTTAACTGATGGATGTTCTAA*
AGCATTATTCATGGCTTTGGAAAGCTGTTGTAAACAAAATATCACATAATGTGATTAACATTCTGTAG
AGTTCTTTCTTATATTTTAGAAGGCTTATTGCAAGAGATCATTTTTATATTTGTTTGTAAAACTTGGA
AAGATAATAGCACTCCTCTTTATTATAACTCATTTTCTCAGCATCTTTAAGAATTAGACACCTTACTT
*ACCTTACCCTCTTTCATACATTTCAG*CGTGATAATGTTGTACGGGCATTCAGAGAAAGAAAAGTCTGG
GTTCTGATATGCACAGAACTCATGGCCCGTGGTATTGACTTCAAGGGGGTGAACCTTGTCATCAACTA
TGACTTCCCTCCTTCAGCAATCAGTTACATTCACAGAGTAGGAAGGACTGGCAGAGCTCATCATCAAG
GAAGAGCGGTCACATTCT**GGACAATGGCAGATAAGCC**

Figure 3.10 (A) A 1.5% ethidium bromide-stained agarose gel showing the amplification product of *DEAD-box protein 52* using primers Deadbox-3F/R. Lane M = 100 bp DNA ladder, lanes 1-3 = the amplified product from genomic DNA of *P. monodon* and lane N = negative control (without genomic DNA template). (B) The partial genomic DNA sequence of *DEAD-box protein 52* amplified from primers Deadbox-3F/R. An intron is italicized and underlined. Primer sequences are illustrated in boldface and underlined.

3.2.4 *Thyroid hormone binding protein (THBP)*

A 595 bp fragment (Figure 3.11A) amplified by primer RT-THBP-F/R was cloned and sequenced. One intron (245 bp) was found within the amplification product along with the partial sequence of 2 exons (134 and 216 bp; Figure 3.11B).

A



B

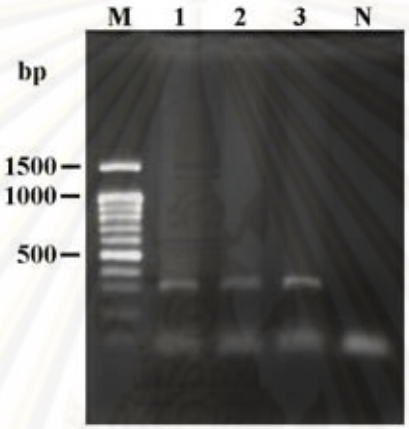
TACAAGCCCGATACTTATGACCTCAGCGAGAGTGGGCTTACCGGCTTCGTCAAGAGTTTCCTTGATGGA
AAGCTGAAGCAGCACCTTCTGTCTCAGGATCTGCCAGAAGACTGGGATAAAGAGCCAGTCAAGGT*AATA*
TCTTTTTCAGGATGTTTGCAATATGGAACTGGATAGATGCATTTGGTATTAAAAGTAATCAGTTTGAAC
ATGTGTGAACAAGTAATAGATGGTGTATATTTTGTGGATTTGGTTTTGTGGTAACTAGGAGTCAACATG
AAAAAGATTTCGGGAGTTTTCATTTTACAAATATAAAGGTTGAAAAAACAAACTCAATGGCCAAATGAA
*ATTGGGGAATATAATATAATTGAACTATTCAGGT*TCTGGTGTCCTCCAACTTCGATGAAGTTGCCTTGA
ACAAGGAGAAGGATGTCCTTGTAGAGTTCTATGCACCATGGTGTGGACACTGCAAACAGCTGGCACCTA
TCTATGATCAGCTAGGTGAAAAATACAAGGACCACGATACTATTGTTGTGGCGAAAATGGATGCTACAG
TCAATGAGCTTGAGCATACGAAG**ATTCAGTCCTTCCCAACCTT**

Figure 3.11 (A) A 1.5% ethidium bromide-stained agarose gel showing the amplification product of *thyroid hormone binding protein* using primers RT-THBP-F/R. Lane M = 100 bp DNA ladder, lanes 1-3 = the amplified product from genomic DNA of *P. monodon* and lane N = negative control (without genomic DNA template). (B) The partial genomic DNA sequence of *thyroid hormone binding protein* amplified from primers RT-THBP-F/R. An intron is italicized and underlined. Primer sequences are illustrated in boldface and underlined.

3.2.5 *Insulin degrading enzyme (IDE)*

A 313 bp fragment (Figure 3.12A) was amplified by primer pairs RT-IDE-F/R was cloned and sequenced. One intron (110 bp) was found within the amplification product along with the partial sequence of 2 exons (106 and 97 bp; Figure 3.12B).

A



B

GGAGTGGGTGTTTGAAGAGTGTAAAGACCTCAGTGCCATGACATTCCGCTTTAAGAATAAGGAGCGTCG
CCAGAGTTACACTTGTGCTCTTGCTGAACAGCTGCAT*GTATGTGTAGGATCTTTTGGTGTATTTTATTG*
GTTTTATTTTGTGTTTACTCATAGGATCATGCAATTATTTACATGAAATTTGGTATGTAATGAAAATTC
*TCCTTTCAG*TATTACCCACTGGAAGAGGTGCTGTGTGGTGGATACCTGTTGAGTGAATACCAGCCACAT
CTTATTGACATGGT**TCTAACTCATCTTCAGCCTACCA**

Figure 3.12 The partial genomic DNA sequence of *insulin degrading enzyme* amplified from primer RT-IDE-F/R. An intron is italicized and underlined. Primer sequences are illustrated in boldface and underlined.

3.2.6 *Epidermal growth factor (EGF)*

A 266 bp fragment (Figure 3.13) was amplified by primer pairs RT-EGF-F/R. The expected size was observed indicating that the amplified product did not contain any exon.

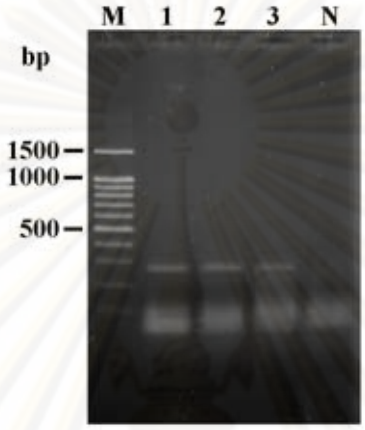


Figure 3.13 A 1.5% ethidium bromide-stained agarose gel showing the amplification product of *epidermal growth factor* using primer pairs RT-EGF-F/R. Lane M = 100 bp DNA ladder, lanes 1-3 = the amplified product from genomic DNA of *P. monodon* and lane N = negative control (without genomic DNA template).

3.3 Genotyping of commercially cultivated shrimp using microsatellite analysis

Broodstock-sized and juvenile *P. momodon* used for SNP analysis were collected from Broodstock Multiplication Center, Burapha University, Chanthaburi. Representative individuals of these samples were genotyped to evaluate whether each shrimp group was established from a single family or not.

3.3.1 Fourteen-month-old broodstock (BU14F)

Genotypes of 14-month-old shrimp were examined at *CUPmo02*, *CUPmo13* and *CUPmo15* loci. A total of 3, 4 and 3 alleles generating 4 (A = 6, B = 5, C = 4 and D = 4), 4 (A = 10, B = 3, C = 4 and D = 3) and 3 (A = 10, B = 2, and C = 8) genotypes were found from 20 individuals examined (Figures 3.14 - 3.16). Considering allelic segregation, this group of samples should be offspring of a full-sib family.

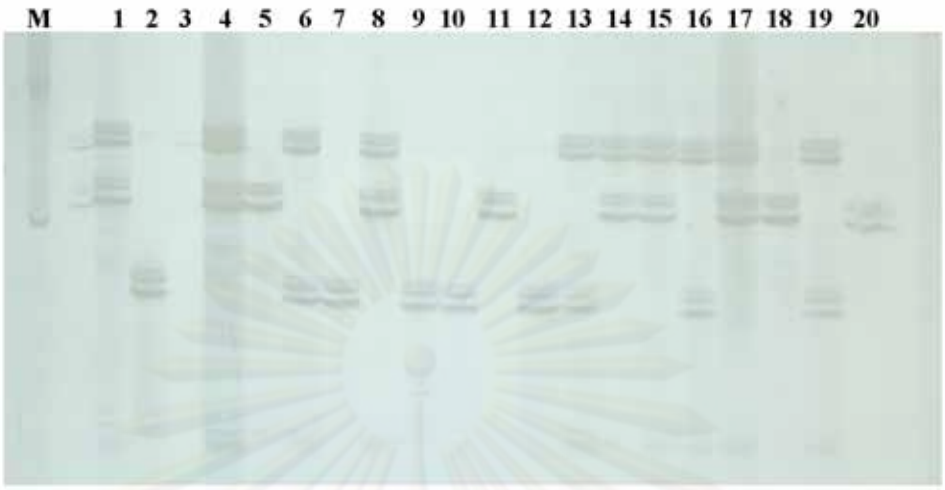


Figure 3.14 Microsatellite patterns of twenty individuals of 14-month-old shrimp (Lanes 1 - 20) cultured in the same pond at the *CUPmo02* locus. The PCR product was denatured, size-fractionated through 6% denaturing polyacrylamide gel and silver-stained. Lane M is a 100 bp DNA marker.

Figure 3.15 Microsatellite patterns of twenty individuals of 14-month-old shrimp (Lanes 1 - 20) cultured in the same pond at the *CUPmo13* locus. The PCR product was denatured, size-fractionated through 6% denaturing polyacrylamide gel and silver-stained.

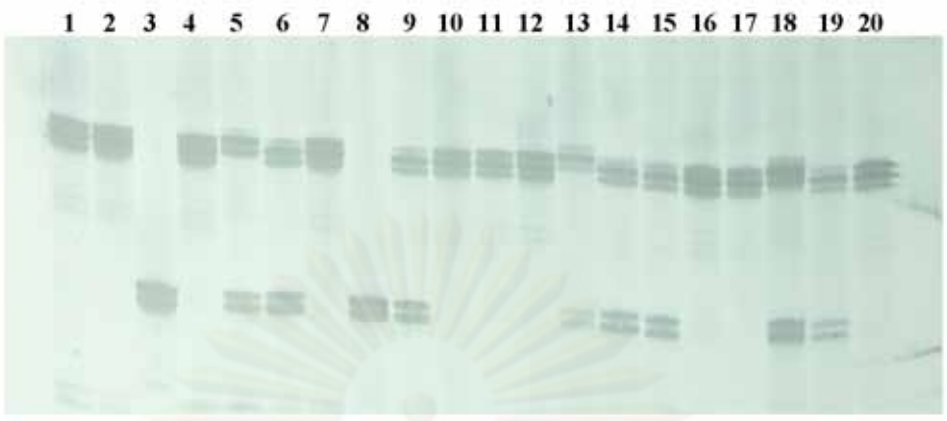


Figure 3.16 Microsatellite patterns of twenty individuals of 14-month-old shrimp (Lanes 1 - 20) cultured in the same pond at the *CUPmo15* locus. The PCR product was denatured, size-fractionated through 6% denaturing polyacrylamide gel and silver-stained.

3.3.2 Three-month-old juveniles (BUM03)

In addition, forty individuals of 3-month-old shrimp were genotyped at three loci. *CUPmo13* loci provided 3 alleles generating 4 (A = 6, B = 12, C = 3 and D = 4) genotypes (Figure 3.17). Four and seven alleles were found in *CUPmo02* and *CUPmo15* generating 3 and 15 genotypes, respectively (Figures 3.18 and 3.19). Results from allele segregation analysis suggested that juvenile shrimp should be originated from a few closely related families.

Figure 3.17 Microsatellite patterns of forty individuals of 3-month-old shrimp (Lanes 1 - 40) cultured in the same pond at the *CUPmo13* locus. The PCR product was denatured, size-fractionated through 6% denaturing polyacrylamide gel and silver-stained. Lane M is a 100 bp DNA marker.

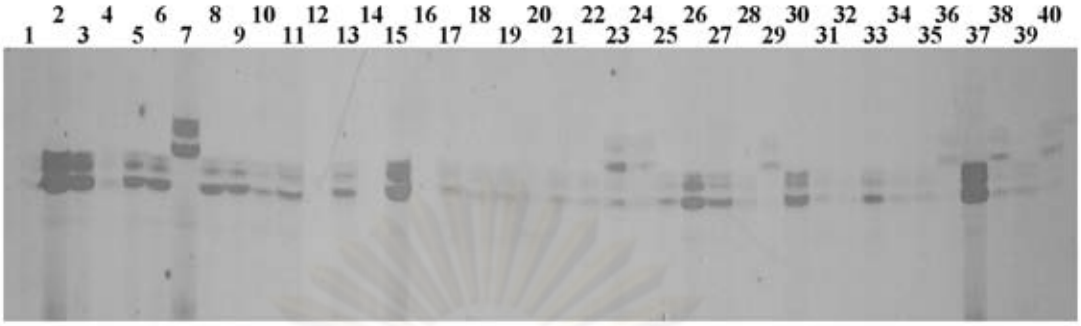


Figure 3.18 Microsatellite patterns of forty individuals of 3-month-old shrimp (Lanes 1 - 40) cultured in the same pond at the *CUPmo02* locus. The PCR product was denatured, size-fractionated through 6% denaturing polyacrylamide gel and silver-stained.

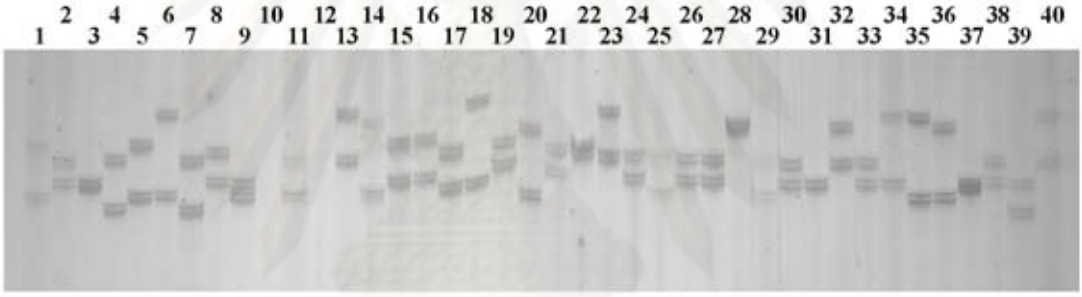


Figure 3.19 Microsatellite patterns of forty individuals of 3-month-old shrimp (Lanes 1 - 40) cultured in the same pond at the *CUPmo15* locus. The PCR product was denatured, size-fractionated through 6% denaturing polyacrylamide gel and silver-stained.

3.4 Identification of SNP in the 5′UTR and genomic sequences of the coding region of reproduction- and growth-related genes using SSCP analysis

3.4.1 Reproduction-related genes

In total, sixty-six individuals of 14-month-old shrimp (BU14F) were used to identify SNPs in the 5′UTR region of *NADP-dependent leukotriene B4 12-hydroxydehydrogenase (LTB4DH)* and *nuclear autoantigenic sperm protein (NASP)*

genes. A primer pair of each genes was designed and successfully generated the fragments of 592 and 316 bp in size against genomic DNA of *P. monodon*, respectively (Figures 3.20 and 3.21).

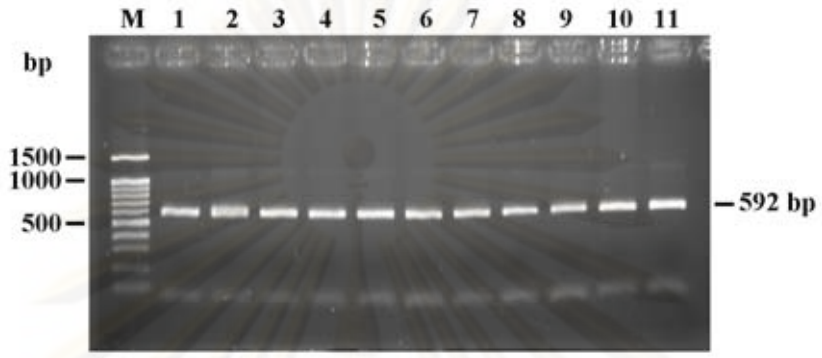


Figure 3.20 A 1.5% ethidium bromide-stained agarose gel showing the amplification product of *LTB4DH* against genomic DNA of different individuals of 14-month-old *P. monodon*. Lane M is a 100 bp DNA ladder.

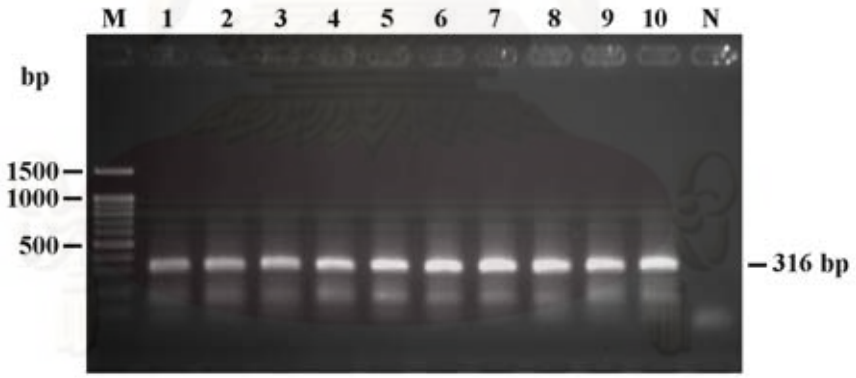


Figure 3.21 A 1.5% ethidium bromide-stained agarose gel showing the amplification product of *NASP* against genomic DNA of different individuals of 14-month-old *P. monodon* (Lanes 1 - 10). Lanes M and N are a 100 bp DNA ladder and the negative control (without genomic DNA template), respectively.

A monomorphic pattern was observed from that of *NASP* (Figure 3.22) when the amplification product of this reproduction-related gene was further analyzed by SSCP.

In contrast, polymorphic SSCP genotypes were observed from that of *LTB4DH* (Figure 3.23). Two SSCP genotypes (A and B) were found in 37 and 29 individuals accounting for 56.06 and 43.84% of examined specimens (N = 66, the average body weight = 87.32 ± 12.47 g), respectively. Association analysis between SSCP genotypes of *LTB4DH* and the GSI values of broodstock-sized shrimp was examined but the results was not significant (t value = 1.432, df = 64 and P = 0.157).

Therefore, the correlation between SSCP genotypes of this gene (genotype) and its expression levels in ovaries of cultured 14-month-old shrimp (phenotype) was subsequently examined (see below).

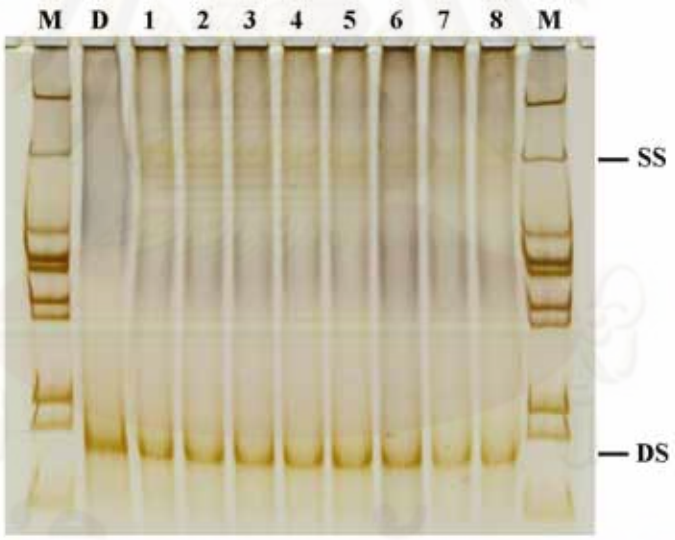


Figure 3.22 SSCP pattern of the 5′UTR of *NASP* amplified from different individuals of 14-month-old shrimp (Lanes 1 - 8). Lanes M and D are a 100 bp DNA marker and non-denaturing PCR product (double strand control), respectively.

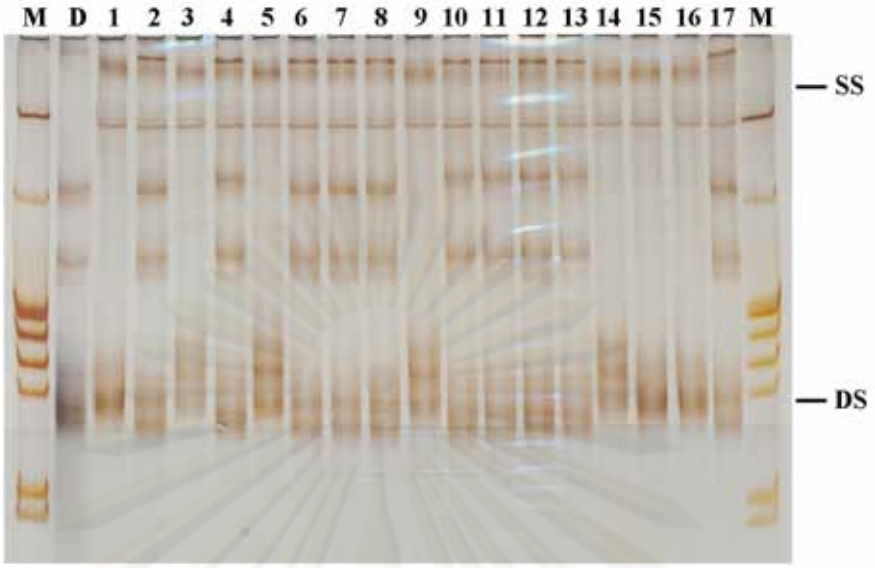


Figure 3.23 SSCP patterns of the 5′UTR of *LTB4DH* amplified from different individuals of 14-month-old shrimp (Lanes 1 - 17). Two polymorphic genotypes (A and B) were obtained. Lanes M and D are a 100 bp DNA marker and non-denaturing PCR product (double strand control). Lanes 2 and 5 are examples of genotypes A and B, respectively.

3.4.2 Growth related genes

For identification of SNPs in candidate growth-related genes, primers designed from the coding region of *thyroid hormone binding protein (THBP)*, *epidermal growth factor (EGF)*, *dolichyl diphosphooligosaccharyl protein glycotransferase (DDPG)* and *insulin degrading enzyme (IDE)* were used for amplifying the genomic fragments of these genes. The PCR product of 595, 240, 376 and 203 bp in size were obtained, respectively (Figures 3.24 - 3.27).

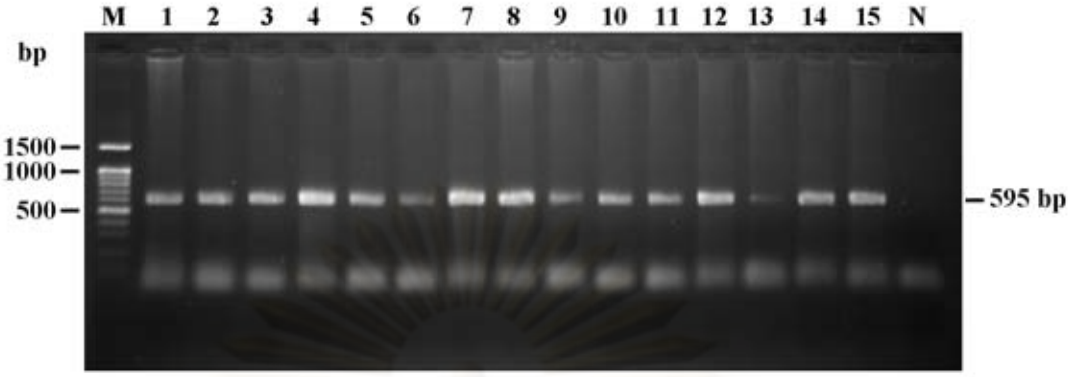


Figure 3.24 A 1.5% ethidium bromide-stained agarose gel showing the amplification product of *THBP* against genomic DNA of different individuals of 3-month-old *P. monodon* (Lanes 1 – 15). Lanes M and N are a 100 bp DNA ladder and the negative control (without genomic DNA template), respectively.

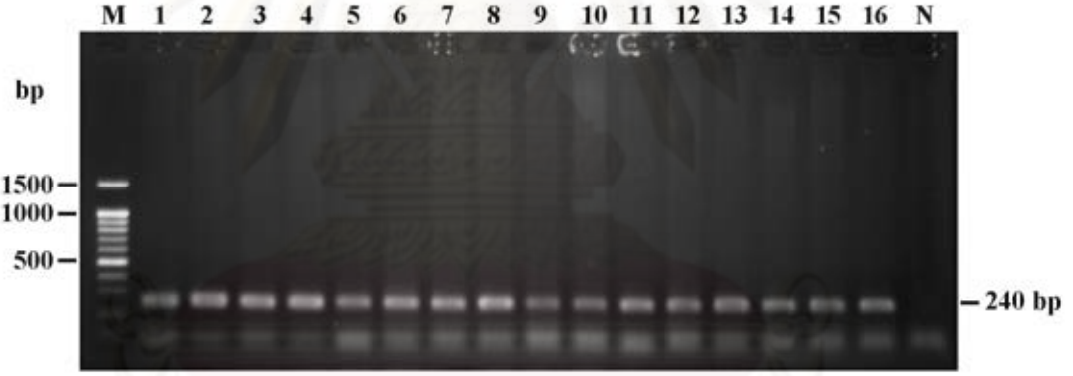


Figure 3.25 A 1.5% ethidium bromide-stained agarose gel showing the amplification product of *EGF* against genomic DNA of different individuals of 3-month-old *P. monodon* (Lanes 1 – 16). Lanes M and N are a 100 bp DNA ladder and the negative control (without genomic DNA template), respectively.

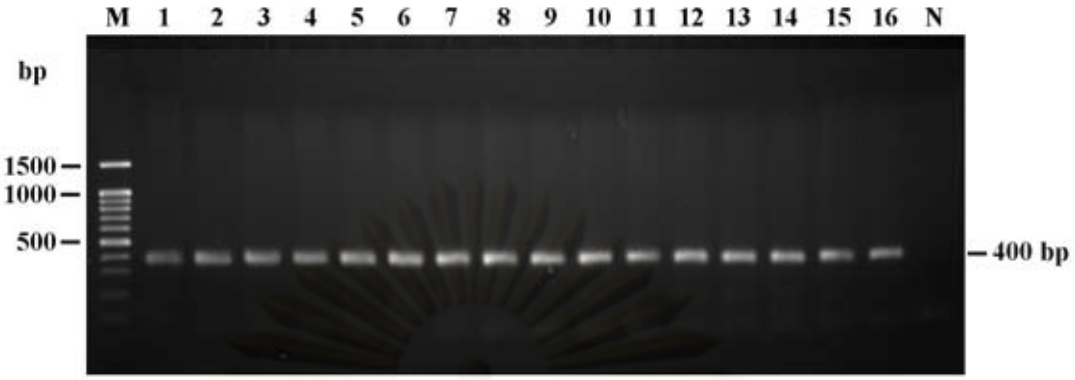


Figure 3.26 A 1.5% ethidium bromide-stained agarose gel showing the amplification product of *DDPG* against genomic DNA of different individuals of 3-month-old *P. monodon* (Lanes 1 – 16). Lanes M and N are a 100 bp DNA ladder and the negative control (without genomic DNA template), respectively.

Figure 3.27 A 1.5% ethidium bromide-stained agarose gel showing the amplification product of *IDE* against genomic DNA of different individuals of 3-month-old *P. monodon* (Lanes 1 – 12). Lane M are a 100 bp DNA ladder.

Subsequently, PCR-SSCP of these genes were then examined using specimens exhibiting approximately 16% of the body weight from the top (N = 54) and the bottom (N = 56) of three-month-old *P. monodon* (BUM03).

Only one SSCP genotype was found from the amplified genomic DNA segment of *DDPG* and *IDE* (Figures 3.28 and 3.29). Accordingly, these candidate growth-related genes were not analyzed further to covering the entire sample set but

the expression level of these genes in hepatopancreas of fast and slow-growing cultured juveniles (3-month-old) was subsequently examined.

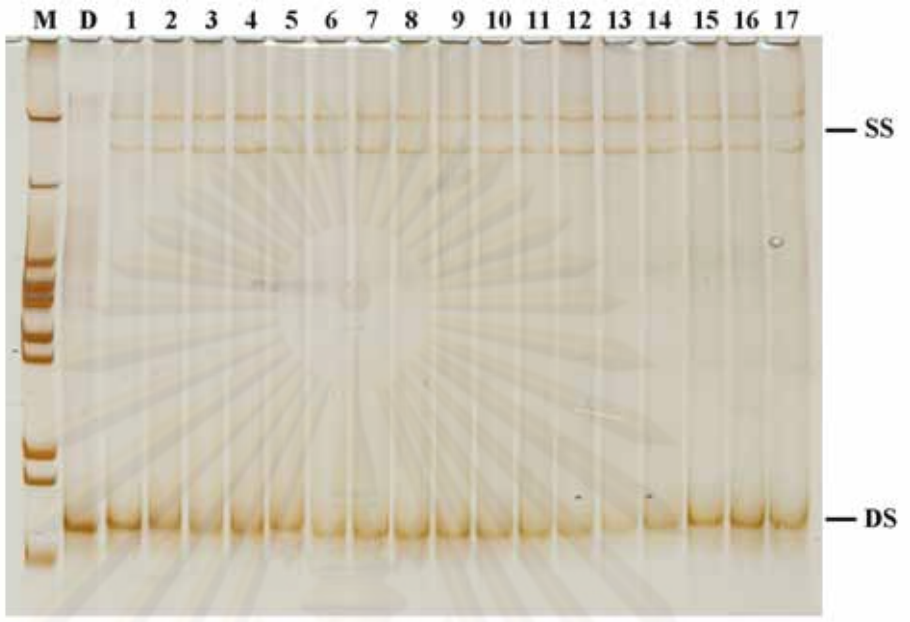


Figure 3.28 SSCP pattern of *DDPG* amplified from genomic DNA of different individuals of 3-month-old *P. monodon* (Lanes 1 - 17) Lanes M and D are a 100 bp DNA marker and non-denaturing PCR product (double strand control), respectively.

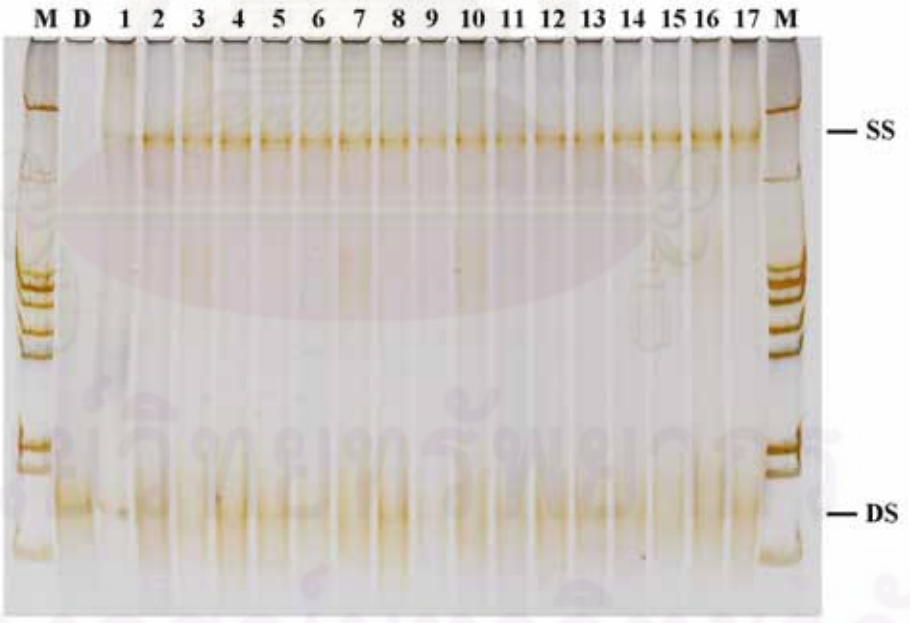


Figure 3.29 SSCP pattern of *IDE* amplified from genomic DNA of different individuals of 3-month-old *P. monodon* (Lanes 1 - 17). A monomorphic genotype was obtained. Lanes M and D are a 100 bp DNA marker and non-denaturing PCR product (double strand control), respectively.

In contrast, four SSCP genotypes (A, B, C and D) were observed from the partial genomic fragment of *THBP* (Figure 3.30). These SSCP genotypes were found in 113, 4, 89 and 134 individuals accounting for 33.24, 1.18, 26.18 and 39.40% of examined specimens, respectively.

Likewise, two polymorphic SSCP genotypes (A and B) were also observed from the partial genomic sequence of *EGF* (Figure 3.31). These genotypes were found in 266 and 74 individuals accounting for 78.24 and 21.76% of examined specimens, respectively.

PCR-SSCP of *THBP* and *EGF* was further examined using the entire sample set (340 individuals). In addition, the correlation between SSCP genotypes of these genes and their expression level in hepatopancreas of cultured 3-month-old shrimp was also examined.

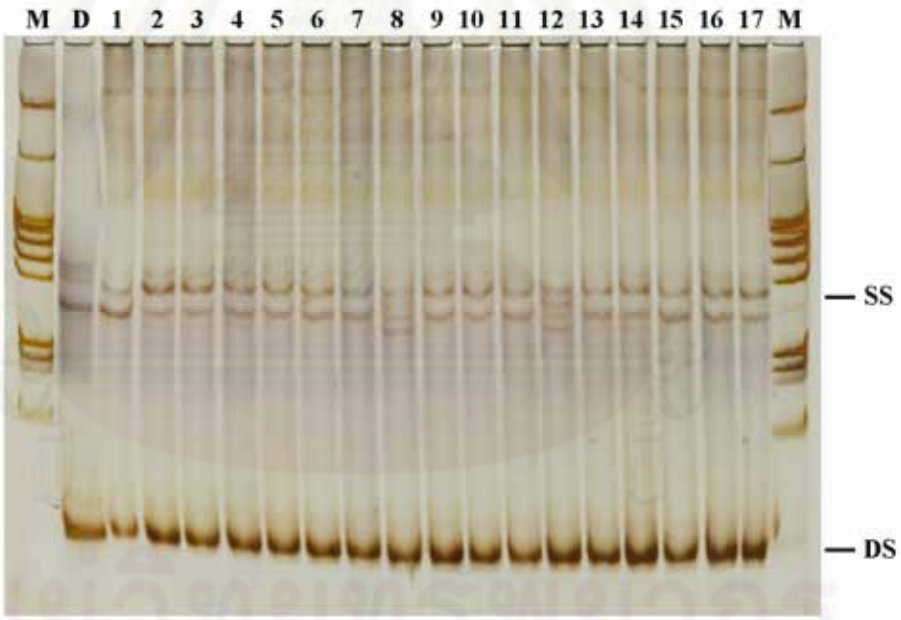


Figure 3.30 SSCP patterns of *THBP* amplified from genomic DNA of different individuals of 3-month-old *P. monodon* (Lanes 1 - 17). Four polymorphic SSCP genotypes (A, B, C and D) were found. Lanes M and D are a 100 bp DNA marker and non-denaturing PCR product (double strand control). Lanes 1, 2, 3 and 5 are examples of genotypes A, B, C and D, respectively.

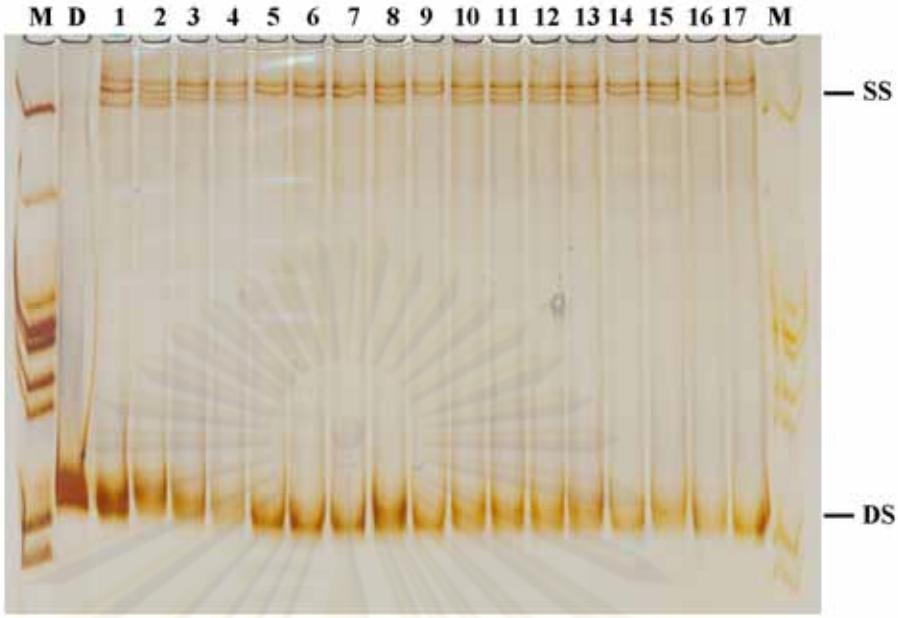


Figure 3.31 SSCP patterns of *EGF* amplified from genomic DNA of different individuals of 3-month-old *P. monodon* (Lanes 1 - 17). Two polymorphic SSCP genotypes were observed. Lanes M and D are a 100 bp DNA marker and non-denaturing PCR product (double strand control). Lanes 2 and 8 are examples of genotypes A and B, respectively.

Table 3.1 Number of SSCP genotype of growth- and reproduction-related genes examined using 14-month-old (BU14F) and 3-month-old (BUM03) *P. monodon*

Gene	No. of SSCP genotypes
Nuclear autoantigenic sperm protein	1 (A = 66)
NADP-dependent leukotriene B4 12-hydroxydehydrogenase	2 (A = 37, B = 29)
Dolichyl diphosphooligosaccharyl protein glycotransferase	1 (A = 110)
Insulin degrading enzyme	1 (A = 110)
Thyroid hormone binding protein	4 (A = 113, B = 4, C = 89, D = 134)
Epidermal growth factor	2 (A = 266, B = 74)

3.5 Identification and characterization of SNP by DNA sequencing of the PCR product

3.5.1 Identification of SNPs in *LTB4DH* and *THBP* gene segments by direct sequencing

Direct sequencing of the amplification fragment was carried out to examine the accuracy of SSCP analysis in comparison to DNA sequencing on detection of polymorphic nucleotides of the same gene fragments. Accordingly, the amplified product of an individual representing each SSCP genotype of *LTB4DH* and *THBP* was gel-eluted and direct sequenced. Nucleotide sequences of a particular gene were multiple-aligned.

Two SSCP patterns were found from SSCP analysis of the *LTB4DH* gene segment (592 bp) against genomic DNA of BU14F ($N = 66$). Representatives of these genotypes: LTB4DH-A ($N = 37$) and LTB4DH-B ($N = 29$) were direct sequenced and 545 bp sequences were obtained after multiple alignments. As can be seen from Figure 3.32, a large indel of 22 bp (AAATAATTTCTCACTCTACTCT) along with 31 SNPs and indels (with one or two bases) were found and could be used to distinguish genotypes A and B of this gene unambiguously. Nevertheless, the amplification product of *LTB4DH* was approximately 592 bp in length but direct sequencing of this amplified gene product still did not cover the entire target fragment. Moreover, too many different polymorphic sites between genotypes A and B were found and may not be accepted as the real SNPs of this gene.

Likewise, 4 SSCP genotypes of *THBP* (595 bp) were found. A representative individual of each SSCP pattern (THBP-A, THBP-B, THBP-C and THBP-D) was direct sequenced and 566 bp sequences were obtained after multiple alignments. As can be seen from Figure 3.33, 4 SNPs that can be used to distinguish these SSCP genotypes were found.

Results suggested that SSCP and direct sequencing are equally potential for identification of polymorphic DNA fragments but the former is more convenient and cost-effective.

```
LTB4DH-A        ATATTTTCAGTTCTAGTTTTTTATAGCATTTCCTCAGCTACCTTATATATTAGCCTTTTC
LTB4DH-B        ATATTTTCAGCTCTAGTTTTTTAT-GCATTTCCTCAGCTACCTTATATATT-ACCTTTTC
                ********** ************* *************************  *******

LTB4DH-A        ATTGTCTGTCATGGCCATTTTTTATGACTTCCCAGACCGCTATTTTAACTAACTTCATTA
LTB4DH-B        ATTGTCTGTCATG-CCATTTTTTATGACTTCCCAGACCGCTATTTTAACTAACTTCATT-
                ************* ******************************************

LTB4DH-A        GCTTTTCAGTTTAGCTTCTTTACACCTTATTTCATCTTCCGAATTTTTGCTTGTTTCCAT
LTB4DH-B        GCTTTTCAGTTT-GCTTCTTTACACCTTATTTCATCTTCCTTATTTTTGCTTGTTTCCAT
                ************ ************************  ******************

LTB4DH-A        TTATTTTTTTGTTTCATAGATTTCTTATTCCACCACCAAGTCGGTATTTTTTGCTCTTGT
LTB4DH-B        TTATTTTTTTGTTTCATAGATTTCTTATTCCACCACCAAGTCGGTATTTTTTGCTCTTGT
                ************************************************************

LTB4DH-A        ATCCTTTGCTGTTATAACCTTTGTCTCCTTCATATTTTGCCATCCTTCCTCTCTCTTCTG
LTB4DH-B        ATCCTTTGCTGTTATAACCTTTGTCTCCTTCATATTTTGCCATCCTTCCTCTCTCTTCTG
                ************************************************************

LTB4DH-A        TCTTCTCTTTATATTGTTTCAGAGGATCATCTGGCTTTGTAAATATTTATAGAAAATACC
LTB4DH-B        TCTTCTCTTTATATTGTTTCAGAGGATCATCTGGCTTTGTAAATATTTATAGAAAATACC
                ************************************************************

LTB4DH-A        AGACTCTGCGCT----------------------TGACCTGTTTTTTTTCTCCTCAAGAC
LTB4DH-B        AGACTCTGTGCTAAATAATTTCTCACTCTACTCTTGACCTGCTTTTTTTCTCCACAAAGC
                ******** ***                     ******* *********** ***  *

LTB4DH-A        CTGCTCCCTTCTTTCCAAAATGCCAATTTAATCCTCCTTTCATGGCCATTCAAACCC--C
LTB4DH-B        CAACTCCCCTCTTTCCAACATGTCCATTTAAGTCTCCTTTCACTGTCACACACACACACC
                •  ***** ********* *** * ******  *********  * **  ** ** *  *
                •
LTB4DH-A        CTTTCCGGGACTTTATTATCTGTACAGATGGGGGTGTTTTTTGTCTTGCAGGAAATCGGG
LTB4DH-B        CCTTCCGGGACTTTATTATCTGTAGAGATGCATGTGGCTTTTGTCTTGCAGGAAATCTGT
                * ********************** *****   ***  ******************* *

LTB4DH-A        GCTTC
LTB4DH-B        AGTCC
                  * *
```

Figure 3.32 Multiple alignments of *LTB4DH* amplified from genomic DNA of representative individuals exhibiting different SSCP genotype found in 14-month-old shrimp which using 5′UTR *LTB4DH* forward (F) and reverse (R) primers as the direct sequencing primers.

```
THBP-A          TATGACCTCAGCGAGAGTGGGCTTACCGGCTTCGTCAAGAGTTTCCTTGATGGAAAGCTG
THBP-B          TATGACCTCAGCGAGAGTGGGCTTACCGGCTTCGTCAAGAGTTTCCTTGATGGAAAGCTG
THBP-C          TATGACCTCAGCGAGAGTGGGCTTACCGGCTTCGTCAAGAGTTTCCTTGATGGAAAGCTG
THBP-D          TATGACCTCAGCGAGAGTGGGCTTACCGGCTTCGTCAAGAGTTTCCTTGATGGAAAGCTG
                ************************************************************

THBP-A          AAGCAGCACCTTCTGTCTCAGGATCTGCCAGAAGACTGGGATAAAGAGCCAGTCAAGGTA
THBP-B          AAGCAGCACCTTCTGTCTCAGGATCTGCCAGAAGACTGGGATAAAGAGCCAGTCAAGGTA
THBP-C          AAGCAGCACCTTCTGTCTCAGGATCTGCCAGAAGACTGGGATAAAGAGCCAGTCAAGGTA
THBP-D          AAGCAGCACCTTCTGTCTCAGGATCTGCCAGAAGACTGGGATAAAGAGCCAGTCAAGGTA
                ************************************************************

THBP-A          ATATCTTTTTCAAGATGTTTGCAATATGGAACTGGATAGATGCATTTGGTATTAAAAGTA
THBP-B          ACATCTTTTTCAAGATGTTTGCAATATGGAACTGGATAGATGCATTTGGTATTAAAAGTA
THBP-C          ATATCTTTTTCAAGATGTTTGCAATATGGAACTGGATAGATGCATTTGGTATTAAAAGTA
THBP-D          ATATCTTTTTCAAGATGTTTGCAATATGGAACTGGATAGATGCATTTGGTATTAAAAGTA
                * ************************************************************

THBP-A          ATCAGTTTGAACATGTGTGAACAAGTAATAGATGGTGTATATTTTGTGGATTTGGTTTTG
THBP-B          ATCAGTTTGAACATGTGTGAACAAGTAATAGATGGTGTATATTTTGTGGATTTGGTTTTG
THBP-C          ATCAGTTTGAACATGTGTGAACAAGTAATAGATGGTGTATATTTTGTGGATTTGGTTTTG
THBP-D          ATCAGTTTGAACATGTGTGAACAAGTAATAGATGGTGTATATTTTGTGGATTTGGTTTTG
                ************************************************************

THBP-A          TGGTAACTAGGAGTCAACATGAAAAAGATTTCGGGAGTTTTCATTTTACAAATATAAAGG
THBP-B          TGGTAACTAGGAGTCAACATGAAAAAGATTTCGGGAGTTTTCATTTTACAAATATAAAGG
THBP-C          TGGTAACTAGGAGTCAACATGAAAAAGATTTCGGGAGTTTTCATTTTACAAATATAAAGG
THBP-D          TGGTAACTAGGAGTCAACATGAAAAAGATTTCGAGAGTTTTCATTTTACAAATATAAAGG
                ********************************* ************************

THBP-A          TTGAAAAAACAAACTCAATGGCCAAATGAAATTGGGGAATATAATATAATTGAACTATTC
THBP-B          TTGAAAAAACAAACTCAATGGCCAAATGAAATTGGGGAATATAATATAATTGAACTATTC
THBP-C          TTGAAAAAACAAACTCAATGGCCAAATGAAATTGGGGAATATAATATAATTGAACTATTC
THBP-D          TTGAAAAAACAAACTCAATGGCCAAATGAAATTGGGGAATATAATATAATTGAACTATTC
                ************************************************************

THBP-A          AGGTTCTGGTGTCCTCCAACTTCGATGAAGTTGCCTTGAACAAGGAAAAGGATGTCCTTG
THBP-B          AGGTTCTGGTGTCCTCCAACTTCGATGAAGTTGCCTTGAACAAGGAAAAGGATGTCCTTG
THBP-C          AGGTTCTGGTGTCCTCCAACTTCGATGAAGTTGCCTTGAACAAGGAGAAGGATGTCCTTG
THBP-D          AGGTTCTGGTGTCCTCCAACTTCGATGAAGTTGCCTTGAACAAGGAGAAGGATGTCCTTG
                ********************************************* *************

THBP-A          TAGAGTTCTATGCACCATGGTGTGGACACTGCAAACAGCTGGCACCTATCTATGATCAAC
THBP-B          TAGAGTTCTATGCACCATGGTGTGGACACTGCAAACAGCTGGCACCTATCTATGATCAAC
THBP-C          TAGAGTTCTATGCACCATGGTGTGGACACTGCAAACAGCTGGCACCTATCTATGATCAGC
THBP-D          TAGAGTTCTATGCACCATGGTGTGGACACTGCAAACAGCTGGCACCTATCTATGATCAGC
                ************************************************************ *

THBP-A          TAGGTGAAAAATACAAGGACCACGATACTATTGTTGTGGCGAAAATGGATGCTACAGTCA
THBP-B          TAGGTGAAAAATACAAGGACCACGATACTATTGTTGTGGCGAAAATGGATGCTACAGTCA
THBP-C          TAGGTGAAAAATACAAGGACCACGATACTATTGTTGTGGCGAAAATGGATGCTACAGTCA
THBP-D          TAGGTGAAAAATACAAGGACCACGATACTATTGTTGTGGCGAAAATGGATGCTACAGTCA
                ************************************************************

THBP-A          ATGAGCTTGAGCATACGAAGATTCAG
THBP-B          ATGAGCTTGAGCATACGAAGATTCAG
THBP-C          ATGAGCTTGAGCATACGAAGATTCAG
THBP-D          ATGAGCTTGAGCATACGAAGATTCAG
                *************************
```

Figure 3.33 Multiple alignments of *THBP* amplified from genomic DNA of representative individuals exhibiting different SSCP genotypes found in 3-month-old shrimp using RT-THBP forward (F) and reverse (R) primers as the direct sequencing primers.

3.5.2 Identification of SNPs in *LTB4DH* and *THBP* gene segments by DNA sequencing of the cloned PCR products

To confirm polymorphic SNPs identified by direct sequencing, the PCR product of the same individuals exhibiting different SSCP genotypes of *LTB4DH* and *THBP* was gel-eluted, cloned and sequenced for both directions. Polymorphic sites in *EGF* were also detected by the same approach.

The nucleotide sequence covering the amplified product of *LTB4DH* (595 bp) was obtained. A large indel of 22 bp in length (GTGCTAAATAATTTCTCACTCTA) and 9 SNPs/indel (1-2 bp) were observed (Figure 3.34). The polymorphic sites could be used to differentiate genotypes A and B found in this study.

```
LTB4DH-A        TTCAGCCTCTACTATGTTTCTATTTCCATCCCTTACCATACCTATATTTTCAGTTCTAGT
LTB4DH-B        TTCAGCCTCTACTATGTTTCTATTTCCATCCCTTACCACACCTATATTTTCAGCTCTAGT
                ************************************ ************** ******
LTB4DH-A        TTTTTATACATTTCCTTAGCTACCTTATATATTGCCTTTTCATTGTCTGTCATGCCATTT
LTB4DH-B        TTTTTATGCATTTCCTCAGCTACCTTATATATTGCCTTTTCATTGTCTGTCATGCCATTT
                ******* ******** ***********************************************
LTB4DH-A        TTTATGACTTCCCAGACCGCTATTTTAACTAACTTCATTGCTTTTCAGTTTGCTTCTTTA
LTB4DH-B        TTTATGACTTCCCAGACCGCTATTTTAACTAACTTCATTGCTTTTCAGTTTGCTTCTTTA
                ************************************************************
LTB4DH-A        CACCTTATTTCATCTTCCTAATTTTTGCTTGTTTCCATTTATTTTTTTGTTTCATAGATT
LTB4DH-B        CACCTTATTTCATCTTCCTTATTTTTGCTTGTTTCCATTTATTTTTTTGTTTCATAGATT
                ****************** ****************************************
LTB4DH-A        TCTTACTCCACCACCAAGTCGGTATTTTTTGCTCTTGTATCCTTTGCTGTTATAACCTTT
LTB4DH-B        TCTTATTCCACCACCAAGTCGGTATTTTTTGCTCTTGTATCCTTTGCTGTTATAACCTTT
                ***** ********************************************************
LTB4DH-A        GTCTCCTTCATATTTTGCCATCCTTCCTCTCTCTTCTGTCTTCTCTTTATATTGTTTCAG
LTB4DH-B        GTCTCCTTCATATTTTGCCATCCTTCCTCTCTCTTCTGTCTTCTCTTTATATTGTTTCAG
                ************************************************************
LTB4DH-A        AGGATCATCTGGCTTTGTAAATATTTATAGAAAATACCAGACTCT----------------
LTB4DH-B        AGGATCATCTGGCTTTGTAAATATTTATAGAAAATACCAGACTCTGTGCTAAATAATTTC
                *******************************************
LTB4DH-A        -------ACTCTTGACCTGCTTTTTTTCTCCACAAAGCCAACTCCCCTCTTTCCAACATG
LTB4DH-B        TCACTCTACTCTTGACCTGCTTTTTTTCTCCACAAAGCCAACTCCCCTCTTTCCAACATG
                       ****************************************************
LTB4DH-A        TCCATTTAAGTCTCCTTTCGCTGTCACACACACAC--CCCTTCCGGGACTTTAT-ATCTG
LTB4DH-B        TCCATTTAAGTCTCCTTTCACTGTCACACACACACACCCCTTCCGGGACTTTATTATCTG
                ****************** ***************  ***************** *****
LTB4DH-A        TAGAGATGCATGTGGCTTTTGTCTTGCAGGAAATCTGTAGTCCCTAGCATCCCAG
LTB4DH-B        TAGAGATGCATGTGGCTTTTGTCTTGCAGGAAATCTGTAGTCCCTAGCATCCCAG
                ********************************************************
```

Figure 3.34 Multiple alignments of cloned *LTB4DH* amplified from genomic DNA of representative individuals exhibiting different SSCP genotypes found in 14-month-old shrimp. The amplified fragment was sequenced using M13F and M13R primers.

Eight SNP positions (A → T_{80}, T → C_{137}, G → A_{149}, A → G_{306}, A → G_{312}, T → C_{382}, A → G_{422} and A → G_{494}) were found from multiple alignments of

nucleotide sequences of different SSCP genotypes of *THBP* (Figure 3.35). Results indicated more polymorphic sites of *THBP* identified from the cloned gene segment compared to direct sequencing (4 positions) of the same target fragment.

```
THBP-A          TACAAGCCCGATACTTATGACCTCAGCGAGAGTGGGCTTACCGGCTTCGTCAAGAGTTTC
THBP-B          TACAAGCCCGATACTTATGACCTCAGCGAGAGTGGGCTTACCGGCTTCGTCAAGAGTTTC
THBP-C          TACAAGCCCGATACTTATGACCTCAGCGAGAGTGGGCTTACCGGCTTCGTCAAGAGTTTC
THBP-D          TACAAGCCCGATACTTATGACCTCAGCGAGAGTGGGCTTACCGGCTTCGTCAAGAGTTTC
                ************************************************************
THBP-A          CTTGATGGAAAGCTGAAGCAGCACCTTCTGTCTCAGGATCTGCCAGAAGACTGGGATAAA
THBP-B          CTTGATGGAAAGCTGAAGCAGCACCTTCTGTCTCAGGATCTGCCAGAAGACTGGGATAAA
THBP-C          CTTGATGGAAAGCTGAAGCTGCACCTTCTGTCTCAGGATCTGCCAGAAGACTGGGATAAA
THBP-D          CTTGATGGAAAGCTGAAGCAGCACCTTCTGTCTCAGGATCTGCCAGAAGACTGGGATAAA
                ******************* ****************************************
THBP-A          GAGCCAGTCAAGGTAATATCTTTTTCAAGATGTTTGCAATATGGAACTGGATAGATGCAT
THBP-B          GAGCCAGTCAAGGTAACATCTTTTTCAAGATGTTTGCAATATGGAACTGGATAGATGCAT
THBP-C          GAGCCAGTCAAGGTAATATCTTTTTCAAAATGTTTGCAATATGGAACTGGATAGATGCAT
THBP-D          GAGCCAGTCAAGGTAATATCTTTTTCAAGATGTTTGCAATATGGAACTGGATAGATGCAT
                **************** ********** ******************************
THBP-A          TTGGTATTAAAAGTAATCAGTTTGAACATGTGTGAACAAGTAATAGATGGTGTATATTTT
THBP-B          TTGGTATTAAAAGTAATCAGTTTGAACATGTGTGAACAAGTAATAGATGGTGTATATTTT
THBP-C          TTGGTATTAAAAGTAATCAGTTTGAACATGTGTGAACAAGTAATAGATGGTGTATATTTT
THBP-D          TTGGTATTAAAAGTAATCAGTTTGAACATGTGTGAACAAGTAATAGATGGTGTATATTTT
                ************************************************************
THBP-A          GTGGATTTGGTTTTGTGGTAACTAGGAGTCAACATGAAAAAGATTTCGGGAGTTTTCATT
THBP-B          GTGGATTTGGTTTTGTGGTAACTAGGAGTCAACATGAAAAAGATTTCGGGAGTTTTCATT
THBP-C          GTGGATTTGGTTTTGTGGTAACTAGGAGTCAACATGAAAAAGATTTCGGGAGTTTTCATT
THBP-D          GTGGATTTGGTTTTGTGGTAACTAGGAGTCAACATGAAAAAGATTTCGGGAGTTTTCATT
                ************************************************************
THBP-A          TTACAAATATAAAGGTTGAAAAAACAAACTCAATGGCCAAATGAAATTGGGGAATATAAT
THBP-B          TTACAAATATAGAGGTTGAAAAAACAAACTCAATGGCCAAATGAAATTGGGGAATATAAT
THBP-C          TTACAGATATAAAGGTTGAAAAAACAAACTCAATGGCCAAATGAAATTGGGGAATATAAT
THBP-D          TTACAAATATAAAGGTTGAAAAAACAAACTCAATGGCCAAATGAAATTGGGGAATATAAT
                ***** ***** ************************************************
THBP-A          ATAATTGAACTATTCAGGTTCTGGTGTCCTCCAACTTCGATGAAGTTGCCTTGAACAAGG
THBP-B          ATAATTGAACTATTCAGGTTCTGGTGTCCTCCAACTTCGATGAAGTTGCCTTGAACAAGG
THBP-C          ATAATTGAACTATTCAGGTTCCGGTGTCCTCCAACTTCGATGAAGTTGCCTTGAACAAGG
THBP-D          ATAATTGAACTATTCAGGTTCTGGTGTCCTCCAACTTCGATGAAGTTGCCTTGAACAAGG
                ********************* ************************************
THBP-A          AAAAGGATGTCCTTGTAGAGTTCTATGCACCATGGTGTGGACACTGCAAACAGCTGGCAC
THBP-B          AAAAGGATGTCCTTGTAGAGTTCTATGCACCATGGTGTGGACACTGCAAACAGCTGGCAC
THBP-C          AGAAGGATGTCCTTGTAGAGTTCTATGCACCATGGTGTGGACACTGCAAACAGCTGGCAC
THBP-D          AGAAGGATGTCCTTGTAGAGTTCTATGCACCATGGTGTGGACACTGCAAACAGCTGGCAC
                * ************************************************************
THBP-A          CTATCTATGATCAACTAGGTGAAAAATACAAGGACCACGATACTATTGTTGTGGCGAAAA
THBP-B          CTATCTATGATCAACTAGGTGAAAAATACAAGGACCACGATACTATTGTTGTGGCGAAAA
THBP-C          CTATCTATGATCAGCTAGGTGAAAAATACAAGGACCACGATACTATTGTTGTGGCGAAAA
THBP-D          CTATCTATGATCAGCTAGGTGAAAAATACAAGGACCACGATACTATTGTTGTGGCGAAAA
                ************* ********************************************
THBP-A          TGGATGCTACAGTCAATGAGCTTGAGCATACGAAGATTCAGTCCTTCCCAACCTT
THBP-B          TGGATGCTACAGTCAATGAGCTTGAGCATACGAAGATTCAGTCCTTCCCAACCTT
THBP-C          TGGATGCTACAGTCAATGAGCTTGAGCATACGAAGATTCAGTCCTTCCCAACCTT
THBP-D          TGGATGCTACAGTCAATGAGCTTGAGCATACGAAGATTCAGTCCTTCCCAACCTT
                ********************************************************
```

Figure 3.35 Multiple alignments of cloned *THBP* amplified from genomic DNA of representative individuals exhibiting different SSCP genotypes found in 3-month-old shrimp. The amplified fragment was sequenced using M13F and M13R primers.

Two SSCP patterns were obtained from SSCP analysis of the *EGF* gene segment. Only 2 SNP positions (A → C_{32} and C → A_{75}) were found after cloning, sequencing and multiple alignments of these SSCP genotypes (Figure 3.36).

```
EGF-A1          GTCACTGCTGAAAATAGCCATCTACAGGAAGAGATTGAACACAGCAAGGCAGATGACCAG
EGF-A2          GTCACTGCTGAAAATAGCCATCTACAGGAAGAGATTGAACACAGCAAGGCAGATGACCAG
EGF-B1          GTCACTGCTGAAAATAGCCATCTACAGGAAGAGATTGAACACAGCAAGGCAGATGACCAG
EGF-B2          GTCACTGCTGAAAATAGCCATCTACAGGAAGCGATTGAACACAGCAAGGCAGATGACCAG
                ****************************** ****************************

EGF-A1          AGGAACACAGAAACCAAGACGCAGCTCGAACAAAAAATAAATGCACTGAACACCCAAAAA
EGF-A2          AGGAACACAGAAACCAAGACGCAGCTCGAACAAAAAATAAATGCACTGAACACCCAAAAA
EGF-B1          AGGAACACAGAAACCAAGACGCAGCTCGAACAAAAAATAAATGCACTGAACACCCAAAAA
EGF-B2          AGGAACACAGAAACAAAGACGCAGCTCGAACAAAAAATAAATGCACTGAACACCCAAAAA
                ************** ******************************************

EGF-A1          GCAGAACTGGGAAGAAACTTGGAACAAGTCAAAACACAAAAAAGTCAGGTTTCAAAATCC
EGF-A2          GCAGAACTGGGAAGAAACTTGGAACAAGTCAAAACACAAAAAAGTCAGGTTTCAAAATCC
EGF-B1          GCAGAACTGGGAAGAAACTTGGAACAAGTCAAAACACAAAAAAGTCAGGTTTCAAAATCC
EGF-B2          GCAGAACTGGGAAGAAACTTGGAACAAGTCAAAACACAAAAAAGTCAGGTTTCAAAATCC
                ************************************************************

EGF-A1          ATTTCGAGTGTTCAGAAAGAAATTAAAGACATTGAAACGTATTTCAGTGTTTCAGAACCG
EGF-A2          ATTTCGAGTGTTCAGAAAGAAATTAAAGACATTGAAACGTATTTCAGTGTTTCAGAACCG
EGF-B1          ATTTCGAGTGTTCAGAAAGAAATTAAAGACATTGAAACGTATTTCAGTGTTTCAGAACCG
EGF-B2          ATTTCGAGTGTTCAGAAAGAAATTAAAGACATTGAAACGTATTTCAGTGTTTCAGAACCG
                ************************************************************
```

Figure 3.36 Multiple alignments of cloned *EGF* amplified from genomic DNA of representative individuals exhibiting SSCP genotypes found in 3-month-old shrimp. The amplified fragment was sequenced using M13F and M13R primers.

Results from this study strongly indicated that SSCP analysis is potential and cost-effective for initial screening of SNPs in various candidate genes against a large number of specimens. Due to large sequencing errors found from direct sequencing of the amplified gene segment, DNA sequencing for identification of SNPs in various genes of *P. monodon* should be performed using the cloned gene segments.

3.6 Association between SNPs by SSCP (SBS) and the body weight or GSI of *P. monodon*

SSCP analysis of *THBP* and *EGF* was further carried out covering 340 individuals of 3-month-old shrimp (the average body weight = 13.32 ± 2.43 g). Genotypes A, B, C and D of *THBP* were found in 113, 4, 89 and 134 examined

individuals. Likewise, 2 SSCP genotypes (A and B) of *EGF* were found in 163 and 45 examined individuals, respectively.

Genotype distribution frequencies of candidate growth-related genes were used to determine whether SSCP genotypes (SNPs) and the body weight of 3-month-old shrimp are statistically significant.

Results did not reveal significant correlation between SSCP genotypes of *EGF* and the body weight of cultured juvenile *P. monodon* in this study (t-value = 0.856, *df* = 338, *P* = 0.392, Figure 3.38). *LTB4DH* did not show significant correlation between SSCP genotypes and the GSI values ($P > 0.05$, Figure 3.39).

In contrast, a statistically significant difference in allele frequencies between juvenile shrimp carrying genotypes B (the average body weight of 11.89 ± 2.34 g, *N* = 4) and C (the average body weight of 14.12 ± 2.80 g, *N* = 89) of *THBP* was observed ($P < 0.05$, Figure 3.37). Juvenile *P. monodon* carrying genotypes A (the average body weight of 13.49 ± 2.34 g, *N* = 113) and D (the average body weight of 12.69 ± 2.06 g, *N* = 134) exhibited slightly lower body weight than that of genotype B but statistical analysis was not significant ($P > 0.05$). Accordingly, a larger sample size of juvenile shrimp should be analyzed to eliminate the possible sampling errors, if any, in this study.

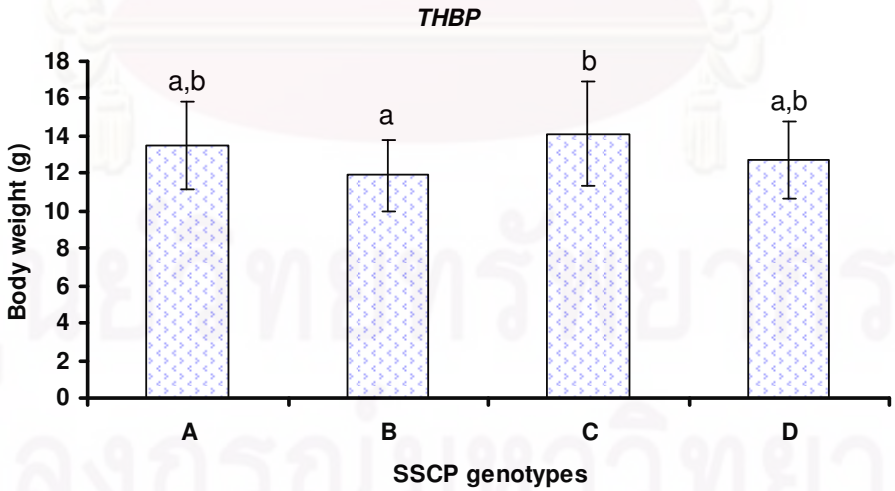


Figure 3.37 Histogram showing association between the body weight of 3-month-old *P. monodon* (BUM03, *N* = 340) and SSCP genotypes of *THBP*.

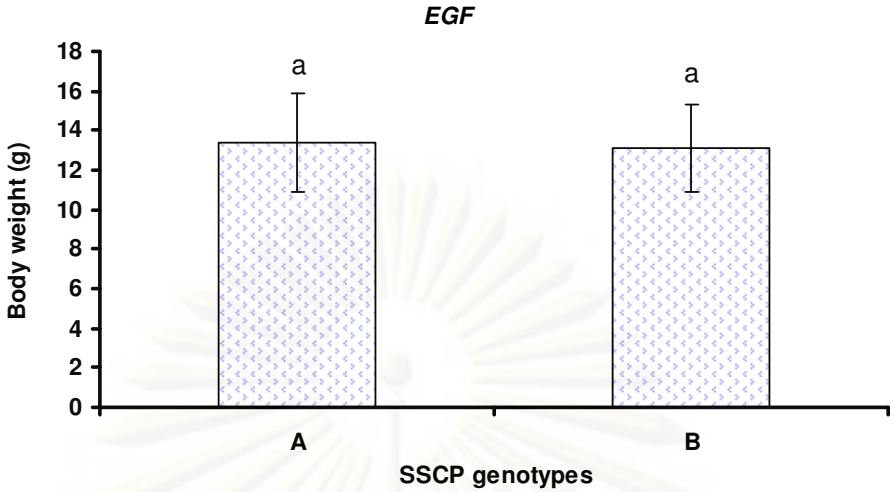


Figure 3.38 Histogram showing association between the body weight of 3-month-old *P. monodon* (BUM03, *N* = 340) and SSCP genotypes of *EGF*.

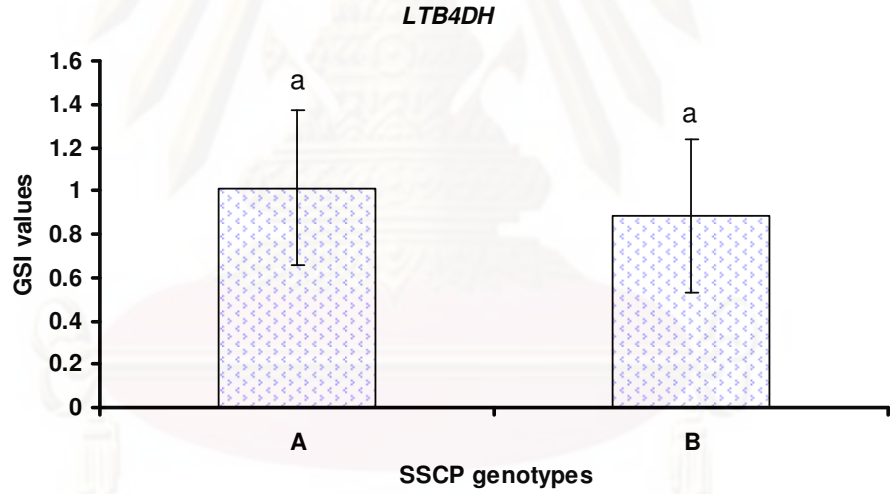


Figure 3.39 Histogram showing association between the GSI values of 14-month-old *P. monodon* (BU14F, *N* = 66) and SSCP genotypes of *LTB4DH*.

3.7 Total RNA extraction

Total RNA was extracted from hepatopancreas of 3-month-old juveniles (BUM03) and ovaries of 14-month-old broodstock (BU14F) of *P. monodon*. The quantity and quality of total RNA was spectrophotometrically and electrophoretically analyzed. The ratio of OD_{260}/OD_{280} of extracted RNA ranged from 2.0 - 2.1 indicating

that RNA samples were relatively pure. Agarose gel electrophoresis indicated smear total RNA with a few discrete bands implying the accepted quality of extracted total RNA (Figure 3.40). The first strand cDNA was reverse-transcribed. Agarose gel electrophoresis indicated that the resulting products ranged between 100 bp - 2 kb indicating the acceptable quality of the synthesized first strand cDNA (Figure 3.41).

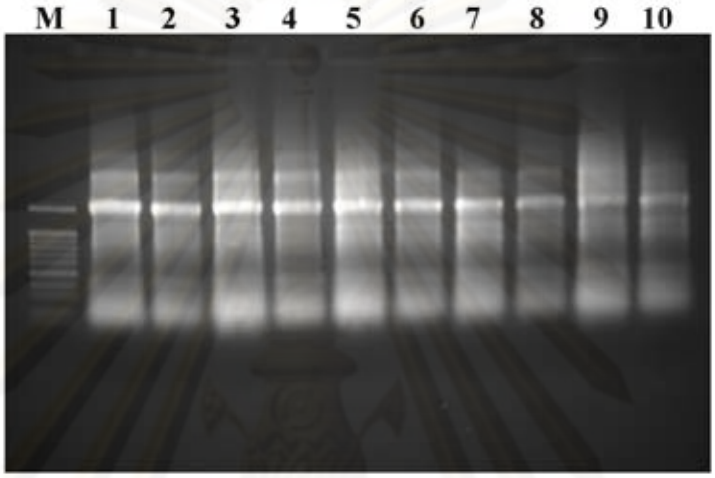


Figure 3.40 A 0.8% ethidium bromide-stained agarose gel showing the quality of total RNA extracted from hepatopancreas of different individuals of juveniles (Lanes 1 - 5) and ovaries of different individuals of 14-month-old broodstock (Lanes 6 - 10) of *P. monodon*. Lane M is a 100 bp DNA ladder.

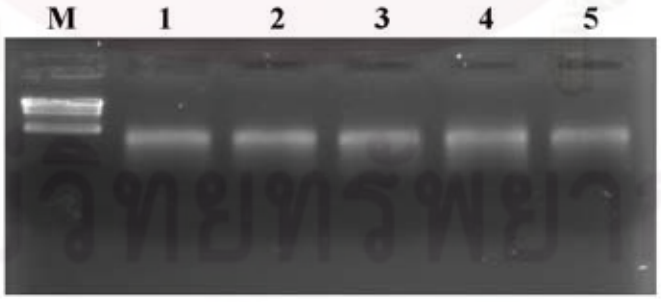


Figure 3.41 A 1.0% ethidium bomide-stained agarose gel showing the first strand cDNA reverse-transcribed from DNA-free total RNA of hepatopancreas (Lanes 1 - 5) of different *P. monodon* juveniles. Lane M is λ-*Hin*d III DNA marker.

3.8 Determination of the expression patterns of functionally important genes in *P. monodon* by RT-PCR

Expression profiles of *LTB4DH* and *NASP* in ovaries of 14-month-old broodstock and those of *insulin degrading enzyme (IDE), epidermal growth factor (EGF)* and *thyroid hormone binding protein (THBP)* in hepatopancreas of 3-month-old juveniles were examined by RT-PCR.

The expected product of each gene was obtained from the cDNA template of all examined individuals (Figures 3.42 - 3.46). The positive product of *EF-1α* was successfully amplified from the same template.

Figure 3.42 RT-PCR of *LTB4DH* (panel B) using the first strand cDNA of ovaries of 14-month-old broodstock (BU14F, Lanes 1 - 17). *EF-1α* was included as the positive control and successfully amplified in all specimens (Lanes 1 - 17, panel A). Lanes M are a 100 bp DNA ladder.

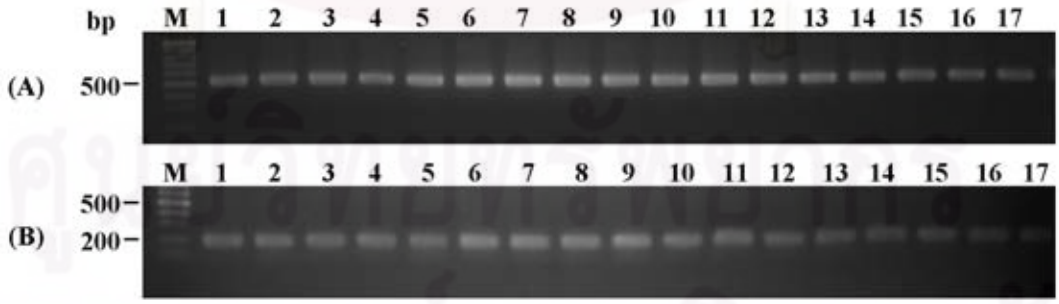


Figure 3.43 RT-PCR of homologue of *NASP* (panel B) using the first strand cDNA of ovaries of 14-month-old broodstock (BU14F, Lanes 1 - 17). *EF-1α* was included as the positive control and successfully amplified in all specimens (Lanes 1 - 17, panel A). Lanes M are a 100 bp DNA ladder.

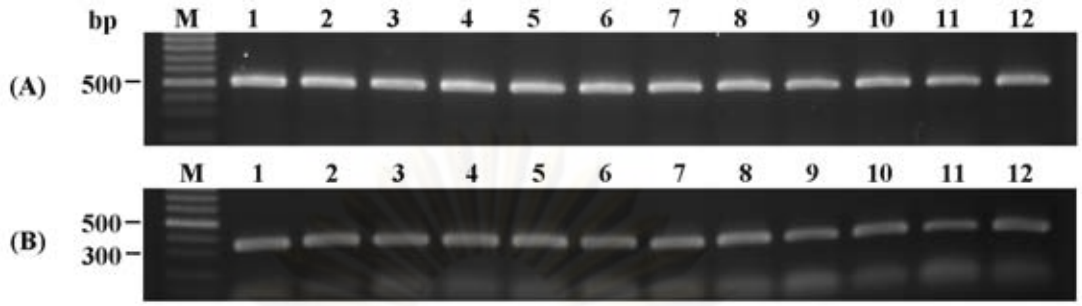


Figure 3.44 RT-PCR of *THBP* (panel B) using the first strand cDNA of hepatopancrease of 3-month-old juveniles (BUM03, Lanes 1 - 12). *EF-1α* was included as the positive control and successfully amplified in all specimens (Lanes 1 - 12, panel A). Lanes M are a 100 bp DNA ladder.

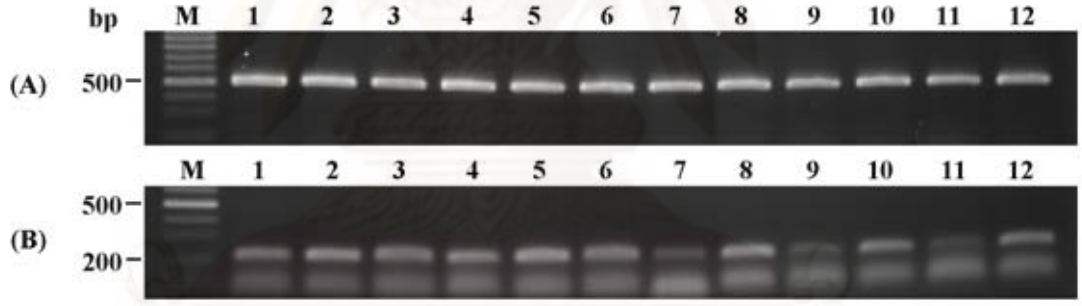


Figure 3.45 RT-PCR of homologue of *IDE* (panel B) using the first strand cDNA of hepatopancrease of 3-month-old juveniles (BUM03, Lanes 1 - 12). *EF-1α* was included as the positive control and successfully amplified in all specimens (Lanes 1 - 12, panel A). Lanes M are a 100 bp DNA ladder.

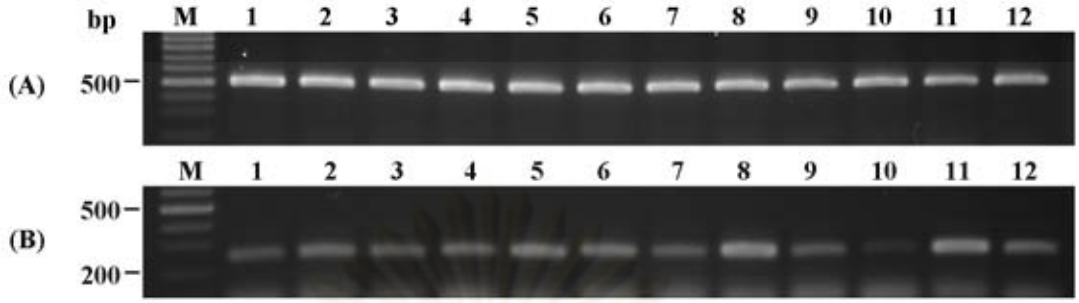


Figure 3.46 RT-PCR of homologue of *EGF* (panel B) using the first strand cDNA of hepatopancrease of 3-month-old juveniles (BUM03, Lanes 1 - 12). *EF-1α* was included as the positive control and successfully amplified in all specimens (Lanes 1 - 12, panel A). Lanes M are a 100 bp DNA ladder.

3.9 Tissue distribution analysis of *insulin degrading enzyme (IDE)* and *epidermal growth factor (EGF)* in *P. monodon* juveniles

Tissue distribution analysis was carried out to examine the expression of *IDE* and *EGF* in hemocytes, gill, heart, ovaries, hepatopancreas, stomach, intestine, lymphoid, thoracic ganglion, eyestalk and pleopods of a female and testes of a male of juveniles.

IDE was expressed in all tissues. Relatively high expression level of this gene was found in hemocytes, stomach, lymphoid organ, thoracic ganglion, ovaries and testes but lower expression level was observed in gill, heart, intestine and eyestalk of juvenile *P. monodon* (Figure 3.47).

Likewise, *EGF* was abundantly expressed in all examined tissues where that in hemocytes, stomach, pleopods and testes were greater than other tissues (Figure 3.48). Expression levels of these genes in different tissues are illustrated in Table 3.2.

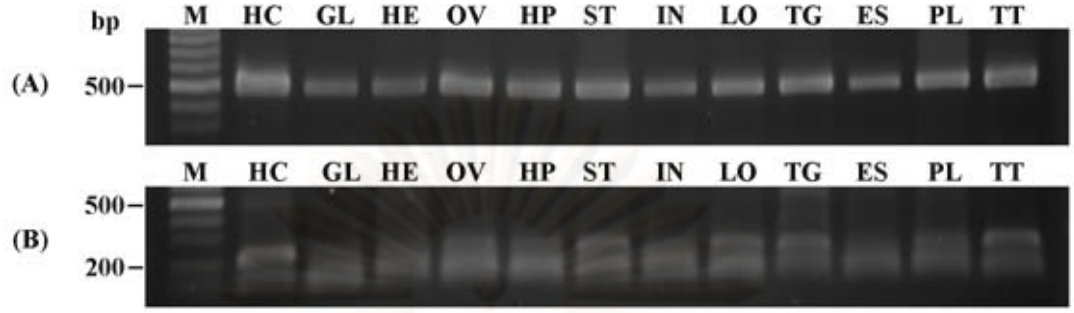


Figure 3.47 A 1.5% ethidium bromide-stained agarose gel showing results from tissue distribution analysis of *insulin degrading enzyme* (*IDE*) using the first strand cDNA of various tissues (B) of 3-month-old *P. monodon* juveniles. *EF-1α* was successfully amplified from the same template (A). Lane M are a 100 bp DNA ladder. HC = hemocytes, GL = gill, HE = heart, OV = ovaries, HP = hepatopancreas, ST = stomach, IN = intestine, LO = lymphoid, TG = thoracic ganglion, ES = eyestalk, PL = pleopods and TT = testes.

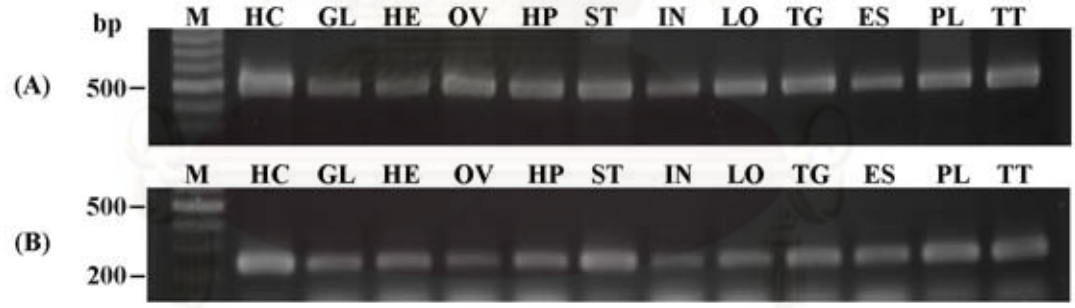


Figure 3.48 A 1.5% ethidium bromide-stained agarose gel showing results from tissue distribution analysis of a homologue of *epidermal growth factor* using the first strand cDNA of various tissues (B) of 3-month-old *P. monodon* juveniles. *EF-1α* was successfully amplified from the same template (A). Lane M are a 100 bp DNA ladder. HC = hemocytes, GL = gill, HE = heart, OV = ovaries, HP = hepatopancreas, ST = stomach, IN = intestine, LO = lymphoid, TG = thoracic ganglion, ES = eyestalk, PL = pleopods and TT = testes

Table 3.2 Tissue distribution analysis of *insulin degrading enzyme* and *epidermal growth factor* in varies tissues of 3-month-old *P. monodon* juveniles

Gene homologues	Size (bp)	Tissue											
		HC	GL	HE	OV	HP	ST	IN	LO	TG	ES	PL	TT
Insulin degrading enzyme (IDE)	204	++ +	+	+	++	++	++	+	++	++	+	++	++
Epidermal growth factor (EGF)	240	++ +	++	++	+	++	++ +	+	++	++	++	++ +	++ +

+ = low level of expression, ++ = moderate level of expression, +++ = abundant level of expression

3.10 Isolation and characterization of the partial cDNA sequence of *insulin degrading enzyme (IDE)*

Several discrete bands were obtained from 5′ and 3′ RACE-PCR of *IDE*. A 700 bp and 2.0 kb fragment generated from nested 5′ and 3′ RACE-PCR was cloned and sequenced (Figure 3.49). Nucleotide sequences of these fragments and EST (Figure 3.50) were assembled.

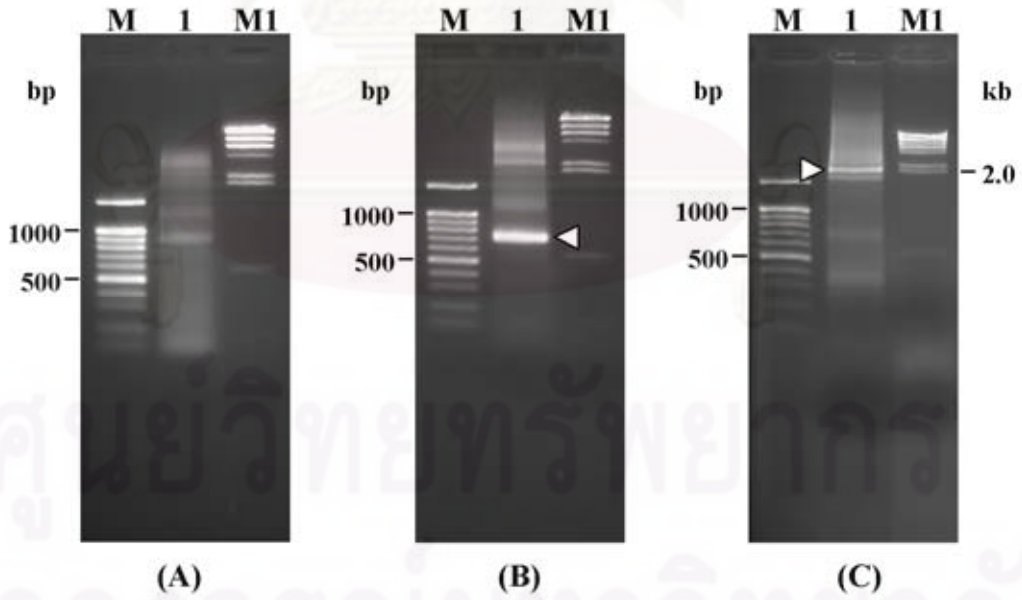


Figure 3.49 The primary 5′ (A), 3′ (C) RACE-PCR and secondary 5′ RACE-PCR products (B) of *insulin degrading enzyme*. Arrowheads indicate RACE-PCR products that were cloned and sequenced. Lanes M and M1 are a 100 bp DNA ladder and λ-*Hin*d III DNA marker, respectively.

A

ACGCGGGATCCAGTAATGCCTACACAGCAGCTGACCACACCAACTATTATTTTGACGTAGCTCCTGAGG
CATTCTCGGGAGCCTTAGACAGATTTGCACAGTTCTTCCTGACACCACTGTTTACTGAAAGTGCAGTGG
ACCGTGAGGTGAATGCAGTTAATTCTGAACACGAAAAAAATTTACAAAATGATTACTGGAGGTTAACAC
AGCTTGAAAAGGCAACTGCTGATCCAACTCATGACTTCAGCAAGTTTAACACTGGAAACAAGGAAACAT
TGGATACAATTCCTAAAGAGAAAGGTGTAAATGTCCGAGATGCTCTGCTGGAATTTCACAAAAAATGGT
ACTCTTCAAACATCATGGCCTTAGCTGTCCTGGGCAAAGAAAGTTTAGAAGAGCTGCAAGCAATGGTGC
TGGAATTATTTTCTGAGGTTGAGAATAAAAATGTGACCGTACCTGAGTGGAAGAGCCACCCCTTTGGCC
CCGAACAGTGTCGTAAATATGGGCAGGTTG

B

TGCCTGTGAAGGATATTCGGAACTTGTATATTACTTTTCCGATCCCAGATCTTCATCCTCATTATAAGA
CAGCTCCTGGCCACTATTTGGGGCATCTCATCGGGCATGAAGGTCCAGGGTCATTGCTCTCCTATCTTA
5′RACE-PCR(nested)
AGGGTTG**TGGATGGGTGAACTCTCTTGTTGG**TGGTCAGAAATCTGGTGCTAAAGGCTTTGCCTTCTTTG
TGGTCAATGTTGATCTCACAGAAGAAGGGATTGAGCATGTGCAAGATATTGTGACAGCTGTTTTCCAGT
3′RACE-PCR
ACCTGAACTTGCTGAGAAGAGAGGGTTCTCA***GGAGTGGGTGTTTGAAGAGTGT*AAAGA**CCTCAGTGCCA
TGACATTCCGCTTTAAGGATAAGGAGCGTCCCCAGAGTTACACTTGTGCTCTTGCTGAACAGCTGCATT
ATTACCCACTGGAAGAGGTGCTGTGTGGTGGATACCTGTTGAGTGAATACCAGCCACATCTTATTGACA
5′RACE-PCR
TGGTTCTAACTCATCTTCAGCCTACCAATATTCGAATTGCTGTGGTAGGGAAAGCTCTAGCAGAGAAGG
CAACGGATGTTGAACAATGGTATGGGACACAGTACAAGATGGATAACATTCAGCAAGACCAACTTGAAC
AGTGGGAAAAAGCTGGGTTCACTGAGTTGTTGAAGTTACCAACCAAAAATGACTTTGTTCCTACAAACT
TTGAGTTGTTCACCGATGAAAAGATTGGAAATACTTTGCCAGAGATGATTAGTGAGTCTCCTTGGCTAG
AATTGGGTCAGCAAGATGACGAA

C-I

TTCAAGCTGCCAAAATCTGTTGTGTATGCAGAACTTTTTAGTCCTTTGGCGTACCTGGATCCTCACCAC
ACCAATCTACTGCACATGTTTGCACAACTGTTTCGGGATGCACTCACTGAATACACTTATGCTGCTGAG
TTGGCTGGACTTGTCTATTCCTTGTCCAATACCAAGTATGGTCTTACACTGAACGTCAAGGGTTACAAT
3′RACE-PCR(internal)
GACAAACAACATGTACTGCTAGAGAAAATCATGGATCGCATGA**CTACATTCACTGTTGATCCGA**AGCGT
TTTGAGATTCTTAAAGATGCTTATGTCCGTGCCTTACGTAACTTCAGAGCAGATCAACCACATCAACAT
GTTGTGTTTTATACATCCTTATTGCTGGCAGAAAATGGTTGGACGAAGGAGGAGCT

C-II

TTTGAATGCCACTTCAGAGCTGACAGTGGAGGCTCTAGAGGCCTTTATCCCAAGGTTCCTTTCAAATCT
CCATATAGAAATGTTAGTGCATGGAAATGCAACAAGAGATGCCGCAGTGCGCATGTCTTCTATCATCCA
AAAACATCTGACTGACAAGGCACACACCAAACCTCTCCTAAATTCCCAGTTGACTCGACAGAGAGAGTA
CCAGCTGAAGGATGGAAGCAGCTTAGTGTACCAGGCTGAGAATACTGTCCATCGTGACTCGGCTATTGA
AACCTTTTTCCAGTGTGGCATGCAGGGAACACACCCAAATATGTTGCTGGAATTACTCTGTCAAATCTT
TGCTGAGCCTGCATTTGATGAGCTAAGAACCAAGGAACAACTGGGCTACATTATGTGGTGTGGTGTCCG
AAGGGCTAATGGAACCCAGGGCCTAAGGATCATAGTCCAAGGTGACCGTCACCCACAGTACTTAGACTC
CAGGATAGAGGCCTTCCTTCACAAAATGGGTGAACGCTTGGAGAGTATCTCAGATGAGGAATTCATGCG
TCACAGAGATGCATTAGCCAGCCGCCGTTTAGAGAGACCAAAGAAGCTGGCACATCTGACGTCCATCTG
GTGGGCTGAGATCACCTCCAATCAGTACAACTTTGACCGAGACACGATTGAGGTGGCGCACTTGAAGAC
GCTTACCAAACAGAATATTGTTGATTTCTACAAGGCCTGTATTGCAAGTTATGCCGAAGAGAAAGAGCT
GTCAGTGCAAGTGGTGTCGGTTTCCCCTGAGGAG

C-III

CAGGGACGTGAGCGCGCCGGCCAGCCGTCCACCCAGCCAGATGACGGACTCTCGCAGCCTCCGCCACTC
GCTGAGGCTGAGAACATCACGGATATATCGGAATTCAAGCAGGGTCTCCCACTCTATCCTCTGATGAGA
CCCTACACACCTTTCTCAAAGACCTCGAAGTCCAAGCTGTAATACCGAATTTAATTTTTAAAATAGAAC
CACTTGTTTCTGTCAAGTTAATCACTTTCTCTACATATTGTACTTGGAGAACTATGAAATAAATAATGA
AATTGATTGTTTGTTTGATATATCCAGATTCGTAGGTATTTATGGATGCGGCGTCATGTTTTTTGCACC

```
CATTAGATTATTTTTTGTTTCCACTATACTGTTATTTAGATAAAAAGGTGTTTGATATTTATGGATACA
TTTCGTACCTTCTTTTCTATATGTATTTTTTGGAAGGAGTAGTTGTTTTGAGGAAGTCACAAATCCTGC
AGCCATGACACTTTTCTCATCTCCTGTCTGTTTTTTAAGGATTGTTAAAATAATGTTATCTGTCAGTAA
GTTTTATTTGTATAGTTGGAAATTTTATTTTCTCCTGAATCCCAATTAGATATGTAAATAGGAGCTTCT
TTTATGATGGGTTTATATATTTTTATATTAGAAATGTTTGTGTTATTTTGATGGTCAGAATTATTTTCA
ATATGTTATTCCTCACAGCAGACTGTAATTGGAATATTGAAAATAAACAGATAAATCATAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAA
```

Figure 3.50 The sequence of nested 5′ RACE-PCR (A), EST (B), 3′ RACE-PCR (C-I and C-III) and internal sequencing of the 3′ RACE-PCR product (C-II) of *IDE* of *P. monodon*. RT-PCR primers (IDE-F and IDE-R, highlighted) were initially used for 5′ and 3′ RACE-PCR followed by nested 5′RACE-PCR primers for isolation of the partial cDNA sequence of this gene.

The combined sequence of *IDE* is 3269 bp in length containing an open reading frame (ORF) of 2727 bp corresponding to 908 amino acids with the 3′UTR of 569 bp (excluding the poly A tail, Figure 3.51). It significantly matched *insulin degrading enzymes* of the zebrafish (*Danio rerio*) with the *E*-value of 0.0 (Figure 3.52).

```
ACGCGGGATCCAGTAATGCCTACACAGCAGCTGACCACACCAACTATTATTTTGACGTAG    60
  A  G  S  S  N  A  Y  T  A  A  D  H  T  N  Y  Y  F  D  V  A   20
CTCCTGAGGCATTCTCGGGAGCCTTAGACAGATTTGCACAGTTCTTCCTGACACCACTGT   120
  P  E  A  F  S  G  A  L  D  R  F  A  Q  F  F  L  T  P  L  F   40
TTACTGAAAGTGCAGTGGACCGTGAGGTGAATGCAGTTAATTCTGAACACGAAAAAAATT   180
  T  E  S  A  V  D  R  E  V  N  A  V  N  S  E  H  E  K  N  L   60
TACAAAATGATTACTGGAGGTTAACACAGCTTGAAAAGGCAACTGCTGATCCAACTCATG   240
  Q  N  D  Y  W  R  L  T  Q  L  E  K  A  T  A  D  P  T  H  D   80
ACTTCAGCAAGTTTAACACTGGAAACAAGGAAACATTGGATACAATTCCTAAAGAGAAAG   300
  F  S  K  F  N  T  G  N  K  E  T  L  D  T  I  P  K  E  K  G  100
GTGTAAATGTCCGAGATGCTCTGCTGGAATTTCACAAAAAATGGTACTCTTCAAACATCA   360
  V  N  V  R  D  A  L  L  E  F  H  K  K  W  Y  S  S  N  I  M  120
TGGCCTTAGCTGTCCTGGGCAAAGAAAGTTTAGAAGAGCTGCAAGCAATGGTGCTGGAAT   420
  A  L  A  V  L  G  K  E  S  L  E  E  L  Q  A  M  V  L  E  L  140
TATTTTCTGAGGTTGAGAATAAAAATGTGACCGTACCTGAGTGGAAGAGCCACCCCTTTG   480
  F  S  E  V  E  N  K  N  V  T  V  P  E  W  K  S  H  P  F  G  160
GCCCCGAACAGTGTCGTAAATATGGGCAGGTTGTGCCTGTGAAGGATATTCGGAACTTGT   540
  P  E  Q  C  R  K  Y  G  Q  V  V  P  V  K  D  I  R  N  L  Y  180
ATATTACTTTTCCGATCCCAGATCTTCATCCTCATTATAAGACAGCTCCTGGCCACTATT   600
  I  T  F  P  I  P  D  L  H  P  H  Y  K  T  A  P  G  H  Y  L  200
TGGGGCATCTCATCGGGCATGAAGGTCCAGGGTCATTGCTCTCCTATCTTAAGGGTTGTG   660
  G  H  L  I  G  H  E  G  P  G  S  L  L  S  Y  L  K  G  C  G  220
GATGGGTGAACTCTCTTGTTGGTGGTCAGAAATCTGGTGCTAAAGGCTTTGCCTTCTTTG   720
  W  V  N  S  L  V  G  G  Q  K  S  G  A  K  G  F  A  F  F  V  240
TGGTCAATGTTGATCTCACAGAAGAAGGGATTGAGCATGTGCAAGATATTGTGACAGCTG   780
  V  N  V  D  L  T  E  E  G  I  E  H  V  Q  D  I  V  T  A  V  260
TTTTCCAGTACCTGAACTTGCTGAGAAGAGAGGGTTCTCAGGAGTGGGTGTTTGAAGAGT   840
  F  Q  Y  L  N  L  L  R  R  E  G  S  Q  E  W  V  F  E  E  C  280
GTAAAGACCTCAGTGCCATGACATTCCGCTTTAAGGATAAGGAGCGTCCCCAGAGTTACA   900
```

```
  K  D  L  S  A  M  T  F  R  F  K  D  K  E  R  P  Q  S  Y  T  300
CTTGTGCTCTTGCTGAACAGCTGCATTATTACCCACTGGAAGAGGTGCTGTGTGGTGGAT   960
  C  A  L  A  E  Q  L  H  Y  Y  P  L  E  E  V  L  C  G  G  Y  320
ACCTGTTGAGTGAATACCAGCCACATCTTATTGACATGGTTCTAACTCATCTTCAGCCTA  1020
  L  L  S  E  Y  Q  P  H  L  I  D  M  V  L  T  H  L  Q  P  T  340
CCAATATTCGAATTGCTGTGGTAGGGAAAGCTCTAGCAGAGAAGGCAACGGATGTTGAAC  1080
  N  I  R  I  A  V  V  G  K  A  L  A  E  K  A  T  D  V  E  Q  360
AATGGTATGGGACACAGTACAAGATGGATAACATTCAGCAAGACCAACTTGAACAGTGGG  1140
  W  Y  G  T  Q  Y  K  M  D  N  I  Q  Q  D  Q  L  E  Q  W  E  380
AAAAAGCTGGGTTCACTGAGTTGTTGAAGTTACCAACCAAAAATGACTTTGTTCCTACAA  1200
  K  A  G  F  T  E  L  L  K  L  P  T  K  N  D  F  V  P  T  N  400
ACTTTGAGTTGTTCACCGATGAAAAGATTGGAAATACTTTGCCAGAGATGATTAGTGAGT  1260
  F  E  L  F  T  D  E  K  I  G  N  T  L  P  E  M  I  S  E  S  420
CTCCTTGGCTAGAATTGGGTCAGCAAGATGACGAATTCAAGCTGCCAAAATCTGTTGTGT  1320
  P  W  L  E  L  G  Q  Q  D  D  E  F  K  L  P  K  S  V  V  Y  440
ATGCAGAACTTTTTAGTCCTTTGGCGTACCTGGATCCTCACCACACCAATCTACTGCACA  1380
  A  E  L  F  S  P  L  A  Y  L  D  P  H  H  T  N  L  L  H  M  460
TGTTTGCACAACTGTTTCGGGATGCACTCACTGAATACACTTATGCTGCTGAGTTGGCTG  1440
  F  A  Q  L  F  R  D  A  L  T  E  Y  T  Y  A  A  E  L  A  G  480
GACTTGTCTATTCCTTGTCCAATACCAAGTATGGTCTTACACTGAACGTCAAGGGTTACA  1500
  L  V  Y  S  L  S  N  T  K  Y  G  L  T  L  N  V  K  G  Y  N  500
ATGACAAACAACATGTACTGCTAGAGAAAATCATGGATCGCATGACTACATTCACTGTTG  1560
  D  K  Q  H  V  L  L  E  K  I  M  D  R  M  T  T  F  T  V  D  520
ATCCGAAGCGTTTTGAGATTCTTAAAGATGCTTATGTCCGTGCCTTACGTAACTTCAGAG  1620
  P  K  R  F  E  I  L  K  D  A  Y  V  R  A  L  R  N  F  R  A  540
CAGATCAACCACATCAACATGTTGTGTTTTATACATCCTTATTGCTGGCAGAAAATGGTT  1680
  D  Q  P  H  Q  H  V  V  F  Y  T  S  L  L  L  A  E  N  G  W  560
GGACGAAGGAGGAGCTTTTGAATGCCACTTCAGAGCTGACAGTGGAGGCTCTAGAGGCCT  1740
  T  K  E  E  L  L  N  A  T  S  E  L  T  V  E  A  L  E  A  F  580
TTATCCCAAGGTTCCTTTCAAATCTCCATATAGAAATGTTAGTGCATGGAAATGCAACAA  1800
  I  P  R  F  L  S  N  L  H  I  E  M  L  V  H  G  N  A  T  R  600
GAGATGCCGCAGTGCGCATGTCTTCTATCATCCAAAAACATCTGACTGACAAGGCACACA  1860
  D  A  A  V  R  M  S  S  I  I  Q  K  H  L  T  D  K  A  H  T  620
CCAAACCTCTCCTAAATTCCCAGTTGACTCGACAGAGAGAGTACCAGCTGAAGGATGGAA  1920
  K  P  L  L  N  S  Q  L  T  R  Q  R  E  Y  Q  L  K  D  G  S  640
GCAGCTTAGTGTACCAGGCTGAGAATACTGTCCATCGTGACTCGGCTATTGAAACCTTTT  1980
  S  L  V  Y  Q  A  E  N  T  V  H  R  D  S  A  I  E  T  F  F  680
TCCAGTGTGGCATGCAGGGAACACACCCAAATATGTTGCTGGAATTACTCTGTCAAATCT  2040
  Q  C  G  M  Q  G  T  H  P  N  M  L  L  E  L  L  C  Q  I  F  700
TTGCTGAGCCTGCATTTGATGAGCTAAGAACCAAGGAACAACTGGGCTACATTATGTGGT  2100
  A  E  P  A  F  D  E  L  R  T  K  E  Q  L  G  Y  I  M  W  C  720
GTGGTGTCCGAAGGGCTAATGGAACCCAGGGCCTAAGGATCATAGTCCAAGGTGACCGTC  2160
  G  V  R  R  A  N  G  T  Q  G  L  R  I  I  V  Q  G  D  R  H  740
ACCCACAGTACTTAGACTCCAGGATAGAGGCCTTCCTTCACAAAATGGGTGAACGCTTGG  2220
  P  Q  Y  L  D  S  R  I  E  A  F  L  H  K  M  G  E  R  L  E  760
AGAGTATCTCAGATGAGGAATTCATGCGTCACAGAGATGCATTAGCCAGCCGCCGTTTAG  2280
  S  I  S  D  E  E  F  M  R  H  R  D  A  L  A  S  R  R  L  E  780
AGAGACCAAAGAAGCTGGCACATCTGACGTCCATCTGGTGGGCTGAGATCACCTCCAATC  2340
  R  P  K  K  L  A  H  L  T  S  I  W  W  A  E  I  T  S  N  Q  800
AGTACAACTTTGACCGAGACACGATTGAGGTGGCGCACTTGAAGACGCTTACCAAACAGA  2400
  Y  N  F  D  R  D  T  I  E  V  A  H  L  K  T  L  T  K  Q  N  820
ATATTGTTGATTTCTACAAGGCCTGTATTGCAAGTTATGCCGAAGAGAAAGAGCTGTCAG  2460
  I  V  D  F  Y  K  A  C  I  A  S  Y  A  E  E  K  E  L  S  V  840
TGCAAGTGGTGTCGGTTTCCCCTGAGGAGCAGGGACGTGAGCGCGCCGGCCAGCCGTCCA  2520
  Q  V  V  S  V  S  P  E  E  Q  G  R  E  R  A  G  Q  P  S  T  860
CCCAGCCAGATGACGGACTCTCGCAGCCTCCGCCACTCGCTGAGGCTGAGAACATCACGG  2580
  Q  P  D  D  G  L  S  Q  P  P  P  L  A  E  A  E  N  I  T  D  880
ATATATCGGAATTCAAGCAGGGTCTCCCACTCTATCCTCTGATGAGACCCTACACACCTT  2640
  I  S  E  F  K  Q  G  L  P  L  Y  P  L  M  R  P  Y  T  P  F  900
TCTCAAAGACCTCGAAGTCCAAGCTGTAATACCGAATTTAATTTTTAAAATAGAACCACT  2700
  S  K  T  S  K  S  K  L  *                                    908
```

```
TGTTTCTGTCAAGTTAATCACTTTCTCTACATATTGTACTTGGAGAACTATGAAATAAAT  2760
AATGAAATTGATTGTTTGTTTGATATATCCAGATTCGTAGGTATTTATGGATGCGGCGTC  2820
ATGTTTTTTGCACCCATTAGATTATTTTTTGTTTCCACTATACTGTTATTTAGATAAAAA  2880
GGTGTTTGATATTTATGGATACATTTCGTACCTTCTTTTCTATATGTATTTTTTGGAAGG  2940
AGTAGTTGTTTTGAGGAAGTCACAAATCCTGCAGCCATGACACTTTTCTCATCTCCTGTC  3000
TGTTTTTTAAGGATTGTTAAAATAATGTTATCTGTCAGTAAGTTTTATTTGTATAGTTGG  3060
AAATTTTATTTTCTCCTGAATCCCAATTAGATATGTAAATAGGAGCTTCTTTTATGATGG  3120
GTTTATATATTTTTATATTAGAAATGTTTGTGTTATTTTGATGGTCAGAATTATTTTCAA  3180
TATGTTATTCCTCACAGCAGACTGTAATTGGAATATTGAAAATAAACAGATAAATCATAA  3240
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA                                 3269
```

Figure 3.51 The partial cDNA and deduced amino sequences of *IDE* of *P. monodon*. The stop codon is illustrated in boldfaced and underlined.

```
>ref|NP_001082994.1|  insulin-degrading enzyme [Danio rerio]
 gb|AAI39608.1|  Zgc:162603 protein [Danio rerio]
Length=978

 GENE ID: 561390 zgc:162603 | zgc:162603 [Danio rerio]
(10 or fewer PubMed links)

 Score = 1031 bits (2666),  Expect = 0.0
 Identities = 497/881 (56%), Positives = 651/881 (73%), Gaps = 2/881 (0%)
 Frame = +3

Query  3     AGSSNAYTAADHTNYYFDVAPEAFSGALDRFAQFFLTPLFTESAVDREVNAVNSEHEKNL  182
             AGSSNA+T+ +HTNYYFDV+ E   GALDRFAQFFL PLF ES  DREVNAV+SEHEKNL
Sbjct  94    AGSSNAFTSGEHTNYYFDVSHEHLQGALDRFAQFFLCPLFDESCKDREVNAVDSEHEKNL  153

Query  183   QNDYWRLTQLEKATADPTHDFSKFNTGNKETLDTIPKEKGVNVRDALLEFHKKWYSSNIM  362
              ND WRL QLEKAT +P H FSKF TGNK TL+T P ++G+++R+ LL+FH  +YSSN+M
Sbjct  154   MNDAWRLFQLEKATGNPKHPFSKFGTGNKLTLETRPSQQGIDIREELLKFHSTYYSSNLM  213

Query  363   ALAVLGKESLEELQAMVLELFSEVENKNVTVPEWKSHPFGPEQCRKYGQVVPVKDIRNLY  542
              L VLG+E+L+EL +MV++LF EVENKNV VPE+ +HPF  E  R++ +VVP+KDIRNLY
Sbjct  214   GLCVLGRETLDELTSMVVKLFGEVENKNVPVPEFPTHPFQEEHLRQFYKVVPIKDIRNLY  273

Query  543   ITFPIPDLHPHYKTAPGHYLGHLIGHEGPGSLLSYLKGCGWVNSLVGGQKSGAKGFAFFV  722
             +TFPIPDL  +YK+ PGHYLGHLIGHEGPGSLLS LK  GWVN+LVGGQK GA+GF FF+
Sbjct  274   VTFPIPDLQKYYKSNPGHYLGHLIGHEGPGSLLSELKSKGWVNTLVGGQKEGARGFMFFI  333

Query  723   VNVDLTEEGIEHVQDIVTAVFQYLNLLRREGSQEWVFEECKDLSAMTFRFKDKERPQSYT  902
             +NVDLTEEG+ HV+DI+  +FQY+  LR EG QEWVF+ECKDL+ + FRFKDKERP+ YT
Sbjct  334   INVDLTEEGLLHVEDIIFHMFQYIQKLRTEGPQEWVFQECKDLNTVAFRFKDKERPRGYT  393

Query  903   CALAEQLHYYPLEEVLCGGYLLSEYQPHLIDMVLTHLQPTNIRIAVVGKALAEKATDVEQ  1082
               +A  LHYYPLEE+L   YLL E++P LI+MVL  L+P N+R+AVV K+   +    E+
Sbjct  394   SKVAGLLHYYPLEEILAAEYLLEEFRPDLIEMVLDKLRPENVRVAVVSKSFEGQTDRTEE  453

Query  1083  WYGTQYKMDNIQQDQLEQWEKAGFTELLKLPTKNDFVPTNFELFTDEKIGNTLPEMISES  1262
             WYGTQYK + I  + +++W+ A      KLP KN+F+PTNFE++  EK   + P +I ++
Sbjct  454   WYGTQYKQEAITDEAIKKWDNADLNGKFKLPMKNEFIPTNFEIYPLEKDSPSAPTLIKDT  513

Query  1263  PWLELG-QQDDEFKLPKSVVYAELFSPLAYLDPHHTNLLHMFAQLFRDALTEYTYAAELA  1439
               ++  +QDD+F LPK+ +  E FSP AY+DP H N+ +++ +L +D+L EY YAAELA
Sbjct  514   AMSKVWFKQDDKFFLPKACLNFEFFSPFAYVDPLHCNMAYLYLELLKDSLNEYAYAAELA  573

Query  1440  GLVYSLSNTKYGLTLNVKGYNDKQHVLLEKIMDRMTTFTVDPKRFEILKDAYVRALRNFR  1619
             GL Y L NT YG+ L+VKGYNDKQH+LL+KI+++M TF +D KRF+I+K+AY+R+L NFR
Sbjct  574   GLSYDLQNTVYGMYLSVKGYNDKQHILLKKIIEKMATFEIDEKRFDIIKEAYMRSLNNFR  633

Query  1620  ADQPHQHVVFYTSLLLAENGWTKEELLNATSELTVEALEAFIPRFLSNLHIEMLVHGNAT  1799
             A+QPHQH ++Y  LL+ E  WTK+EL +A  ++T+  L+AFIP+ LS LHIE L+HGN T
Sbjct  634   AEQPHQHAMYYLRLLMTEVAWTKDELRDALDDVTLPRLKAFIPQLLSRLHIEALLHGNIT  693
```

```
Query  1800  RDAAVRMSSIIQKHLTDKAHTKPLLNSQLTRQREYQLKDGSSLVYQAENTVHRDSAIETF  1979
             + +A+ M  +++  L + AHTKPLL SQL R RE Q+ DG   VYQ  N VH +  IE +
Sbjct  694   KQSALEMMQMLEDTLIEHAHTKPLLPSQLIRYREVQVPDGGWYVYQQRNEVHNNCGIEIY  753

Query  1980  FQCGMQGTHPNMLLELLCQIFAEPAFDELRTKEQLGYIMWCGVRRANGTQGLRIIVQGDR  2159
             +Q  MQ TH NMLLEL CQI +EP F+ LRTKEQLGYI++ G RRANG QGLR I+Q ++
Sbjct  754   YQTDMQNTHENMLLELFCQIISEPCFNTLRTKEQLGYIVFSGPRRANGVQGLRFIIQSEK  813

Query  2160  HPQYLDSRIEAFLHKMGERLESISDEEFMRHRDALASRRLERPKKLAHLTSIWWAEITSN  2339
              P YL+SR+EAFL  M + +E + DE F +H  ALA RRL++PKKLA   + +W EI S
Sbjct  814   APHYLESRVEAFLKTMEKSVEEMGDEAFQKHIQALAIRRLDKPKKLAAECAKYWGEIISQ  873

Query  2340  QYNFDRDTIEVAHLKTLTKQNIVDFYKACIASYAEEKELSVQVVSVSPEEQGRERAGQPS  2519
             QYNFDRD IEVA+LKTLTK++I+ FY+  +A  A  +   V V  +S E       G+
Sbjct  874   QYNFDRDNIEVAYLKTLTKEHIMQFYRDLLAIDAPRRH-KVSVHVLSREMDSCPLVGEFP  932

Query  2520  TQPDDGLSQPPPLAEAENITDISEFKQGLPLYPLMRPYTPF  2642
              Q D  L+  P L +   + D++EFK+ LPL+PL +P+  F
Sbjct  933   AQNDVNLAPAPSLPQPSLVQDMTEFKRSLPLFPLTKPHINF  973
```

Figure 3.52 Similarity analysis based on Blast*X* of *IDE* of *P. monodon* against previously deposited sequences in the GenBank.

3.11 Examination of expression levels of various functionally important genes of *P. monodon* by quantitative real-time PCR

The expression levels of *NASP* and *LTB4DH* in ovaries of 14-month-old shrimp (GSI < 0.5%, *N* = 6; GSI = 0.5 - 1.0%, *N* = 12 and GSI > 1.0%, *N* = 12) and *insulin degrading enzyme (IDE)*, *epidermal growth factor (EGF)* and *thyroid hormone binding protein (THBP)* in hepatopancreas of 3-month-old shrimp (a large group with the average body weight of 16.78 ± 2.24 g, *N* = 22 and a small group with the average body weight of 10.43 ± 1.08 g, *N* = 22 groups) were quantitatively estimated by real-time PCR.

The standard curves used for quantitative real-time PCR analysis of interesting transcripts are shown in Figures 3.53 and 3.54. High R^2 values and amplification efficiency of examined transcripts were found. These standard curves were used for quantitative estimation of the mRNA levels of *NASP*, *LTB4DH, IDE*, *EGF* and *THBP*.

The expression levels of *NASP* in ovaries of 14-month-old broodstock between shrimp exhibiting different GSI values: < 0.5% (group I). 0.5-1.0% (group II) and >1.0% (group III), were not significantly different ($P > 0.05$, Figure 3.55).

In contrast, the expression level of *LTB4DH* in ovaries of group II shrimp was greater that that of group I ($P < 0.05$) but not significantly different from that of group III shrimp ($P > 0.05$, Figure 3.56A). However, expression level of this gene in shrimp carrying different SSCP (SNP) genotypes was not significantly different ($P > 0.05$, Figure 3.56B).

The expression level of *IDE* in hepatopancreas of small (10.43 ± 1.08 g, N = 22) and large (16.78 ± 2.24 g, N = 22) groups of 3-month-old shrimp was not significantly ($P > 0.05$, Figure 3.57).

The expression level of *EGF* in hepatopancreas of small (N = 22) and large (N = 22) groups of 3-month-old shrimp was not significantly ($P > 0.05$, Figure 3.58A). Interestingly, the expression level of *EGF* in shrimp carrying the SSCP (SNP) genotypes B was greater than that of genotype A but the results were not statistically significant due to the large standard deviation within each group ($P > 0.05$, Figure 3.58B).

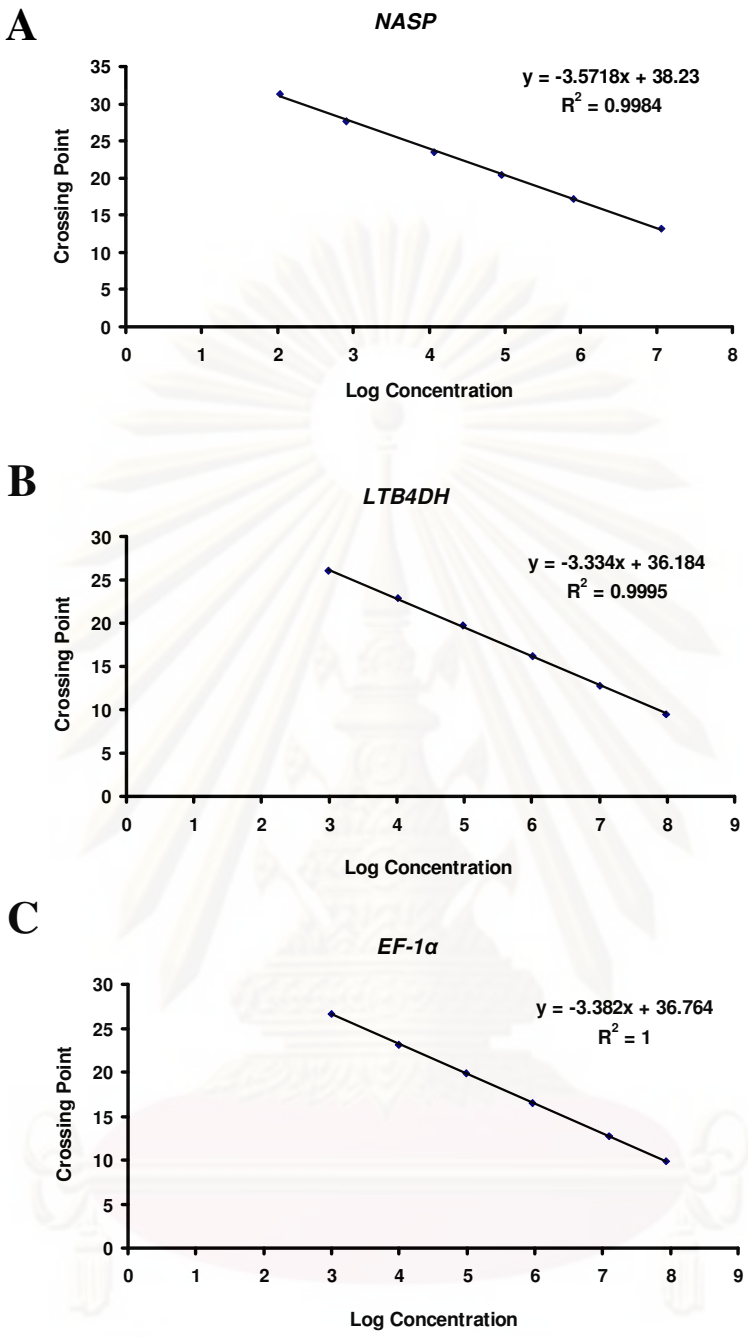


Figure 3.53 The standard amplification curves of various genes examined by quantitative real-time PCR. The standard curves of *NASP* (A; R2 = 0.9984, amplification efficiency = 96.7% or (log2) 1.967 and the equation Y = -3.404 * log(X) + 38.30), *LTB4DH* (A; R^2 = 0.9995, amplification efficiency = 95.8% or (log_2) 1.958 and the equation Y = -3.426 * log(X) + 36.16) and *EF-1α* (A; R^2 = 1, amplification efficiency = 98.4% or (log_2) 1.984 and the equation Y = -3.362 * log(X) + 36.74) using 10-fold dilution of plasmid DNA of *NASP, LTB4DH* and *EF-1α* (10^3 - 10^8 copy). *EF-1α* was used as a reference transcript.

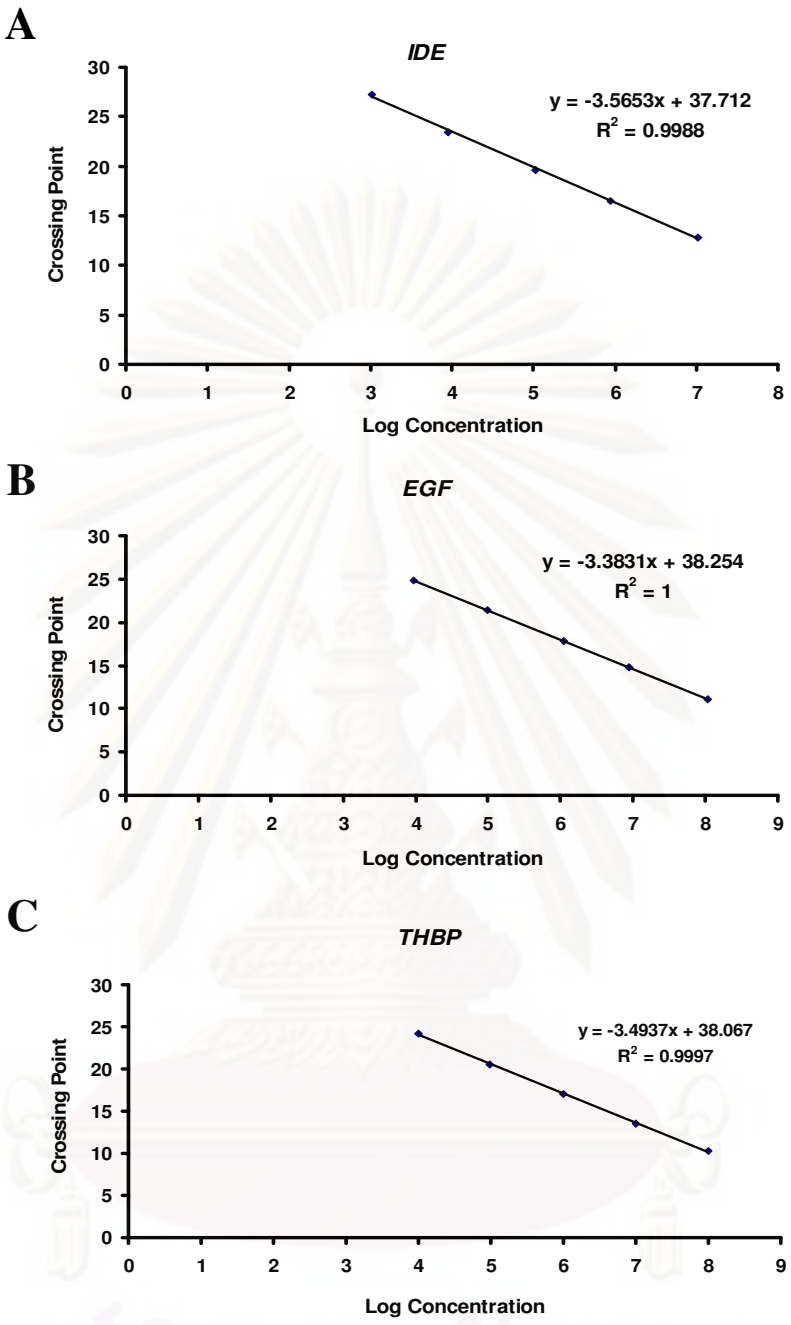


Figure 3.54 The standard amplification curves of various genes examined by real-time PCR. The standard curves of *IDE* (A; R^2 = 0.9988, amplification efficiency = 97.3% or (log_2) 1.973 and the equation Y = -3.389 * log(X) + 37.78), *EGF* (A; R2 = 1, amplification efficiency = 97.4% or (log2) 1.974 and equation; Y = -3.386 * log(X) + 38.39) and *THBP* (A; R^2 = 0.9997, amplification efficiency = 97.9% or (log_2) 1.979 and the equation Y = -3.374 * log(X) + 38.08) using 10-fold dilution of plasmid DNA of *IDE*, *EGF* and *THBP* (10^3 - 10^8 copy). *EF-1α* was used as a reference transcript.

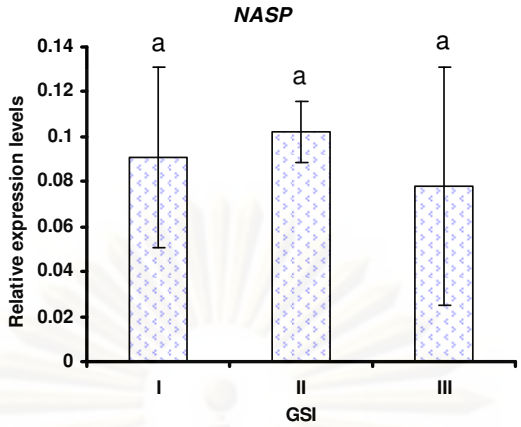


Figure 3.55 Histograms showing the relative expression levels of *NASP* in ovaries of 14-month-old broodstock exhibiting GSI < 0.5% (I). >0.5-1.0% (II) and >1.0% (III). Expression levels were measured as the absolute copy number of *NSAP* and *LTB4DH* mRNA (50 ng template) and normalized by that of *EF-1α* mRNA (50 ng template). The same letters indicate non-significant differences between relative expression levels of different groups of samples.

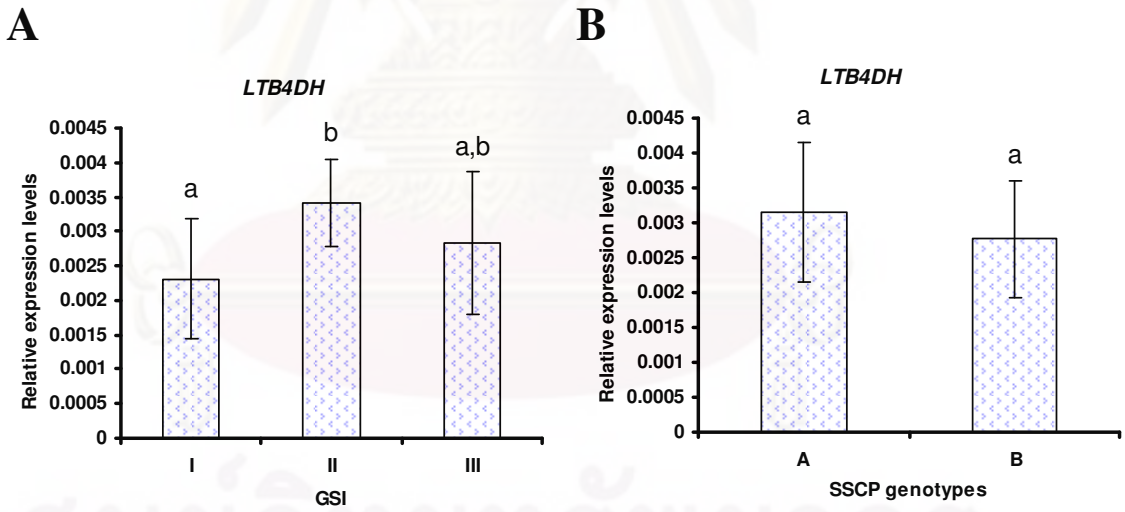


Figure 3.56 Histograms showing the relative expression levels of *LTB4DH* in ovaries 14-month-old broodstock considering those of exhibiting GSI < 0.5% (I). >0.5-1.0% (II) and >1.0% (III) (panel A) or those possessing different SSCP genotypes (panel B). Expression levels were measured as the absolute copy number of *LTB4DH* mRNA (50 ng template) and normalized by that of *EF-1α* mRNA (50 ng template). The same letters indicate non-significant differences between relative expression levels of different groups of s

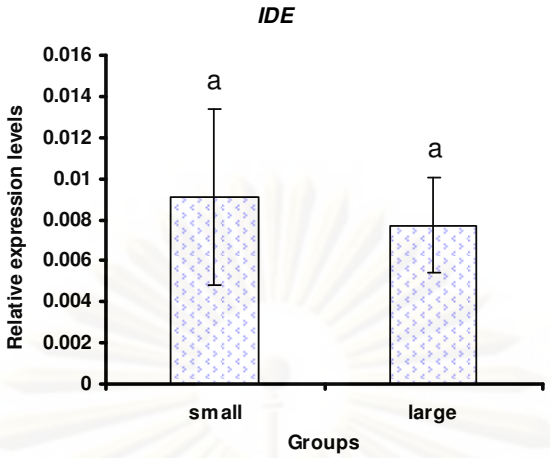


Figure 3.57 Histograms showing the relative expression levels of *IDE* in hepatopancreas of 3-month-old juveniles exhibiting 16% from the top (large) and bottom (small) of the body weight. Expression levels were measured as the absolute copy number of *IDE* mRNA (100 ng template) and normalized by that of *EF-1α* mRNA (10 ng template). The same letters indicate non-significant differences between relative expression levels of different groups of samples.

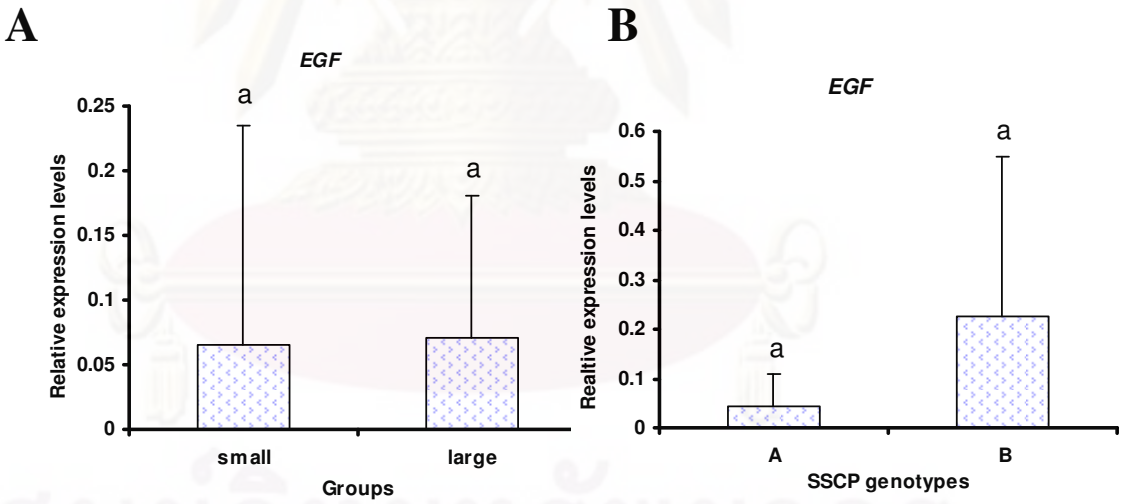


Figure 3.58 Histograms showing the relative expression levels of *EGF* in hepatopancreas of 3-month-old juveniles considering those exhibiting 16% from the top (large) and bottom (small) of the body weight (A) or those possessing different SSCP genotypes (B). Expression levels were measured as the absolute copy number of *EGF* mRNA (100 ng template) and normalized by that of *EF-1α* mRNA (10 ng template). The same letters indicate non-significant differences between relative expression levels of different groups of samples.

The expression level of *THBP* in hepatopancreas of small (N = 22) and large (N = 22) groups of 3-month-old shrimp was not significantly ($P > 0.05$, Figure 3.59A). Similarly, the expression levels of *EGF* in shrimp carrying different SSCP (SNP) genotypes were not significantly different ($P > 0.05$, Figure 3.59B).

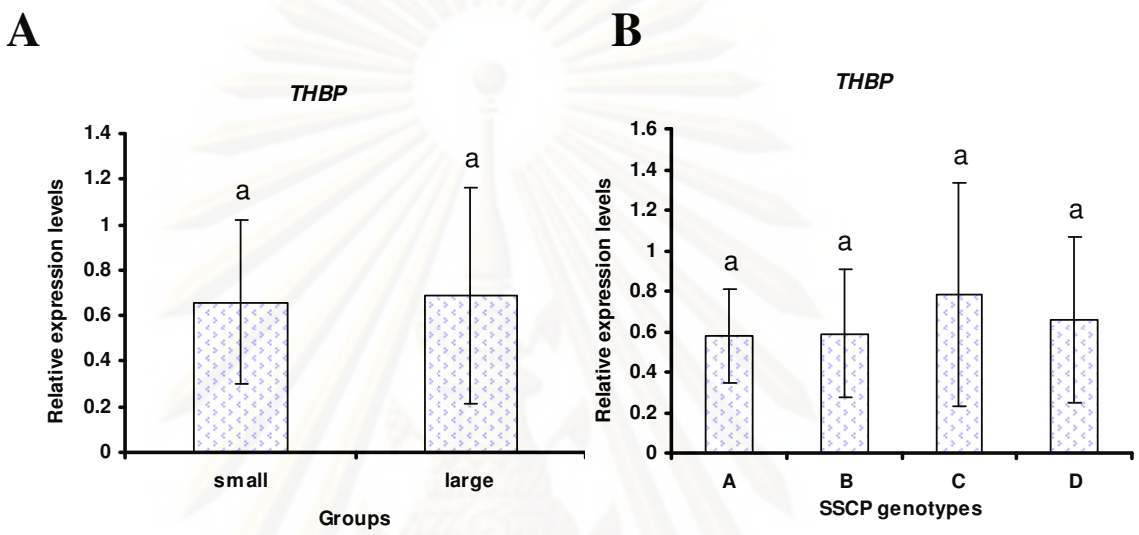


Figure 3.59 Histograms showing the relative expression levels of *THBP* in hepatopancreas of 3-month-old juveniles considering those exhibiting 16% from the top (large) and bottom (small) of the body weight (A) or those possessing different SSCP genotypes (B). Expression levels were measured as the absolute copy number of *THBP* mRNA (100 ng template) and normalized by that of *EF-1α* mRNA (10 ng template). The same letters indicate non-significant differences between relative expression levels of different groups of samples.

CHAPTER IV

DISCUSSION

Genes significantly matched growth- and reproduction related genes were chosen for association analysis between their polymorphism and the body weight of juveniles or GSI of domesticated broodstock of *P. monodon*, respectively. SSCP is favored for examining genetic diversity of various species owing to its convenience and cost effectiveness (Orita *et al*., 1989). Previously, SNP in the coding sequence of the bovine *transferrin* gene was reported. Two alleles (SSCP1 and SSCP2) were detected and the mutation point was identified and confirmed by direct sequencing of the amplified gene segment. The relationship between deduced proteins and DNA polymorphism was established. Protein variants A, D1 and E correspond to SSCP allele 1 and a variant D2 corresponds to SSCP allele 2. DNA sequences from genotypes AA, AE, AD2, D1E, D2E and D2D2 reveal an A/G substitution at position 1455 of the cDNA which causes a Gly/Glu substitution. Identification of multiple variants suggested that other SNPs also exist in the bovine *transferrin* gene (Laurent and Rodellar, 2001).

No validated marker linked to performance and production traits has been identified in penaeid shrimp up to date. Recently, a genetic linkage map of a full-sib F2 intercross family of *M. japonicus* was constructed for identification of QTL influencing growth (weight, total length, and carapace length) using AFLP analysis. A homologue of gene encoding the elongation of very long chain fatty acids-like (ELOVL) protein family resided within the QTL peak was cloned and sequenced. However, association between SNP and growth rates or expression levels of this gene has not been examined (Lyons *et al*., 2007).

The fundamental controls of growth in *P. monodon* have not been reported. Several gene homologues encoding vertebrate-like growth factors (e.g. *growth factor receptor bound protein* and *endothelial cell growth factor 1*) and cell cycle regulating proteins (e.g. *cyclin A*, *cyclin dependent kinase 1* and *cell division cycle 2*) were

identified and characterized by EST analysis (Preechaphol *et al*., 2007; Leelatanawit *et al*., 2009).

Likewise, several reproduction-related genes (e.g. profilin, ubiquitin-conjugating enzyme, egalitarian CG4051-PA isoform 1, cofilin/actin-depolymerizing factor homolog, checkpoint kinase 1 and serine/threonine-protein kinase LATS1) were also found from EST analysis of ovarian cDNA libraries of *P. monodon* (Preechaphol, 2008). However, correlation between SNP in growth- and reproduction-related genes and commercial important phenotypes has not been reported in this species.

In this study, polymorphism at the 5′UTR and exon (and intron) of various genes in *P. monodon* was examined. Correlations between genotypes (SNP through SSCP patterns) and phenotypes of juveniles or domesticated broodstock were searched and, if it is significant, possibly applied for selective breeding programs of *P. monodon*.

Identification of polymorphic SNP in reproduction-related genes of *P. monodon* and association analysis

The 5′UTR of *LTB4DH* (767 bp) and *NASP* (670 bp) of *P. monodon* were successfully identified by genome walking analysis. Polymorphism of the amplified 5′UTR segments of these reproduction-related genes was examined by SSCP analysis.

NASP was monomorphic in 14-month-old shrimp but two SSCP genotypes (A and B for which their sequences were different by several SNP positions). Association between SNP by SSCP genotypes of *LTB4DH* and GSI values (<0.5, $N = 6$; 0.5-1.0, of $N = 12$ and >1.0%, of $N = 12$) of domesticated *P. monodon* (14-month-old with the average body weight of 87.32 ± 12.47 g, $N = 66$). Nevertheless, SSCP genotypes of *LTB4DH* and the GSI values of broodstock-sized shrimp were not significant ($P > 0.05$).

Several publications revealed significant association of SNP (and SSCP patterns) at the 5′UTR and phenotypes (e.g. growth rates) and/or levels of gene expression. Morganti *et al*., (2005) reported relationships between promoter

polymorphism in the thymidylate synthase gene and mRNA levels in colorectal cancer.

Aoki-Suzuki *et al*., (2005) revealed specific haplotypes encompassing alternatively spiced exons of the nitrin-G1 (*NTNG1*) were associated with schizophrenia. The mRNA of different isoforms was significantly different between schizophrenic and control brains. An association between *NTNG2* and schizophrenia was also observed with SNP and haplotypes of the 5′UTR.

Liang *et al*., (2005) studied polymorphism of 5′ flanking region in chicken prolactin (*cPRL*) and provided the possibility that polymorphic SNP site might be related to the broodiness in chickens via modulating the transcriptional level of *cPRL*. The dissociation among *cPRL* gene transcription, mRNA storage and hormone was observed.

Trakooljul *et al*., (2004) showed that the androgen receptor (*AR*) was highly polymorphic and polymorphism affect the predicted amino acid sequence and the transcription factor binding sites and are associated with allele-specific differences of the *AR* mRNA transcript level in liver of the porcine.

Villette *et al*., (2002) studied SNP in the 3'UTR of membrane associated phospholipid hydroperoxide glutathione peroxidase (*GPX4*) of human. SNP was identified as a C/T variation at the 718th position. Individuals of different genotypes exhibited significant differences in the level of lymphocyte 5-lipoxygenase total products, with C_{718} showing increase levels of those products compared to T_{718} and T/C_{718} (36% and 44% increases, respectively). This suggests that SNP_{718} has functional effects and support the hypothesis that *GPX4* plays a regulatory role in leukotriene biosynthesis.

Therefore, the correlation between SSCP genotypes of *NASP* and *LTB4DH* and its expression levels in ovaries of cultured 14 months olds shrimp (phenotype) was subsequently examined by quantitative real-time PCR.

Generally, the efficiency to detect single-base substitutions by SSCP was approximately 80% of those verified by DNA sequencing (Shastry, 2002). Therefore, SNP in polymorphic SSCP patterns A and B of *LTB4DH* were examined by

sequencing of the amplified PCR products and cloned PCR products. Like results previously reported by Buaklin (2005), the level of sequencing errors by direct sequencing of the PCR products was far greater than the cloned fragments. Therefore, SSCP is suitable for screening of the target gene segments and polymorphism between SSCP genotypes of genes exhibiting significant association with an interesting phenotype (s) should be verified by sequencing of the cloned gene segments.

Association analysis between SNP through SSCP patterns of growth-related genes and the body weight of *P. monodon*

Recently, correlations between SNPs in the actin and crustacean hyperglycemic hormone (CHH) genes with individual growth performance in the giant freshwater prawn (*Macrobrachium rosenbergii*) were examined by DNA sequencing. Thirty SNPs were detected in the actin and CHH genes in broodstock of three *M. rosenbergii* strains (Dong Nai, Mekong, and Hawaiian) and their offspring. A preliminary study that evaluated all SNPs suggested an association between SNPs in intron 3 of the CHH gene and individual growth performance. A larger number of offspring (N = 243) then were genotyped and tested for the associations between SNPs in intron 3 and individual growth performance. Four intronic SNPs were associated significantly with growth traits (body weight, carapace length, and standard length). Of these, CH3-2402 and CH3-2561 were highly correlated with all three traits, while CH3-2407 and CH3-2409 were correlated significantly only with body weight. A further haplotype-trait association analysis confirmed that these four SNP markers were in linkage disequilibrium, and the specific haplotype TGAA had significant associations with high growth ($P < 0.01$). The findings indicate the possibility to developing improved culture lines of *M. rosenbergii* (Thanh *et al*., 2010).

A particular gene product may have multiple functions in physiological processes. Recently, polymorphism of an ATP-dependent DNA helicase gene, *RuvB-like 2* in correlated with body weight and its expression during ovarian development of the giant tiger shrimp (*Penaeus monodon*) was examined. The full length cDNA of *P. monodon RuvB-like 2* (*Pm-RuvBL2*) was 3791 bp in length containing an ORF of 1395 bp corresponding to a polypeptide of 464 amino acids. Polymorphism of the

amplified *Pm-RuvBL2* gene segment (484 bp containing an intron of 259 bp) was initially examined in wild *P. monodon* (N = 15) by single strand conformational polymorphism (SSCP) analysis. Subsequently, correlation between *Pm-RuvBL2* genotypes and the body weight of *P. monodon* was tested using commercially cultivated shrimp from the same pond (approximately 3 months old, average BW = 17.39 ±4.36 g, N = 359). Disregarding sexes, the body weight of shrimp carrying genotypes A (average BW = 19.277 ±3.640 g, N = 37) and B (average BW = 19.293 ±4.548 g, N = 79) was significantly greater than that of shrimp carrying C (average BW = 16.528 ±3.847 g, N = 93) and D (average BW = 16.365 ±4.378 g, N = 124) genotypes. One exonic (G-A_{81}) and two intronic (A-T_{196} and G-T_{248}) SNPs corresponding to ATG, A[T/A]G, GAG and GAT for respective SSCP genotypes were found. *Pm-RuvBL2* was more abundantly expressed in ovaries than testes and other tissues of *P. monodon* broodstock ($P < 0.05$). The relative expression levels of *Pm-RuvBL2* in ovaries of juvenile shrimp carrying different SSCP genotypes were not significantly different ($P < 0.05$). Interestingly, expression of *Pm-RuvBL2* in ovaries of *P. monodon* juveniles was significantly up-regulated upon 5-HT administration (50 μg/g body weight, $P < 0.05$). The expression level of *Pm-RuvBL2* in ovaries of broodstock was greater than that of juvenile shrimp ($P < 0.05$). Results critically suggested that *Pm-RuvBL2* should play an important role in growth and reproduction of *P. monodon* (Prasertlux *et al*., 2010).

It is not possible to obtain single-parent families of *P. monodon* that have been domesticated for several generations at present. Accordingly, juvenile and broodstock-sized *P. monodon* was collected from the early generation of domesticated shrimp with unknown pedigrees. Microsatellite analysis indicated that 14-month-old shrimp used in this study was established from a single pair of parents but 3-month-old shrimp were composed of a few genetically related families. Therefore, the power of the test for identification of growth-related markers in juvenile *P. monodon* was reduced due to the discontinuous distribution of the body weight among different shrimp families.

In this study, correlation between SNP through SSCP genotypes and the body weight the cultured juvenile *P. monodon* (3 months old shrimp, the average body weight of 13.32 ± 2.43 g, N = 340) were examined. Initially, only shrimp exhibiting

approximately 16% from the top (N = 54, the average body weight of 16.78 ± 2.23 g) and the bottom (N = 56, the average body weight of 10.43 ± 1.08 g) according to the body weight were examined.

THBP and *EGF* were polymorphic whereas *DDPG* and *IDE* were monomorphic. Correlation of SNP through SSCP of *THBP* and *EGF* and the body weight of 340 juvenile individuals were examined. Significant result was observed in *THBP* ($P < 0.05$) but not in *EGF* ($P > 0.05$). It should be noted that the average body weight of shrimp carrying SSCP genotype A and B of *EGF* seem to be different but non-significant resulted were form large standard deviation with each group.

Eight SNP positions ($A \rightarrow T_{80}$, $T \rightarrow C_{137}$, $G \rightarrow A_{149}$, $A \rightarrow G_{306}$, $A \rightarrow G_{312}$, $T \rightarrow C_{382}$, $A \rightarrow G_{422}$ and $A \rightarrow G_{494}$) were unambiguously differentiated 4 SSCP genotypes of *THBP*. Likewise, two SNP positions ($A \rightarrow C_{32}$ and $C \rightarrow A_{75}$) were unambiguously differentiated 2 SSCP genotypes of *EGF*.

Recently, the full length cDNA of *P. monodon cellular thyroid hormone-binding protein* was successfully characterized. It was 2220 bp in length composing of an ORF of 1506 bp corresponding to a polypeptide of 501 amino acids and the 5′ and 3′UTRs of 71 and 643 bp, respectively (Preechaphol, 2008).

Analysis of gene-based SNP is one of the efficient approaches for discovery of genes which are significantly contributed in complex traits of *P. monodon*. Although larger sample sizes are required for association analysis of SNP in functional important genes and age-specific growth rates, this preliminary study demonstrated the possibility to a lovus responding for quantitative traits of *P. monodon* where the information on correlations of genotypes and phenotypes through genetic linkage maps in this species are not available at present.

However, *P. monodon* has not yet domesticated therefore specimens from the same families and/or lineages are not available at present. Screening of various gene homologues were then carried out using juvenile and broodstock-sized shrimp from a few families cultured in the same cultured environments. Therefore, correlation between SNP in the *THBP* gene and protein and age-specific growth rates of *P. monodon* should be confirmed using a growth-selected stock.

In addition, the correlation between SSCP genotypes of *THBP*, *EGF DDPG* and *IDE* and its expression levels in ovaries of cultured 3-month-old shrimp (phenotype) was subsequently examined by quantitative real-time PCR.

Characterization of the cDNA sequence of *insulin degrading enzyme (IDE)* of *P. monodon*

Insulin degrading enzyme (IDE) also known as insulysin, is a zinc metalloprotease first described based on its ability to cleave insulin. Mutations in *IDE* are linked to a type II diabetic phenotype in the GK rat. In addition *IDE* has been shown to play a key role in Alzheimer's disease in that it is one of the major enzymes responsible for amyloid β peptide (Aβ) clearance in the brain (Song *et al*., 2010).

Several different transcripts of the *IDE* gene have been reported in different organisms and different tissues, suggesting that multiple proteins may be produced by this gene and it may be differentially expressed and developmentally regulated. In rat, two transcripts of 3.7 and 5.5 kb were found in various tissues, consistent with other studies and consistent with the production of at least two *IDE*-related proteins but three different transcripts were found in testis (3.7, 4.1, and 6.1 kb). The 4.1- and 6.1-kb transcripts correlated with testis-specific gene activation and sperm cell differentiation, supporting a role for *IDE*-related proteins in differentiation and development in specific tissues (Baumeister *et al*., 1993).

The partial cDNA sequences of *IDE* combined from EST and RACE-PCR sequences was 3269 bp in length containing an ORF of 2727 bp corresponding to 908 amino acids and significantly matched *insulin degrading enzymes* of the zebrafish (*Danio rerio*) with the *E*-value of 0.0. Although the full length cDNA of *P. monodon IDE* was not successfully characterized, SNP in this gene could be further determined and tested for correlation with its expression levels in fast and slow-growing juveniles cultured in the same pond.

Association analysis between SNP through SSCP patterns of functionally important genes of *P. monodon* and their expression levels

Correlation between expression levels of *NASP* in ovaries of 14 months old shrimp in shrimp exhibiting different GSI values; group I (< 0.5, $N = 6$), group II (0.5

- 1.0, $N = 12$) and group III (> 1.0, $N = 12$) were not significantly different ($P > 0.05$). In contrast, the expression level of *LTB4DH* in ovaries of group II shrimp was greater that that of group I ($P < 0.05$) but not significantly different from that of group III shrimp ($P > 0.05$).

LTB4DH is the key enzymes responsible for biological inactivation of prostaglandins and related eicosanoids. Prasertlux (2006) identified and characterized *LTB4DH* of *P. monodon*. The full length cDNA of *LTB4DH* was 2485 bp in length with the ORF of 1038 bp encoding a polypeptide of 345 amino acids. *LTB4DH* was preferentially expressed in ovaries than testes of *P. monodon* broodstock. Expression of *LTB4DH* in ovaries of normal *P. monodon* broodstock was temporally lowered at the stage II (vitellogenic stage, $P < 0.05$), returned to the normal level at the stage III (early cortical rod stage, $P > 0.05$) and significantly down-regulated again at the stage IV of ovarian development (mature stage, $P < 0.05$). In eyestalk-ablated female broodstock, *LTB4DH* was significantly up-regulated at the stage IV of ovarian development ($P < 0.05$). More importantly, relative expression levels of *LTB4DH* in each ovarian stage of eyestalk-ablated females were lower than those in normal *P. monodon* females ($P < 0.05$). This strongly suggested that eyestalk ablation potentially suppresses expression of *LTB4DH* possible leading to the activation of prostaglandins and related eicosanoids during ovarian maturation of *P. monodon*.

The expression levels of *IDE*, *EGF* and *THBP* in hepatopancreas of fast- ($N = 22$) and slow-growing ($N = 22$) 3-month-old juveniles were not significantly ($P > 0.05$). Non-significantly different levels of these transcripts suggested that the steady state amounts of these mRNA may be sufficient to maintain their translations in hepatopancreas of *P. monodon*. It is interesting to further examined the correlation between SNP and the expression level of *IDE*, *EGF* and *THBP* proteins in the appropriate sample set of *P. monodon*.

The information on genome-wide sequences is important for a large scale studies about correlations between genotypes and commercially important phenotypes in shrimp. A large number of gene homologues in *P. monodon* identified by an EST approach (Tassanakajon *et al*., 2006; Preechaphol *et al*., 2007; Leelatanawit *et al*., 2009) provide gene catalogues for further polymorphic studies of SNPs and indels in EST-derived markers. The availability of appropriate families with a particular

selected trait of domesticated *P. monodon* will allow the direct application of phenotype-associated SNPs to assist genetic selection of this economically important species.

CHAPTER V

CONCLUSIONS

1. 5′UTR of *NASP* (670 bp) and *LTB4DH* (767 bp) was successfully identified by genome walking analysis. In addition, the partial cDNA sequence of *IDE* was further identified by 3′RACE-PCR and it was 3269 bp in length containing an ORF of 2727 bp corresponding to 908 amino acids.

2. Nucleotide polymorphism in reproduction-related genes (*NASP* and *LTB4DH*) was examined by SSCP analysis against genomic DNA of 14-month-old *P. monodon* (N = 66). *NASP* was monomorphic whereas 2 SSCP patterns were observed in *LTB4DH*. DNA sequencing revealed that these SSCP genotypes could be differentiated by 7 SNPs and 3 indels. For nucleotide polymorphism in growth-related genes (*DDPG*, *IDE*, *EGF* and *THBP*), only one SSCP pattern was found in *DDPG* and *IDE* (N = 110). In contrast, two and four patterns were observed in *EGF* and *THBP*, respectively (N = 340). Two and eight SNPs were found between different genotypes of these gene segments, respectively.

3. Association between SNP through SSCP genotypes of *LTB4DH* and the GSI values of 14-month-old *P. monodon* (N = 66) was determined. Non-significant correlation between genotypes and a phenotype was observed ($P > 0.05$).

4. Association between SNP through SSCP genotypes of *EGF* and *THBP* and the body weight of the domesticated juvenile *P. monodon* was determined (N = 340). Significant correlation between genotypes and phenotype was observed in *THBP* ($P < 0.05$) but not in *EGF* ($P > 0.05$).

5. The expression levels of *NASP* in ovaries of 14-month-old shrimp exhibiting different GSI values were not significantly different ($P > 0.05$). In contrast, *LTB4DH* in ovaries of group II (GSI = 0.5 – 1.0%, N = 12) was significantly greater than that of group I (GSI < 0.5%, N = 6; $P < 0.05$) but not different from that of group III (GSI > 1.0%, N = 12; $P > 0.05$) shrimp. However, the expression levels of ovarian *LTB4DH*

in shrimp exhibiting SSCP genotypes A and B were not significantly different ($P > 0.05$).

6. The expression levels of *IDE*, *EGF* and *THBP* in hepatopancreas of 3-month-old shrimp exhibiting fast (large, $N = 22$) and slow (small, $N = 22$) growth rates were not significantly different ($P > 0.05$). Likewise, the expression levels of these genes in hepatopancreas of juveniles exhibiting different SSCP genotypes were not significantly different ($P > 0.05$).

7. Polymorphism of *THBP* may be applied as DNA markers for selection of fast growing families of domesticated *P. monodon*. However, correlation between SNP in the *THBP* gene and age-specific growth rates of *P. monodon* should be confirmed using a growth-selected stock.

REFERENCES

Aoki-Suzuki, M., Yamada, K., Meerabux, J., Iwayama-Shigeno, Y., Ohba, H., Iwamoto, K., Takao, H., Toyota, T., Suto, Y., Nakatani, N., Dean, B., Nishimura, S., Seki, K., Kato, T., Itohara, S., Nishikawa, T. and Yoshikawa, T. 2005. A family-based association study and gene expression analyses of netrin-G1 and -G2 genes in schizophrenia. *Biol Psychiatry*. 57: 382-393.

Bauchau, A. G. 1981. *Crustaceans. In: Ratcliffe N.A., and Rowley, A.F. (editors). Invertebrate blood cells*. London and New York : Academic Press.

Baumeister, H., Miiller, D., Rehbein, M. and Richter, D. 1993. The rat insulin-degrading enzyme. *FEBS Lett*. 317: 250-254.

Benzie, J. A. H. 1998. Penaeid genetics and biotechnology. *Aquaculture*. 164: 23-47.

Browdy, C. L. 1998. Recent developments in penaeid broodstock and seed production technologies: improving the outlook for superior captive stocks. *Aquaculture*. 164: 3-21.

Brusca, R. C. and Brusca, G. J. 1990. *Invertebrates*. Massachusetts : Sinauer Associates.

Buaklin, A. 2005. *Identification of SNP markers in the giant tiger shrimp Penaeus monodon*. Master's Thesis, Chulalongkorn University.

FAO, *Fishery Statistic* [Online]. 2009. Available from: http://www.fao.org/fishery/statistics/en [2009, February 20].

Frohman, M. A., Dush, M. K. and Martin, G. R. 1988. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc Natl Acad Sci*. 85: 8998-9002.

Glenn, K. L., Grapes, L., Suwanasopee, T., Harris, D. L., Li, Y., Wilson, K. and Rothschild, M. F. 2005. SNP analysis of AMY2 and CTSL genes in *Litopenaeus vannamei* and *Penaeus monodon* shrimp. *Anim Genet*. 36: 235-236.

Gross, P. S., Bartlett, T. C., Browdy, C. L., Chapman, R. W. and Warr, G. W. 2001. Immune gene discovery by expressed sequence tag analysis of hemocytes and hepatopancreas in the Pacific white shrimp, *Litopenaeus vannamei* and the Atlantic white shrimp, *L. setiferus*. *Dev Comp Immunol*. 25: 565-577.

Hayes, B., Laerdahl, J. K., Lien, S., Moen, T., Berg, P., Hindar, K., Davidson, W. S., Koop, B. F., Adzhubei, A. and Høyheim, B. 2007. An extensive resource of single nucleotide polymorphism markers associated with Atlantic salmon (*Salmo salar*) expressed sequences. *Aquaculture*. 265: 82-90.

He, C., Chen, L., Simmons, M., Li, P., Kim, S. and Liu, Z. J. 2003. Putative SNP discovery in interspecific hybrids of catfish by comparative EST analysis. *Anim Genet*. 34: 445-448.

He, F., Wen, H. S., Dong, S. L., Shi, B., Chen, C. F., Wang, L. S., Yao, S., Mu, X. J., Zhou, Y. G. 2008. Identification of single nucleotide polymorphism cytochrome P450-c19a and its relation to reproductive traits in Japanese flounder (*Paralichthys olivaceus*). *Fish Physiol Biochem*. 279: 177-181.

Immervoll, T., Loesgen, S., Dütsch, G., Gohlke, H., Herbon, N., Klugbauer, S., Dempfle, A., Bickeböller, H., Becker-Follmann, J., Rüschendorf, F., Saar, K., Reis, A., Wichmann, H. E. and Wjst, M. 2001. Fine mapping and single nucleotide polymorphism association results of candidate genes for asthma and related phenotypes. *Hum Mutat*. 18: 327-336.

Johnson, P. T. 1980. *Histology of the blue crab, Callinectes sapidus*. New York : A model for the Decapoda. Praeger.

Kenway, M., Macbeth, M., Salmon, M., McPhee, C., Benzie, J., Wilson, K. and Knibb W. 2006. Heritability and genetic correlations of growth and survival in black tiger prawn *Penaeus monod*on reared in tanks. *Aquaculture*. 259: 138-145.

Kirby, L. 1992. *DNA extraction. In: DNA Fingerprinting*. New York : W.H. Freeman and Company.

Klinbunga, S. 1996. *Genetic variation and population structure of the giant tiger shrimp (Penaeus monodon Fabricius)*. Doctoral dissertation, University of Stirling, Scotland.

Klinbunga, S., Penman, D. J., McAndrew, B. J. and Tassanakajon, A. 1999. Mitochondrial DNA diversity in three populations of the giant tiger shrimp, *P. monodon*. *Mar Biotechnol*. 1: 113-121.

Klinbunga, S., Sittikankaew, K., Yuvanatemiya, V., Preechaphol, R., Prasertlux, S., Yamano, K. and Menasveta, P. 2009. Molecular cloning and expression analysis of ovary-specific transcript 1 (Pm-OSTI) of the giant tiger shrimp, *Penaeus monodon*. *Zoolog Sci*. 26: 783-790.

Laurent, P. and Rodellar, C. 2001. Characterization of a single nucleotide polymorphism in the coding sequence of the bovine transferrin gene. *Muta Res*. 458: 1-5.

Leelatanawit, R., Sittikankeaw, K., Yocawibun, P., Klinbunga, S., Roytrakul, S., Aoki, T., Hirono, I. and Menasveta, P. 2009. Identification, characterization and expression of sex-related genes in testes of the giant tiger shrimp *Penaeus monodon*. *Comp Biochem Physiol A Mol Integr Physiol*. 152: 66-76.

Liang, Y., Cui, J., Yang, G., Leung, F. C. and Zhang, X. 2005. Polymorphisms of 5' flanking region of chicken prolactin gene. *Domest Anim Endocrinol*. 30: 1-16.

Limsuwan, C. 2004. *Diseases of Pacific white shrimp (Litopenaeus vannamei) cultured in Thailand*. Proceeding of the JSPS-NRCT International Symposium Joint Seminar 2004: Management of Food Safety in Aquaculture and HACCP. pp. 36-41, Kasetsart University.

Liu, Z. J. and Cordes, J. F. 2004. DNA marker technologies and their applications in aquaculture genetics. *Aquaculture*. 238: 1-37.

Lyons, R. E. and Li, Y. 2002. *Crustacean genomics: implications for the future in aquaculture*. AgBiotechNet. 4: 1-6.

Lyons, R. E., Dierens, L. M., Tan, S. H., Preston, N. P. and Li, Y. 2007. Characterization of AFLP markers associated with growth in the Kuruma prawn, *Marsupenaeus japonicus*, and identification of a candidate gene. *Mar Biotechnol*. 9: 712-721.

Maxam, A. M. and Gilbert, W. 1977. A new method for sequencing DNA. *Proc Natl Acad Sci*. 74: 560-564.

Morganti, M., Ciantelli, M., Giglioni, B., Putignano, A. L., Nobili, S., Papi, L., Landini, I., Napoli, C., Valanzano, R., Cianchi, F., Boddi, V., Tonelli, F., Cortesini, C., Mazzei, T., Genuardi, M. and Mini, E. 2005. Relationships between promoter polymorphisms in the thymidylate synthase gene and mRNA levels in colorectal cancers. *Eur J Cancer*. 41: 2176-2183.

Mullis, K. B. and Faloona, F. A. 1987. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol*. 155: 335-350.

Orita, M., Iwahana, H., Kanazawa, H., Hayashi, K., and Sekiya, T. 1989. Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Proc Natl Acad Sci*. 868: 2766-2770.

Prasertlux, S. 2006. *Identification of single nucleotide polymorphism of genes in ovaries and testes of the giant tiger shrimp Penaeus monodon*. Master's Thesis, Chulalongkorn University.

Prasertlux, S., Khamnamtong, B., Chumtong, P., Klinbunga, S. and Menasveta, P. 2010. Expression levels of *RuvBL2* during ovarian development and association between its single nucleotide polymorphism (SNP) and growth of the giant tiger shrimp *Penaeus monodon*. *Aquaculture* (in press).

Preechaphol, R., Leelatanawit, R., Sittikankeaw, K., Klinbunga, S., Khamnamtong, B., Puanglarp, N. and Menasveta, P. 2007. Expressed sequence tag analysis for identification and characterization of sex-related genes in the giant tiger shrimp *Penaeus monodon*. *Biochem Mol Biol*. 40: 501-510.

Preechaphol, R. 2008. *Identification of genes related to ovarian development of the giant tiger shrimp Penaeus monodon*. Doctoral dissertation, Chulalongkorn University.

Primavera, J. H. 1990. *External and internal anatomy of adult penaeid prawns/shrimps*. SEAFDEC, Aquaculture Department, The Philippines, Poster.

Rojtinnakorn, J., Hirono, I., Itami, T., Takahashi, Y. and Aoki, T. 2002. Gene expression in haemocytes of kuruma prawn, *Penaeus japonicus* in response to infection with WSSV by EST approach. *Fish Shellfish Immunol*. 13: 69-83.

Sambrook, J. and Russell, D. W. 2001. *Molecular Cloning: A Laboratory Manual, 3rd ed*., New York : Cold Spring Harbor Laboratory Press.

Sanger, F., Nicklen, S., and Coulson, A. R. 1977. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci*. 74: 5463-5467.

Shastry, B. 2002. SNP alleles in human disease and evulaution. *Hum Genet*. 47: 561-566.

Shen, Y. Q., Xiang, J. H., Wang, B., Li, F. H. and Tong, W. 2004. Discovery of immune related factors in *Fenneropenaeus chinensis* by annotation of ESTs. *Prog Nat Sci*. 14: 47-54.

Song, E. S., Rodgers, D. W. and Hersh, L. B. 2010. A Monomeric Variant of Insulin Degrading Enzyme (IDE) Loses Its Regulatory Properties. *PLoS One*. 5: 9719-9726.

Supungul, P., Klinbunga, S., Pichyangkura, R., Hirono, I., Aoki, T. and Tassanakajon, A. 2004. Antimicrobial peptides discovered in the black tiger shrimp *Penaeus monodon* using the EST approach. *Dis Aquat Organ*. 61: 123-135.

Tan, F. K., Wang, N., Kuwana, M., Chakraborty, R., Bona, C. A., Milewicz D. M. and Arnett, F. C. 2001. Association of fibrillin 1 single-nucleotide polymorphism haplotypes with systemic sclerosis in Choctaw and Japanese populations. *Arthritis Rheum*. 44: 893-901.

Tao, W. and Boulding, E. 2003. Associations between single nucleotide polymorphisms in candidate genes and growth rate in Arctic charr (*Salvelinus alpinus* L.). *Heredity*. 91: 60-69.

Tassanakajon, A., Klinbunga, S., Paunglarp, N., Rimpanichayakit, V., Udomkit, A., Jitrapakdee, S., Dangtip, K., Pongdara, A., Pongsomboon, S., Supungul, P., Tang, S., Kuphanumart, K, Pichyangkura, R. and Lursinsap, C. 2006. *Penaeus monodon* gene discovery project: generation of an EST collection and establishment of a database. *Gene*. 384: 104-112.

Thanh, N. M., Barnes, A. C., Mather, P. B., Li, Y. and Lyons, R. E. 2010. Single nucleotide polymorphisms in the actin and crustacean hyperglycemic hormone genes and their correlation with individual growth performance in giant freshwater prawn *Macrobrachium rosenbergii*. *Aquaculture*. 301: 7-15.

Thompson, J. D., Higgins, D. G. and Gibson, T. J. 1994. CLUSTAL W improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*. 22: 4673-4680.

Tong, J., Lehnert, S. A., Byrne, K., Kwan, H. S. and Chu, K. H. 2002. Development of polymorphic EST markers in *Penaeus monodon* applications in penaeid genetics. *Aquaculture*. 208: 69-79.

Trakooljul, N., Ponsuksili, S., Schellander, K. and Wimmers, K. 2004. Polymorphisms of the porcine androgen receptor gene affecting its amino acid sequence and expression level. *Biochim Biophys Acta*. 25: 94-101.

Vignal, A., Milan, D., Cristobal, M. S. and Eggen, A. 2002. A review on SNP and other types of molecular markers and their use in animal genetics. *Genet Sel Evol*. 34: 275-305.

VanGuilder, H. D., Vrana, K. E. and Freeman, W. M. 2008. Twenty-five years of quantitative PCR for gene expression analysis. *Biotechniques*. 44: 619-626.

Villette, S., Kyle, J. A. M., Brown, K. M., Pickard, K., Milne, J. S., Nicol, F., Arthur, J. R. and Hesketh, J. E. 2002. A novel single nucleotide polymorphism in the 3' untranslated region of human glutathione peroxidase 4 influences lipoxygenase metabolism. *Academic Press*. 29: 174-178.

Withyachumnarnkul, B., Boonsaeng, V., Flegel, T. W., Panyim, S. and Wongteerasupaya, C. 1998. Domestication and selective breeding of *Penaeus monodon* in Thailand. In Felgel, T. (Ed.), *Proceedings to the Special Session on Advances in Shrimp Biotechnology, The Fifth Asian Fisheries Forum: International Conference on Fisheries and Food Security Beyond the Year 2000*, pp. 73-77. 11-14 November 1998. Thailand: Chiangmai.

Xu, Y. X., Zhu, Z. Y., Lo, L. C., Wang, C. M., Lin, G., Feng, F. and Yue, G. H. 2006. Characterization of two parvalbumin genes and their association with growth traits in Asian seabass (*Lates calcarifer*). *Anim Genet*. 37: 266-268.

APPENDICES

APPENDIX A

Table A1 Raw data of 3-month-old shrimp of *P. monodon* and SSCP genotypes of growth related genes.

No.	Name	Sex	Weight (g)	Length (cm)	Size	SSCP genotypes			
						THBP	***DDPG***	***IDE***	***EGF***
1	BUM03-001	F	16.23	13.0	medium	A	-	-	B
2	BUM03-002	F	14.96	14.0	medium	A	-	-	A
3	BUM03-003	F	16.54	13.1	medium	D	-	-	A
4	BUM03-004	F	11.31	11.0	small	A	A	A	A
5	BUM03-005	F	12.03	11.5	small	A	A	A	A
6	BUM03-006	F	13.76	12.2	medium	D	-	-	A
7	BUM03-007	F	16.90	13.0	medium	A	-	-	A
8	BUM03-008	F	16.28	12.8	medium	D	-	-	A
9	BUM03-009	F	13.61	12.1	medium	D	-	-	A
10	BUM03-010	F	11.70	11.5	small	B	A	A	A
11	BUM03-011	F	14.43	12.2	medium	-	-	-	-
12	BUM03-012	F	12.00	11.5	small	D	A	A	A
13	BUM03-013	F	10.12	11.0	small	C	A	A	A
14	BUM03-014	F	15.18	12.5	medium	A	-	-	B
15	BUM03-015	F	13.39	12.0	medium	A	-	-	B
16	BUM03-016	F	12.93	11.8	medium	C	-	-	B
17	BUM03-017	F	17.88	13.0	large	C	A	A	A
18	BUM03-018	F	12.00	11.5	medium	D	-	-	A
19	BUM03-019	F	17.31	13.0	large	A	A	A	A
20	BUM03-020	F	12.30	11.7	medium	A	-	-	A
21	BUM03-021	F	17.88	13.0	large	C	A	A	A
22	BUM03-022	F	13.17	12.0	medium	C	-	-	A
23	BUM03-023	F	16.90	13.0	large	A	A	A	A
24	BUM03-024	F	16.83	13.0	medium	A	-	-	A
25	BUM03-025	F	16.63	12.7	medium	C	-	-	B
26	BUM03-026	F	12.68	11.8	medium	A	-	-	A

126

Table A1 (Cont.)

No.	Name	Sex	Weight (g)	Length (cm)	Size	SSCP genotypes			
						THBP	***DDPG***	***IDE***	***EGF***
27	BUM03-027	F	11.64	11.4	small	A	A	A	A
28	BUM03-028	F	16.05	12.5	medium	A	-	-	B
29	BUM03-029	F	13.94	12.2	medium	C	-	-	B
30	BUM03-030	F	11.03	11.3	small	D	A	A	A
31	BUM03-031	F	11.20	11.3	small	D	A	A	A
32	BUM03-032	F	15.34	12.4	medium	D	-	-	A
33	BUM03-033	F	17.90	13.0	large	A	A	A	A
34	BUM03-034	F	17.51	13.0	large	D	A	A	A
35	BUM03-035	F	11.79	11.4	small	D	A	A	A
36	BUM03-036	F	14.21	12.1	medium	A	-	-	A
37	BUM03-037	F	12.31	11.7	medium	D	-	-	A
38	BUM03-038	F	11.96	11.6	small	A	A	A	A
39	BUM03-039	F	11.02	11.2	small	D	A	A	A
40	BUM03-040	F	15.18	12.7	medium	C	-	-	A
41	BUM03-041	F	15.45	12.6	medium	C	-	-	A
42	BUM03-042	F	24.33	14.1	large	C	A	A	A
43	BUM03-043	F	15.47	12.5	medium	D	-	-	A
44	BUM03-044	F	16.91	13.5	large	A	A	A	A
45	BUM03-045	F	16.85	13.0	large	C	A	A	A
46	BUM03-046	F	15.37	12.5	medium	C	-	-	A
47	BUM03-047	F	14.16	12.3	medium	A	-	-	A
48	BUM03-048	F	15.21	12.4	medium	D	-	-	A
49	BUM03-049	F	14.91	12.3	medium	C	-	-	A
50	BUM03-050	F	13.34	11.9	medium	B	-	-	A
51	BUM03-051	F	15.64	12.5	medium	D	-	-	A
52	BUM03-052	F	13.12	11.7	medium	D	-	-	A
53	BUM03-053	F	11.82	11.5	small	D	A	A	A
54	BUM03-054	F	11.77	11.3	small	C	A	A	B
55	BUM03-055	F	7.45	10.0	small	D	A	A	A
56	BUM03-056	F	13.15	12.0	medium	D	-	-	A

127

Table A1 (Cont.)

No.	Name	Sex	Weight (g)	Length (cm)	Size	SSCP genotypes			
						THBP	***DDPG***	***IDE***	***EGF***
57	BUM03-057	F	11.24	11.4	small	A	A	A	A
58	BUM03-058	F	14.00	12.0	medium	C	-	-	A
59	BUM03-059	F	14.16	12.3	medium	D	-	-	A
60	BUM03-060	F	16.23	12.9	medium	C	-	-	A
61	BUM03-061	F	13.44	12.1	medium	C	-	-	A
62	BUM03-062	F	18.93	13.0	large	C	A	A	A
63	BUM03-063	F	14.03	12.3	medium	C	-	-	A
64	BUM03-064	F	13.30	12.0	medium	D	-	-	B
65	BUM03-065	F	15.07	12.5	medium	C	-	-	B
66	BUM03-066	F	14.88	12.5	medium	A	-	-	A
67	BUM03-067	F	12.14	11.8	medium	C	-	-	A
68	BUM03-068	F	13.76	12.3	medium	C	-	-	B
69	BUM03-069	F	13.71	12.0	medium	D	-	-	A
70	BUM03-070	F	13.99	12.0	medium	A	-	-	B
71	BUM03-071	F	12.06	11.6	medium	D	-	-	B
72	BUM03-072	F	8.77	10.5	small	D	A	A	A
73	BUM03-073	F	13.03	12.0	medium	A	-	-	A
74	BUM03-074	F	23.47	14.5	large	A	A	A	A
75	BUM03-075	F	20.45	13.6	large	C	A	A	A
76	BUM03-076	F	15.23	12.4	medium	A	-	-	A
77	BUM03-077	F	14.09	12.5	medium	C	-	-	B
78	BUM03-078	F	15.46	12.0	medium	D	-	-	A
79	BUM03-079	F	11.74	11.6	small	A	A	A	A
80	BUM03-080	F	13.65	11.8	medium	D	-	-	A
81	BUM03-081	F	14.05	12.1	medium	A	-	-	A
82	BUM03-082	F	13.60	11.9	medium	D	-	-	A
83	BUM03-083	F	11.88	12.0	small	D	A	A	A
84	BUM03-084	F	12.40	11.9	medium	D	-	-	A
85	BUM03-085	F	13.72	12.1	medium	A	-	-	A
86	BUM03-086	F	13.30	12.0	medium	D	-	-	A

128

Table A1 (Cont.)

No.	Name	Sex	Weight (g)	Length (cm)	Size	SSCP genotypes			
						THBP	***DDPG***	***IDE***	***EGF***
87	BUM03-087	F	13.92	12.0	medium	C	-	-	B
88	BUM03-088	F	12.10	11.6	medium	D	-	-	A
89	BUM03-089	F	12.71	11.6	medium	C	-	-	B
90	BUM03-090	F	11.69	11.5	small	C	A	A	B
91	BUM03-091	F	16.00	12.5	medium	D	-	-	A
92	BUM03-092	F	13.30	12.0	medium	A	-	-	A
93	BUM03-093	F	11.86	11.6	small	A	A	A	A
94	BUM03-094	F	16.64	13.1	medium	A	-	-	A
95	BUM03-095	F	13.07	11.8	medium	C	-	-	B
96	BUM03-096	F	14.92	12.3	medium	A	-	-	A
97	BUM03-097	F	15.36	12.5	medium	C	-	-	A
98	BUM03-098	F	12.90	11.9	medium	A	-	-	A
99	BUM03-099	F	12.78	11.6	medium	D	-	-	A
100	BUM03-100	F	12.22	11.9	medium	A	-	-	B
101	BUM03-101	F	17.10	13.0	large	C	A	A	A
102	BUM03-102	F	11.82	11.7	small	D	A	A	A
103	BUM03-103	F	11.68	11.4	small	C	A	A	B
104	BUM03-104	F	11.73	11.5	small	D	A	A	A
105	BUM03-105	F	14.22	12.4	medium	D	-	-	A
106	BUM03-106	F	14.56	12.5	medium	D	-	-	A
107	BUM03-107	F	10.50	11.1	small	D	A	A	A
108	BUM03-108	F	12.40	11.8	medium	D	-	-	A
109	BUM03-109	F	18.44	13.4	large	D	A	A	A
110	BUM03-110	F	14.52	12.3	medium	A	-	-	A
111	BUM03-111	F	11.74	11.5	small	A	A	A	B
112	BUM03-112	F	12.95	11.9	medium	D	-	-	A
113	BUM03-113	F	17.33	13.2	large	C	A	A	B
114	BUM03-114	F	18.30	13.3	large	C	A	A	B
115	BUM03-115	F	15.91	13.0	medium	D	-	-	A
116	BUM03-116	F	16.17	13.0	medium	C	-	-	A

129

Table A1 (Cont.)

No.	Name	Sex	Weight (g)	Length (cm)	Size	SSCP genotypes			
						THBP	***DDPG***	***IDE***	***EGF***
117	BUM03-117	F	16.28	12.8	medium	C	-	-	B
118	BUM03-118	F	21.02	13.6	large	C	A	A	A
119	BUM03-119	F	19.19	13.4	large	A	A	A	A
120	BUM03-120	F	18.14	13.0	large	C	A	A	A
121	BUM03-121	F	17.18	12.9	large	D	A	A	A
122	BUM03-122	F	18.34	13.1	large	C	A	A	A
123	BUM03-123	F	17.18	13.1	large	D	A	A	A
124	BUM03-124	F	16.06	12.5	medium	C	-	-	B
125	BUM03-125	F	17.15	12.8	large	C	A	A	B
126	BUM03-126	F	13.54	11.8	medium	A	-	-	A
127	BUM03-127	F	15.53	12.8	medium	D	-	-	A
128	BUM03-128	F	15.75	12.6	medium	D	-	-	A
129	BUM03-129	F	13.36	12.0	medium	A	-	-	A
130	BUM03-130	F	12.61	11.8	medium	D	-	-	A
131	BUM03-131	F	16.51	13.0	medium	C	-	-	A
132	BUM03-132	F	13.46	12.0	medium	D	-	-	A
133	BUM03-133	F	15.95	12.5	medium	C	-	-	A
134	BUM03-134	F	18.25	13.5	large	A	A	A	B
135	BUM03-135	F	15.75	12.6	medium	A	-	-	B
136	BUM03-136	F	15.30	12.3	medium	C	-	-	A
137	BUM03-137	F	13.48	12.4	medium	A	-	-	A
138	BUM03-138	F	15.34	12.3	medium	A	-	-	B
139	BUM03-139	F	14.97	12.5	medium	A	-	-	A
140	BUM03-140	F	17.34	13.2	large	A	A	A	B
141	BUM03-141	F	15.39	12.8	medium	A	-	-	A
142	BUM03-142	F	12.68	11.9	medium	D	-	-	A
143	BUM03-143	F	14.08	12.5	medium	D	-	-	B
144	BUM03-144	F	15.17	12.3	medium	C	-	-	A
145	BUM03-145	F	15.77	13.0	medium	A	-	-	A
146	BUM03-146	F	13.55	12.1	medium	A	-	-	A

130

Table A1 (Cont.)

No.	Name	Sex	Weight (g)	Length (cm)	Size	SSCP genotypes			
						THBP	***DDPG***	***IDE***	***EGF***
147	BUM03-147	F	16.03	12.8	medium	A	-	-	A
148	BUM03-148	F	15.38	13.0	medium	A	-	-	B
149	BUM03-149	F	15.06	12.7	medium	A	-	-	A
150	BUM03-150	F	13.16	12.0	medium	D	-	-	A
151	BUM03-151	M	11.75	12.8	medium	C	-	-	A
152	BUM03-152	M	14.73	12.3	large	C	A	A	A
153	BUM03-153	M	10.99	12.5	medium	D	-	-	B
154	BUM03-154	M	10.53	11.6	medium	C	-	-	A
155	BUM03-155	M	12.45	11.8	medium	D	-	-	A
156	BUM03-156	M	12.46	11.8	medium	A	-	-	A
157	BUM03-157	M	12.50	11.8	medium	A	-	-	A
158	BUM03-158	M	14.78	12.5	large	C	A	A	A
159	BUM03-159	M	12.05	11.8	medium	A	-	-	A
160	BUM03-160	M	9.60	10.9	medium	C	-	-	A
161	BUM03-161	M	16.88	13.2	large	A	A	A	A
162	BUM03-162	M	9.99	11.2	small	D	A	A	A
163	BUM03-163	M	13.64	12.3	medium	A	-	-	B
164	BUM03-164	M	15.27	12.6	large	A	A	A	A
165	BUM03-165	M	14.61	12.7	large	A	A	A	A
166	BUM03-166	M	12.08	11.8	medium	A	-	-	B
167	BUM03-167	M	12.00	11.8	medium	C	-	-	B
168	BUM03-168	M	11.91	11.6	medium	A	-	-	A
169	BUM03-169	M	12.42	12.0	medium	A	-	-	A
170	BUM03-170	M	13.83	12.2	medium	D	-	-	A
171	BUM03-171	M	15.29	12.8	medium	D	-	-	A
172	BUM03-172	M	14.50	12.3	large	A	A	A	A
173	BUM03-173	M	14.55	12.5	large	C	A	A	A
174	BUM03-174	M	15.19	12.3	large	C	A	A	A
175	BUM03-175	M	14.18	12.5	large	D	A	A	A
176	BUM03-176	M	14.46	12.5	large	A	A	A	A

131

Table A1 (Cont.)

No.	Name	Sex	Weight (g)	Length (cm)	Size	SSCP genotypes			
						THBP	***DDPG***	***IDE***	***EGF***
177	BUM03-177	M	10.47	11.2	small	C	A	A	A
178	BUM03-178	M	10.68	11.2	medium	D	-	-	A
179	BUM03-179	M	9.98	11.1	small	D	A	A	A
180	BUM03-180	M	14.20	12.2	large	A	A	A	A
181	BUM03-181	M	11.65	11.6	medium	D	-	-	A
182	BUM03-182	M	14.65	12.5	large	D	A	A	A
183	BUM03-183	M	12.96	12.0	medium	D	-	-	A
184	BUM03-184	M	13.86	12.0	medium	C	-	-	B
185	BUM03-185	M	9.37	11.0	small	B	A	A	A
186	BUM03-186	M	12.72	11.9	medium	D	-	-	B
187	BUM03-187	M	11.58	11.8	medium	A	-	-	A
188	BUM03-188	M	11.31	11.3	medium	D	-	-	A
189	BUM03-189	M	10.91	11.5	medium	D	-	-	B
190	BUM03-190	M	10.84	11.3	medium	A	-	-	B
191	BUM03-191	M	12.58	12.0	medium	A	-	-	B
192	BUM03-192	M	12.50	12.1	medium	C	-	-	A
193	BUM03-193	M	11.41	11.5	medium	D	-	-	B
194	BUM03-194	M	10.44	11.2	small	D	A	A	A
195	BUM03-195	M	13.50	12.3	medium	D	-	-	A
196	BUM03-196	M	16.54	13.0	large	C	A	A	A
197	BUM03-197	M	9.90	10.9	small	C	A	A	A
198	BUM03-198	M	10.31	11.3	small	D	A	A	A
199	BUM03-199	M	11.02	11.4	medium	D	-	-	A
200	BUM03-200	M	14.34	12.2	large	C	A	A	A
201	BUM03-201	M	14.00	12.0	medium	A	-	-	B
202	BUM03-202	M	11.84	11.8	medium	D	-	-	A
203	BUM03-203	M	13.08	12.1	medium	D	-	-	A
204	BUM03-204	M	11.80	11.8	medium	C	-	-	A
205	BUM03-205	M	11.09	11.4	medium	C	-	-	B
206	BUM03-206	M	14.18	12.1	medium	D	-	-	A

132

Table A1 (Cont.)

No.	Name	Sex	Weight (g)	Length (cm)	Size	SSCP genotypes			
						THBP	***DDPG***	***IDE***	***EGF***
207	BUM03-207	M	11.18	11.6	medium	A	-	-	B
208	BUM03-208	M	13.10	12.0	medium	A	-	-	B
209	BUM03-209	M	10.90	11.5	medium	D	-	-	A
210	BUM03-210	M	11.05	11.5	medium	A	-	-	B
211	BUM03-211	M	12.70	11.9	medium	A	-	-	A
212	BUM03-212	M	13.75	12.0	medium	C	-	-	A
213	BUM03-213	M	9.49	10.7	small	D	A	A	A
214	BUM03-214	M	10.70	11.6	medium	C	-	-	B
215	BUM03-215	M	14.16	12.3	medium	D	-	-	A
216	BUM03-216	M	11.11	11.5	medium	C	-	-	B
217	BUM03-217	M	14.56	12.5	large	D	A	A	A
218	BUM03-218	M	13.10	12.1	medium	C	-	-	B
219	BUM03-219	M	9.75	11.0	small	D	A	A	A
220	BUM03-220	M	11.71	11.8	medium	D	-	-	A
221	BUM03-221	M	12.50	11.9	medium	A	-	-	A
222	BUM03-222	M	12.85	12.0	medium	D	-	-	A
223	BUM03-223	M	10.18	11.0	small	C	A	A	A
224	BUM03-224	M	12.92	11.8	medium	A	-	-	A
225	BUM03-225	M	12.26	11.9	medium	D	-	-	A
226	BUM03-226	M	12.70	11.8	medium	D	-	-	A
227	BUM03-227	M	13.31	12.1	medium	D	-	-	A
228	BUM03-228	M	11.79	11.5	medium	C	-	-	B
229	BUM03-229	M	14.65	12.4	large	D	A	A	A
230	BUM03-230	M	11.45	11.4	medium	D	-	-	A
231	BUM03-231	M	14.90	12.4	large	A	A	A	A
232	BUM03-232	M	12.51	11.9	medium	C	-	-	A
233	BUM03-233	M	11.70	11.7	medium	D	-	-	A
234	BUM03-234	M	12.25	11.9	medium	A	-	-	B
235	BUM03-235	M	14.88	12.6	large	A	A	A	B
236	BUM03-236	M	14.28	12.3	large	A	A	A	B

133

Table A1 (Cont.)

No.	Name	Sex	Weight (g)	Length (cm)	Size	SSCP genotypes			
						THBP	***DDPG***	***IDE***	***EGF***
237	BUM03-237	M	10.60	11.4	medium	A	-	-	A
238	BUM03-238	M	13.28	12.2	medium	D	-	-	A
239	BUM03-239	M	10.00	11.0	small	D	A	A	A
240	BUM03-240	M	8.96	10.6	small	C	A	A	A
241	BUM03-241	M	11.65	11.8	medium	A	-	-	A
242	BUM03-242	M	11.83	11.6	medium	D	-	-	A
243	BUM03-243	M	9.80	10.9	small	A	A	A	A
244	BUM03-244	M	10.51	11.3	small	A	A	A	A
245	BUM03-245	M	13.15	12.1	medium	D	-	-	A
246	BUM03-246	M	9.26	10.8	small	A	A	A	B
247	BUM03-247	M	15.33	12.5	medium	C	-	-	A
248	BUM03-248	M	10.55	11.2	medium	A	-	-	A
249	BUM03-249	M	9.10	10.8	small	C	A	A	B
250	BUM03-250	M	14.75	12.3	large	C	A	A	A
251	BUM03-251	M	14.82	12.5	large	A	A	A	A
252	BUM03-252	M	10.94	11.0	medium	A	-	-	B
253	BUM03-253	M	16.89	13.0	large	A	A	A	A
254	BUM03-254	M	13.91	12.1	medium	D	-	-	A
255	BUM03-255	M	12.49	11.8	medium	A	-	-	A
256	BUM03-256	M	13.99	12.2	medium	C	-	-	A
257	BUM03-257	M	12.02	11.9	medium	A	-	-	A
258	BUM03-258	M	12.89	11.9	medium	A	-	-	B
259	BUM03-259	M	11.99	11.8	medium	D	-	-	A
260	BUM03-260	M	10.91	11.5	medium	A	-	-	A
261	BUM03-261	M	11.82	11.8	medium	D	-	-	A
262	BUM03-262	M	8.63	10.4	small	A	A	A	A
263	BUM03-263	M	14.21	12.2	large	D	A	A	A
264	BUM03-264	M	11.77	11.7	medium	C	-	-	A
265	BUM03-265	M	11.45	11.6	medium	A	-	-	A
266	BUM03-266	M	12.93	12.1	medium	D	-	-	A

134

Table A1 (Cont.)

No.	Name	Sex	Weight (g)	Length (cm)	Size	SSCP genotypes			
						THBP	***DDPG***	***IDE***	***EGF***
267	BUM03-267	M	11.18	11.5	medium	D	-	-	A
268	BUM03-268	M	11.45	11.6	medium	D	-	-	A
269	BUM03-269	M	15.45	12.8	large	C	A	A	A
270	BUM03-270	M	12.81	12.0	medium	C	-	-	B
271	BUM03-271	M	13.10	11.9	medium	A	-	-	A
272	BUM03-272	M	9.76	11.0	small	A	A	A	A
273	BUM03-273	M	10.10	11.1	small	D	A	A	A
274	BUM03-274	M	12.70	11.8	medium	C	-	-	B
275	BUM03-275	M	13.05	12.0	medium	D	-	-	A
276	BUM03-276	M	14.31	12.4	large	C	A	A	A
277	BUM03-277	M	10.08	11.0	small	D	A	A	A
278	BUM03-278	M	11.62	11.7	medium	C	-	-	A
279	BUM03-279	M	10.31	11.2	small	C	A	A	A
280	BUM03-280	M	12.86	12.0	medium	D	-	-	A
281	BUM03-281	M	14.05	12.3	medium	C	-	-	A
282	BUM03-282	M	10.75	11.3	medium	D	-	-	B
283	BUM03-283	M	9.62	10.8	small	D	A	A	A
284	BUM03-284	M	10.35	11.1	small	D	A	A	A
285	BUM03-285	M	12.80	11.2	medium	-	-	-	-
286	BUM03-286	M	18.50	13.4	large	C	A	A	A
287	BUM03-287	M	12.89	11.6	medium	C	-	-	A
288	BUM03-288	M	13.54	12.0	medium	A	-	-	B
289	BUM03-289	M	13.57	12.3	medium	A	-	-	A
290	BUM03-290	M	12.11	11.8	medium	A	-	-	A
291	BUM03-291	M	11.61	11.7	medium	D	-	-	A
292	BUM03-292	M	12.24	11.9	medium	C	-	-	A
293	BUM03-293	M	11.70	11.8	medium	A	-	-	A
294	BUM03-294	M	18.18	13.5	large	D	A	A	A
295	BUM03-295	M	12.73	11.9	medium	D	-	-	A
296	BUM03-296	M	8.92	10.8	small	A	A	A	B

135

Table A1 (Cont.)

No.	Name	Sex	Weight (g)	Length (cm)	Size	SSCP genotypes			
						THBP	***DDPG***	***IDE***	***EGF***
297	BUM03-297	M	10.38	11.2	small	D	A	A	A
298	BUM03-298	M	12.49	12.0	medium	D	-	-	A
299	BUM03-299	M	12.19	12.0	medium	D	-	-	A
300	BUM03-300	M	13.23	12.1	medium	D	-	-	A
301	BUM03-301	F	13.37	12.0	medium	D	-	-	A
302	BUM03-302	F	13.35	12.0	medium	B	-	-	A
303	BUM03-303	F	12.27	11.8	medium	D	-	-	B
304	BUM03-304	F	11.80	11.6	medium	A	-	-	B
305	BUM03-305	F	14.37	12.5	medium	D	-	-	A
306	BUM03-306	F	11.08	11.5	medium	D			A
307	BUM03-307	F	12.92	12.0	medium	A	-	-	A
308	BUM03-308	F	12.85	11.8	medium	C	-	-	A
309	BUM03-309	F	13.24	12.0	medium	D	-	-	A
310	BUM03-310	F	13.92	12.2	medium	A	-	-	B
311	BUM03-311	F	12.79	11.4	medium	C	-	-	A
312	BUM03-312	F	13.40	12.3	medium	D	-	-	B
313	BUM03-313	F	8.67	10.5	small	D	A	A	B
314	BUM03-314	F	13.48	11.6	medium	A	-	-	A
315	BUM03-315	F	11.46	11.8	medium	A	-	-	A
316	BUM03-316	F	13.30	12.3	medium	A	-	-	A
317	BUM03-317	F	11.50	11.6	medium	D	-	-	A
318	BUM03-318	F	13.66	12.3	medium	C	-	-	B
319	BUM03-319	F	11.88	11.4	medium	D	-	-	A
320	BUM03-320	F	11.74	11.5	medium	D	-	-	A
321	BUM03-321	M	10.08	11.3	small	D	A	A	A
322	BUM03-322	M	12.05	11.8	medium	D	-	-	A
323	BUM03-323	M	12.08	11.7	medium	A	-	-	A
324	BUM03-324	M	12.49	11.8	medium	A	-	-	B
325	BUM03-325	M	13.51	12.2	medium	D	-	-	A
326	BUM03-326	M	10.88	11.2	medium	C	-	-	A

136

Table A1 (Cont.)

No.	Name	Sex	Weight (g)	Length (cm)	Size	SSCP genotypes			
						THBP	***DDPG***	***IDE***	***EGF***
327	BUM03-327	M	14.08	12.2	medium	D	-	-	B
328	BUM03-328	M	11.76	11.5	medium	D	-	-	A
329	BUM03-329	M	10.94	11.2	medium	D	-	-	A
330	BUM03-330	M	10.08	11.3	small	D	A	A	B
331	BUM03-331	M	11.92	11.6	medium	A	-	-	A
332	BUM03-332	M	11.48	11.3	medium	A	-	-	A
333	BUM03-333	M	11.69	11.4	medium	C	-	-	A
334	BUM03-334	M	12.52	11.8	medium	C	-	-	A
335	BUM03-335	M	9.27	10.8	small	D	A	A	A
336	BUM03-336	M	9.58	10.8	small	D	A	A	A
337	BUM03-337	M	11.23	11.2	medium	A	-	-	B
338	BUM03-338	M	9.81	11.0	small	A	A	A	B
339	BUM03-339	M	12.85	12.0	medium	D	-	-	A
340	BUM03-340	M	12.39	11.8	medium	D	-	-	A
341	BUM03-341	M	16.38	13.0	large	A	A	A	A
342	BUM03-342	M	17.54	13.0	large	D	A	A	A

137

Table A2 Raw data of 14-month-old shrimp of *P. monodon* and SSCP genotypes of reproductive related genes.

No.	Name	Sex	Weight (g)	Length (cm)	Ovary (g)	GSI	Group	SSCP genotypes	
								LTB4DH	*NASP*
1	BU14F-01	F	82.82	20.50	0.88	1.06	GSI-III	B	A
2	BU14F-02	F	82.68	21.00	1.42	1.72	GSI-III	A	A
3	BU14F-03	F	81.02	20.70	1.20	1.48	GSI-III	B	A
4	BU14F-04	F	96.57	22.10	1.17	1.21	GSI-III	A	A
5	BU14F-05	F	104.66	23.70	1.15	1.10	GSI-III	B	A
6	BU14F-06	F	90.42	21.10	1.37	1.52	GSI-III	A	A
7	BU14F-07	F	89.78	21.60	0.82	0.91	GSI-II	A	A
8	BU14F-08	F	80.56	22.00	0.58	0.72	GSI-II	A	A
9	BU14F-09	F	90.32	21.50	0.71	0.79	GSI-II	B	A
10	BU14F-10	F	85.53	22.50	0.83	0.97	GSI-II	A	A
11	BU14F-11	F	86.40	21.00	1.32	1.53	GSI-III	B	A
12	BU14F-12	F	93.32	21.60	1.61	1.73	GSI-III	A	A
13	BU14F-13	F	76.39	20.00	0.27	0.35	GSI-I	A	A
14	BU14F-14	F	99.63	22.10	0.84	0.84	GSI-II	A	A
15	BU14F-15	F	62.18	18.60	0.76	1.22	GSI-III	A	A
16	BU14F-16	F	103.66	21.80	1.44	1.39	GSI-III	A	A
17	BU14F-17	F	83.23	20.60	1.10	1.32	GSI-III	B	A
18	BU14F-18	F	67.32	19.30	0.76	1.13	GSI-III	B	A
19	BU14F-19	F	99.87	21.90	1.63	1.63	GSI-III	B	A
20	BU14F-20	F	80.20	20.60	0.94	1.17	GSI-III	A	A
21	BU14F-21	F	90.10	21.00	1.14	1.27	GSI-III	A	A
22	BU14F-22	F	69.07	20.00	0.60	0.87	GSI-II	A	A
23	BU14F-23	F	77.01	20.40	0.39	0.51	GSI-II	B	A
24	BU14F-24	F	86.10	21.50	1.51	1.75	GSI-III	A	A
25	BU14F-25	F	85.80	21.00	0.25	0.29	GSI-I	B	A
26	BU14F-26	F	59.64	19.30	0.66	1.11	GSI-III	B	A
27	BU14F-27	F	83.64	21.00	0.37	0.44	GSI-I	B	A
28	BU14F-28	F	76.40	20.50	0.56	0.73	GSI-II	A	A
29	BU14F-29	F	96.74	22.00	0.81	0.84	GSI-II	A	A
30	BU14F-30	F	100.95	22.00	1.44	1.43	GSI-III	A	A

Table A2 (Cont.)

No.	Name	Sex	Weight (g)	Length (cm)	Ovary (g)	GSI	Group	SSCP genotypes	
								LTB4DH	*NASP*
31	BU14F-31	F	89.40	21.50	0.56	0.63	GSI-II	A	A
32	BU14F-32	F	95.40	21.60	0.92	0.96	GSI-II	B	A
33	BU14F-33	F	83.21	20.50	0.37	0.44	GSI-I	B	A
34	BU14F-34	F	72.75	12.50	1.29	1.77	GSI-III	A	A
35	BU14F-35	F	78.76	20.20	0.42	0.53	GSI-II	A	A
36	BU14F-36	F	93.58	22.00	0.78	0.83	GSI-II	B	A
37	BU14F-37	F	71.64	19.50	0.65	0.91	GSI-II	A	A
38	BU14F-38	F	65.87	19.50	0.60	0.91	GSI-II	A	A
39	BU14F-39	F	84.19	21.00	0.65	0.77	GSI-II	A	A
40	BU14F-40	F	55.96	18.70	0.58	1.04	GSI-III	A	A
41	BU14F-41	F	101.04	23.50	1.22	1.21	GSI-III	B	A
42	BU14F-42	F	99.47	23.50	0.99	1.00	GSI-II	B	A
43	BU14F-43	F	93.78	21.00	0.61	0.65	GSI-II	B	A
44	BU14F-44	F	99.67	23.00	1.12	1.12	GSI-III	B	A
45	BU14F-45	F	104.21	24.00	0.79	0.76	GSI-II	A	A
46	BU14F-46	F	96.03	21.70	0.82	0.85	GSI-II	A	A
47	BU14F-47	F	88.54	20.20	0.84	0.95	GSI-II	A	A
48	BU14F-48	F	98.93	21.50	0.90	0.91	GSI-II	A	A
49	BU14F-49	F	89.46	21.00	0.68	0.76	GSI-II	B	A
50	BU14F-50	F	84.86	20.70	0.42	0.49	GSI-I	B	A
51	BU14F-51	F	100.66	22.00	0.89	0.88	GSI-II	A	A
52	BU14F-52	F	75.88	20.00	0.57	0.75	GSI-II	A	A
53	BU14F-53	F	110.76	22.80	0.55	0.50	GSI-II	B	A
54	BU14F-54	F	81.07	20.90	0.66	0.81	GSI-II	B	A
55	BU14F-55	F	101.52	22.80	1.01	0.99	GSI-II	A	A
56	BU14F-56	F	86.91	20.10	0.60	0.69	GSI-II	A	A
57	BU14F-57	F	108.16	22.40	0.79	0.73	GSI-II	A	A
58	BU14F-58	F	84.51	21.20	0.71	0.84	GSI-II	A	A
59	BU14F-59	F	82.26	20.90	0.57	0.69	GSI-II	B	A
60	BU14F-60	F	76.34	20.30	0.61	0.80	GSI-II	B	A

139

Table A2 (Cont.)

No.	Name	Sex	Weight (g)	Length (cm)	Ovary (g)	GSI	Group	SSCP genotypes	
								LTB4DH	*NASP*
61	BU14F-61	F	102.48	21.90	0.86	0.84	GSI-II	B	A
62	BU14F-62	F	79.55	20.70	0.28	0.35	GSI-I	B	A
63	BU14F-63	F	115.65	22.30	1.21	1.05	GSI-III	B	A
64	BU14F-64	F	91.35	22.50	1.15	1.26	GSI-III	A	A
65	BU14F-65	F	80.51	21.50	0.54	0.67	GSI-II	A	A
66	BU14F-66	F	76.63	21.00	0.62	0.81	GSI-II	B	A

GSI- I = < 0.5%, GSI-II = 0.5 – 1.0%, GSI-III = > 1.0%

140

APPENDIX B

Table B1 Raw data of standard real-time PCR

Sample	Mean Cp	Mean concentration	Log mean concentration
NADP-dependent-leukotriene B4 12-hydroxydehydrogenase (*LTB4DH*)			
1E8	9.44	97800000	7.990339
1E7	12.80	10165000	7.007107
1E6	16.20	1035000	6.014940
1E5	19.73	97000	4.986772
1E4	22.90	10400	4.017033
1E3	25.99	989.5	2.995416
Nuclear autoantigenic sperm protein (*NASP*)			
1E7	13.20	11727723.96	7.069213
1E6	17.15	810004.02	5.908487
1E5	20.37	91289.84	4.960422
1E4	23.43	11573.84	4.063477
1E3	27.61	822.69	2.915236
1E2	31.37	107.09	2.029762
Elongation factor - 1 alpha (*EF-1α*)			
1E8	9.91	88367251.53	7.946291
1E7	12.79	12314792.62	7.090427
1E6	16.56	936369.35	5.971447
1E5	19.84	98984.31	4.995566
1E4	23.18	10034.94	4.001515
1E3	26.67	1005.56	3.002408
Thyroid hormone binding protein (*THBP*)			
1E8	10.21	100000086.80	8.000000
1E7	13.58	10001144.18	7.000049
1E6	17.01	996909.03	5.998655
1E5	20.59	96919.82	4.986412
1E4	24.17	10144.43	4.006227
Insulin degrading enzyme (*IDE*)			
1E7	12.83	10583429.93	7.024626
1E6	16.48	884092.26	5.946497
1E5	19.59	106883.72	5.028912
1E4	23.46	8866.31	3.947743
1E3	27.21	1032.63	3.013945
Epidermal growth factor (*EGF*)			
1E8	11.08	107196434.00	8.030180
1E7	14.78	8739493.30	6.941486
1E6	17.83	1096103.64	6.039852
1E5	21.37	97861.71	4.990613
1E4	24.80	9445.68	3.975233

Table B2 Raw data of real-time PCR of *EF-1α* of 14-month-old shrimp

Sample	Mean Cp	Mean concentration	GSI
BU14F-13	12.54	22972048.20	I
BU14F-25	12.21	28614311.73	I
BU14F-27	11.75	39429997.76	I
BU14F-33	13.14	15221587.72	I
BU14F-50	12.98	16947504.57	I
BU14F-62	11.85	36688280.30	I
BU14F-14	12.64	21366231.68	II
BU14F-22	12.55	22792386.67	II
BU14F-31	12.21	28622387.78	II
BU14F-32	12.45	24394600.31	II
BU14F-36	12.06	31837268.83	II
BU14F-42	12.33	26393259.15	II
BU14F-47	12.85	18498510.16	II
BU14F-48	12.18	29237545.40	II
BU14F-54	12.28	27415938.61	II
BU14F-55	12.64	21405485.65	II
BU14F-65	11.94	34595007.64	II
BU14F-66	12.71	20388228.66	II
BU14F-11	12.79	19257376.01	III
BU14F-15	13.32	13439034.74	III
BU14F-16	14.26	7030225.79	III
BU14F-17	12.64	21449209.72	III
BU14F-19	11.92	34914304.27	III
BU14F-21	12.67	20982545.00	III
BU14F-24	12.44	24447040.84	III
BU14F-26	13.24	14137092.12	III
BU14F-30	14.29	6912354.58	III
BU14F-34	12.83	18740992.90	III
BU14F-40	12.33	26510179.82	III
BU14F-41	12.44	24570630.43	III

Table B3 Raw data of real-time PCR of *LTB4DH* of 14-month-old shrimp

Sample	Mean Cp	Mean concentration	GSI
BU14F-13	19.63	92049.73	I
BU14F-25	20.48	51428.86	I
BU14F-27	19.97	72893.11	I
BU14F-33	21.25	30202.56	I
BU14F-50	21.26	29954.21	I
BU14F-62	19.68	89062.73	I
BU14F-14	20.28	58982.81	II
BU14F-22	19.88	77775.23	II
BU14F-31	19.60	94115.58	II
BU14F-32	20.64	45847.51	II
BU14F-36	19.14	129142.02	II
BU14F-42	19.57	96085.14	II
BU14F-47	20.05	68922.06	II
BU14F-48	19.24	121215.45	II
BU14F-54	19.77	83863.20	II
BU14F-55	19.87	78008.70	II
BU14F-65	19.03	140085.63	II
BU14F-66	20.09	67168.58	II
BU14F-11	20.05	69090.57	III
BU14F-15	21.28	29508.07	III
BU14F-16	23.56	6119.06	III
BU14F-17	19.69	88659.87	III
BU14F-19	19.38	109593.39	III
BU14F-21	20.52	49985.24	III
BU14F-24	19.49	101515.23	III
BU14F-26	21.05	34615.82	III
BU14F-30	22.36	14025.96	III
BU14F-34	19.86	79005.02	III
BU14F-40	20.17	63803.62	III
BU14F-41	20.26	59732.86	III

Table B4 Raw data of real-time PCR of *NASP* of 14-month-old shrimp

Sample	Mean Cp	Mean concentration	GSI
BU14F-13	15.51	3000807.47	I
BU14F-25	15.43	3169676.73	I
BU14F-27	16.04	2076662.02	I
BU14F-33	16.32	1716566.68	I
BU14F-50	18.01	532505.33	I
BU14F-62	14.91	4536372.34	I
BU14F-14	15.68	2676798.91	II
BU14F-22	15.76	2528086.42	II
BU14F-31	15.75	2539451.82	II
BU14F-32	15.66	2700088.12	II
BU14F-36	15.50	3016505.59	II
BU14F-42	15.99	2153678.41	II
BU14F-47	16.05	2066603.57	II
BU14F-48	15.72	2602647.04	II
BU14F-54	15.66	2705787.02	II
BU14F-55	15.78	2495314.43	II
BU14F-65	15.13	3909064.92	II
BU14F-66	15.81	2430536.36	II
BU14F-11	15.88	2326089.39	III
BU14F-15	18.99	269650.77	III
BU14F-16	22.07	31842.35	III
BU14F-17	15.72	2596145.84	III
BU14F-19	15.01	4234247.66	III
BU14F-21	18.38	410483.62	III
BU14F-24	15.25	3599823.47	III
BU14F-26	18.23	456318.61	III
BU14F-30	19.92	140745.55	III
BU14F-34	15.90	2287624.44	III
BU14F-40	15.31	3442925.95	III
BU14F-41	15.77	2502513.95	III

Table B5 Raw data real-time PCR of *EF-1α* of 3-month-old shrimp

Sample	Mean Cp	Mean concentration	Group
BUM03-004	21.26	153137.78	small
BUM03-057	23.44	34475.78	small
BUM03-249	21.31	147856.99	small
BUM03-074	21.45	133935.64	small
BUM03-134	21.73	110563.51	small
BUM03-286	21.57	123575.55	small
BUM03-010	20.93	191814.10	small
BUM03-107	21.63	118249.05	small
BUM03-262	21.12	168824.35	small
BUM03-075	21.24	154853.79	small
BUM03-161	21.09	172098.05	small
BUM03-294	21.40	139456.15	small
BUM03-013	21.63	118559.95	small
BUM03-313	25.09	11376.29	small
BUM03-283	21.59	121665.10	small
BUM03-109	22.94	48317.92	small
BUM03-164	21.43	136328.71	small
BUM03-341	23.05	44925.31	small
BUM03-030	20.95	188886.35	small
BUM03-185	21.90	98281.17	small
BUM03-296	21.58	122651.62	small
BUM03-114	21.72	111644.12	small
BUM03-174	20.50	256532.94	large
BUM03-342	23.05	44923.62	large
BUM03-031	20.02	357829.06	large
BUM03-213	20.60	239712.20	large
BUM03-335	25.70	7659.80	large
BUM03-118	21.64	117532.69	large
BUM03-182	21.18	161514.92	large
BUM03-039	20.94	190336.20	large
BUM03-219	20.88	198596.14	large
BUM03-336	26.44	4766.25	large
BUM03-119	20.70	224010.69	large
BUM03-196	21.78	106990.58	large
BUM03-055	21.63	118543.01	large
BUM03-240	21.29	149973.88	large
BUM03-042	21.25	156440.35	large
BUM03-120	20.99	183349.61	large
BUM03-253	23.83	26406.08	large
BUM03-072	21.13	166978.24	large
BUM03-246	20.75	217447.16	large
BUM03-062	21.25	153453.87	large
BUM03-122	21.58	122606.36	large
BUM03-269	25.46	8922.36	large

Table B6 Raw data real-time PCR of *THBP* of 3-month-old shrimp

Sample	Mean Cp	Mean concentration	Group
BUM03-004	20.70	93305.34	small
BUM03-057	22.79	23642.28	small
BUM03-249	21.18	68352.31	small
BUM03-074	19.88	160875.78	small
BUM03-134	21.60	51692.89	small
BUM03-286	21.46	56536.35	small
BUM03-010	19.92	155981.51	small
BUM03-107	20.89	82611.70	small
BUM03-262	20.86	84313.06	small
BUM03-075	19.92	156804.62	small
BUM03-161	20.72	92132.72	small
BUM03-294	21.10	71832.60	small
BUM03-013	20.65	96684.83	small
BUM03-313	25.55	3820.82	small
BUM03-283	21.92	41991.95	small
BUM03-109	22.83	22937.12	small
BUM03-164	21.14	69962.92	small
BUM03-341	23.37	16085.54	small
BUM03-030	19.63	189739.16	small
BUM03-185	22.14	36185.39	small
BUM03-296	22.21	34661.55	small
BUM03-114	20.67	95709.54	small
BUM03-174	20.63	98050.85	large
BUM03-342	23.21	17846.99	large
BUM03-031	18.94	298998.79	large
BUM03-213	20.63	97864.31	large
BUM03-335	26.70	1796.50	large
BUM03-118	21.61	51433.87	large
BUM03-182	21.04	74736.42	large
BUM03-039	19.77	172102.33	large
BUM03-219	20.08	140760.76	large
BUM03-336	27.09	1384.67	large
BUM03-119	19.93	155856.29	large
BUM03-196	21.85	43703.06	large
BUM03-055	19.45	212756.34	large
BUM03-240	20.47	109105.45	large
BUM03-042	20.56	102655.28	large
BUM03-120	20.39	114367.94	large
BUM03-253	23.49	14870.76	large
BUM03-072	19.59	194106.70	large
BUM03-246	20.49	107066.01	large
BUM03-062	19.96	152633.92	large
BUM03-122	21.14	69924.28	large
BUM03-269	22.84	22865.88	large

Table B7 Raw data real-time PCR of *IDE* of 3-month-old shrimp

Sample	Mean Cp	Mean concentration	Group
BUM03-004	26.64	1273.84	small
BUM03-057	29.55	194.77	small
BUM03-249	26.62	1294.30	small
BUM03-074	26.28	1609.19	small
BUM03-134	27.44	761.52	small
BUM03-286	27.38	792.29	small
BUM03-010	26.56	1338.45	small
BUM03-107	27.27	849.20	small
BUM03-262	26.50	1393.45	small
BUM03-075	27.04	983.66	small
BUM03-161	26.92	1067.19	small
BUM03-294	26.68	1241.61	small
BUM03-013	26.99	1012.35	small
BUM03-313	30.22	126.47	small
BUM03-283	26.89	1087.84	small
BUM03-109	28.57	366.34	small
BUM03-164	27.12	933.76	small
BUM03-341	28.26	449.23	small
BUM03-030	26.22	1674.11	small
BUM03-185	28.20	464.53	small
BUM03-296	26.82	1131.76	small
BUM03-114	28.51	381.26	small
BUM03-174	27.44	760.94	large
BUM03-342	27.86	580.25	large
BUM03-031	25.31	3018.50	large
BUM03-213	26.06	1856.80	large
BUM03-335	29.50	200.81	large
BUM03-118	27.05	976.47	large
BUM03-182	26.73	1204.46	large
BUM03-039	26.62	1289.19	large
BUM03-219	26.35	1540.46	large
BUM03-336	31.25	64.99	large
BUM03-119	26.23	1663.40	large
BUM03-196	27.59	688.38	large
BUM03-055	26.57	1336.53	large
BUM03-240	27.01	1004.96	large
BUM03-042	26.88	1091.14	large
BUM03-120	26.42	1465.99	large
BUM03-253	29.15	252.29	large
BUM03-072	26.42	1468.00	large
BUM03-246	26.50	1396.01	large
BUM03-062	26.58	1324.37	large
BUM03-122	26.77	1174.83	large
BUM03-269	31.27	63.99	large

Table B8 Raw data real-time PCR of *EGF* of 3-month-old shrimp

Sample	Mean Cp	Mean concentration	Group
BUM03-004	29.74	327.91	small
BUM03-057	28.74	647.44	small
BUM03-249	25.01	8218.59	small
BUM03-074	27.66	1357.36	small
BUM03-134	22.58	43002.23	small
BUM03-286	24.58	11043.27	small
BUM03-010	27.12	1955.79	small
BUM03-107	30.84	155.88	small
BUM03-262	24.76	9737.65	small
BUM03-075	34.16	16.19	small
BUM03-161	27.01	2106.30	small
BUM03-294	23.08	30564.16	small
BUM03-013	27.61	1401.14	small
BUM03-313	24.83	9307.02	small
BUM03-283	25.95	4320.65	small
BUM03-109	25.26	6934.09	small
BUM03-164	27.64	1369.37	small
BUM03-341	27.59	1416.23	small
BUM03-030	28.86	597.03	small
BUM03-185	25.18	7305.24	small
BUM03-296	25.82	4722.30	small
BUM03-114	26.70	2602.36	small
BUM03-174	25.40	6301.52	large
BUM03-342	24.14	14898.88	large
BUM03-031	26.26	3512.57	large
BUM03-213	23.88	17791.36	large
BUM03-335	32.17	63.05	large
BUM03-118	27.76	1261.64	large
BUM03-182	25.43	6157.20	large
BUM03-039	26.67	2657.42	large
BUM03-219	24.47	11903.02	large
BUM03-336	31.46	101.83	large
BUM03-119	26.08	3979.26	large
BUM03-196	28.40	815.27	large
BUM03-055	27.15	1916.01	large
BUM03-240	26.81	2421.76	large
BUM03-042	28.91	576.47	large
BUM03-120	27.36	1654.72	large
BUM03-253	26.29	3431.44	large
BUM03-072	24.90	8832.71	large
BUM03-246	25.25	7002.96	large
BUM03-062	25.79	4822.27	large
BUM03-122	28.62	704.49	large
BUM03-269	30.93	146.39	large

APPENDIX C

Chemicals for Preparation of Polyacrylamide Gels and Silver Staining

1. 4.5% Denaturing acrylamide solution(crosslink = 19:1), 500 ml

Acrylamide	21.375	g
Bis-aclylamide	1.125	g
7 M urea	210	g
1X TBE to	500	ml

2. 40% acrylamide solution (crosslink = 37.5:1), 500 ml

Acrylamide	194.80	g
Bis-aclylamide	5.19	g
1X TBE to	500	ml

3. 40% acrylamide solution (crosslink = 75:1), 500 ml

Acrylamide	78.94	g
Bis-aclylamide	1.052	g
1X TBE to	500	ml

4. Fix/stop solution (10% glacial acetic acid), 2 liters

Glacial acetic acid	200	ml
Ultrapure or deionized water	1800	ml

5. Staining solution, 1.5 liters

$AgNO_3$	1.5	g
37% Formaldehyde	2.25	ml
Ultrapure or deionized water to	1500	ml

6. Developing solution, 3 liters

Na_2CO_3	90	g
37% Formaldehyde	5	ml
Sodium thiosulfate (10 mg/ml)	600	µl
Ultrapure or deionized water to	3000	ml

BIOGRAPHY

Mr. Sayan Prakobpetch was born on July 6, 1980 in Bangkok, Thailand. He graduated with degree of Bachelor of Science (Biotechnology) from Ramkhamhaeng University in 2002. He has studied for the degree of Master of Science (Biotechnology) at the Program in Biotechnology, Faculty of Science, Chulalongkorn University since 2007.

A publication from this thesis

1. **Prakobpetch, S.**, Khamnamtong, B., Klinbunga, S. and Menasveta, S. (2009). SNPs and expression levels of genes involved in growth and reproductive maturation of the giant tiger shrimp *Penaeus monodon*. 21st Annual Meeting and International Conference of the Thai Society for Biotechnology, 24–25 September 2009, Queen Sirikit National Convention Center, Bangkok, Thailand (Poster presentation).