การพัฒนาเครื่องหมายพันธุกรรมเพื่อเพิ่มประสิทธิภาพผลผลิตของหอยเป๋าฮื้อเขตร้อน *Haliotis asinina* ในประเทศไทย

ว่าที่ ร.ต. ปิติ อ่ำพายัพ

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต
สาขาวิชาเทคโนโลยีชีวภาพ
คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2547
ISBN 974-53-1056-5

DEVELOPMENT OF GENETIC MARKERS FOR INCREASING PRODUCTION EFFICIENCY OF THE TROPICAL ABALONE *Haliotis asinina* IN THAILAND

Acting 2Lt. Piti Amparyup

A Dissertation Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in Biotechnology
Faculty of Science
Chulalongkorn University
Academic Year 2004

ISBN 974-53-1056-5

Thesis Title	Development of genetic markers for increasing production efficiency of the tropical abalone *Haliotis asinina* in Thailand
By	Acting 2Lt. Piti Amparyup
Field of study	Biotechnology
Thesis Advisor	Associate Professor Padermsak Jarayabhand, Ph.D.
Thesis Co-advisor	Sirawut Klinbunga, Ph.D.

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

……. Dean of the Faculty of Science

(Professor Piamsak Menasveta, Ph.D.)

THESIS COMMITTEE

……. Chairman

(Assistant Professor Charoen Nitithamyong, Ph.D.)

……. Thesis Advisor

(Associate Professor Padermsak Jarayabhand, Ph.D.)

……. Thesis Co-advisor

(Sirawut Klinbunga, Ph.D.)

……. Member

(Associate Professor Anchalee Tassanakajon, Ph.D.)

……. Member

(Assistant Professor Thaithaworn Lirdwitayaprasit, Ph.D.)

……. Member

(Professor Uthairat Na-Nakorn, Ph.D.)

ปิติ อ่ำพายัพ : การพัฒนาเครื่องหมายพันธุกรรมเพื่อเพิ่มประสิทธิภาพผลผลิตของหอยเป๋าฮื้อเขตร้อน *Haliotis asinina* ในประเทศไทย. (DEVELOPMENT OF GENETIC MARKERS FOR INCREASING PRODUCTION EFFICIENCY OF THE TROPICAL ABALONE *Haliotis asinina* IN THAILAND) อ. ที่ปรึกษา : รศ.ดร. เผดิมศักดิ์ จารยะพันธุ์, อ.ที่ปรึกษาร่วม : ดร.ศิราวุธ กลิ่นบุหงา 253หน้า. ISBN 974-53-1056-5.

จากการพัฒนาเครื่องหมายทางพันธุกรรมที่จำเพาะกับหอยเป๋าฮื้อชนิด *Haliotis asinina , H. ovina* และ *H. varia* ในประเทศไทย โดยการวิเคราะห์ด้วยเทคนิคอาร์เอพีดี พบแถบอาร์เอพีดีที่จำเพาะต่อ *H. asinina*, *H. ovina* และ *H. varia* จำนวน 10, 3 และ 2 แถบ ตามลำดับ จึงทำการโคลนหาลำดับนิวคลีโอไทด์ และออกแบบไพรเมอร์จำนวน 20 คู่ และทำการตรวจสอบความจำเพาะเบื้องต้นของไพรเมอร์ จากนั้นจึงเลือกไพรเมอร์ (CUHA2, CUHA12, CUHA13, CUHO3, และ CUHV1) จำนวน 5 คู่ มาทดสอบความจำเพาะของไพรเมอร์ กับตัวอย่างหอยเป๋าฮื้อจำนวน 216 ตัวอย่าง พบว่า CUHA2, CUHA12, และ CUHA13 มีความจำเพาะต่อ *H. asinina* และ CUHV1 มีความจำเพาะต่อ *H. varia* จากการตรวจสอบความว่องไวของไพรเมอร์กับดีเอ็นเอต้นแบบ พบว่า CUHA2, CUHA12, CUHA13 และ CUHV1 สามารถให้ผลบวกเมื่อใช้ดีเอ็นเอต้นแบบประมาณ 100, 100, 500 และ 20 pg ตามลำดับ และพบว่าไพรเมอร์เหล่านี้สามารถตรวจสอบชนิดของหอยเป๋าฮื้อที่เก็บในแอลกอฮอล์ หอยเป๋าฮื้อตากแห้ง และหอยเป๋าฮื้อที่ผ่านการต้มได้ ดังนั้นเครื่องหมาย SCAR ที่พัฒนานี้สามารถนำมาใช้ตรวจสอบหอยเป๋าฮื้อชนิด *H. asinina* ในรูปตากแห้ง และผลิตภัณฑ์กระป๋องได้

ทำการค้นหาเครื่องหมายพันธุกรรมที่เกี่ยวข้องกับเพศใน *H. asinina* ด้วยเทคนิค AFLP, EST, subtractive cDNA, RT-PCR และ RACE-PCR สำหรับเทคนิค AFLP ได้ใช้ไพรเมอร์ทั้งหมดจำนวน 214 คู่ผสม กับดีเอ็นเอของหอยเป๋าฮื้อ 4 กลุ่ม พบแถบดีเอ็นเอที่มีความจำเพาะต่อเพศผู้ และเมีย อย่างละ 7 แถบ หลังจากการหาลำดับนิวคลีโอไทด์ของชิ้น AFLP ดังกล่าวแล้ว ได้ออกแบบ SCAR ไพรเมอร์ จำนวน 5 และ 6 คู่ จากแถบAFLP ของเพศผู้ และเมีย พบว่าไพรเมอร์เหล่านี้สามารถเพิ่มจำนวนดีเอ็นเอได้ทั้งในเพศผู้ และเพศเมีย เมื่อทำการวิเคราะห์ผลิตผลพีซีอาร์ด้วยเทคนิค SSCP พบว่าเครื่องหมาย HaMale1, 5, และ 6 และ HaFemale2, 3, 4, และ 5 มีความหลากหลายของรูปแบบ SSCP แต่ไม่แสดงความจำเพาะต่อเพศ นอกจากนี้ได้ทำการสร้างห้องสมุด normal cDNA และ subtractive cDNA จากรังไข่ และอัณฑะของ *H. asinina* เพื่อค้นหาเครื่องหมาย ESTที่มีประโยชน์ และทำการหาลำดับนิวคลีโอไทด์ทั้งหมดจำนวน 588 โคลน ประกอบด้วย 200 และ 118 โคลน จากห้องสมุด normal cDNA และ 110 และ 160 โคลน จากห้องสมุด subtractive cDNA ของรังไข่ และอัณฑะ ตามลำดับ พบยีนใน normal cDNA ของ รังไข่ และอัณฑะที่เหมือนกับยีนที่มีรายงานใน GenBank ไว้จำนวน 109 (54.5%) และ 73 (61.9%) โคลนตามลำดับ ในจำนวนนี้พบยีน vitelline coat proteins (VCP) จำนวน 40 โคลน (20.5%) ในรังไข่ ส่วนในอัณฑะพบ sperm lysin จำนวน 9 โคลน (8.5%) ส่วนห้องสมุด subtractive cDNA จากรังไข่ และอัณฑะ พบยีนที่เหมือนกับยีนที่มีรายงานไว้จำนวน 71 โคลน (64.5%) และ 56 (35%) โคลน ตามลำดับ

เมื่อทำการตรวจสอบการแสดงออกของยีนที่เกี่ยวของกับเพศด้วย RT-PCR พบว่ายีน VCP1, 3, 7, 49, 75 และ VTG-1 มีการแสดงออกจำเพาะต่อหอยเป๋าฮื้อเพศเมียเต็มวัย และยีน axonemal p66.0, tektinA1, sperm lysin, FP และ DMRT1 แสดงออกจำเพาะต่อหอยเป๋าฮื้อเพศผู้เต็มวัย และได้ตรวจการแสดงออกของยีนในลูกหอยเป๋าฮื้ออายุ 2, 3 และ 5 เดือน พบการแสดงออกของยีน tektinA1, FP, sperm lysin, VCP1, 2, 3, 7, 49, และ 75 ในขณะที่ไม่พบการแสดงออกของยีน axonemal p66.0, DMRT1 และ VTG-1 ในลูกหอยเป๋าฮื้อระยะดังกล่าว จากการตรวจระดับการแสดงออกของยีนในรังไข่ และอัณฑะของหอยขนาดตัวเต็มวัย พบว่ายีน axonemal p66.0, tektinA1 และ DMRT1 มีการแสดงออกเพิ่มขึ้นจากระยะการพัฒนาการของอัณฑะระหว่าง ระยะ 1 กับ ระยะ 2, 3 และ 4 ($p<0.05$) ในขณะที่พบการแสดงออกที่น้อยลงของ VCP1, 2, 3, 7, 49 และ 75 จากระยะการพัฒนาการของรังไข่ระยะ 1 กับระยะ 3 ($p<0.05$) ส่วนยีน sperm lysin, FP และVTG-1 ไม่มีการแสดงออกที่แตกต่างกันในอัณฑะ และรังไข่ที่มีระยะพัฒนาการที่แตกต่างกัน

จากการหาความยาวสมบูรณ์ของยีนด้วยเทคนิค RACE-PCR พบว่าสามารถหาความยาวสมบูรณ์ของยีน tektinA1 (2166 bp, 1350 bp ORF), axonemal p66.0 (2038 bp, 1683 bp ORF), และ DMRT1 (1727 bp, 732 bp ORF) ได้ โดยในส่วน DMRT1 พบว่ายีนนี้มี 2 รูปแบบ ประกอบด้วย ยีนที่มีส่วนของ DM domain (1727 bp) และ ยีนที่ไม่มีส่วน DM domain (1351 bp) และยีนทั้ง 2 รูปแบบมีการแสดงออกจำเพาะในอัณฑะของหอยเป๋าฮื้อ *H. asinina* เพศผู้เต็มวัย

สาขาวิชา.....เทคโนโลยีชีวภาพ..... ลายมือชื่อนิสิต.....

ปีการศึกษา.....2547.....ลายมือชื่ออาจารย์ที่ปรึกษา.....

ลายมือชื่ออาจารย์ที่ปรึกษาร่วม.....

4373875223 : MAJOR BIOTECHNOLOGY
KEY WORD: ABALONE / MOECULAR MARKERS / RAPD / AFLP / EST

PITI AMPARYUP: DEVELOPMENT OF GENETIC MARKERS FOR INCREASING PRODUCTION EFFICIENCY OF THE TROPICAL ABALONE *Haliotis asinina* IN THAILAND. THESIS ADVISOR: ASSOC. PROF. PADERMSAK JARAYABHAND, Ph.D., THESIS COADVISOR : SIRAWUT KLINBUNGA, Ph.D. 253 pp. ISBN 974-53-1056-5.

Species-specific markers were developed in the tropical abalone (*Haliotis asinina*, *H. ovina* and *H. varia*) found in Thai waters. RAPD analysis generated 10, 3 and 2 candidate species-specific markers in *H. asinina*, *H. ovina*, and *H. varia*, respectively. All RAPD fragments were cloned and sequenced. Twenty pairs of primers were designed and preliminary tested. Specificity of five primer sets (CUHA2, CUHA12, CUHA13, CUHO3 and CUHV1) was further examined against a large sample size (N = 216). Results indicated species-specific nature of all except CUHO3. Sensitivity of detection was approximately 100 pg, 100 pg, 500 pg and 20 pg of genomic DNA of the target species for CUHA2, CUHA12, CUHA13 and CUHV1, respectively. Identification of species-origin of ethanol-preserved, dried and boiled *H. asinina* specimens were sucessfully carried out. SCAR markers developed in this study can be used for quality control of various forms of *H. asinina* products from Thailand.

Sex-related markers of *H. asinina* were identified and characterized by AFLP, EST, subtractive cDNA, RT-PCR and RACE-PCR approaches. Seven candidate female-specific and seven male-specific AFLP fragments were identified from screening 214 AFLP primer combinations against 4-bulked genomic DNA of male and female *H. asinina.* Six and five SCAR markers were developed from those fragments and tested with genomic DNA of male and female abalone. Results indicated the positive amplification product in both male and female *H. asinina*. Further analysis using SSCP indicated polymorphic patterns/fragments in most of the markers (HaMale1, 5, and 6 and HaFemale2, 3, 4, and 5) but they were not sex-linked. In addition, normal and subtractive cDNA libraries from ovaries and testes of *H. asinina* were established. A total of 588 randomly selected clones (200 and 118 transcripts for normal and 110 and 160 transcripts for subtractive libraries from ovaries and testes, respectively) were unidirectional sequenced. Results indicated that 109 (54.5%) and 73 (61.9%) of normal cDNA libraries from *H. asinina* ovaries and testes significantly match with known genes in the GenBank (E value $<10^{-4}$). Of these, vitelline coat proteins (VCPs; 40 clones, 20.0%) were predominant in the former library, but sperm lysin (9 clones, 7.6%) was the most abundant transcript in the latter library. For subtractive cDNA library, 71 clones (64.5%) and 56 clones (35.0%) of ovary (forward subtraction) and testis (reverse subtraction) libraries were known transcripts.

Using RT-PCR, sex-specific expression of VCP1, VCP3, VCP7, VCP49, VCP75, and VTG-1 and axonemal p66.0, tektin A1, sperm lysin, FP, and DMRT1 was found in adult females and males of *H. asinina*, respectively. TektinA1, FP, sperm lysin, VCP1, VCP2, VCP3, VCP7, VCP49, and VCP75 were expressed in 2, 3, 5-month-old juvenile abalone whereas axonemal p66.0, DMRT1 and VTG-1 were not expressed in those stages. Expression of sex-related transcripts was semi-quantitatively estimated. The relative expression level of axonemal p66.0, tektinA1 and DMRT1 in testes was significantly different between adults having stage 1 and stages 2, 3, and 4 of testicular development ($p < 0.05$). Likewise, expression levels of VCP1, VCP2, VCP3, VCP7, VCP49 and VCP75 in ovaries were significantly different between adults having stages 1 and 3 of ovarian development ($p < 0.05$). No significant expression levels of sperm lysin and FP in testes and VTG-1 in ovaries were observed at all stages of testicular and ovarian development.

RACE-PCR was carried out for characterization of full-length cDNA of tektinA1 (2166 bp, 1350 bp ORF), axonemal p66.0 (2038 bp, 1683 bp ORF), and DMRT1 (1727 bp, 732 bp ORF). An EST containing a DM domain (1727 bp) and its variant (1351 bp) without the DM domain were also successfully isolated and characterized. These transcripts were specifically expressed in testes of adult *H. asinina*.

Field of study……Biotechnology……..… Student's signature..…….……

Academic year……...2004……...…… Advisor's signature……...

Co-advisor's signature…...……...

ACKNOWLEDGEMENT

I would like to express my deepest gratitude and grateful appreciation to my advisor, Associate Professor Dr. Padermsak Jarayabhand for his valuable advice, encouragement, supervision, support and understanding throughout the period of my study.

My sincere appreciation is extended to my co-advisor, Dr. Sirawut Klinbunga for his excellent advice, encouragement, supervision and kindness throughout my study. I am deeply grateful for the time and effort he has given to the research, teaching me all practicalities of molecular biology. Without his support and understanding, this work would not have been accomplished.

My sincere gratitude is expressed to Associate Professor Dr. Anchalee Tassanakajon for her valuable suggestion, support, kindness and encouragement throughout my study.

My gratitude is also extended to Professor Dr. Uthairat Na-Nakorn, Assistant Professor Dr. Charoen Nitithamyong, Associate Professor Dr. Anchalee Tassanakajon, and Assistant Professor Dr. Thaithaworn Lirdwitayaprasit for their valuable suggestions and their kindness in serving as the thesis committee.

I am grateful to Professor Dr. Takashi Aoki and Ikuo Hirono at Laboratory of Genome Science, Graduate school of Marine Science and Technology, Tokyo University of Marine Science and Technology in Japan for providing the research facilities and the possibility to carry out this work.

Many thanks are expressed to all colleagues at Center of Excellence for Marine Biotechnology (CEMB) for their friendly assistance as well.

This work was supported by research grants from Royal Golden Jubilee Ph. D. Program (RGJ) and a Basic Research Grant for the Royal Golden Jubilee Ph. D. Program, The Thailand Research Funds (TRF).

Finally, I would like to express my deepest appreciation to my parents and my sisters for their unlimited love, understanding and encouragement.

CONTENTS

CONTENTS (cont.) **Page**

CONTENTS (cont.) **Page**

CONTENTS (cont.) **Page**

CONTENTS (cont.)

LIST OF TABLES

LIST OF TABLES (cont.)

LIST OF FIGURES

LIST OF FIGURES (cont.)

LIST OF FIGURES (cont.)

LIST OF FIGURES (cont.)

LIST OF FIGURES (cont.)

LIST OF FIGURES (cont.)

LIST OF FIGURES (cont.)

LIST OF FIGURES (cont.)

LIST OF FIGURES (cont.)

LIST OF ABBREVIATIONS

bp	base pair
°C	degree Celcius
dATP	deoxyadenosine triphosphate
dCTP	deoxycytosine triphosphate
dGTP	deoxyguanosine triphosphate
dTTP	deoxythymidine triphosphate
DNA	deoxyribonucleic acid
HCl	hydrochloric acid
IPTG	isopropyl-thiogalactoside
Kb	kilobase
M	Molar
$MgCl_2$	magnesium chloride
mg	Milligram
ml	Millilitre
mM	Millimolar
ng	Nanogram
OD	optical density
PCR	polymerase chain reaction
RNA	Ribonucleic acid
RNase A	ribonuclease A
rpm	revolution per minute
SDS	sodium dodecyl sulfate
Tris	tris (hydroxyl methyl) aminomethane
µg	Microgram
µl	Microlitre
µM	Micromolar
UV	ultraviolet

CHAPTER I

INTRODUCTION

Abalone are economically important archeogastropods currently being cultured worldwide. They are 75-100 extant species and all of them are allocated into a single family Haliotidae (Hahn, 1989). Over 20 species of abalone are commercially farmed (Table 1.1; Jarayabhand and Paphavasit 1996). Sizes and geographic distribution of abalone are temperature-dependent; generally those inhabiting the temperature and semi-tropical regions have larger sizes than those in the tropical and arctic regions (Lindberg, 1992). The red abalone, *Haliotis rufescens*, is the largest abalone. It attain a shell length of 30 cm when reaches the adult stage (Cox, 1962). In contrast, adult of *H. discus hannai* and *H. diversicolor* (tropical species) attain only 5 cm in length (Hahn, 1989, Fallu, 1991 and Ino, 1980).

Abalone products are purchased in fresh, frozen, canned and dried forms. The shells of several abalone are used for decorative purposes and as a traditional medicine in Asia with a high market value.

The quantity of wild-caught abalone had remained relatively stable between 1998-2002. The main producers are Australia, Japan, New Zealand, Mexico and South Africa that produced 5,974, 2,682, 1,553, 1,616, and 1,281 metric tons (MT) of wild caught abalone in 2002, respectively.

Farming of abalone began in Japan and China, since the late 1950's to the early 1960's respectively. The rapid development of abalone cultivation has taken place in those countries during the 1990s. The largest producer of cultured abalone is China where over 300 farms are established with an annual production of approximately 3500 tons in 1998. Currently, Taiwan has over 400 farms with the annual production is 2312 tons in the same year. Most of which is consumed locally.

The aquaculture industry of abalone has still consistently expanded in Taiwan and China (FAO 2000a,b; Gordon and Cook, 2003). The total world production of abalone was 22677 MT in 2002 (Table 1.2) and 8696 MT of those were from the aquaculture sector accounting for 62.2 % of the fisheries production (13,981 metric tons).

Table 1.1 Commercial important abalone species (Jarayabhand and Paphavasit, 1996)

No.	Scientific name	Common name	Shell length (mm)
1	*H. rufescens*	Red abalone	>275
2	*H. fulgens*	Green, southern green or blue abalone	125-200
3	*H. corrugata*	Pink or corrugated abalone	150-175
4	*H. sorenseni*	White or sorensen abalone	125-200
5	*H. assimilis*	Threaded abalone	<100
6	*H. cracherodii*	Black abalone	75-125
7	*H. walallensis*	Flat or northern green abalone	75-125
8	*H. kamtschatkana*	Pinto abalone	100
9	*H. discus hannai*	Ezo awabi abalone	180-200
10	*H. discus* Kuro	Awabi, oni or onigai abalone	200
11	*H. diversicolor supertexta**	Tokobushi abalone	50
12	*H. gigantea*	Madaka abalone	250
13	*H. sieboldii*	Megae abalone	170
14	*H. asinina**	Mimigai, donkey's ear abalone	70-100
15	*H. rubra*	Black lip abalone	120-140
16	*H. laevigata*	Green lip abalone	130-140
17	*H. roei*	Roe's abalone	70-80
18	*H. iris*	Paua or black abalone	170
19	*H. australis*	Silver or queen paua abalone	125
20	*H. virginea*	Virgin abalone	70
21	*H. tuberculata*	Ormer abalone	120
22	*H. midae*	Perlemon abalone	90

*Tropical abalone species

Gordon and Cook (2001 and 2003) reported that approximately 67% of the production farmed in China are *H. diversicolor supertexta* and the rest were from *H. discus hannai* and/or *H. discus discus*. The main species produced in Taiwan is *H. diversicolor supertexta.* The actual production data for China was not reported in the FAO statistics because abalone are sold locally without any official documentation required and all of the production in China is consumed domestically (Gordon and Cook, 2003).

Almost all of the Japanese farming consists of ocean bottom growing from farmed seeds. Major farming operations in Japan have involves in ocean stock enhancement where approximately 30 - 40 million seeds are planted annually. Other countries including South Africa, Australia, New Zealand, Chile, Mexico, USA and South Korea are presently farmed abalone commercially. In Thailand, two commercial farms for *H. asinina* namely Phuket Abalone Farm, located in Phuket (west of peninsular Thailand) and Thai Abalone Farm located in Chumphon (east of peninsular Thailand) were established

Table 1.2 The World abalone production in 2002 (Gordon and Cook, 2003)

Country	**World Fisheries Abalone Landings (MT)**	**World Illegal abalone Catch (MT)**	**World Cultured Abalone Production (MT)**	**World Supply of Abalone from all Sources (MT)**
Australia	4974	1000	162	6224
China	-	-	4500	4500
Taiwan	-[*]	-[*]	3000	3000
Japan	2146	536	200	2882
South Africa	431	850	450	1731
Mexico	1066	550	53	1669
New Zealand	1,153	400	2	1555
USA	0	250	169	412
Chile	-	-	150	150
Others	442[a]	110[a]	10[b]	554
Total	10212	3696	8696	22677

Others [a]: includes Korea, Philippines, Solomons, Oman and Taiwan [*]
Others [b]: includes Korea, Philippines, Europe, Thailand and others

Three species of tropical abalone, *H. asinina* (Linnaeus, 1758) *H. ovina* (Gmelin, 1791) and *H. varia* (Linnaeus, 1758) are found in Thai waters (Jarayabhand and Paphavasit 1996). They inhabit in the coral reefs at 1-8 m in depth. Both *H. asinina* and *H. ovina* are distributed along the east coast of the upper Gulf of Thailand and all three species are found in the Andaman Sea (Tookwinas et al. 1986; Nateewatana and Bussarawit 1988). On the basis of their size and meat textures, it is potential to culture these species commercially and market as the cocktail-sized abalone (Jarayabhand and Paphavasit 1996).

At the macrogeographic scale**,** *H. asinina* is widely distributed in the southern Pacific Ocean in sub-tropical and tropical region (Thailand, Australia, Japan and the Philippines) at the depths around 3-7 meters (Sungthong, 1991). *H. ovina* are usually found in the same area with *H. asinina* and they are wider distributed than *H. asinina* (Kakhai, 1992). *H. ovina* prefer the depths between 1.5 - 4 meters. It selects rock crevices rather than coral substrate and prefers substrate size larger than 0.1 m^2 (Ngow, 1993). Among these three species, *H. asinina* has the highest percentage of a ratio between the meat weight and the total weight (85%) compared to that of 40% and 30% for *H. ovina* and *H. varia*, respectively (Singhagraiwan and Doi, 1993). Moreover, the spawning cycle of *H. asinina* is highly predictable allowing convenient incorporation of genetic-based breeding programs to increase management and culture efficiency of this species. It is currently promoted for commercially culture in Thailand at present.

1.1 Taxonomy of Thai abalone

Taxonomic definition of Thai abalone is as followed (Cox, 1960);

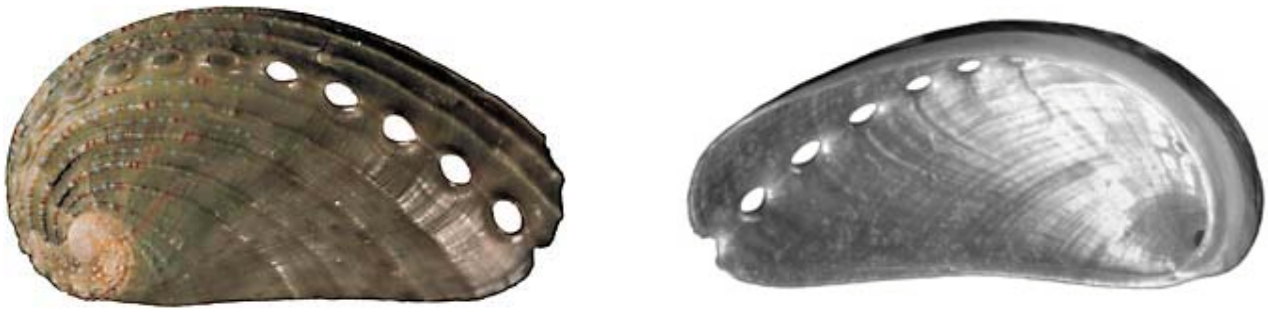
Phylum Mollusca

Class Gastropoda

Subclass Prosobranchia

Order Archaeogastropoda

Suborder Pleurotomariina



***Haliotis asinina* (Linnaeus, 1758)**

***Haliotis ovina* (Gmelin, 1791)**

***Haliotis varia* (Linnaeus, 1758)**

Figure 1.1 Shell Morphology of *H. asinina*, *H. ovina*, and *H. varia* (Geiger, 2000).

Superfamily Pleurotomariacea

Family Haliotidae

Haliotis asinina (Linnaeus, 1758)

Haliotis ovina (Gmelin, 1791)

Haliotis varia (Linnaeus, 1758)

1.2 Distributions of *H. asinina*, *H. ovina* and *H. varia*

Abalone are globally distributed in the major oceans, especially in the temperate zones. They are generally existent in the hard rock or coral areas. Geiger (2000) reported that *H. asinina* and *H. ovina* were found in more than 100 localities whereas *H. varia* Linnaeus, 1758 had the highest occurrence, 317 data points (Figure 1.2). Distribution of *H. asinina* are reported in South East Asia, Japan and Australia but *H. ovina* and *H. varia* are wider distributed. Their geographic distribution includes the Indian Ocean and the East Coast of Africa.

In Thailand, *H. ovina* had been reported to be more common than *H. asinina* along the upper eastern Gulf of Thailand (Kakhai and Petjamrat, 1992), even though Sungthong et al. (1991) found that *H. asinina* is more common at Samet Island (Rayong , eastern Thailand . Only small numbers of *H. ovina* have been found along the lower eastern coast. There has been no report on the finding of *H. varia* in the Gulf of Thailand but a few literatures reported their availability in Peninsular Malaysia (Geiger, 1999 and 2000). Along the Andaman coastline, the relative abundance of each species was 81%, 17.3% and 1.7% for *H. varia*, *H. ovina* and for *H. asinina*, respectively (Jarayabhand and Paphavasit, 1996).

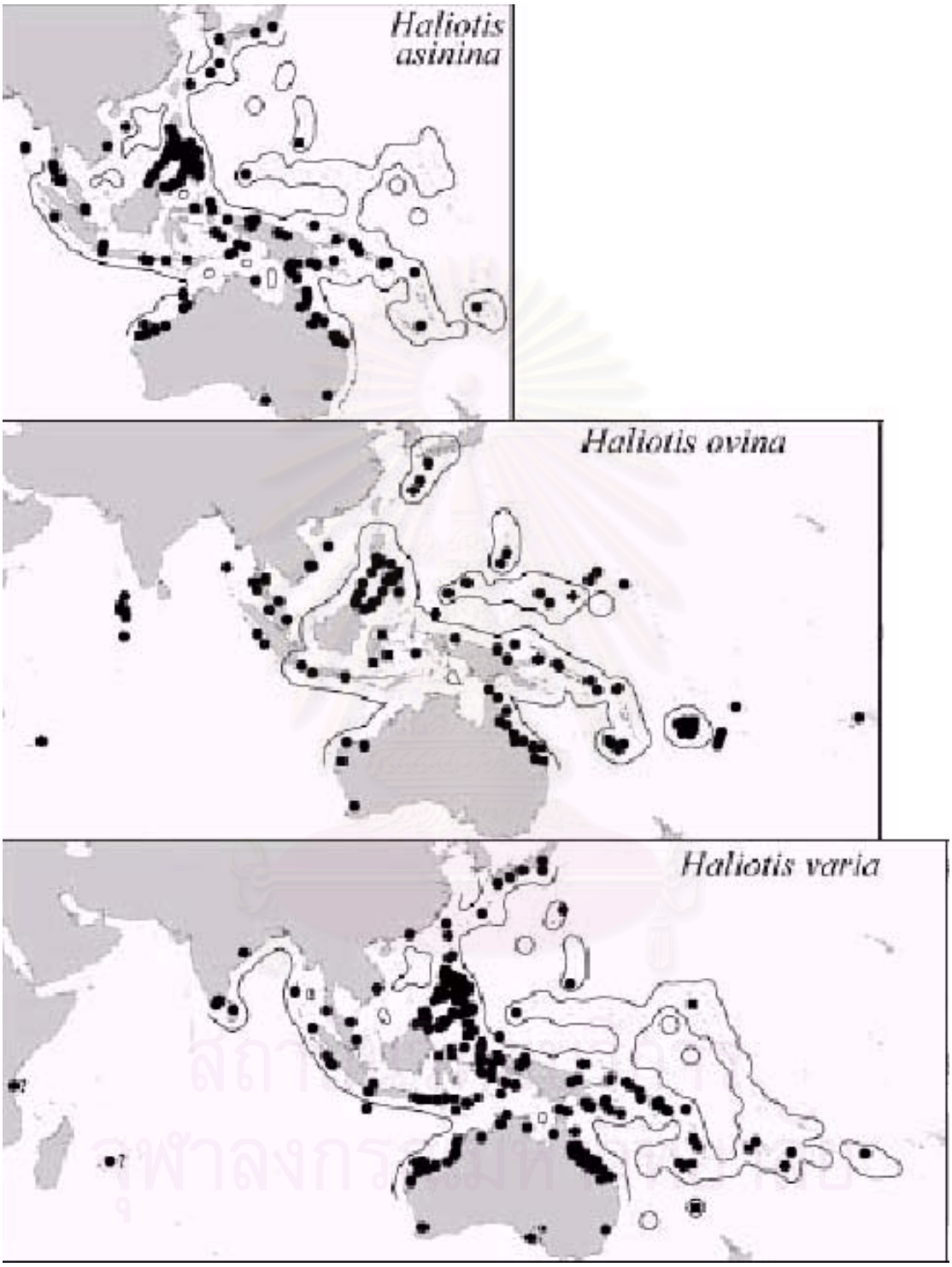


Figure 1.2 Macrogeographic distributions of *H. asinina, H. ovina* and *H. varia* (Geiger, 2000)

1.3 Molecular genetic techniques used in this thesis

Molecular genetic approaches are useful and can be applied for genetic studies of various organisms. Several molecular approaches were used for identification of molecular markers at genomic DNA and cDNA levels of *H. asinina*. The former included randomly amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), single strand conformational polymorphism (SSCP), and DNA sequencing. The latter included reverse transcription (RT)-PCR, Expressed Sequence Tags (ESTs), cDNA subtraction and rapid amplification of cDNA ends-polymerase chain reaction (RACE) analysis.

1.3.1 Polymerase chain reaction (PCR)

The polymerase chain reaction (PCR) is an effective approach applied for molecular genetic studies. The method is based on an enzymatic amplification of DNA *in vitro*. Amplification of a particular target DNA fragment is carried out from a very low amount of the starting template DNA (mostly in the nanogram range). The PCR reaction components are constituting of DNA template, a pair of primers for the target sequence, dNTPs (dATP, dCTP, dGTP and dTTP), the suitable buffer and heat-stable DNA polymerase (usually *Taq* polymerase). The amplification reaction consists of three major steps; a denaturation step of double stranded DNA at high temperature, an annealing step to allow forming of hybrid primers-template at the optimal temperature, and an extension step of the annealed primers by the heat-stable DNA polymerase. The cycles are repeated for 30-40 times (Figure 1.3). The amplification product is usually determined by electrophoresis.

1.3.2 Randomly Amplified Polymorphic DNA (RAPD) analysis

RAPD analysis was concurrently developed by Williams et al. (1990) and Welsh and McClelland (1990). RAPD is a simple method for identification of molecular genetic markers at different taxonomic levels using an arbitrarily primed PCR-based technique. RAPD markers are usually generated by the amplification of random DNA segments with single short arbitrary primers (e.g. 10-12 mer with GC content usually at least 50%) at a low annealing temperature (e.g. 36°C).

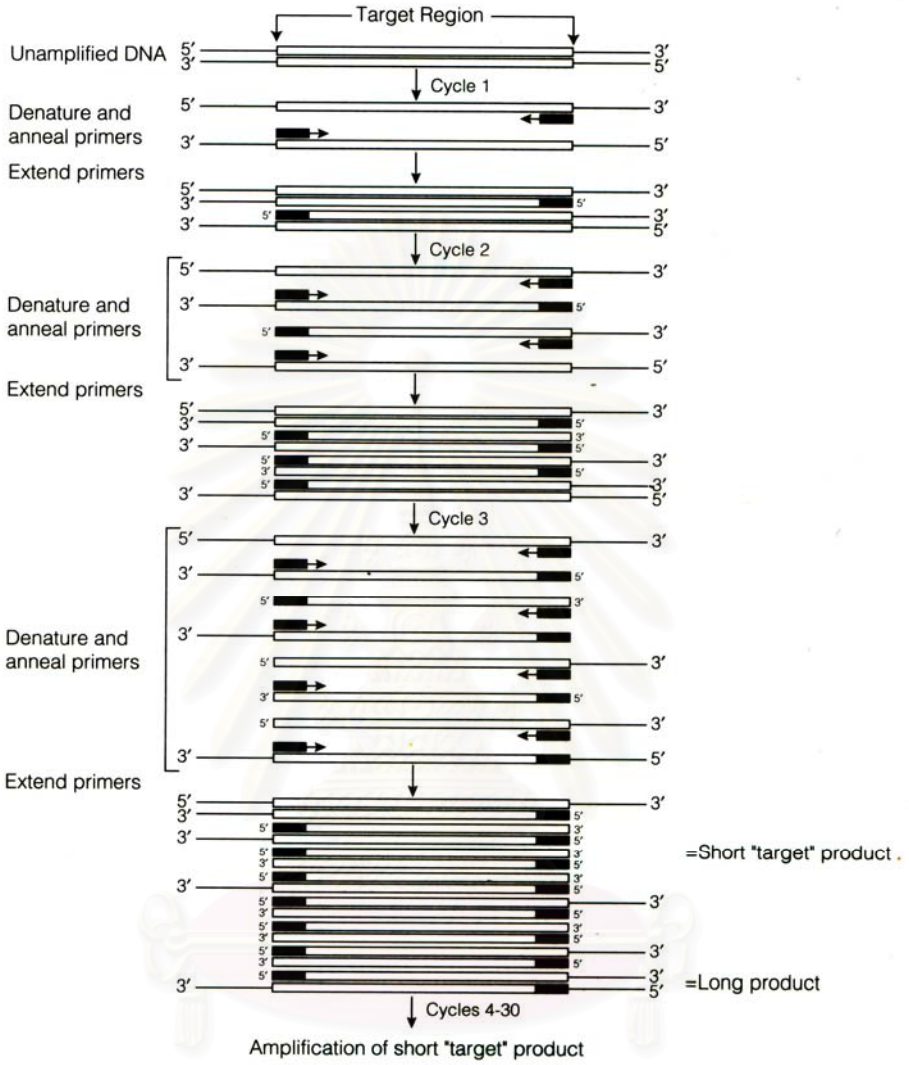


Figure 1.3 General illustration for amplifying DNA using PCR (Promega)

Accordingly, the RAPD primer is utilized to scan genome for the small inverted sequences resulting in the amplified DNA segments of variable length. The amplification products are separated on polyacrylamide or agarose gels and detected by either silver or ethidium bromide staining.

Polymorphisms detected by RAPD are due to point mutations, which affect the binding of the primer to the DNA and insertions and deletions (indels) between primer binding sites.

Benefits of RAPD analysis mainly include its simplicity because it does not require any previous sequence data. The technique is also less time consuming and cost-effective than other methods, for example RFLP and/or conventional multi-locus DNA fingerprinting. Nevertheless, the disadvantage of RAPD is due mainly to its low reproducibility. Changes in enzyme manufacturers, primer or enzyme concentrations, or thermal cyclers, can give inconsistent results (MacPherson *et al.*, 1993; Meunier and Grimont, 1993).

Accordingly, RAPD markers found from the experiments are usually converted to sequence-characterized amplified region (SCAR) markers through cloning and sequencing of the original RAPD fragment. A pair of locus-specific primer is designed and used for amplification of the target fragment more specifically.

1.3.3 Amplified Fragment Length Polymorphism (AFLP) analysis

AFLP was developed by Vos *et al*.(1995) to increase the potential of RFLP and to compensate reproducibility of RAPD by selective amplification of a random array of restriction fragments ligated to linkers of known sequence.

AFLP approach is based on selective amplification of digested genomic DNA by a series of extended primers. Initially, genomic DNA is digested with two enzymes (a frequent and a rare cutter, for example *Mse* I and *Eco* RI). Adaptors, with a core sequence and the appropriate overhangs for the enzymes used are ligated to the ends of the digested fragments.

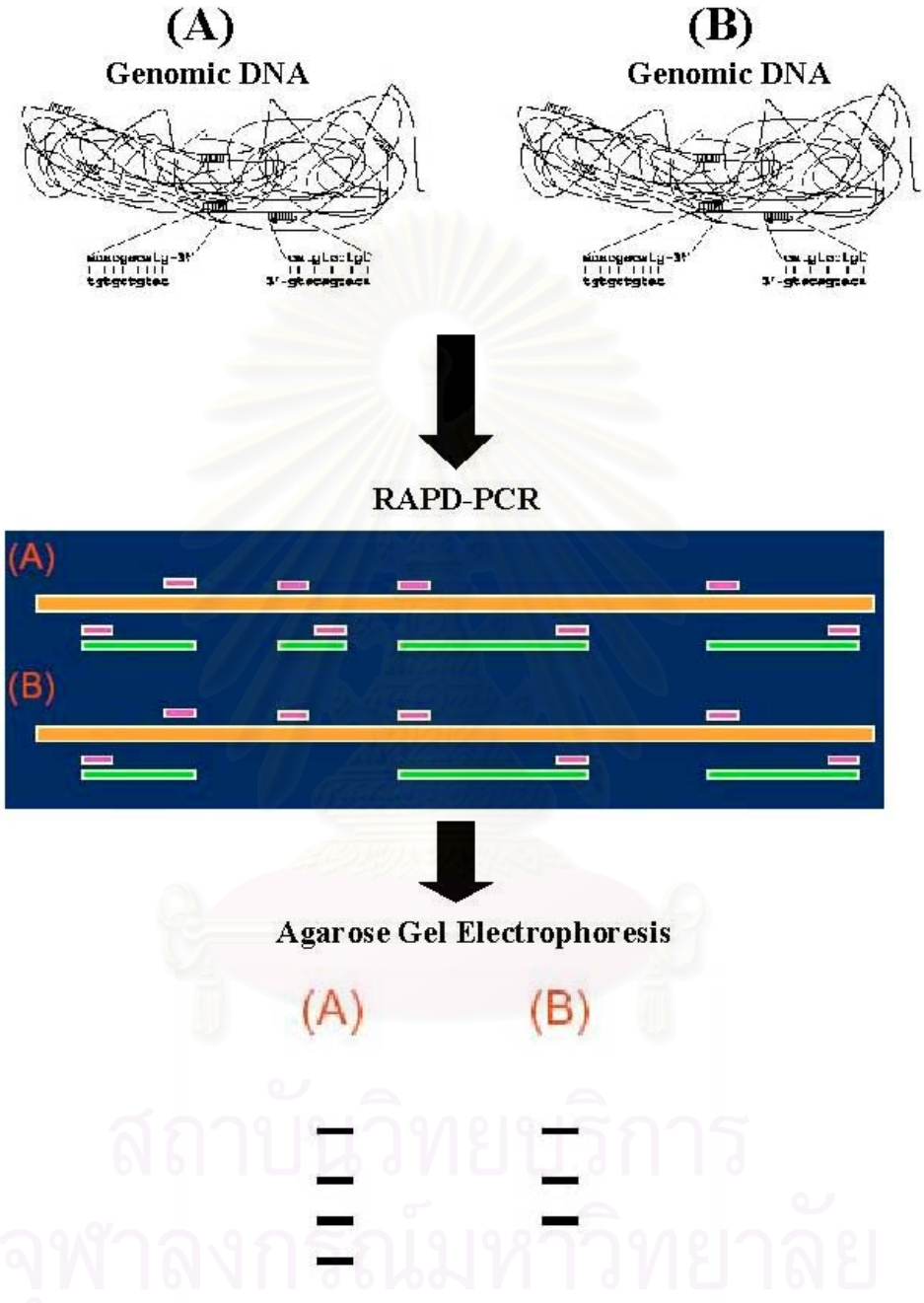


Figure 1.4 Schematic presentation of RAPD-PCR.

Pre-amplification of these fragments is achieved by PCR using primers with a core sequence (an enzyme specific sequence) and an extension nucleotide at the 3/ end. This generates a pool of fragments, which are re-amplified by two selective primers of core sequence, usually 3 base extension primers. Optimal amplified genomic DNA fragments are analyzed by denaturing polyacrylamide gels electrophoresis (Figure 1.5)

The main advantages of AFLP are its more reproducibility than RAPD-PCR due to specificity of the PCR primer and high stringency of the amplification reaction. Like RAPD-PCR, AFLP analysis does not require the prior knowledge about genome sequences of species under investigation. The technique offers great flexibility in the number of loci that can be co-amplified in one PCR reaction. Typically, 50 to 100 restriction fragments are co-amplified in one fingerprint. This technique is therefore extremely powerful for identification of DNA polymorphism. AFLP markers can be further converted to SCAR markers by sequencing, cloning and designing of locus-specific primers for direct application for the developed markers.

1.3.4 Single-Stranded Conformational Polymorphism (SSCP) analysis

SSCP analysis is one of the effective techniques used for the detection of mutations and variation of the DNA due to single nucleotide polymorphism (SNP) and/or insertions/deletions (indels). SSCP relies on the principle that the electrophoretic mobility of a single-stranded DNA molecule in a non-denaturing gel is dependent on its structure and size (Orita et al., 1989; Hayashi, 1991).

Single-stranded 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. Hence, molecules differing in their conformations (e.g., due to a single nucleotide change) can be separated in low cross-link non-denaturing polyacrylamide gels (with or without glycerol supplementation) . The method has effectively been used to analyze point mutations of small amplicons (100–400 bp) (Hayashi, 1996).

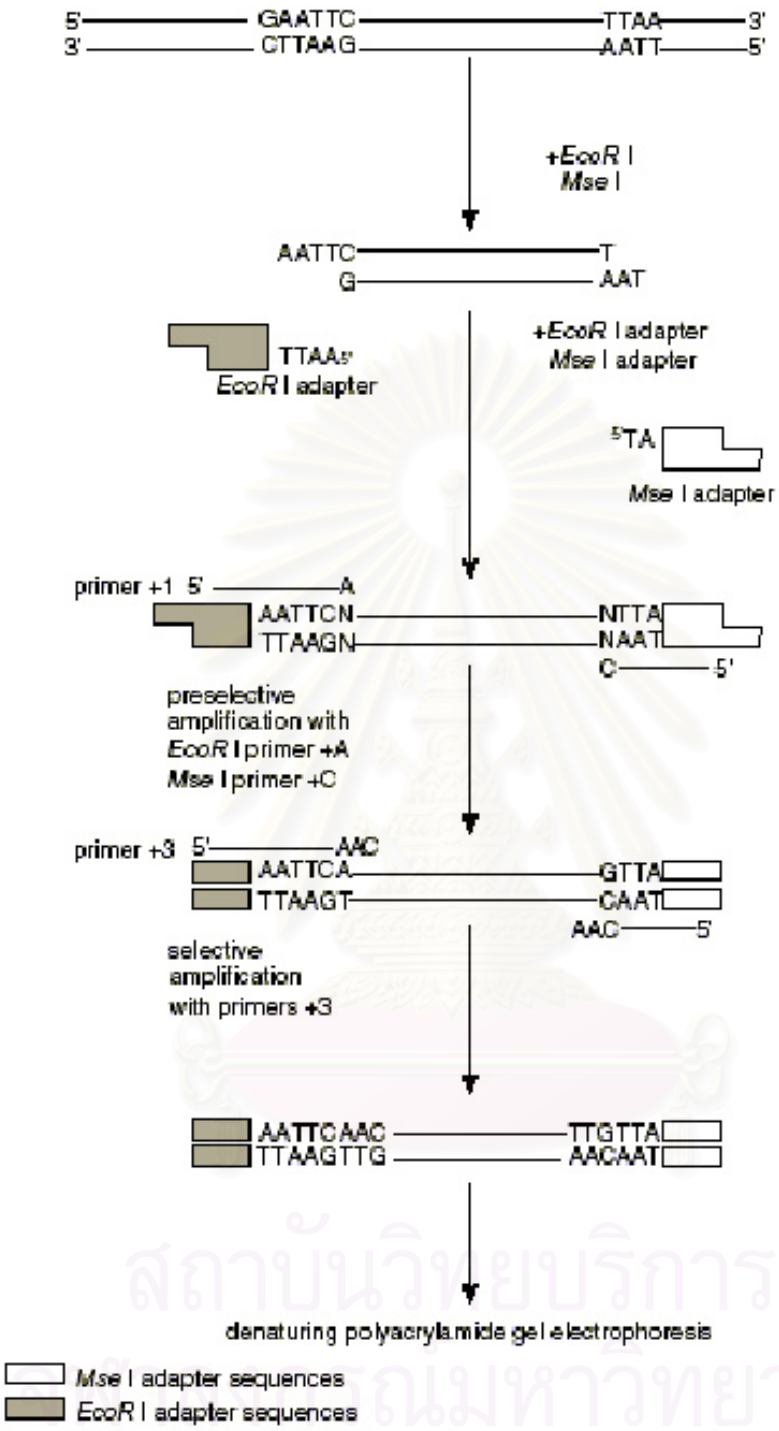


Figure 1.5 A schematic diagram illustrating the general protocol for AFLP analysis (AFLP® Analysis System I, USA).

SSCP is simple, and sensitive enough to detect one or a few base differences in the sequence of short DNA fragments (Hayashi, 1996). The disadvantage of SSCP is that it requires nucleotide sequence data for the design of the specific primers. Moreover, reproducibility of the technique seems to be problematic because SSCP patterns are affected by temperature and degree of cross-linking. Additionally, multiallelic patterns of some nuclear DNA markers may cause the SSCP patterns complicate for estimation of allele frequencies precisely.

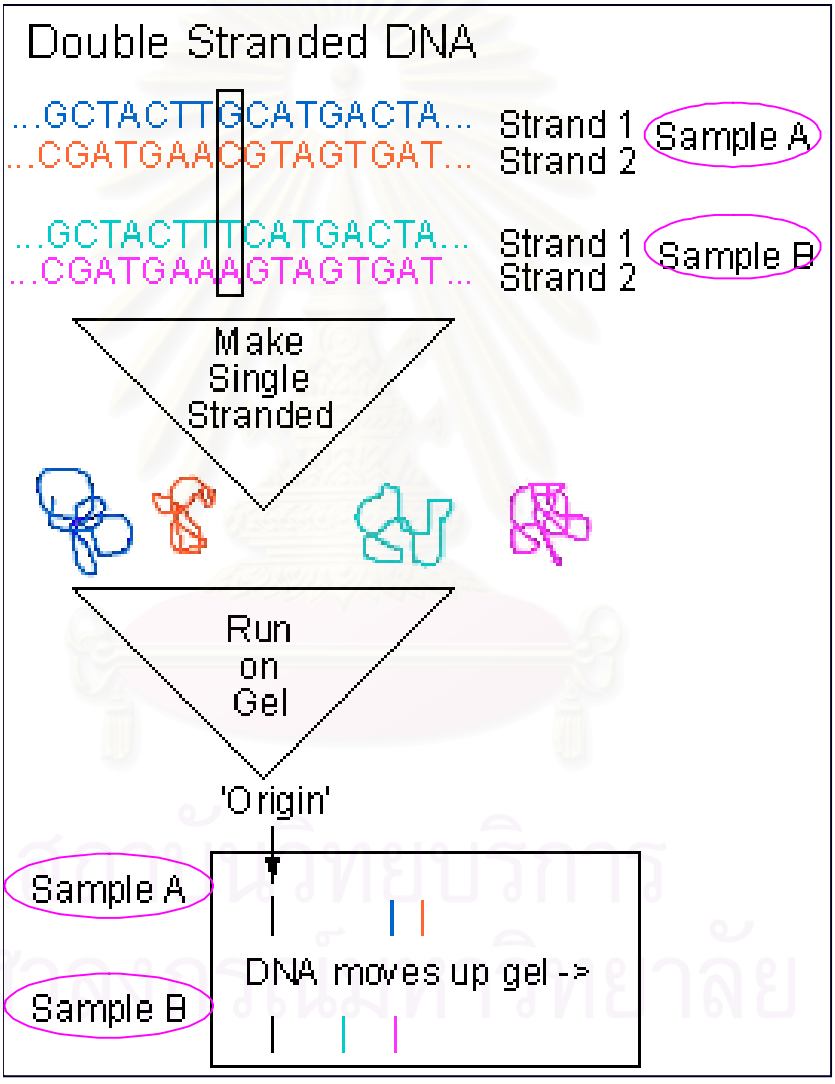


Figure 1.6 A schematic diagram for analysis of DNA polymorphism using SSCP analysis.

1.3.5 DNA sequencing

Polymorphism at the DNA level can be studied by several methods but the most direct strategy is determination of nucleotide sequences of a defined region. The sequences obtained can be aligned by comparing an orthologous region in the genome of related organisms (or populations). PCR-amplified fragments can be directly sequenced using the typical chain termination reaction (Sanger, 1977) or alternatively by cycle-sequencing. Nevertheless, more accurate sequences of DNA fragments are obtained through cloning prior to sequencing of the interesting fragments.

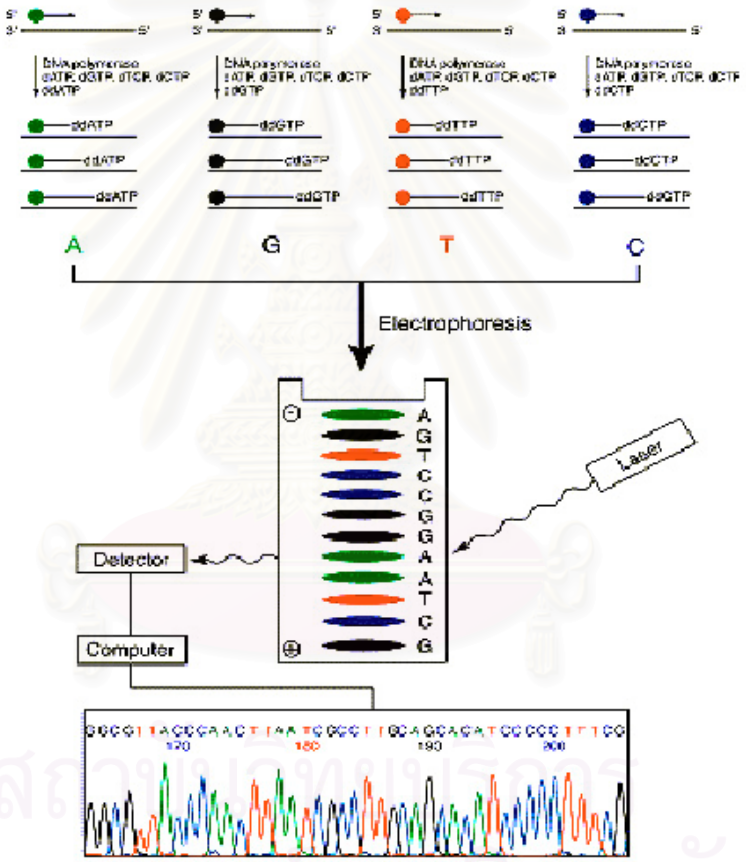


Figure. 1.7 A diagram showing the general protocol for DNA sequencing

Previously, DNA sequencing is tedious, time consuming and expensive. The number of bases, which can be, determined by sequencing usually cover a few hundred bases. Increasing the length of investigated DNA with a large number of samples are prohibited by several factors mentioned above. With an introduction of automated DNA sequencers, the experiments can be carried out much faster than that based on manual sequencing allowing its wider applications for genetic studies at present.

1.3.6 Reverse transcription-polymerase chain reaction (RT-PCR)

RT-PCR is a comparable method of conventional PCR but the first strand cDNA rather than genomic DNA was used as the template in the amplification reaction. It is a basic technique for determination of gene expression in a particular RNA population. The template for RT-PCR can be the first stranded cDNA synthesized from total RNA or poly A^+ RNA. Reverse transcription of total RNA can be performed with oligo(dT) or random primers using a reverse transcriptase (Figure 1.8). The product is then subjected to the second strand synthesis using a gene specific primer. The resulting product is used as the typical PCR.

Semi-quantitative RT-PCR is an quantitative approach where the target genes and the internal control (e.g. a housekeeping gene) were separately or simultaneously amplified using the same template. The internal control (such as β-actin, elongation factor EF-1α or G3PDH) is used under the assumption that those coding genes are transcribed constantly and independently from the extracellular environment stimuli and that their transcripts are reverse transcribed with the same efficiency as the product of interesting transcript.

1.3.7 Expressed Sequence Tags (EST) analysis

ESTs are large-scale single-pass sequences of randomly picked clones from a cDNA library usually constructed from mRNA at a particular developmental stage and/or tissue (Adams et al., 1991). This method has been widely employed for discovering novel and uniquely expressed genes, and for characterizing the gene expression profiles of several tissues (Adams et al., 1993; Affara et al., 1994; Pawlak et al., 1995; Liew et al., 1994; Takeda et al., 1993).

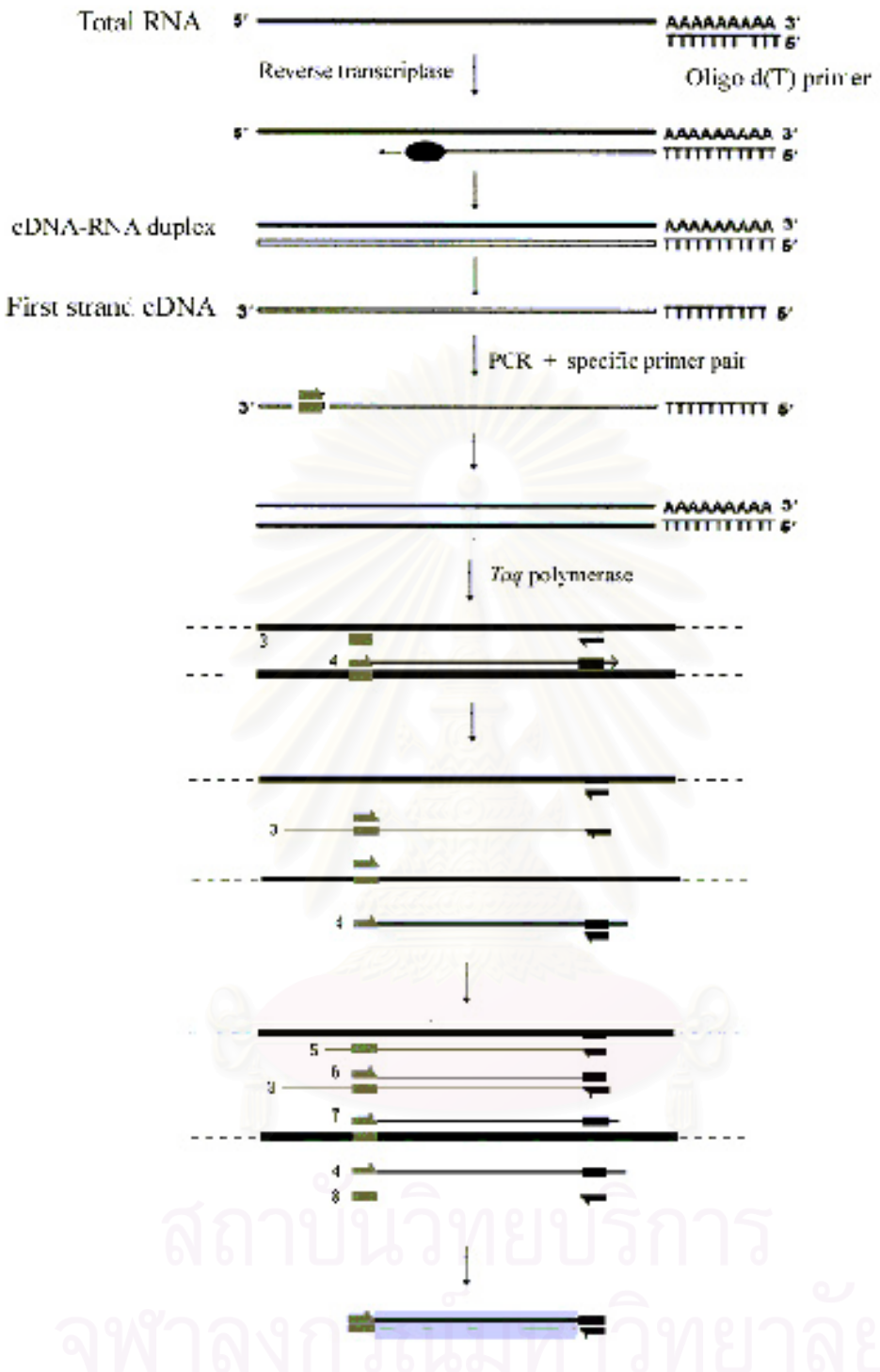


Figure 1.8 Overall concepts of the RT-PCR procedure. During first-strand cDNA synthesis, an oligo d (T) primer anneals and extends from sites present within the total RNA (or poly A^+ RNA). Second strand cDNA synthesis primed by the 18 – 25 base specific primer proceeds during a single round of DNA synthesis catalyzed by *Taq* polymerase. These DNA fragments serve as templates for subsequent PCR amplification.

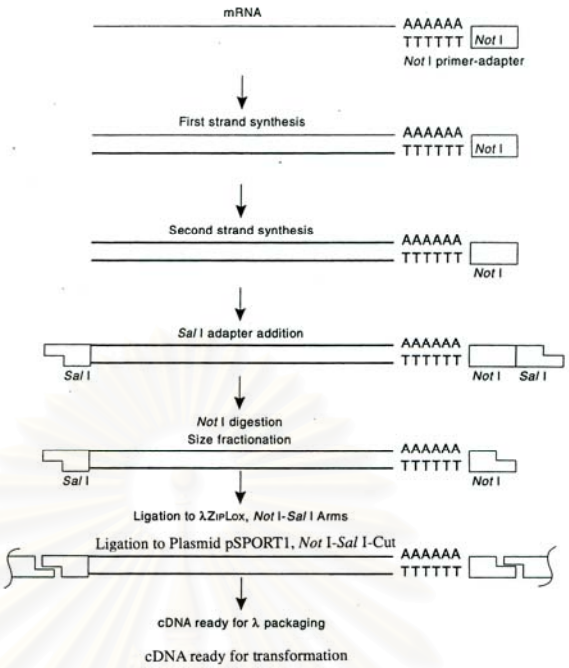
The general principles for construction of a cDNA library begin with purification of the target mRNA that is transcribed into the first-strand cDNA. This step is catalyzed by reverse transcriptase using the oligo (dT) primer as the synthesizing primer. The second-strand DNA is then copied from the first-strand cDNA using *E. coli* DNA polymerase I. The double-strand cDNA is ligated to adapter and subsequently to an appropriate vector using T4 DNA ligase. The recombinant vector-cDNA molecules are packaged (λ vector) *in vitro* and transfected to the appropriate host. If a plasmid is used recombinant plasmid is transformed into *E. coli* host cells to generate a cDNA library.

ESTs can be sequenced from either 5 / or 3 / ends of cloned cDNA. The 3/ end of the cloned insert is usually marked by the poly A stretch which is often problematic for thermostable polymerase sequencing, and sequencing through poly T can reduce the length and quality of the subsequent sequence. Nevertheless, 3/ UTR usually exhibit high polymorphism and is a promising location for SNP identification. The 5/ ESTs have the advantage of being more likely to include some of the open reading frame (ORF) of the cDNA and thus facilitate identification of the encoded product.

EST sequences are used as a tag to homology search through the sequence data in the GenBank (Altschl et al., 1997). The Blast*N* program uses nucleotide sequence to compare against the NCBI nucleotide database whereas the BLAST*X* uses the translated protein products to compare against the NCBI protein database in all possible 6 reading frames. Sequences are considered to be significantly matched when the possibility value (E-value) is less than 10^{-4} and the match length is > 100 nucleotides for Blast*N* and a match length is > 10 amino acid residues for Blast*X*, respectively, (Anderson and Brass, 1998).

EST analysis is an important tool for several applications. They have mainly applied for rapid gene discovery (Adams, 1992, Aoki, 1999 and Zeng, 2002) of several animal genome and for studying a large scale gene expression (Waterston, 1992, Zeng, 2002 and Satou, 2003), comparative genomics and functional genomics (Waterston, 1992; McCombie, 1992; Tugendreich, 1994 and Mironov, 1999) in various organisms.

A.



B

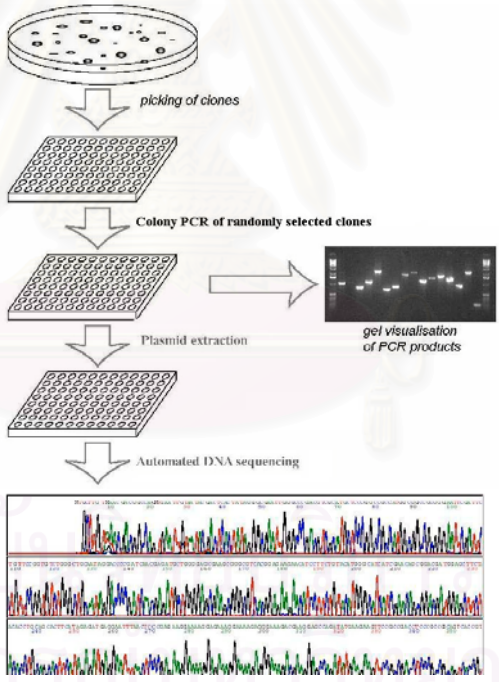


Figure 1.9 Overview for construction of cDNA insert (A) and automated DNA sequencing (single-pass) of randomly selected cDNA clones (the entire process simply called EST analysis).

1.3.9 cDNA subtractive hybridization

Subtractive hybridization is one of powerful technique for detecting differences of gene expression between two populations of RNAs or cDNAs. It is based on base pair complementary that nucleic acid sequences in common with the two populations can from hybrids. After hybridization, hybrids are removed and the sequences in only one population are identified. This technique can be used for detecting differences between the mRNA in different cells, tissues or organisms under two different conditions and/or various stages of development and growth (Lin and Ying, 2003).

Although traditional subtractive cDNA hybridization approaches have been applied and successfully used in several cases, They require several rounds of hybridization and are not well suited for identification of rare transcripts. Suppression cDNA subtraction is a method based on selective amplification of differentially expressed sequences, which overcomes technical limitations of traditional subtraction methods (Diatchenko et al., 1996; Gurskaya et al., 1996).

Two types of mRNA populations that contains specific (or differentially expressed) transcripts (the tester), and the reference cDNA (the driver) were reverse-transcribed into two cDNA pools by the reverse transcriptase with poly dT as the primer. The cDNAs from the two pools (tester and driver) are digested with *Rsa* I to generate shorter fragments.

The tester cDNA is divided to two portions, and each of which is ligated with different cDNA adaptor. The ends of adaptor do not have a phosphate group, therefore only one strand of each adaptor attaches to the 5/ ends of cDNA. The two adaptors have stretches of identical sequences to allow annealing of the PCR primer once recessed ends have been filled in.

The tester and driver cDNAs are then hybridized. In the first hybridization, an excess of driver is added to each tester. The samples are heat-denatured and allow to anneal. Differential expressed transcripts are then enriched. During the second hybridization, template for PCR amplification is generated from differentially

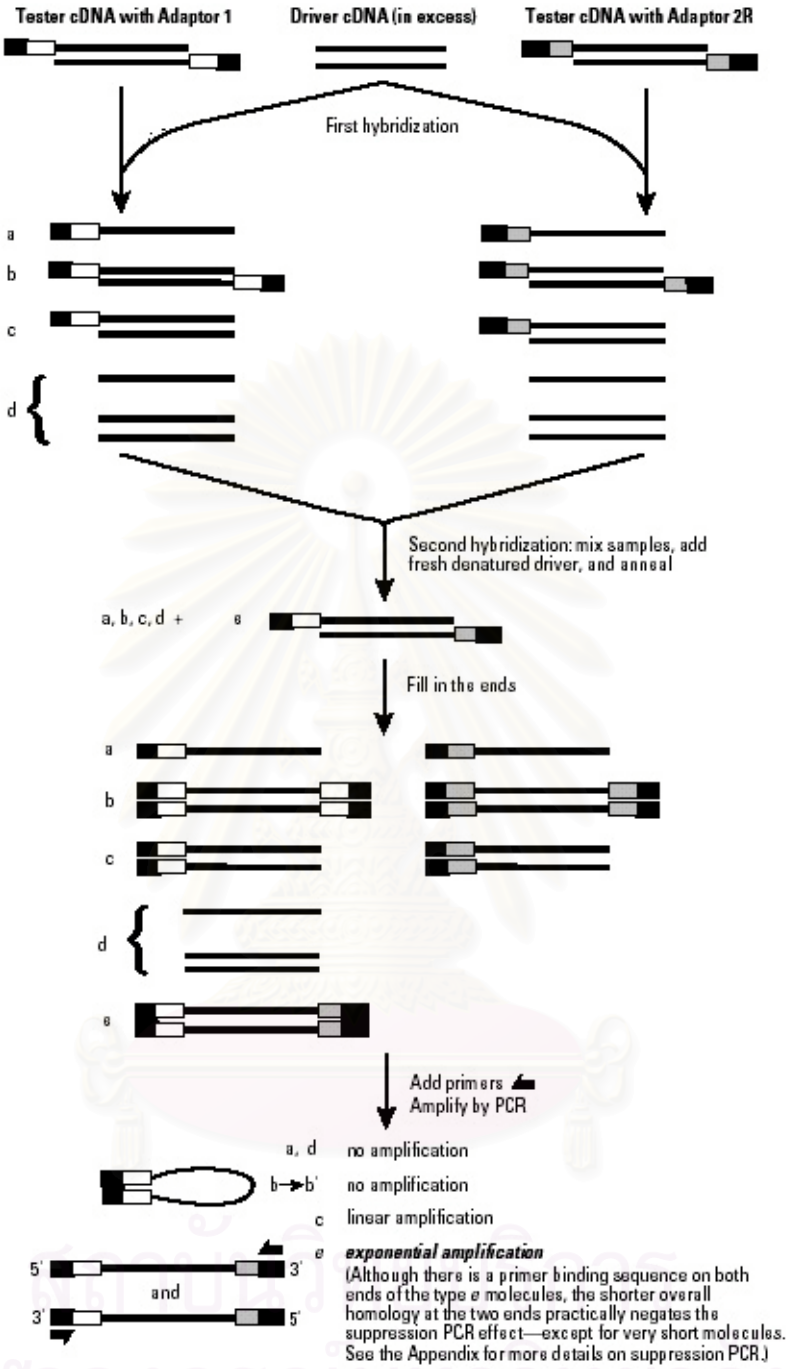


Figure 1.10 Overview of a suppressive cDNA subtraction procedure. The cDNA in which specific transcripts are to be found is called “the tester” and the reference cDNA is called “the driver”.

expressed transcripts. Two rounds of suppression PCR was carried out and only differentially expressed transcripts are amplified exponentially. Subtractive cDNA products are then cloned into the T-vector. Positive clones are characterized by hybridization or sequencing.

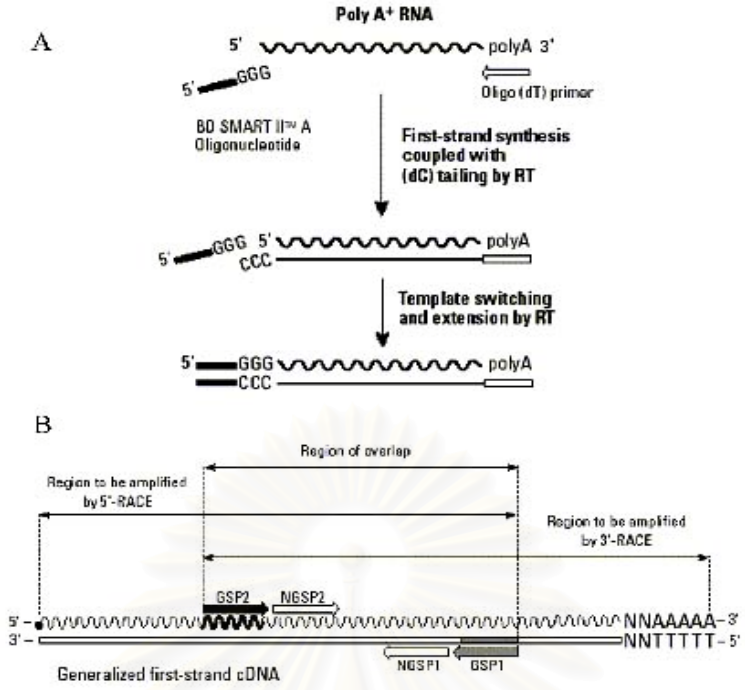
1.3.10 Rapid amplification of cDNA ends-polymerase chain reaction (RACE-PCR)

RACE-PCR is the common approach used for isolation of the full length of characterized cDNA. Using SMART (Switching Mechanism At 5′ end of RNA Transcript) 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 at the 5′ end is synthesized with the additional SMART sequence at the end (Figure 1.11).

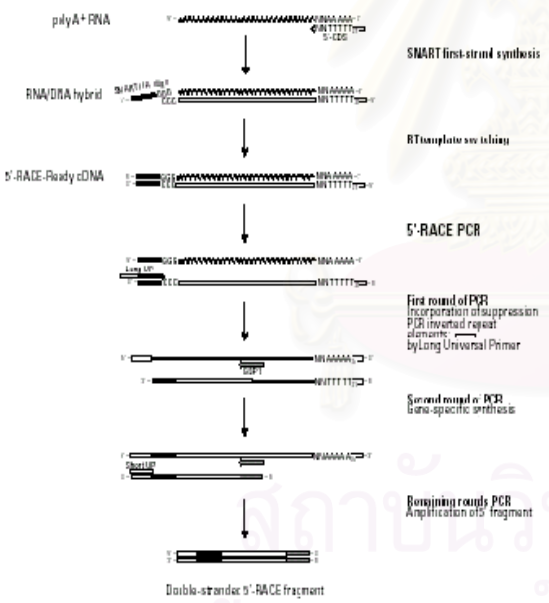
The first strand cDNA of 3′ RACE is synthesized using a modified oligo dT primers and serve as the template for RACE PCR reactions. Gene specific primers (GSPs) are designed from interesting gene sequence for 5′- RACE PCR (antisense primer) and 3′-RACE PCR (sense primer) and used with the universal primer (UPM) that recognize the SMART sequence. RACE products are usually characterized by cloning and sequencing. Finally, the full-length cDNA is amplified and used for subsequent applications.

1.4 Molecular Genetic Studies in Abalone

Molecular genetic markers can be used to elevate culture and management efficiency of abalone. The sustainable success of aquacultural activity of commercially important species requires the basic knowledge on population, molecular and quantitative genetics and the use of suitable molecular genetic markers to establish effective management programmes of exploited species. Genetic markers commonly used in aquaculture are those from allozymes, mitochondrial DNA, RFLP, RAPD, AFLP, and microsatellites.



C: Detailed flow chart of 5′-RACE-PCR



D: Detailed flow chart of 3′-RACE-PCR

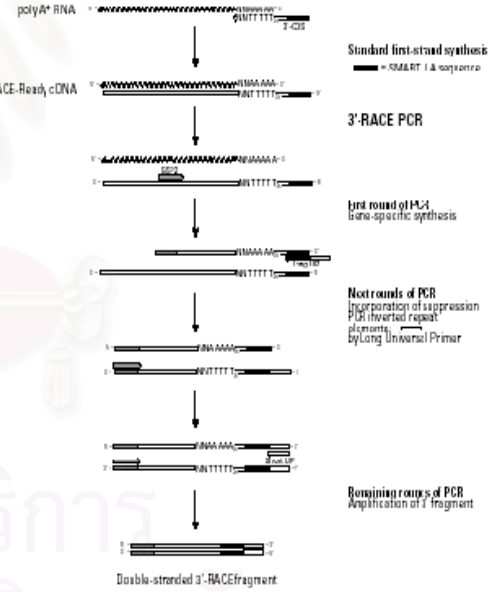


Figure 1.11 Overview of the SMART^{TM} RACE cDNA Amplification reaction.

The applications of DNA markers has allowed rapid progress for investigations of genetic variability and inbreeding, parentage assignments, species and strain identification, and the construction of high-resolution genetic linkage maps for aquaculture species (Liu and Cordes, 2004).

1.4.1 Identifying species origin of abalone

Muchmore *et al* (1998) identified a tandemly repeated satellite DNA (290-291 bp in length) by digestion of genomic DNA of five species of the Eastern Pacific (California) abalone including the red abalone (*H. rufescens*), the white abalone (*H. sorenseni*), the flat abalone (*H. walallensis*), the pinto abalone (*H. kamtschatkana*), and the pink abalone (*H. corrugata*) with *Sal* I. The fragment was cloned and sequenced. Satellite-specific primers were synthesized from a repeat unit of *H. rufescens.* The consensus sequences of satellite DNA were determined in all five species by direct sequencing of the PCR product using the satellite-specific primers. In *H. rufescens*, the 290 bp *Sal* I satellite DNA represents approximately 0.5% of total DNA, equivalent to approximately 28,000 copies per haploid genome.

Sweijd et al (1998) developed species-specific markers (species-specific PCR and PCR-RFLP) for identification of *H. midae* (only commercially exploited abalone in South Africa) and *H. spadicea* (a sympatrically congeneric species). Species-specific PCR primers were designed from a portion of lysin cDNA sequences. The PCR primers successfully amplify approximately 1,300 bp of genomic DNA from dried, cooked, and fresh *H. midae* tissue. A smaller fragment of 146 bp product was used for identification of canned *H. midae*. Additionally, PCR-RFLP analysis revealed interspecific polymorphism that discriminate between these two species by digestion of a 1300 bp fragment with *Cfo* I, *Dra* I, *Taq* I and *Hinf* I.

Naganuma *et al*. (1998) compared partial 18S rDNA sequences of closely related abalone, *H. discus discus* Reeve (from Izu Peninsula, Shizuoka Prefecture, central Japan) and *H. discus hannai* Ino (from Kesennuma, Iwate Prefecture, northeastern Japan). The PCR product of 18S rDNA (forward 5′-AAC CTG GTT GAT CCT GCC AGT-3′ and reverse 5′ -TGA TCC TCC TGC AGG TTC A-3′) were directly sequenced. Sequences were multiple-aligned with those from other abalone (*H. madaka* Haba and *H. gigantea* Gmelin, both from Naruto, Tokushima Prefecture, western Japan). A land gastropod

(*Limicolaria kambeul*) was also included as an outgroup. The inferred 18S rDNA phylogenies indicated that the *H. discus discus* and *H. discus hannai* are closely related but distinguishable presumably at the subspecies level.

Popongviwat (2001) examined genetic diversity and identified species-specific markers of three abalone species in Thailand; *Haliotis asinina*, *H. ovina* and *H. varia* using RAPD analysis. Five decanucleotide primers (OPB11, UBC101, UBC195, UBC197 and UBC271) were selected from one hundred and thirty primers screened for genetic analysis of abalone in Thailand. High levels of genetic diversity between three abalone species were observed when UBC101 and OPB11 were used. The average similarity index within samples of *H. asinina, H. ovina* and *H. varia* were 0.7927-0.8496, 0.6010-0.7032 and 0.5259-0.6102, respectively. Genetic differences within 3 species were 0.2995, 0.4328 and 0.4295 for *H. asinina, H. ovina* and *H. varia,* respectively. The average genetic distance of *H. asinina* in the Gulf of Thailand was 0.0243, which indicate that they are closely related. Large genetic distances were observed between each of the Gulf of Thailand and the Andaman and the Philippine samples. A neighbor-joining tree constructed from the average genetic distance between paired geographic samples indicated phylogenetically clear separation between investigated abalone species (using two primers) and geographic samples of *H. asinina* (using five primers). Species-specific RAPD markers were found in *H. asinina* (1700 and 320 bp from UBC101, 1030 and 650 bp from UBC195, 1450 and 750 bp from UBC197 and 680 bp from UBC271, respectively). In addition, a population-specific marker was observed in the Philippine sample (380 bp from UBC101) and the Talibong Island (880 bp from UBC271).

Klinbunga et al (2003) determined genetic diversity of three abalone (*H. asinina*, *H. ovina*, and *H. varia*) in Thailand using RAPD-PCR and PCR-RFLP of 18S and 16S rDNAs. Species-specific RAPD markers were found in each abalone. Restriction analysis of 18S (nuclear) ribosomal DNA with *Alu* I, *Taq* I, and *Hae* III and 16S (mitochondrial) rDNA with *Bam* HI, *Eco* RI, *Hae* III, and *Alu* I gave 12 and 13 digestion patterns, respectively. A total of 49 composite haplotype were found. A UPGMA dendogram constructed from divergence between pairs of composite haplotypes revealed separate gene pools of these abalones and indicated that *H. asinina* and *H. ovina* are genetically closer than *H. varia.* Geographic heterogeneity analysis and F_{ST} estimate indicated clear genetic differentiation between *H. ovina* originating from the Andaman Sea (west) and the

Gulf of Thailand (east, P<0.0001) but did not reveal genetic heterogeneity of *H. asinina* and *H. varia* in Thai waters (P > 0.0021). The amplified 16S rDNAs of individuals representing composite haplotypes found in their study (AAAA and AAAE for *H. asinina*, ABBB, AAAB and AABB for *H. ovina* and BABG, BABC, BABD, BABF, and AABG for *H. varia*) were cloned and sequenced. A neighbor-joining tree constructed from sequence divergence of 16S rDNA accurately allocated those sequences according to species origins of abalone. Species-specific PCR based on 16S rDNA polymorphism was successfully developed in *H. asinina* and *H. varia* but not in *H. ovina*.

1.4.2 Identification of gene expression in abalone

Lee and Vacquire (1995) compared cDNA sequences of the sperm lysin of 27 species of abalone from California, Japan, Australia, New Zealand, Taiwan, Borneo, Madagascar, South Africa, Greece, France, Italy, and the Azores. The lysin cDNA sequences revealed that 22 of 27 investigated taxa were clearly distinguishable by at least 20 nucleotide differences. For the remaining taxa, *H. coccinea* may be a subspecies of *H. tuberculata* and the lysin sequences are almost identical between *H. makada* and *H. discus hannai*, *H. conicopora* and *H. rubra*, *H. diversicolor supertexta* and *H. diversicolor aquatilis*, and *H. tuberculata lamellosa* and *H. tuberculata tuberculata.*

Phylogeny of lysin cDNA suggests that there are three abalone groups among the 27 species-group taxa: (1) all California species (*H. rufescens*, *H. sorenseni*, *H. kamtschatkana*, *H. walallensis*, *H. cracherodii*, *H. corrugata*, *H. fulgens*) and 3 Japanese species (*H. gigantea*, *H. discus hannai*, *and H. madaka*); (2) The New Zealand species (*H. iris*); and (3) 1 Japanese species (*H. diversicolor aquatilis*), Indo-West Pacific species (*H. roei*, *H. scalaris*, *H. laevigata*, *H. cyclobates*, *H. rubra*, *H. ovina*, *H. conicopora* from Australia; *H. australis* from New Zealand; *H. diversicolor supertexta* from Taiwan; *H. varia* from Borneo), and European species (*H. pustulata* from Madagascar; *H. midae* from South Africa; *H. tuberculata lamellosa* from Greece; *H. tuberculata tuberculata* from France; *H. coccinea* from Azores). These groups were assigned with three previously recognized subgenera (*Nordotis, Paua* and *Padollus*) in the genus *Haliotis*.

Degnan el al (1997) isolated a homologue of Mox homeobox cDNA from *H. rufescens* (HruMox). In vertebrates, Homeobox genes (*Mox1* and *Mox2*) appear to be responsible for early mesoderm development, being expressed in the paraxial mesoderm

and its derivatives, and during somitogenesis and limb muscle development. The derived *Haliotis* homeodomain sequence is 85% similar to mouse and frog Mox-2 homeodomains and 88.9% similar to the partial cnidarian cnox5-Hm homeodomain. Quantitative RT-PCR analysis of mRNA accumulation reveals that this gene is expressed in the larva, but not in the early embryo. Transcripts were most prevalent during larval morphogenesis from trochophore lavae to veligers. There are also transient increases in transcript prevalence at 1 and 3 days after the initiation of metamorphosis from veliger to juvenile. The identification of a molluscan Mox homeobox gene that is more closely related to vertebrate genes than other protostome (e.g. *Drosophila*) genes suggests the Mox class of homeobox genes may consist of several different families that have been conserved through evolution.

Hinman and Degnan (2002) studied expression of *Mox gene* in *H. asinina* (*Has Mox*). During larval development, *Has-Mox* expression remains restricted to mesodermal cells destined to form adult muscle in the foot. This restricted expression of *Has-Mox* in *Haliotis* is similar to those observed in vertebrates, suggesting a conserved role in myogenesis in deuterostomes and lophotrochozoans.

O'Brien and Degnan (2000) identified sets of *POU*, *Sox*, and *Pax* transcription factor genes that are expressed in the anterior ganglia in *H. asinina*. *POU* multigene family regulates the development of central and peripheral nervous systems and involves in the activation of growth genes. It has been implicated in the neuroendocrine control of puberty and ovulation and in the regulation of neurotransmitter synthesis. *Sox,* a high mobility group (HMG) gene family where its gene products known to interact with *POU* proteins and to facilitate DNA binding, as well as acting as transcription factors in the CNS. *Pax* genes are essential in developing CNS and sensory systems, such as eyes. They are also expressed during morphogenesis of a variety of organs. Using highly degenerate oligonucleotide primers designed to anneal conserved codons in each of these gene families.

Two *POU* genes (Has*POU*-III and Ha*sPOU*-IV), two *Sox* genes (Has*Sox*-B and Has*Sox*-C), and two *Pax* genes (Has*Pax*-258 and Has*Pax*-6) were amplified by RT-PCR. The expression analyses with gene-specific primers indicated that these genes are expressed in the cerebral and pleuropedal ganglia of both reproductively active and spent adults, in a number of sensory structures, and in a subset of other adult tissues.

Giusti el al (2000) examined developmental expression of a *Scr*/*Hox*5 gene in *H. rubra*. *Hox* genes encode a set of evolutionarily conserved transcription factors that regulate anteroposterior patterning mechanisms in insects and vertebrates and are expressed along this axis in a range of bilaterians. In *Haliotis*, embryogenesis yields non-feeding trochophore larvae that subsequently develop into the veliger larvae, which possesses many characteristics of the adult body plan. Quantitative RT-PCR analysis reveals that Hru-*Hox*5 is first expressed in the trochophore larva. Hru-*Hox*5 transcript prevalent increases continually through larval developmental stages until competent development in the veligers and then again over the first four days of metamorphosis. *In situ* hybridization reveals that larval expression of Hru-*Hox*5 is restricted primarily to the primordial and newly formed branchial ganglia, located between the anterior cerebral-pleuropedal ganglionic complex and the posterior visceral ganglia. The expression of Hru-*Hox*5 in the central region of the abalone CNS is similar to that observed for its orthologue (*Lox*20) in the leech, suggesting that *Hox*5 genes were used, along with other *Hox* genes, to pattern the CNS of the ancestral spiralian lophotrochozoan.

Galindoa el al (2002) isolated the full-length cDNA sequence (11,166 bp) of the VE receptor for lysin (VERL) from the red abalone (*H. rufescens*). There are 42 amino acids from the start Met residue to the beginning of the first 'VERL repeat'. Most of VERL (9981 bp; 89.4%) consists of 22 tandem repeats of a 153 amino acid sequence that is structurally predicted to be β-sheet. The last VERL repeat is followed by 353 non-repeat amino acid residues containing a furin cleavage site (RTRR), a ZP domain and a hydrophobic COOH-terminus with a 3′ UTR of only 10 nucleotides. VERL repeats 3–22 have been subjected to concerted evolution and consequently have almost identical sequences.

1.5 Sex determination markers and sex differentiation cascades

An understanding of sexual biology of any sexual-reproducing species is important for designing breeding programs in that species. Unfortunately, the genetic sex of many species cannot be identified from external morphology. This problem is usually arisen when dealing with embryonic or juvenile forms of interesting species. One effective solution is to exploit molecular genetic markers at genomic DNA and , if necessary, cDNA levels to diagnose sex.

The criteria of sex determination and sex differentiation were often difficult to distinct. Sex determination is used to describe the genetic and environmental processes and variables that influence sex differentiation, whereas sex differentiation (morphological, cellular, and molecular) is used to describe the physical realization of these events in terms of development of testis and ovary (Devlin and Nagahama, 2002).

Sex determination is problematic in researches of many species. This can usually be solved by the application of DNA-based technology but this is only possible if a sex-specific markers located on the unique sex chromosomes is available. Several molecular genetic approaches can be used to identify sex-identification/differentiation markers. Basically, molecular genetic markers at genomic DNA level are ideal for sex identification purposes but these markers, in many cases, are not available. Therefore, sex-specific/differential expression markers have been identified at cDNA level in several species.

For species exhibiting well-defined sex chromosomes, molecular approaches at the genomic DNA levels for example, RAPD and AFLP analysis are useful for isolation of sex-specific markers. The isolated marker can be converted to sequence-characterized amplified region (SCAR) markers by sequencing, cloning and designing of specific primers for direct application for sex determination purposes.

For species that do not exhibit sex chromosomes, both molecular approaches at the genomic DNA and cDNA levels; RT-PCR, cDNA subtraction, RACE-PCR, Differential display (DD)-PCR and RNA-arbritary primed (RAP)-PCR may be used for isolation and characterization of sex identification DNA markers and/or sex-specific/differential expression markers.

1.5.1 Sex chromosomal systems and sex determination/differentiation genes

The most common sex-determination systems in animals involve a genetic mechanism based on the chromosomal system of that organism. However, sex chromosomal systems in some organisms may by can influencally other variables such as temperature. Therefore, details of some sex-determination systems are not yet fully understood. Generally, three different sex chromosomal systems have been described, XX/XY, XX/XO and ZZ/ZW systems.

The XX/XY system is found in mammals. Males possess heterogametic sex chromosomes (XY) and females exhibit homogametic sex chromosomes (XX). The XX/XO system is found in grasshoppers, crickets, roaches and some insects. Males exhibit a single set of X chromosome (XO) but females possess homogametic sex chromosomes (XX). In addition, ZZ/ZW system is found in birds, some fishes and some insects including butterflies and moths. Sex heterogametic chromosomes (ZW) are found in females but sex homogametic chromosomes (ZZ) are found in males.

Studies of sex determining and chromosome systems in fish are well advanced compared to other aquatic species. Nine sex determination systems have been reported in fish. Sex is controlled by sex chromosome in eight: XX/XY, ZW/ZZ, XX/XO, ZO/ZZ, X1X1X2X2/X1X2Y, ZW1W2/ZZ, XX/XY1Y2 and the WXY (where XY and YY fish are males while XX, WX and WY are females) of these systems.

Molluscs collectively display diverse reproductive systems comprising of parthenogenesis sequential and/or simultaneous hermaphroditism and dioecy (Coe, 1943 and Heller, 1993). Sex chromosomes in gastropod species were found in many species. An XX/XY sex chromosome has been observed in member of the order Mesogastropoda (family :Viviparidae, *Viviparus subpurpureus*; family: Hydrobiidae, *Benedictia baicalensis*; family: Littorinidae, *Littorina saxatilis*; family: Rissoidae (*Rissoa ventricosa*), family: Carinariidae, *Pterosoma planum*) and the order Neogastropoda (Family: Buccinidae ,*Pisania maculosa*, and family : Fasciolariidae , *Fasciolaria lignaria*).

The XO/XX sex-determining mechanism has been reported in member of the order Archaeogastropoda (family: Neritidae, *Neritina cornucopia*, *N. violacea*, *Theodoxus baeticus, T. fluviatilis*, *T. meridionalis*, *T. valentina*, and *T. velascoi*, and family:

Trochidae, *Monodonta vermicularis*) and the order Mesogastropoda (family: Pomatiopsidae, *Neotricula aperta* and family: Littorinidae, *Melarhaphe neritoides*).

Moreover, ZW/ZZ sex chromosome system has found in member of the order Mesogastropoda (family: Vivipararus, *Viviparus ater*, *V. acerosus*, and *V. mamillatus*). In contrast, *Carinaria japonica* in family: Carinariidae, exhibits a multiple sex-determining mechanism (male: XY_1Y_2) as in family: Pterotrachaeidae (*Firoloida desmaresti*, *Pterotrachea coronata*, *P. scutat*, and *P. hippocampus*) of orders Mesogastropoda. No sex chromosomes have been found in the order Opisthobranchia and Pulmonata and of Archaeogastropoda (family Haliotidae).

Sex determination gene was classified two categories of gene: those, which are well conserved in divergent groups of organisms and those which are only found in a specific branch of the phylogenetic tree.

In mammals, sex-determining region on the Y chromosome (*SRY*) was isolated from the non-recombinant region on the human Y. *SRY* is a member of a large gene family. *SRY* encodes a protein of a 79-amino acid of conserved high mobility group (HMG) box, which can bind to and bend DNA. This protein has a role as a transcription factor in the developing gonads leading to sex differentiation cascades of testes. Many genes that encode proteins with *SRY*-related HMG boxes have been identified and named *Sox* genes. *SRY* is conserved and functional only in mammals but *Sox-9* gene involve in sex determination among mammals, chickens and fish (Zhou et al., 2001).

Zinc-finger protein on Y- and X-chromosomes (*Zfy* and *Zfx*) are transcription factors implicated in mammalian sex determination. Multiple copies of a C2H2 Zinc finger are found in the N-terminus. (Schneider-Gadicke et al. 1989; Koopman, Ashworth, and Lovell-Badge 1991). Most mammalian species possess a simply *Zfx*/*Zfy* systems. In contrast, murine zinc-finger gene family is more complex. Two Y-linked genes (*Zfy-1* and *Zfy-2*), an X-linked gene (*Zfx*), and an autosomal homolog on chromosome 10 (*Zfa*) are found (Koopman, Ashworth, and Lovell-Badge 1991). The localization of *ZFX* and *ZFY* on the X and Y chromosomes suggests that they have evolved from a single ancestral gene separated by the evolution of an ancestral autosomal chromosome pair into heteromorphic sex chromosomes. *Zfy*-type sequences have also been identified in the channel catfish (Tiersch et al. 1992), *Xenopus* (Connor and Ashworth 1992).

Although, *Zfy* and *Sry*-like genes have been shown to be present in birds and reptiles such as snapping turtles (Chelydra serpentina), alligator (*Alligator mississippiensis*), lizards, amphibians, and fishes, these genes do not seem to be sex-linked in non-mammalian species (Tiersch et al., 1992; Spotila et al., 1994).

1.5.2 Identification of sex determination markers at genomic DNA lavel

Griffiths and Tiwari (1993) used 16 primers of 21 – 29 mer and 20 of 10 mer to isolate candidate sex-specific RAPD marker of the great tit (*Parus major*), the jackdaw (*Corvus monedula)* and the zebra finch (*Taenopygia guttata)*. Of these, ss2 (5′-CGG TCG GGA GGT TTC AAG GAA TG-3′) amplified a female-specific 724 bp fragment of *P. major* whereas AB09 (GGGCGACTAC) and AB18 (CTGGCGTGTC) yielded a candidate female-specific fragments in *C. monedula* (1100 bp) and *T. guttata* (900 bp), respectively. Each of these amplified bands was used as the probe against genomic DNA of males and females of each species using Southern Blot analysis. Results confirmed their W-linked and also indicated the occurrence of related sequences that do not link with the sex chromosome.

Griffiths and Orr (1999) carried out AFLP analysis for isolation of sex-specific markers in the ostrich (*Struthio camelus*) previously report to posess a ZW heterogamy for females. Screening 3 male and 3 female ostriches with a single E-primer (E_{AGG}) in combinations of each of the eight of M primers 2 W-linked markers from E_{AGG}/M_{CAG} and E_{AGG}/M_{CAA} primer combinations were obtained, respectively. These two primer combination were tested against 8 individuals of female ostriches and can be identified their gender correctly. ScW1F/R and ScW2F/R derived from cloned AFLP fragments of the ostriches exhibited female specificity when further tested with 12 individuals of ostriches.

Griffiths et al. (2000) isolated sex-specific markers of the three-spined stickleback (*Gasterostus aculeatus*) by AFLP. Specimens from geographically different locations were collected from Edinburgh, Scotland (*N* = 16), Milngavie, Scotland (*N* = 5), Silverdale, England (*N* = 8) west coast of Canada (*N* = 6) and various population in Japan (*N* = 10) and sexed. The nine-spined *Pungitus pungitus* (*N* = 8) and 15-spined sticklebacks *Spinachia spinachia* (*N* = 6) collected from Sweden were included as the control. Males (*N*=3) and female (*N*=3) individuals were screened for identification of candidate sex-

specific AFLP markers. Primers produced bands in one sex were re-screened with additional 6 individuals. Two candidate male-specific markers were found from primers E_{AAG}/M_{CAA} and E_{AAG}/M_{CAG}, respectively. SCAR markers were developed. Primers GalF/R yielded a 600 bp fragment in females and both 600 bp and 371 bp fragments in males. Therefore, it was concluded that this species possesses the XX/XY sex chromosomal system rather than the ZW/ZZ system previously reported by Beatty (1964).

Kovacs et al. (2001) identified sex-specific DNA markers in the African catfish, *Clarias gariepinus* (Burchell, 1822). RAPD analysis was carried out against pooled genomic DNA of males and females of *C. gariepinus*. Two sex-linked RAPD markers (2.6 kb for CgaY1 and 458 bp for CgaY2) were identified from the pooled DNA of male *C. gariepinus*. The specificity was further confirmed on individual specimens. Both RAPD markers were cloned and sequenced. Southern blot analysis using the CgaY1 probe under high or low stringency conditions showed strong hybridized fragments only in males but not in females. Patterns indicated the existence of indicating the presence of multiple copies of CgaY1 in the male genome. CgaY2 produced a similar hybridization pattern in both sexes of *C. gariepinus.* Sequence-pecific primers were designed. The specific marker bands were co-amplified with a band commonly found in all individuals. Three band (male) versus one band (female) patterns were found allowing rapid molecular sexing of *C. gariepinus*.

Karyotypes of abalone have been extensively studies. On the basis of chromosome numbers, abalone can be divided into 3 groups; ones having 2n = 28 (*H. tuberculata* and *H. lamellosa*) distributed in the European-Mediterranean area, others (including *H. asinina*) having 2n = 32 (*H. asinina*, *H. diversicolor*, *H. exigua*, *H. planata*, *H. ovina* and *H. varia*) distributed in the Indo-Pacific region and the rest having 2n = 36 (*H. cracherodii*, *H. discus discus*, *H. discus hannai* and *H. madaka*) distributed in the North Pacific region (Jarayabhand et al., 1998) Neither sex chromosome nor temperature sex determination (TSD) have been reported in abalone.

1.5.3 Sex specific/differential expression analysis

The EST approach is powerful for identification of expressed transcripts from cDNA libraries. Zeng and Gong (2002) studied gene expression in cDNA libraries established from testes and ovaries of zebrafish (*Danio reio*). A total of 1025 ESTs (501

from the testis and 524 from the ovary libraries) were sequenced. A total of 641 EST matched with known sequences in the GenBank. Both libraries have a higher portion of clones for nuclear proteins and a lower portion for proteins in translational machinery, cytoskeleton and mitochondria than the previously characterized whole-adult cDNA library. Abundant cDNA clones in the two gonad libraries were identified and over 10% of ovary clones were egg membrane encoded proteins (zona pellucida or ZP proteins). Furthermore, the testis library showed a more even distribution of cDNA clones with relatively fewer abundant clones in the library. Eleven potential cDNA clones were used as molecular markers for the analysis of the gonad development in zebrafish by Northern blot hybridization. Most of them showed a specific or predominant expression in testes or ovaries as expected. Four of these were further analyzed by *in situ* hybridization and found to be expressed specifically in the germ cells of testes and ovaries. Therefore, they are suitable markers for analyses of spermatogenesis and oogenesis in the zebrafish.

Inaba et al (2002) isolated and characterized gene expression in testes of the ascidian, *Ciona intestinalis* using an EST approach. A set of 5,461 ESTs was analyzed and grouped into 2,806 independent clusters. Approximately 30% of the clusters showed significant sequence matches to the proteins reported in DDBJ/GenBank/EMBL database including a set of proteins closely related to the gene regulation during functional and morphological changes of spermatogenic cells during spermiogenesis, and physiological functions of sperm. Some clones show similarities to the proteins present in vertebrate lymphocytes, suggesting a primitive immune system in ascidians. Other are known to participate in hormonal regulation of spermatogenesis in vertebrates. The large majority of the genes expressed in *Ciona* testis show no significant matches to known proteins and the further analysis of these genes may shed new light on the molecular mechanism of spermatogenesis and sperm functions in invertebrates.

The subtractive hybridization technique was used to identify differences in gene expression between males and females of medaka (*Oryzias latipes*) during sex differentiation. Fifty female-specific cDNA fragments were cloned. They were classified into three groups by their expression at 1, 5, or 30 days after hatching. Fifteen EST showing near full-length cDNAs belonging to the first two groups were cloned. Many of these female-specific genes are coordinately expressed in oocytes at the earliest stages of oogenesis. Some gene homologues including egg envelope proteins, oocyte-specific RNA

binding proteins, and a transcription factor containing a basic helix-loop-helix motif were identified (Kanamori, 2000).

Ota et al (2002) used an mRNA subtraction approach for construction of a testis-specific cDNA library of the silk worm (*Bombyx mori*). Several clones were randomly sequences. *BmTST*, contained a 3/-part of an ORF homologous to tektin (a the protein known to form filamentous polymers in the walls of ciliary and flagellar microtubules) was isolated. In addition, a genomic fragment containing the promoter region and the 5/-part of the coding sequence of *BmTST* was also isolated. The ORF of BmTST was predicted to encode a protein of 508 amino acid residues, having sequence similarity of 28%, 28% and 30% identities with tektins A1, B1 and C1 of *Strongylocentrotus purpuratus*, respectively. Expression analysis by RT-PCR and Western blotting analysis indicated that the *BmTST* gene was expressed specifically in the testis during sperm maturation. The protein was immunologically detected exclusively in the fraction expected to contain the 9q2 flagellar axonemes of sperms. Results indicated that the *BmTst* protein is possibly involved in the spermatogenesis of *B. mori*.

Sex-lethal (*Sxl*) protein gene encodes the RNA binding protein that regulates genes expression in sex differentiation and dosage compensation of *Drosophila* (Bell et al 1988). *Sxl* acts as a regulator of both alternative pre-mRNA splicing and translation. *Sxl* controls splicing of the *transformer* (*tra*) pre-mRNA, its downstream target in sex determination cascades, by binding to and thereby blocking an optimal splice acceptor site (Sosnowski et al., 1989; Inoue et al., 1990). This allows female *D. melanogaster* to produce the functional *tra* protein through female-specific splicing. The female *tra* protein induces female-specific splicing of the *doublesex* (*dsx*) pre-mRNA in cooperation with the *tra*-2 gene product. However, *sxl* found in *Drosophila melanogaster* are not well conserved across different groups of organisms.

DMRT1 is one of the conserved genes involved with sex differentiation found from invertebrates to vertebrates. DMRT1 has been fully characterized in human and shows similarity with doublesex (*dsx*) and *mab-3* genes in *D. melanogaster* and *Caenorhabditis elegans*, respectively. These genes also share a number of properties and contain the conserved DM domain (a new zinc finger DNA binding motif, C2H2C4). Characterization of DMRT1 in different species has provided additional support for its involvement in testis development and sex differentiation processes. (Marchand et al.,

2000; Raymond et al., 2000; Zhu et al., 2000). Orthologues of DMRT1 have been discovered on the Z chromosome in birds (Nanda et al., 2000), reptiles and other lower vertebrates (Raymond et al., 1999a; Shan et al., 2000; Marchand et al., 2000; Guan et al., 2000).

The DM domain was first identified in the *doublesex* (*dsx*) gene, the final gene of sex-determining cascades in *D. melanogaster*. The *dsx* is alternatively spliced to encode male- and female-specific proteins. The latter protein regulates the expression of sex-specific differentiation gene such as *yolk protein* 1 (YP1) (Burtis et al., 1991; Jursnich et al., 1993; An et al., 1995a, b). The *dsx* protein binds to a specific site of an enhancer sequence in the 5′ flanking region of the yolk protein genes and regulates their transcription (Garabedian et al., 1986; Coschigano et al., 1993).

Male abanormal-3 *(mab-3)* of the nematode *C. elegans* encodes a protein with two copies of a non-classical DM domain (Raymond et al., 1998). The *mab-3* gene functions downstream of *tra-1* to promote male-specific development of two tissues. It directs the development of sensory rays (V rays) that are required for mating, in the peripheral nervous system of the male tail. It also prevents expression of vitellogenin (yolk proteins) in the male intestine (Shen and Hodgkin, 1998).

Both *dsx* and *mab-3* genes are direct transcriptional regulators of yolk protein genes (Coschigano and Wensink, 1993; Yi and Zarkower, 1999) and are required for differentiation of sex-specific sense organs (Baker and Ridge, 1980; Shen and Hodgkin, 1988). The male-specific isoform of *dsx* can substitute for *mab-3* in the *C. elegans* male peripheral nervous system, indicating that the two proteins are functionally similar (Raymond et al., 1998). The *dsx* appears to be more widely required than *mab-3*, by regulating all external sexually dimorphic features (Baker and Ridge, 1980) and playing a role in the central nervous system that is essential for mating behavior (Villella and Hall, 1996). Based on their extensive similarities, *mab-3* and *dsx* may represent the first example of evolutionary conservation between distantly related sex-determination pathways (Raymond et al., 1998).

Marchand et al. (2000) identified a DMRT1 homologue in the rainbow trout (Oncorhynchus mykiss, rtDmrt1) by RT-PCR using degenerate primers designed against conserved regions of different DM domains of *D. melanogaster dsx*, *C. elegans mab-3* and

human DMRT1. Semi-quantitative RT-PCR analysis showed that rtDMRT1 is expressed during testicular differentiation, but not during ovarian differentiation. After 10 days of steroid treatment, expression was shown to be decreased in estrogen-treated male differentiating gonads but not to be restored in androgen-treated female gonads. In adults, a single 1.5 kb transcript was detected by Northern blot analysis in the testis, and its expression was found to be sustained throughout spermatogenesis and declined at the end of spermatogenesis (stage VI). RT-PCR analysis detected a very low level of rtDMRT1 expression in the ovaries of *O. mykiss*.

1.5.4 Male- and female-specific proteins

Tektins are microtubule-associated cytoskeletal proteins that are primarily expressed in the male germ cell-lineage in centrioles and basal bodies and within ciliary and flagellar doublet microtubules. They are proposed to be important for axonemal architecture and microtubule stability in the sperm tail and in other ciliated and flagellar structures (Roy, 2004). Tektins have so far been described in the sea urchin *Strongylocentrotus purpuratus*, human *Homo sapiens* (Iguchi et al., 2002), the canine (Zhiyong et al., 2000), and the mouse *Mus musculus* (Wolkowicz et al., 2002).

Tektins form filaments of 2–3 nm in diameter and are composed of equal amounts of A, B and C subunits. (Norrander et al., 1996). The filaments composed of tektin AB heterodimers are associated with the A-microtubule wall of cilia flagellar doublet microtubules. The C homodimers form separate tektin filaments assembled onto the surface of AB filaments (Pirner and Linck, 1994). At present the roles of tektins are suggested to stabilize tubulin protofilaments, to form the attachment points between A and B tubules of ciliary flagellar doublet microtubules and the C-tubules of centrioles and to generate the complex, long-range patterns of binding sets for axonemal components, such as nexins, radial spokes and dynein arms (Norrander et al 1998).

Xu el al (2001) cloned and characterized a new member of the tektin gene family from the cDNA library established from human fetal brain. The tektin1 cDNA consists of 1375 bp and has a putative open reading frame that encodes a protein of 418 amino acids. The predicted protein is 48.3 kDa in size, and its deduced amino acid sequence is 82% identical to that of the mouse, rat, and dog. One conserved peptide domain RPNVELCRD was observed at position number 323–331 of the amino acid sequence, which is a

prominent feature of tektins and is likely to represent a functionally important protein domain.

Tektin 1 gene was mapped to the human chromosome 17 and at least eight exons were found. Northern blot analysis indicated that tektin 1 was predominantly expressed in testis. Using *in-situ* hybridization analysis, tektin1 mRNA was localized in spermatocytes and round spermatids in the seminiferous tubules of the mouse testis, indicating that it may play an important hold in spermatogenesis.

Roy (2004) isolated a new member of the tektin gene family (tektin 3) from mice. RT-PCR and northern blot analyses confirmed that tektin3 encodes a 1.7 kb transcript detectable preferentially in the testes of adult mice. *In situ* hybridization analysis in the testes revealed that tektin3 mRNA was expressed exclusively in late pachytene spermatocytes and early round spermatids. The putative tektin 3 protein shares 83.5% overall sequence identity with its ortholog in human and includes a fully conserved carboxy terminal nonapeptide signature sequence (RPNVELCRD) as is present in all family members. Using data mining, highly conserved tektin3 orthologs were identified in the puffer fish (*Fugu rubripes*) and rats (*Rattus norvegicus*) which shared 60.9 and 91.4% identity, with mouse tektin3, respectively. Results demonstrate that tektin3 is a novel male germ cell-enriched protein, which is evolutionarily conserved, and it might perform important roles in male reproductive development and physiology.

Axonemal protein is a protein associated with axonemal structure of the sperm flagella. Eukaryotic cilia and flagella are elaborate molecular machines constructed of more than 250 proteins (Dutcher, 1995). They are centered by precisely organized, microtubule-based structures, the axonemes. The axoneme consists of two central singlet microtubules, called the central pair, and nine outer doublet microtubules. These structures are well conserved during evolution. The outer doublet microtubules, each composed of A and B subfibers, are connected to each other by nexin links, while the central pair is held at the center of the axoneme by radial spokes. Motility in cilia and flagella is generated by sliding of outer doublet microtubules driven by inner and outer dynein arms that protrude from the A tubule (Gibbons, 1981).

Many marine invertebrate species spawn gametes into seawater, where fertilization and embryogenesis occur. In abalone, congeneric species may have overlapping breeding

seasons and habitats. Such species usually exhibit species selectivity (specificity) to prevent cross-fertilizations.

Abalone sperm possess a large acrosomal granule containing roughly equal amounts of sperm lysin (16 kDa) and fertilization protein (FP, 18 kDa). Contaction between sperm and egg glycoproteinaceous vitelline envelope (VE) triggers an acrosomal reaction. The dimeric lysin is deposited onto the surface of the egg VE where it binds to the VE receptor for lysin (VERL) and creates a hole in the VE by a species-specific mechanism. FP coats the membrane covering the acrosomal process where it is thought to mediate sperm and egg cell membrane fusion. Sperm lysin and fertilization protein were generated by gene duplication from a shared ancestral protein. Although lysin retains some ability to mediate membrane fusion, the FP is unable to dissolve the VE (Kresge et al, 2001).

VERL is a long unbranched glycoprotein of about 1000 KD and is the major component (about 30%) of the egg VE protein. VERL contains approximately 50% carbohydrate by weight. Binding of lysin to isolated VERL shows high affinity, positive cooperatively, and species specificity (Swanson and Vacquier, 1997). Cloning and sequencing of VERL shows that it is composed of approximately 26 repeats of a 153 amino acid sequence that evolves by concerted evolution. Amino acid substitution in VERL repeats between species shows a pattern in which hydrophilic residues are replaced by hydrophobic residues (Swanson and Vacquier, 1998).

Vitellogenin (Vg) is the major precursor of the egg-yolk proteins, vitellin (Vn), which provide energy reserves for embryonic development in oviparous organisms. They are high-density (300–700 kDa, according to species) glycolipophosphoproteins having Ca^{2+} and Zn^{2+} ligands. Vg synthesis occurs in response to endogenous estrogens, which are released into the bloodstream and stored in developing oocytes of mature females (Wahli et al., 1981). Although liver is generally recognized as the organ responsible for vitellogenesis in fish, identification of the organ(s) involved in Vg production in molluscs appears to be more controversial. On the basis of histological observations, Pipe (1987) suggested that Vg are synthesised within the oocytes (autosynthesis) of the mussel, *Mytilus edulis*, whereas Eckelbarger and Davis (1996) proposed a heterosynthetic process for the oyster, *Crassostres virginica* where yolk protein precursors being synthesised in connective tissue surrounding gonads.

Matsumoto et al (2003) used a differential display method to identify genes restrictively expressed in ovaries of the Pacific oyster (*C. gigas)*. One of the isolated fragments was a homologue of vitellogenin. A full-length cDNA encoding putative vitellogenin was isolated by RACE –PCR. The ORF encoded a protein of 1583 amino acid residues. The deduced primary structure of putative vitellogenin in *C. gigas* was similar to that of various mollusc, fish, crustacean and nematode species, especially at the *N*-terminus. RT-PCR revealed that mRNA encoding putative vitellogenin was expressed only in the ovaries. *In situ* hybridization analysis revealed that putative vitellogenin mRNA was expressed strongly in the follicle cells in the ovaries of *C. gigas*. The follicle cells are proposed as the site of vitellogenin synthesis in *C. gigas*.

A described previously *H. asinina* is the most promising species for the aquaculture industry in Thailand. Currently, artificially propagated breeding programs and culture techniques of *H. asinina* are well developed and indicated that the growth rate of females in this species is greater than that of males for approximately 25% (P. Jarayabhand unpublished data). Additional, Shephred and Hearn (1983) reported that female *H. laevigata* in natural populations has approximately 25% greater growth rate than does the natural male *H. laevigata*. As a result, the basic information on sex-specific expression markers and genes involving with testicular and ovarian development is necessry for understanding of sex-differentiation process in this species in *H. asinina*

Sex chromosomes in abalone are not yet cytological identifiable. The effective method for sex determination of abalone is development of sex-specific DNA (genomic DNA or cDNA) markers in this economically important species. Generally, sex-specific markers are present in the genome whenever sex determination are genetically, not environmentally, controlled and sex determining genes are located on sex chromosomes. A lack of the basic knowledge on systems of sex chromosomes (XY, ZW or others) and their segregation patterns in abalone prohibits the possibility to elevate culture efficiency of *H. asinina* through farming of all *H. asinina* females. Sex determination and differentiation markers can be used to verify the genetic system of sex chromosomes in *H. asinina*. Moreover, the markers can be used to evaluate the possibility to use and to follow the success of genetic manipulation (e.g. gynogenesis, androgenesis, and polyploidy production) in *H. asinina.*

1.6 Objectives of the thesis

Appropriate genetic markers can be used to increase culture and management efficiency of abalone in Thailand. In addition, species-specific markers play important roles for quality control to prevent incorrect use of abalone larvae for the culture industry and abalone species in canning. In addition, the application of sex-specific DNA/cDNA markers is probably useful for primary understanding of sex determination system or sex differentiation cascades of *H. asinina*. The aims of this thesis is identification and development of species-specific markers of *H. asinina*, *H. ovina* and *H. varia* in Thailand using RAPD and SCAR markers and to isolation and characterized sex-specific DNA/cDNA markers of *H. asinina* using AFLP, EST, cDNA subtraction and RT-PCR approaches.

CHAPTER II

MATERIALS AND METHODS

2.1 General Materials and Methods

The experimental procedures of this thesis were divded to 4 major parts including

1 : Development of species-specific markers of the tropical abalone (*H. asnina*, *H. ovina*, and *H. varia*) in Thailand using RAPD and SCAR markers.

2 : Isolation and characterization of sex-specific markers in *H. asinina* using AFLP analysis.

3 : Isolation and characterization of expressed sequence taq (EST) in ovaries and testes of *H. asinina.*

4 : Isolation and characterization of sex-related genes in ovaries and testes of *H. asinina*.

The experiments were carried out from isolation and characterization of molecular markers that can be applied for genetic management of *H. asinina* at the genomic DNA level. In addition, knowled about transcripts expressed in ovaries and testes were studied. The basic knowledge obtained can be applied for culture and management of this economically important species.

2.2 Development of species-specific markers of the tropical abalone (*H. asnina*, *H. ovina*, and *H. varia*) in Thailand using RAPD and SCAR markers

2.2.1 Sampling

The tropical abalone (*H. asinina*) were collected from 7 samples (Figure. 1) including natural abalone from Talibong Island located in Trang Province (HATRAW, N = 23), Samet Island located in Rayong Province (HASAME, N = 10), Cambodia (HACAME, N = 20) and Indonesia (HAINDW, N = 12) and hatchery samples composing of the first (P_0) and second (F_1) generation samples which founders were originated from

Cambodia (HACAMHE, *N* = 12) and Samet Island (HASAMHE, *N* = 15) and from Philippines (HAPHIE, *N* = 19), respectively (Figure. 2.1 and Table 2.1).

Additional specimens of natural *H. ovina* (*N* = 73) were collected from Sichang Island (HOCHOE, *N* = 20) and Samet Island (HOSAME, *N* = 19) located in the Gulf of Thailand and Churk Island (HOTRAW, *N* = 18) and Similan Island (HOPHAW, *N* = 16) located in the Andaman Sea (Table 2.1). Moreover, *H. varia* which is restrictively distributed in the Andaman Sea was collected from L-Island (HVPHUW, *N* = 29) and Similan Island (HVPHAW, *N* = 3). Speecimens were individually kept in a –30 ^{o}C freezer. Alternatively, the foot muscle of each individual was dissected out and kept in a -80 ^{o}C freezer until further used.

2.2.2 Genomic DNA extraction

2.2.2.1 A proteinase K /phenol-chloroform method

A piece of the foot tissue was dissected out from each specimen as soon as possible after removing from a freezer. The tissue was placed in a prechilled microcentrifuge tube containing 600 µl of the extraction buffer (100 mM Tris-HCl, 100 mM EDTA, 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^{o}C for 1 hour. At the end of the incubation period, proteinase K (10 mg/ml) was added to the final concentration of 300 µg/ml and further incubated at 55 ^{o}C for 3 hours.

An equal volume of buffer-equilibrated phenol: chloroform: isoamyl alcohol (25:24:1) was added and gently mixed for 10 minutes. The mixture 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 and further extracted once with chloroform:isoamylalcohol (24:1). The

Table 2.1 Sample collection sites and sample sizes of abalone (*H. asinina*, *H. ovina* and *H. varia*) used in this study

Geographic origin	Abbreviation	Sample size (*N*)
H. asinina		
Samet Island, Rayong (Gulf of Thailand)*	HASAME	10
P_0 stock, Rayong (Gulf of Thailand)	HASAMHE	15
Cambodia (east of peninsular Thailand)	HACAME	20
P_0 stock, Cambodia (east of peninsular Thailand)	HACAMHE	12
Talibong Island, Trang (Andaman Sea)	HATRAW	23
F_1, Philippines**	HAPHIE	19
Indonesia	HAINDW	12
H. ovina		
Sichang Island, Chon Buri (Gulf of Thailand)	HOCHOE	20
Samet Island, Rayong (Gulf of Thailand)	HOSAME	19
Churk Island, Trang (Andaman Sea)	HOTRAW	18
Similan Island, Phangnga (Andaman Sea)	HOPHAW	16
H. varia		
L-Island, Phuket (Andaman Sea)	HVPHUW	29
Similan Island, Phangnga (Andaman Sea)	HVPHAW	3
Total (*N*)		216

*Hemolymph of each individual was collected and used for DNA extraction, **the whole individual was kept in absolute ethanol at -20° C.

Figure 2.1 Map of Thailand indicating sampling collection sites for abalone (*H. asinina, H. ovina and H. varia*) used in this study. Dots represent geographic locations (excluding *H. asinina* from Indonesia and Philippines) for which at least one abalone species was sampled. Note that Samet Island (SAM) is located in Rayong (RAY) province.

aqueous phase was transferred into a sterile microcentrifuge. One-tenth volume of 3 M sodium acetate, pH 5.2 was added. DNA was precipitated by an addition of two volume of prechilled absolute ethanol. The mixture was mixed thoroughly and incubated at -80^{o} C for 30 minutes. The precipitated DNA was recovered by centrifugation at 12,000 rpm for 10 minutes at room temperature and washed twice with 1 ml of 70% ethanol. The DNA pellet was air-dried and resuspended in 50μl of TE buffer (10 mM Tris-HCl, pH 8.0 and 0.1 mM EDTA). The DNA solution was incubated at 37 ^{o}C for 1 - 2 hours and kept at 4 ^{o}C until further needed.

2.2.2.2 A chelex-based method

A piece of the foot tissue of adult *H. asinina* (or a 2 week-old *H. asinina* larva) was homogenized in 200 μl of a 5% chelex solution; (w/v). The mixture was incubated at 60°C for 3 hours and placed in a bath containing boiling water for 10 minutes. The mixture was briefly centrifuged in a microcentrifuge at 13,000 rpm. The supernatant was transferred to a new tube and kept at 4 ^{o}C. This solution was used as the template for PCR.

2.2.3 Measuring DNA concentrations using spectrophotometer and electrophoresis

The concentration of extracted DNA is spectrophotometrically 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. Therefore, the concentration of DNA samples is estimated in μg/ml by multiplying an OD_{260} value with a dilution factor and 50. The purity of DNA samples can be obtained by calculating a ratio of OD_{260} /OD_{280}. The ratio which is much lower than 1.8 indicates contamination of residual proteins or organic solvents in the DNA solution wheraras that much greater than 1.8 implies contamination of RNA (Kirby, 1992).

Agarose gel electrophoresis (0.8% gel prepared in 1x TBE buffer; 89 mM Tris-HCl, 8.91 mM boric acid and 2.5 mM EDTA, pH 8.3) is an alternative method for rough estimation of high molecular weight DNA on the basis of its direct relationship between the amount of DNA and the level of fluorescence after ethidium bromide staining. DNA

concentration was estimated from the intensity of the fluorescent band by comparing with that of undigested λ DNA.

2.2.4 RAPD-PCR

One hundred and thirteen primers were previously screened for the amplification success against 2-3 representatives of each abalone by Popongviwat (2001). Five selected primer (OPB11, UBC101, UBC195, UBC197 and UBC271) were tested against different populations of *H. asinina* (*N* = 5 per location), *H. ovina* (*N* = 7-8 per location), and *H. varia* (*N* = 15 and 3 for HVPHUW and HVPHAW).

RAPD-PCR were performed in a 25 μl reaction volume containing 10 mM Tris-HCl pH 8.3, 50 mM KCl, 0.001% gelatin, 2-3 mM of $MgCl_2$, 100 μM of each dATP, dCTP, dGTP and dTTP, 0.2 μM of a primer, 1.0 unit of Ampli*Taq* DNA polymerase, and 25 ng of a DNA template. PCR was performed in an Omnigene thermal cycler consisting of a predenaturation at 94 ^{o}C for 3 minutes followed by a 94 ^{o}C denaturation for 10 seconds, a 36 ^{o}C annealing for 30 seconds and a 72 ^{o}C extension for 90 seconds. The final extension was carried out at 72 ^{o}C for 5 minutes

2.2.5. Agarose gel electrophoresis

RAPD products were electrophoretically analyzed by 1.6% agarose gels. An appropriate amount of agarose was weighed out, mixed with 1x TBE and heated in a microwave oven to complete solubilization. After allowing to approximately 60 ^{o}C, the agarose solution was poured into the chamber set with an inserted comb. When the gel had solidified, the comb was carefully removed. The agarose gel was submerged in a chamber containing an enough amount of 1x TBE buffer covering the gel for approximately 0.5 cm.

Ten microliters of the amplification products were mixed with 2 μl of the loading dye (0.25% bromophenol blue and 25% Ficoll in water) and loaded into the well. A 100 bp DNA ladder was used as the standard DNA marker. Electrophoresis was operated at 100 V until bromophenol blue moved to approximately 0.5 cm from the bottom of the gel. The gel was stained with ethidium bromide (2.5 μg/ml) for 5 minutes and destained in deionized water with gently shaking for 20 minutes to remove unbound ethidium bromide from agarose gels. DNA fragments were visualized as fluorescent bands under

an UV transilluminator and photographed through a red filter using Formapan 100 film. The exposure time was usually about 15-18 seconds.

2.2.6 Cloning of RAPD fragment into pGEM®-T Easy vector

2.2.6.1 Preparation of RAPD fragments

The required DNA fragment was fractionated through an agarose gel in duplicate. One was run side-by-side with a 100 bp DNA markers (usually on the left). The other was loaded into another well (usually on the right) of the gel. After electrophoresis, agarose gel containing the DNA standard and its proximal DNA sample were cut and stained with ethidium bromide (0.5 μg/ml). The position of non-stained target DNA fragment was then referenced without exposure to the UV.

Twenty-two RAPD fragments, showing population-, region-, or species-specificity for each abalone were reamplified using the original primers, size-fractionated and excised from a 1.6% agarose gel and eluted out from the gel using a Prep-A-Gene DNA purification kit according to the protocol recommended by the manufacturer (Bio-Rad Laboratories).

2.2.6.2 Ligation of PCR fragment into pGEM®-T Easy vector

The eluted DNA was ligated to pGEM®-T Easy vector in a 10 μl ligation reaction constituting of 5 μl of 2x Rapid Ligation Buffer (60 mM Tris-HCl, pH 7.8, 20 mM $MgCl_2$, 20 mM DTT, 2 mM ATP, 10% polyethylene glycol; MW 8000), 3 weiss units of T4 DNA ligase, 25 ng of pGEM®-T Easy vector and approximately 100 ng of DNA insert. The ligation reaction was incubated overnight at 4 ^{o}C.

2.2.6.3 Transformation of ligated products to *E. coli* host cells by electroporation (Dower et al., 1988)

2.2.6.3.1 Preparation of competent cells

A single colony of *E. coli* XL1-BLUE was inoculated in 10 ml of LB-broth (1% Bacto tryptone, 0.5% Bacto yeast extract and 0.5% NaCl) with vigorous shaking at 37^{o} C overnight. The starting culture was inoculated into 1 liter of LB-broth and continued culture at 37^{o} C with vigorous shaking to the OD_{600} of 0.5 to 0.8. The cells were chilled

briefly on ice for 15 - 30 minutes, and harvested by centrifugation in a prechilled rotor at 4,000 g for 15 minutes at 4^{o} C. The pellet were resuspended in 1 liter of cold water, centrifuged as above and resuspended in 0.5 liter of cold water. The cell suspension was centrifuged at the same speed and the pellet was resuspended in 20 ml of 10% glycerol. The cells were centrifuged, and finally resuspended in 2 - 3 ml of 10% glycerol and divided to 45 µl aliquots. These cells could be used immediately or stored at –70 ^{o}C for subsequently used.

2.2.6.3.2 Electrotransformation

The competent cells were thawed on ice for 5 minutes. One or two microliters of the ligation mixture was added and gently mixed by pipetting. The mixture was left on ice for approximately 1 minute. The mixture was eletroporated in a prechilled 0.2 cm cuvette using a Gene Pulser (Bio-Rad) with the setting parameters of 25 µF 200Ω and 2.5 KV. After electroporation, the mixture were immediately removed from the cuvette and added to a new tube containing 1 ml of SOC medium (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 ^{o}C for 1 - 2 hours. Approximately 10-30 µl of the transformed mixturethis were spread on a LB agar plates supplemented with 50 µg/ml of ampicillin, 25 µg/ml of IPTG and 20 µg/ml of X-Gal and further incubated at 37 ^{o}C overnight (Sambrook et al., 1989). The recombinant clones containing inserted DNA are white whereas those without inserted DNA are blue.

2.2.6.4 Colony PCR

Colony PCR was performed in a 25 µl PCR reaction containing 10 mM Tris-HCl, pH 8.8, 50 mM KCl, 0.1% Triton X - 100, 100 mM of each of each dNTP (dATP, dCTP, dGTP, and dTTP), 2 mM of $MgCl_2$, 0.1 µM each of pUC1 ($5^{/}$-CCG GCT CGT ATG TTG TGT GGA-$3^{/}$) and pUC2 ($5^{/}$-GTG GTG CAA GGC GAT TAA GTT GG-$3^{/}$), 1.0 unit of DyNAzyme^{TM} II DNA Polymerase (Finnzymes). A white colony was gently picked by a pipette tip and mixed well with the mixture. PCR was performed including predenaturation at 94°C for 3 minutes followed by 35 cycles of denaturation at 94°C for 1 minute, annealing at 55°C for 1 minute and extension at 72°C for 1.30 minutes. Final extension was performed at 72°C for 7 min. The colony PCR products were

electrophoresed though 1.2% agarose gels, stained with ethidium bromide and visualized under a UV transilluminator.

2.2.6.5 Isolation of recombinant plasmid DNA

The recombinant plasmid DNA was isolated using NucleoSpin® Plasmid Kit (Macherey-Nagel, Germany). A discrete white colony was inoculated into a sterile tube containing 2 ml of LB broth supplemented with 50 μg/ml of ampicillin and incubated with shaking at 37° C overnight. The culture was transferred into a new 1.5 ml microcentrifuge tube and centrifuged at 10,000 rpm for 1 minute. The supernatant was carefully decanted. Recombinant plasmid was extracted according the conditions recommended by the manufacturer. One or two microliters of the recombinant plasmid was digested with *Eco*RI (2 units) in a 15 μl reaction containing 1X restriction buffer (90 mM Tris-HCl; pH 7.5, 10 mM $MgCl_2$ and 50 mM NaCl) at 37°C for 3 hours. The digested product was analyzed by agarose gel electrophoresis.

2.2.6.6 DNA sequencing and data analysis

The dideoxy chain termination-based method was carried out against 2-6 recombinant clones of each insert using a ThermoSequenase Fluorescent Labeled Primer Sequencing kit (Amersham Pharmacia Biotech) with M13 fluorescent dye labeled primers. The M13 forward and reverse primers were used as the sequencing primer. Briefly, a set of A, C, G and T reaction was prepared by combining 0.6 – 1.0 μl of a fluorescence-labeled primer and 0.5-1.0 μg of double stranded DNA and throughly mixed. Three microlitres of each reaction mixture were aliquoted to four separate reactions and 1 μl of A, C, G or T reagent mixes was added into each reaction.

The sequencing reaction mixtures were then subjected to amplification consisting of a predenaturing at 95 °C for 5 minutes followed by 30 cycles of denaturation at 95 °C for 30 seconds, annealing at 55 °C for 30 seconds and extension at 72 °C for 60 seconds. After the cycle sequencing was finished, 2 μl of formamide loading dye was added into each reaction. The sequencing reactions were loaded onto a gel based-automated DNA sequencer (Li-Cor, Lincoln, USA) and operated for 12-16 hours. Sequences of inserts were analyzed for both directions.

Nucleotide sequences were blasted against those deposited in the GenBank using the Blast*N* and Blanst*X* programs (avalible at http://www.ncbi.nlm.nih.gov). Significant probabilities were considered when the E value was less than 10^{-4}.

2.2.7 Primer design, sensitivity and species-specific tests

Twenty-one pairs of primers (Table 2.2) were designed using OLIGO 4.0. They were preliminarily tested for specificity against genomic DNA of *H. asinina* (*N* = 12), *H. ovina* (*N* = 4) and *H. varia* (*N* = 4).

Five pairs of primers (CUHA2, CUHA12, CUHA13, CUHO3, and CUHV1) were further examined against larger specimens (*N* = 216, Table2.1). The amplification reactions were performed in a 25 µl reaction volume containing 10 mM Tris-HCl pH 8.3, 50 mM KCl, 2 mM $MgCl_2$, 100 µM each of dATP, dCTP, dGTP and dTTP, 0.2 µM of each primer and 1.0 unit of Ampli*Taq* DNA polymerase. The reaction was carried out composing of predenaturation at 94 ^{o}C for 3 minutes followed by 25 cycles of 94 ^{o}C for 45 seconds, 65 ^{o}C for 30 seconds and 72 ^{o}C for 45 seconds. The final extension was carried out at 72 ^{o}C for 5 minutes. Ten microlitres of the amplification products were electrophoretically analyzed by 1.4% agarose gels.

Sensitivity of CUHA2, CUHA12, CUHA13 and CUHV1 primers was examined against varying concentrations of the target DNA template (10 pg - 25 ng), using the same conditions as were used for the specificity test. The stability of CUHA2, CUHA12 and CUHA13 primers was determined using genomic DNA extracted from 10 individuals of the ethanol-preserved larvae (approximately 1 mm in size stored at 4°C for 6 mounts) and frozen (3 years at -30°C), dried (80°C for 72 hour and kept at room temperature for 2 week), and boiled (10 min) *H. asinina* broodstock using the chelex-based method.

Table2.2 Nucleotide sequence, length and melting temperature of primers designed used to develop species-specific PCR of *H. asinina*, *H. ovina* and *H. varia*

Primer	Sequence	Length (bp)	Tm (°C)
CUHA1-F	5′-GAATCCAACATGCGTCAAAG-3′	20	58
CUHA1-R	5′-CTGGAAACAATCGCAGGTCA-3′	20	60
CUHA2-F	5′-TTGTTCAGCATTCTGTGGCAGTTCT-3′	25	72
CUHA2-R	5′-CTTCTTTTTTGCTGACCCTTTGGAG-3′	25	72
CUHA3-F	5′-CCTGGCAAAGGAAAGTGGTG-3′	20	62
CUHA3-R	5′-TGGTTAAAGTGTTCGCTCGT-3′	20	58
CUHA4-F	5′-TCAGCGAAACCAACCAACAC-3′	20	60
CUHA4-R	5′-TTGGACGCAGCTATTCACAT-3′	20	58
CUHA5-F	5′-TATTTCCGTCGCCATCCTAT-3′	20	58
CUHA5-R	5′-GATTCGCCCAGTTTATTTGC-3′	20	58
CUHA6-F	5′-AAGTGAGTTTCGTTTTATGCC-3′	21	58
CUHA6-R	5′-CATGCTGTTGATGACTGGATT-3′	21	60
CUHA7-F	5′-ATTGTGTCATCATTTCGTGCAGTTG-3′	25	70
CUHA7-R	5′-CTTGCTCCTAGTCAGCGATTTTTCC-3′	25	74
CUHA8-F	5′-TCACCATTTAGCATAGTTCCTTTTA-3′	25	66
CUHA8-R	5′-ATGCGTTTTCTGGCTGTTTC-3′	20	58
CUHA9-F	5′-CATGGTTAGTTGTGAGGAGGAC-3′	22	66
CUHA9-R	5′-ACACCAGATACTGTTTCCGTTTAAG-3′	25	70
CUHA10-F	5′-GTCCGTCCCAGGTGCTTCATATC-3′	23	72
CUHA10-R	5′-AGCGCACGGCCAGTCATAGGTTG-3′'	23	74
CUHA11-F	5′-CCCCGAGGAGTATACAACTCTTCC-3′	24	74
CUHA11-R	5′-TCGAGTTCTTTTCCACAATGCACC-3′	24	70
CUHA12-F	5′-CTAATCCCACACAGCCATCACCAG-3′	24	74
CUHA12-R	5′-AAGAAGTGACGAAGAGGTAGGCAG-3′	24	72

Table 2.2 (cont.)

Primer	Sequence	Length (bp)	Tm (°C)
CUHA13-F	5′-TGACCTGTGTTGAGACTCTACGGA-3′	24	72
CUHA13-R	5′-TGAGGGGAGATGGAGTAGCCGC-3′	22	72
CUHA14-F	5′-CGTGAAGACAGTTACTGAAAGTGG-3′	24	70
CUHA14-R	5′-ATCGTTTGTGTTATGTCTCCTCTG-3′	24	68
CUHA15-F	5′-CGAAAGCAACAATAACTAAACC-3′	22	60
CUHA15-R	5′-TAACACGGATCTTAAAGTCG-3′	20	56
CUHO1-F	5′-GGTTCAACCGTTAAGTTCGCAGATG-3′	25	74
CUHO1-R	5′-ACCCGTACCAACAATGGAAGGA-3′	22	66
CUHO2-F	5′-ACCATGGGACTATTTCAATCTG-3′	22	60
CUHO2-R	5′-TCCAAGTCTGGGTGACGATAAC-3′	22	66
CUHO3-F	5′-GGGTATCTTCCCACAACAGC-3′	20	62
CUHO3-R	5′-GCACTTGCCTACATCCTTTCAC-3′	22	66
CUHO4-F	5′-AGAACACGGAACACACGCACGCAC-3′	24	76
CUHO4-R	5′-AATAACCACTCCTCCGCAACTCAC-3′	24	72
CUHO5-F	5′-TTCGTTTGGAGGGAAGTGAGCC-3′	22	68
CUHO5-R	5′-ACAGGGGATGGTGATAGTGTTGAG-3′	24	72
CUHV1-F	5′-CCCCTTGTTTCTCCTTCTTG-3′	20	60
CUHV1-R	5′-CGATGACGCAGGCGGTTTGA-3′	20	64

2.3 Isolation and characterization of sex-specific markers in *H. asinina* using AFLP analysis

2.3.1 Experimental animals

Adult male and female *H. asinina* were collected from the hatchery stock of *H. asinina* established from the Samet Island sample (subsequent generations of HASAMHE) and transported back to the laboratory at Center of Excellence for Marine Biotechnology, Faculty of Science, Chulalongkorn University. The foot muscle of each

abalone was dissected and kept in a -80 ^{o}C freezer. Genomic DNA was extracted from each abalone, pooled and used for AFLP analysis.

2.3.2 Bulked Segregant Analysis (BSA)

Two bulked genomic DNA of male (HaM1, N = 10 and HaM2, N = 8) and female (HaF1, N = 10 and HaF2, N = 8) *H. asinina* were generate by pooling either 25 or 31.25 ng of genomic DNA from each abalone. Bulked genomic DNA was used for screening of candidate sex-specific AFLP markers in this species.

2.3.3 Amplified Fragment Length Polymorphism (AFLP) analysis

2.3.3.1 Restriction enzyme digestion and adaptor ligation

Each bulked DNA (250 ng) was simultaneously digested with 2.5 units of *Eco* RI and *Mse* I in a 25 μl reaction mixture containing 10 mM Tris-HCl, pH 7.5, 10 mM Mg-acetate, 50 mM K-acetate at 37°C for 3 hours. The reaction was heat-terminated at 70°C for 15 minutes.

The *Eco* RI and *Mse* I adaptors (Table 2.2) were ligated to restricted genomic DNA by adding 24 μl of the adaptor-ligation solution (*Eco* RI and *Mse* I adaptors, 0.4 mM ATP, 10 mM Tris-HCl, pH 7.5, 10 mM Mg-acetate and 50 mM K-acetate) and 1 unit of T4 DNA ligase. The ligation reaction was incubated at 16°C for 16 hours.

2.3.3.2 Preamplification

One microliter of the ligation product was ten-fold diluted with TE buffer (10 mM Tris-HCl, pH 8.0 and 0.1 mM EDTA). The preamplification reaction was carried out in a 50 μl reaction volume containing 10 mM Tris-HCl, pH 8.8, 50 mM KCl, 0.1% Triton X-100, 200 μM of each dNTP,1.5 mM $MgCl_2$, 30 ng of E_{+A} (5′-GAC TGC GTA CCA ATT CA-3′) M_{+C} (5′-GAT GAG TCC TGA GTA AC-3′) primers, 1.5 units of DyNAzyme^{TM} II DNA Polymerase (Finnzymes) and 5 μl of the diluted ligation product. PCR was performed in a Perkin Elmer 9700 thermocycler consisting of denaturation at 94°C for 30 seconds, annealing at 56°C for 1 minute and extension at 72°C for 1 minute. Five microliters of the preamplification product was electrophoretically analyzed. The preamplification product was 50 fold diluted and subjected to selective amplification.

2.3.3.3 Selective amplification

Selective amplification was carried out in a 20 μl reaction volume containing 10 mM Tris-HCl, pH 8.8, 50 mM KCl, 0.1% Triton X-100, 200 μM of each dNTP, 1.5 mM $MgCl_2$, 30 ng of a combination pair of E_{+3} and M_{+3} primers (Table 2.3), 1.5 units of DyNAzyme^{TM} II DNA Polymerase (Finnzymes, Finland) and 5 μl of diluted preamplification product. The amplification reaction was carried out by denaturation at 94°C for 30 seconds, annealing at 65°C for 30 seconds and extension at 72°C for 60 seconds for 2 cycles followed by 12 cycles of a touch down phase with lowering of the annealing temperature for 0.7 in every cycle. Additional 25 cycles of 94°C for 30 seconds, 56 °C for 45 seconds and 72°C for 60 seconds was performed. The final extension was carried out at 72°C for 5 minutes.

2.3.4 Denaturing Polyacrylamide Gel Electrophoresis

2.3.4.1 Preparation of Glass Plate

The short and long glass plate was throughly wiped with 1 ml of 95% ethanol (commercial grade). This process was repeated twice. Afterwards, the short glass plate was coated with 1 ml of the Repel silane (2% dimethyldichlorosilane in octa-methylcyclotetrasitoxone) and left for approximately 10 minutes. Excess Repel silane solution was removed with a piece of tissue. The short glass plate was further cleaned with 95% ethanol for 3 times. The long glass plate was treated as described above using freshly prepared Bind silane (5 μl of Bind silane, 995 μl of 95% ethanol and 5 μl of 5% glacial acetic acid). The cleaned glass plates were assembled with a pair of 0.4 mM spacer. The bottom and both sides of assembled glass plates were sealed with the plastic tape.

2.3.4.2 Preparation of denaturing polyacrylamide gel

Six percent denaturing polyacrylamide gels were prepared by combining 40 ml of the degassed acrylamide solution (6% gel; 19:1 acrylamide: bisacrylamide with 7 M urea in TBE buffer) with 240 μl of freshly prepared 10 % ammonium persulphate

Table 2.3 AFLP primers and their sequences used for identification of sex-specific markers in *H. asinina*

Primer	Sequences
Adaptor sequences	
Eco RI adaptor	5′-CTC GTA GAC TGC GTA CC-3′
	5′-AAT TGG TAC GCA GTC TAC-3′
Mse I adaptor	5′-GAC GAT GAG TCC TGA G-3′
	5′-TAC TCA GGA CTC AT-3′
Preamplification primers	
E_{+A}	5′-GAC TGC GTA CCA ATT CA-3′
M_{+C}	5′-GAT GAG TCC TGA GTA AC-3′
Selective amplification primers	
E_{+3}-1	E_{+A}AC
E_{+3}-2	E_{+A}AG
E_{+3}-3	E_{+A}CA
E_{+3}-4	E_{+A}CT
E_{+3}-5	E_{+A}CC
E_{+3}-6	E_{+A}CG
E_{+3}-7	E_{+A}GC
E_{+3}-8	E_{+A}GG
E_{+3}-9	E_{+A}GT
E_{+3}-10	E_{+A}GA
E_{+3}-11	E_{+A}TG
E_{+3}-12	E_{+A}TC
E_{+3}-13	E_{+A}TA
E_{+3}-14	E_{+A}TT
M_{+3}-1	M_{+C}AA
M_{+3}-2	M_{+C}AC
M_{+3}-3	M_{+C}AG
M_{+3}-4	M_{+C}AT
M_{+3}-5	M_{+C}TA
M_{+3}-6	M_{+C}TC
M_{+3}-7	M_{+C}TG
M_{+3}-8	M_{+C}TT
M_{+3}-9	M_{+C}GA
M_{+3}-10	M_{+C}GT
M_{+3}-11	M_{+C}GC
M_{+3}-12	M_{+C}GG
M_{+3}-13	M_{+C}CA
M_{+3}-14	M_{+C}CT
M_{+3}-15	M_{+C}CG
M_{+3}-16	M_{+C}CC

and 24 µl of TEMED. The acrylamide solution was then gently injected into one side of the assembled plates using a 50 ml syringe. 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 paper and left at room temperature for 3-4 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 1x TBE. The sharkstooth comb was rinsed with water.

2.3.4.3 Electrophoresis

The gel sandwich was place 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 sharkstooth comb was reinserted into the gel until the teeth just touched the surface of the gel. Six microlitres of the acrylamide gel loading dye (98 % formamide, 200 ul EDTA, 0.25 % bromophenol blue and 0.25 % xylene cyanol) was loaded into each well. The gel was prerun at 40 W for 30 minutes. After that, the amplification product (6 µl) were mixed with 3 µl of the loading buffer and heated at 95°C for 5 minutes before snapped cooled on ice for 5 minutes. The sample was carefully loaded into the well. Electrophoresis was carried out at 40 W for approximately 2 hours.

2.3.5 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 1.5 litres of the fix/stop solution and agitates well for 40 minutes (25-30 minutes for SSCP gels). The gel was soaked with shaking 3 times for 2 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 4 seconds. The gel was transferred to 0.1% silver nitrate (1.5 litres) and incubated with agitation at room temperature for 30 minutes. The gel was soaked in 1.5 litres off deionized water with shaking (10 forward and 10 backward agitation) and immediately placed in the tray containing 1.5 litres 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 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 litres of chilled developer and shaked until bands from every lane were observed (usually 2 – 3 minutes). One litre 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 left at 80 °C for 2-3 hours.

2.3.6 Cloning of AFLP fragments into pGEM®-T Easy vector

2.3.6.1 Elution of DNA from polyacrylamide gels

Candidate sex-specific AFLP fragments were excised from dried polyacrylamide gels using a sterile razor blade and washed 3 times for 30 minutes each at room temperaturer with 500 µl of sterile deionized H_2O. Twenty microlitres of H_2O was then added and incubated overnight at 37°C. Reamplification of the target fragment was carried out using the same PCR recipes as those for selective amplification with the exception that 100 ng of each primer and 5 µl of the eluted AFLP product were used. The amplification conditions were composed of 35 cycles of 94°C for 1 minute, 55°C for 1 minute and 72°C for 1 minute. The final extension was performed at 72°C for 10 minutes. The reamplified product was electrophoretically analyzed through a 1.8 % agarose gel electrophoresis.

2.3.6.2 Elution of DNA from agarose gels, cloning and sequencing

After electrophoresis, the reamplified fragment was excised from the 1.8% agarose gel with a sterile razor blade. DNA was eluted out from the agarose gels using a NucleoSpin® Extract Kit (Macherey-Nagel, Germany) according to the protocol recommended by the manufacturer.

The gel-eluted AFLP fragment was ligated to pGEM®-T Easy vector and electrotransformed to *E. coli* XL1-Blue. The recombinant clones were identified by colony PCR. The recombinant plasmids were extracted by NucleoSpin® Plasmid Kit (Macherey-Nagel, Germany).The insert size was verified by digestion with *Eco*RI. DNA sequences of 2-5 recombinant clones of each insert were examined from both directions using an automated DNA sequencer (Li-Cor, Lincoln, USA) and blasted against those deposited in the GenBank using the Blast*N* and Blanst*X* programs

(http://www.ncbi.nlm.nih.gov). Significant probabilities were considered when the E value was less than 10^{-4}.

2.3.7 Primer design and PCR amplification

Eleven pairs of primers were designed from sequences of candidate sex-specific AFLP fragments (6 primer pairs for male and 5 primer pairs for female) using Primer Premier 5.0. Nucleotide sequences, length and melting temperatures of primers were shown in Table 2.3. Specificity of primers was preliminarily tested for against genomic DNA of males ($N = 3$ or 4) and females ($N = 3$ or 4) of *H. asinina*.

The PCR reactions were performed in a 25 µl reaction volume containing 10 mM Tris-HCl, pH 8.8, 50 mM KCl, 0.1% Triton X-100, 2.0 mM $MgCl_2$, 1 unit of Dynazyme^{TM} DNA Polymerase (Finnzymes), 100 µM of each dATP, dCTP, dGTP and dTTP, 0.2 µM of each primer (Tables 2.4) and 50 ng of genomic DNA. The amplification profiles were composed of predenaturation at 94°C for 3 minutes followed denaturation at 94°C for 1 minute, annealing temperature at 55°C for 1 minute and extension at 72°C for 1.5 minutes for 35 cycles. The final extension was carried out at the same temperature for 10 minutes. The amplification products were electrophoretically analyzed through 1.8% agarose gels and visualized under a UV transilluminator after ethidium bromide staining. The PCR products were further characterized the fixed single nucleotide polymorphisms (SNPs) of each gender using single-stranded conformation polymorphism analysis (SSCP, see below).

2.3.8 Single strand conformational polymorphism (SSCP) analysis

The glass plates (PROTEAN II xi Cell, BioRad) were cleaned and prepared as described previously. Different concentration of low cross link non-denaturing polyacrylamide gels (37.5:1 or 75:1) were prepared by dilution of a 40% stock solution to required gel concentration. The acrylamide gel solution may be mixed with glycerol (5% or 10% concentration), if desired. After degassed, 240 µl of 10% APS and 24 µl of TEMED were added. The analytical comb was inserted into the prepared gel and allowed to complete polymerization for 4 hours (or overnight).

Table2.4 Nucleotide sequence, length and melting temperature of primers designed from candidate AFLP fragments of *H. asinina*

Gene/Primer	Sequence	Length (bp)	Tm (°C)
HaMale1-F	5′-TAGGTTCTGGTGACACTTTGCCCGA-3′	25	76
HaMale1-R	5′-ATGTGTGAGAACTGGAATGTTGCTG-3′	25	72
HaMale2-F	5′-GCCTCATTGAAGCGATTCCCTAT-3′	23	68
HaMale2-R	5′-CCCTTTCGTCTCGTTTAGATTGGTC-3′	25	74
HaMale3-F	5′-ATTGGGTTCTCCTGTTGTCA-3′	20	58
HaMale3-R	5′-TCTGCCGATTTTTTATCTCTGC-3′	22	62
HaMale4-F	5′-GAAGCAACTGGCACCATCAAT-3′	21	62
HaMale4-R	5′-GTGTCCCCACATCACTATCTG-3′	21	64
HaMale5-F	5'- CATTGCATACTTATGAAGGGAA-3′	22	60
HaMale5-R	5′- TGGTTTTATCACTGAAATAGGGTC-3′	24	66
HaMale6-F	5′-GGCAGGACGCTACGTGATAATG-3′	22	68
HaMale6-R	5′-GGAGCAGACAATCCAGTGAC-3′	22	66
HaFemale1-F	5′-AGCTGGGGCTTTTATGATGAGAC-3′	24	68
HaFemale1-R	5′-AGTACGGGGATCGGTGGAACGGGA-3′	24	78
HaFemale2-F	5′-CACAGGCCCCCAAACGTGGAGCAT-3′	24	78
HaFemale2-R	5′-CAGCGAGCTTGACCACCCCATTCT-3′	24	74
HaFemale3-F	5′-TGGTAAGCTGTGTTCTGATCGGTC-3′	24	72
HaFemale3-R	5′-CCCTCCAAACATTCCCAGTCATTCT-3′	25	74
HaFemale4-F	5′-TTCTTGTGCCGAAAAGTGGAG-3′	21	62
HaFemale4-R	5′-CCTGAGGTGACATTGTTGACGA-3′	22	66
HaFemale5-F	5′-GGTGACTAACGGGATCAAATGGT-3′	23	68
HaFemale5-R	5′-GGAATCTTGCCGACGGGTAA-3′	20	62

Six microlitres of the PCR products were mixed with 24 µl of the 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 10 minutes. The denatured products were electrophoretically analyzed in native polyacrylamide gels at 250 volts for 16 – 24 hours at 4°C. SSCP bands were visualized by silver staining.

2.4 Isolation and characterization of expressed sequence taq (EST) in ovaries and testes of *H. asinina*

2.4.1 Tissue preparation

Fifty adults (approximately 8 cm of shell length and 18-month-old) of male and female *H. asinina* were collected from the hatchery stock of *H. asinina* (subsequent generations of HASAME). Gonads of each male and female were dissected from several stages of ovaries and testes and immediately placed into the liquid N_2 and stored at -80 ^{o}C freezer until required.

2.4.2 Total RNA preparation

Total RNA was extracted from gonads of each abalone using TRI REAGENT®. A piece of tissues (50-100 mg) was placed in a mortar containing liquid nitrogen and ground to the fine powder. The tissue powder was transferred to a microcentrifuge tube containing 500 µl of TRI REAGENT® and homogenized. Additional 500 µl of TRI REAGENT® were added. The homogenate was stored at room temperature and centrifuged at 12,000 rpm for 15 minutes at 4 °C. The aqueous phase was transferred to a new 1.5 ml microcentrifuge tube. Phase separation was performed by the addition of 200 µl of chloroform. The mixture was vortexed for 2 minutes and left at room temperature for 10 minutes. After centrifugation at 12,000 rpm for 15 minutes at 4 °C, the aqueous phase was transferred to a new 1.5 ml microcentrifuge tube. RNA was precipitated by an addition of 500 µl of prechilled isopropanol. The mixture was stored at -80 ^{o}C for 30 minutes. Total RNA was recovered by centrifugation at 12000 rpm for 10 minutes at 4 ° C. The supernatant was removed. The RNA pellet was briefly washed with 1 ml of 75% prechilled ethanol and centrifuged at 10000 rpm for 5 minutes at 4 °C. The supernatant was carefully removed. The RNA pellet was air-dried for 10 minutes. Total RNA was

dissolved in DEPC-treated H_2O for immediately used. Alternatively, the extracted total RNA was kept under absolute ethanol in a -80 ° C freezer for long storage.

2.4.3 mRNA purification

The messenger RNA (mRNA) was purified from total RNA of abalone ovary and testis using a QuickPrep. *Micro* mRNA Purification kit. Oligo (dT)-cellulose was gently swirled and 1 ml of the cellulose suspension was immediately pipetted into a 1.5 ml microcentrifuge tube. This tube was centrifuged for 1 minute and the supernatant was removed.

The mixture containing 50 µl of total RNA, 400 µl of the extraction buffer and 750 µl of elution buffer (10 mM Tris-HCl; pH 7.5, 1 mM EDTA) was added onto the oligo (dT)-cellulose. The tube was inverted gently for 3 minutes to resuspend the oligo (dT)-cellulose. The sample was centrifuged at the top speed of microcantrifuge for 1 minute. The supernatant was removed. The sample was washed with the high-salt buffer (10 mM Tris-HCl; pH 7.5, 1 mM EDTA, 0.5 M NaCl) five times and two times with the low-salt buffer (10 mM Tris-HCl; pH 7.5, 1 mM EDTA, 0.1 M NaCl). After the last washing step, the resin was resuspended with 300 µl of the low salt buffer and transferred to a microspin column placed in a microcentrifuge tube. The column was added with 500 µl of the low-salt buffer and was centrifuged at the full speed for 1 minute.

The effluent in the collection tube was discarded. This step was repeated twice. After that, the column was placed in a sterile 1.5 ml microcentrifuge tubes and 200 µl of prewarmed (65 °C) elution buffer was added to the top of the resin bed. The eluted mRNA was collected by centrifugation at the full speed for 1 minute. The eluted mRNA was precipitated by the addition of one-tenth volume of 3 M sodium acetate and 2 volumes of prechilled absolute ethanol. The mixture was kept at –20 °C for at least 30 minutes before centrifugation at 12000 rpm at 4 °C for 15 minutes. The mRNA pellet was kept under ethanol at –80 °C until used.

2.4.4 Measuring concentrations of RNA

The concentration of total RNA was determined as described for DNA but an OD unit at 260 nm corresponds to approximately 40 μg/ml of RNA (Sambrook et al., 1989).

2.4. 5 Construction of normal cDNA library from *H. asinina* ovaries and testes

Construction of cDNA libraries as well as subtractive cDNA libraries restrictively followed the protocol recommended by the manufacturer of the kit. As a result, only btief protocols are described here.

2.4.5.1 First and second strand cDNA synthesis

Normal cDNA libraries were constructed from mRNA isolated from ovaries and testes of *H. asinina* using a SuperScript Plasmid System with Gateway Technology for cDNA Synthesis and Cloning (Life Technology Inc., USA).

Two microlitres of *Not* I primer-adapter (0.5 μg/μl) was mixed with 2 μg mRNA. The mixture was incubated at 70 °C for 10 minutes and immediately chilled on ice followed by the addition of 5x first strand buffer (250 mM Tris-HCl, pH 8.3, 375 mM KCl, 15 mM $MgCl_2$), 2.0 μl of 0.1 M DTT, 1 μl of dNTP Mix (10 mM each) and 1 μl of DEPC-treated water. The reaction was incubated at 37°C for 2 minutes. Subsequently, 2 μl of SuperScript II Reverse Transcriptase (200 units/μl) was added. The reaction was incubated at 37°C for 1 hour.

The second strand cDNA was synthesized in a total volume of 150 μl by adding 91 μl of DEPC-treated water, 30 μl of 5x second strand buffer, 3 μl of 10 mM dNTP mix, 1 μl of *E. coli* DNA ligase (10 units/μl), 4 μl of *E. coli* DNA polymerase I (40 units/μl) and 1 μl of *E. coli* RNase H (2 units/μl). The mixture was gently mixed and incubated at 16°C for 2 hours. At the end of the incubation period, 2 μl of T4 DNA polymerase (5 units/μl) was added. The reaction mixture was further incubated at 16°C for 5 minutes and then placed on ice. Subsequently, 10 μl of 0.5 M EDTA was added to stop the reaction. An equal volume (150μl) of phenol-chloroform-isoamyl alcohol (25:24:1) was added, thoroughly mixed by vortexing and centrifuged at 12,000 rpm for 5 minutes. The

upper aqueous phase was carefully removed to a new microcentrifuge tube. The cDNA was precipitated by ethanol precipitation and immediately centrifuged at 12,000 rpm for 20 minutes. The supernatant was removed. The pellet was washed with 0.5 ml of 70% ethanol. After centrifugation at 12,000 rpm for 2 minutes, the supernatant was removed. The cDNA was air-dried at 37°C for 10 minutes to evaporate residual ethanol.

2.4.5.2 *Sal* I adapter ligation

Ligation of synthesized cDNA was carried out by the addition of 25 µl of DEPC-treated water, 10 µl of 5x T4 DNA ligase buffer (250 mM Tris-HCl pH 7.6, 50 mM $MgCl_2$, 5 mM ATP, 5 mM DTT and 25 % PEG8000), 10 µl of *Sal* I adapter (Figure. 2.2) and 5 µl of T4 DNA ligase (1 unit/µl), gently mixed and incubated at 16°C for 16 hours. The reaction was then extracted with equal volume of phenol-chloroform-isoamyl alcohol (25:24:1), vortex and centrifuge at 12,000 rpm for 5 minutes. The upper aqueous phase was transferred to new microcentrifuge tube. The cDNA was recovered by ethanol precipitation.

2.4.5.3 *Not* I digestion

The ligated cDNA was digested with *Not* I by combining cDNA resuspended in 41 µl of DEPC-treated water, 5 µl of REact 3 Buffer and 4 µl of *Not* I (15 units/µl) and incubated at 37°C for 2 hours. The reaction was extracted with 50 µl of phenol-chloroform-isoamyl alcohol. The digested cDNA was recovered by ethanol precipitation.

2.4.5.4 Size fractionation of cDNA by column chromatography

The digested cDNA pellet was dissolved in 100 µl of TEN buffer (10 mM Tris-HCl pH 7.5, 0.1 mM EDTA and 25 mM NaCl) and size-fractionated by column chromatography according to the recommended protocol. Fractions were collected. Aliquot of 5 µl from each fraction was electrophoretically analyzed by a 1 % agarose gel to avoid incorpolation of small cDNA and linkers.

2.4.5.5 Ligation of cDNA to pSPORT1 vector and transformation

Selected fractions were pooled ethanol-precipitated and resuspeded in TEN buffer before combined with 5 µl of 5x T4 DNA ligase buffer (250 mM Tris-HCl pH 7.6, 50 mM $MgCl_2$, 5 mM ATP, 5 mM DTT and 25 % PEG8000), 50 ng of *Sal* I-*Not* I-cut pSPORT1 (Figure 2.3) and 1 µl of T4 DNA ligase (1 units/µl) were added, mixed by pipetting and incubated at 4 °C overnight. Ten microlitres of the ligation product were transformed into *E. coli* JM109 and cultured using standard protocols.

After centrifugation at 6,000 rpm for 2 minutes and the supernatant was removed, the pellet was resuspended in 100 µl of SOC medium. Aliquots of the cultured cell (10, 1 and 0.1 µl) were spreaded onto a selective LB agar plates containing 50 µg/ml of ampicillin, 25 µg/ml of IPTG and 20 µg/ml of X-gal and further incubated at 37 °C overnight. White colonies were selected for further analysis.

2.4.6 Construction of subtractive cDNA library from *H. asinina* ovaries and testes

2.4.6.1 First and second strand cDNA Synthesis

Forward and reverse subtractions between cDNAs from ovaries and testes were carried out using a PCR Select cDNA Subtraction Kit (CLONTECH). Two microgram of each tester and driver mRNA was combined with 1 µl of 10 µM of the cDNA synthesis primer ($5^{/}$-TTTTGTACAAGCTT$_{30}N_1N$-$3^{/}$, N_1=A, C and G, N= A, T, C and G) in a 5 µl reaction volume. The reaction tube was incubated at 70 °C in a thermal cycler for 2 minutes, cooled on ice for 5 minutes and briefly centrifuged.

The first strand cDNA synthesis was synthesized by adding 2 µl of 5x First-Strand Buffer (250 mM Tric-HCl, pH 8.5, 40 mM $MgCl_2$, 150 mM KCl and 5 mM Dithiothreitol), 1 µl of dNTP Mix (10 mM each), 1 µl of sterile H_2O, and 1 µl of AMV Reverse Transcriptase (20 units/µl). The reaction mixture was incubated at 42 °C for 1.5 hours in an air incubator. The tubes were placed on ice to terminate the first strand cDNA synthesis.

Each first strand cDNA mixture was combined with 48.4 µl of sterile H_2O, 16.0 µl of 5x second strand buffer (100 mM Tris-HCl, pH 7.5, 25 mM $MgCl_2$, 500 mM KCl, 50 mM ammonium sulfate, 0.75 mM β-NAD and 0.25 mg/ml BSA), 1.6 µl of dNTP mix (10 mM) and 4.0 µl of 20x second strand enzyme cocktail (DNA polymerase I, 6 units/µl; RNase H, 0.25 units/µl; and *E. coli* DNA ligase, 1.2 units/µl) and subjected to thee second strand cDNA synthesis at 16 °C for 2 hours. After that, 2 µl of T4 DNA Polymerase (3 units/µl) were added to and further incubated at 16 °C for 30 minutes. Then, 4 µl of 20x EDTA/Glycogen (0.2 mM EDTA; 1 mg/ml glycogen) was added to the reaction and mixed thoroughly to stop the reaction. The cDNA was extracted twice by phenol/chloroform and recovered by ethanol precipitation.

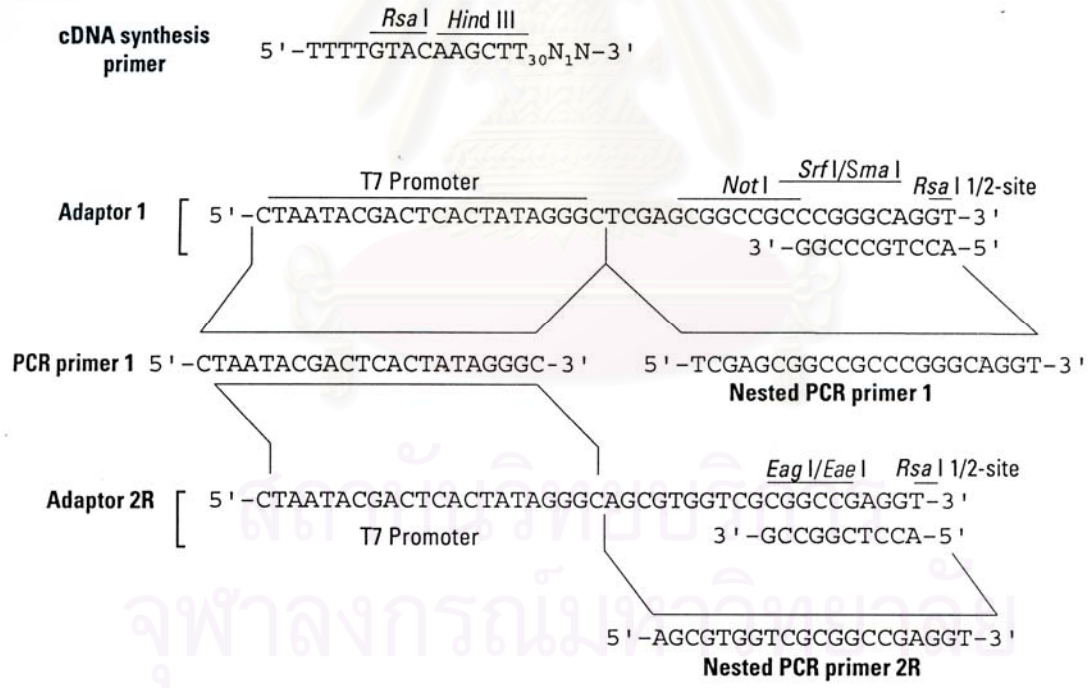


Figure 2.2 Sequences of the PCR-Select cDNA synthesis primer, adaptors and PCR primer.

2.4.6.2 *Rsa* I digestion

The pellet was dissolved in 50 µl of sterile H_2O and subjected to *Rsa* I digestion. The digestion was incubated at 37 °C for 1.5 hours. At the end of the incubation time, 5 µ l of the digested cDNA was collected and electrophoretically analyzed whether the digestion was complete. After that, 2.5 µl of 20x EDTA/glycogen mix was added to terminate the reaction. The digested cDNA was extracted twice with phenol/chloroform and recovered by ethanol precipitation. The pellet was dissolved in 5.5 µl of sterile H_2O and stored at –20 °C. This digested cDNA was served as the experimental driver cDNA. Subsequently, the sample was ligated with adapters to create tester cDNAs for forward and reverse subtraction.

2.4.6.3 Adapter Ligation and hybridization

One microlitre of *Rsa* I– digested tester cDNA was diluted with 5 µl of sterile H_2O. Enough ligation master mix was prepared by combining 3 µl of sterile H_2O, 2µl of 5x ligation buffer (250 mM Tris-HCl; pH 7.8, 50mM $MgCl_2$, 10 mM DTT and 0.25 mg/ml BSA) and 1 µl of T4 DNA ligase (400 units/µl) for multiple reactions .

For each tester, cDNA was divided into two tubes and ligated to either the adaptor 1 or the adaptor 2R. The ligation reactions of cDNA of ovaries (tester 1-1 and 1-2) and testes (tester 2-1 and 2-2) were performed in 10 µl total volume containing 2 µl of diluted tester cDNA, 1 µl of a single required adaptor and 6 µl of master mix. Two microlitres from each adaptor ligation for each tester population were combined to serve as an unsubtracted control. The tubes were incubated at 16 °C overnight. The ligation reactions were stopped by adding 1 µl of the EDTA/glycogen mix, heated at 72 °C for 5 minutes to inactivate the ligase activity. First and second second hybridization were performed according to the instruction manual provided by the manufacturer.

2.4.6.4 PCR Amplification

An a aliquot of 1 µl of each subtracted cDNA samples or unsubtracted testes control into an appropriately labeled tube. A master mix (24 µl), containing 2.5 µl of 10x PCR reaction buffer, 0.5 µl of dNTP mix (10 mM each), 1 µl of the primer1 and 50x Advantage cDNA Polymerase Mix, was dispensed to each tube and overlaid with 50 µl of

mineral oil. The reaction was incubated in a thermal cycler at 75 °C for 5 minutes to extend the adaptors. The amplification reaction was carried out for 27 cycles composing of 94 °C for 30 seconds, 66 °C for 30 seconds and 72 °C for 1.5 minutes. After amplification, 8 µl from each tube were electrophoretically analyzed through a 2.0% agarose gel.

A ten-fold dilution was performed using 3 µl of each primary PCR mixture. One microlitre of each diluted primary PCR product mixture was added into an appropriately labeled tube. An aliquot of 24 µl of the second master mix containing 2.5 µl of 10x PCR reaction buffer, 1 µl of 10 µM nested PCR primer 1, 1 µl of 10 µM nested PCR primer2R, 0.5 µl of dNTP mix (10 mM each) and 0.5 µl of 50x Advantage cDNA Polymerase Mix, was dispensed to each tube and overlaid with a drop of mineral oil. The reaction was carried out for 12 cycles composing of 94 °C for 30 seconds, 68 °C for 30 seconds and 72 °C for 1.5 minutes. Eight microlities from each tube were size-fractionated through a 2.0% agarose gel.

2.4.6.5 Cloning

The secondary PCR products of the forward subtraction (cDNA from ovaries as the tester) and the reverse subtraction (cDNA from testes as the tester) were separately ligated to pGEM®-T Easy vector and transformed to *E. coli* JM109. Recombinant clones were selected. Plasmid DNA was extracted as previously described.

2.4.7 Determination of inserted sizes of recombinant clones by colony PCR

Colony PCR was carried out using pUC1 (5′-CCGGCTCGTATGTTG TGTGGA-3′) and pUC2 (5′-GTGGTGCAAGGCGATTAAGTTGG-3′) to examine sizes of inserts. Positive colonies were randomly picked up and added to a 10 µl reaction volume containing (10 mM Tris-HCl, pH 8.3, 50 mM KCl, 0.001% gelatin, 2 mM Mg Cl_2, 100 µ M of each dNTP, 0.2 µM of each primer and 0.5 units of *Taq* DNA polymerase. The amplification reaction was carried out for 35 cycles consisting of denaturation at 94 °C for 1 minute, annealing at 53 °C for 1 minute and extension at 72 °C for 1.5 minutes. The final extension was carried out at the same temperature for 10 minutes. PCR products were electrophoretically analyzed on 1.0 % agarose gel.

2.4.8 Plasmid DNA preparation

A recombinant clone was inoculated into 2 ml of LB broth containing 50 mg/ml of ampicillin and cultured overnight at 30 °C. The culture was centrifuged at 10,000 rpm for 3 minutes. The supernatant was removed. Plasmid DNA was extracted using an alkaline lysis method (Maniatis et al; 1982).

2.4.9 DNA sequencing and sequence analysis

The dideoxy chain termination-based sequencing method was carried out using a ThermoSequenase Fluorescent Labeled Primer Sequencing kit (Amersham Pharmacia Biotech) with M13 fluorescent dye labeled primers. Four microlitres (0.5-5 μg) of plasmid DNA and one pmol of M13 fluorescent labeled primer were mixed and aliquoed tinto each of 4 tubes. After that 1 μl of A-reagent (Tris-HCl, pH 9.5, $MgCl_2$, Tween 20, Nonidet P-40, 2-mercaptoethanol, dATP, dCTP, 7-deaza-dGTP, dTTP, ddATP, thermostable pyrophosphatase and ThermoSequence DNA polymerase), 1 μl of C-reagent (Tris-HCl, pH 9.5, $MgCl_2$, Tween 20, Nonidet P-40, 2-mercaptoethanol, dATP, dCTP, 7-deaza-dGTP, dTTP, ddCTP, thermostable pyrophosphatatase and ThermoSequence DNA polymerase), 1 μl of G-reagent (Tris-HCl, pH 9.5, $MgCl_2$, Tween 20, Nonidet P-40, 2-mercaptoethanol, dATP, dCTP, 7-deaza-dGTP, dTTP, ddGTP, thermostable pyrophosphatatase and ThermoSequence DNA polymerase) and 1 μl of T-reagent (Tris-HCl pH 9.5, $MgCl_2$, Tween 20, Nonidet P-40, 2-mercaptoethanol, dATP, dCTP, 7-deaza-dGTP, dTTP, ddTTP, thermostable pyrophosphatatase and ThermoSequence DNA polymerase) was add to tubes A, C, G and T, respectively.

Cycle sequencing reactions were carried out in a thermal cycler. Amplification reaction contained a 95 °C initial denaturation for 2 minutes followed by 30 cycles of a 95 °C denaturation for 30 seconds, a 55 °C, annealing for 30 second, a 72 °C, extension for 1 minute for 30 cycles, and a 72 °C, 5 minutes final extension. After amplification, 2 μl of formamide loading dye (formamide, EDTA and funchsin) was added to each reactions and 2.5 μl of each reaction were loaded onto 6% denaturing polyacrylamide gel and were electrophoresed on an automated DNA sequencer (LC4000, LICOR) at 50 W for 14-16 hours.

Sequences of cDNAs were edited and compared with those in the the GenBank (http://www.ncbi.nlm.nih.gov, the National Center for Biotechnology Information; NCBI) using the BLASTN and BLASTX programs (Altschul et al., 1997). Significant probabilities and numbers of matched nucleotide/proteins were considered when E-values < the 10^{-4} .

After homology search, matched ESTs were categorized into 6 broad functional categories; 1) sex-related gene 2) stress response, detoxification and cell defense proteins 3) protein synthesis and ribosomal proteins 4) cell division and DNA replication 5) metabolism 6) unidentified gene function and 7) unknown. Nucleotide sequences of characterized EST were submitted to dbEST of the GenBank (Altschul et al., 1997).

2.5 Isolation and characterization of sex-related genes in ovaries and testes of the tropical abalone (*H. asinina*)

2.5.1 Tissue collection

Ovaries and testes of each abalone were dissected out for further used in isolation and characterization of sex-rerated gene. For tissue distribution studies, foot, hemocyte, mantle, epipodial tentacles, oral tentacles, pairs of eyes, gill, hepatopancreas, digestive gland and left hypobranchial gland was also dissected from each abalone. Dissected tissues were immediately dropped into the liquid N_2 and tranferred to a –80 ^{o}C freezer until required. For hemocyte collection, foot of abalone was cut. Abalone was placed in a sterile petri disc. The haemolymph was withdrawn from that in the petri disc using a 1.0 ml syringe containing 500 µl of an anti-coagulant solution (10% sodium citrate, w/v). The collected haemolymph was immediately centrifuged at 3,000 rpm for 10 min at 4°C to separate haemocytes from the plasma. Haemocyte pellet was resuspended in 200 µl of TRI REAGENT® and kept at –80 °C for RNA extraction.

2.5.2 Cloning and characterization of DMRT1 gene homologue in *Haliotis asinina*

Doublesex **M**ale abnormal-3 **R**elated **T**ranscription factor-**1** (DMRT1) belongs to a gene family of, which shares sequence similarity with the unusual zinc finger DNA binding DM domain. They are *doublesex* (*dsx*) in *Drosophila melanogaster* and *male abnormal*-3 (*mab3*) in *Caenorhabditis elegans* (Raymond et al., 1998). These genes play crucial role in sexual differentiation of those species. DMRT1 involves in testicular development. Expression of DMRT1 in the testes has been shown to be related with testicular differentiation in different vertebrate species including some mammals, birds, reptiles and fish. DMRT1 has been regarded to be the first conserved gene in sex differentiation found from invertebrates to human (Marchand et al., 2000).

2.5.2.1 Designation of degenerate primers

Amino acid sequences of DMRT1 from human, *Homo sapiens* (accession numbers Al162131 and XM_015427) mouse, *Mus musculus* (AL133300 and NM_015826), *Fugu rubribes* (AJ295039), *Tetraodon nigroviridis* (AJ295040*), Oncorhynchus mykiss* (AF209095), *Oryzias latipes* (AF319994) and *Sus scrota* (AF216651), *dsx* gene from *D. melanogaster* (NM_079548 and AY060257), *Bactrocera tryoni* (AF029675 and AF029676) and *Megaselia scalaris* (AF283695 and AF283696), and *mab3* gene from the nematode *C. elegans* (NM_063697 and AF022388), were retrieved from the GenBank (www.ncbi.nlm.nih.gov) and multiple aligned.

Conserved regions in DM domain of these genes were selected for desegregation of primers and nested primers (Table 2.5) and used for RT-PCR (1st strand cDNA) analyses.

2.5.2.2 First strand cDNA synthesis

The first strand cDNA was synthesized from 1.5 µg of total RNA extracted from testes of *H. asinina* using an ImProm – II™ Reverse Transcription System Kit (Promega). Total RNA was combined with 0.5 µg of an oligo$(dT)_{15}$ primer and appropriate DEPC-treated H_2O in a final volume of 5 µl. The reaction was preheated at 70 ºC for 5 minutes and immediately placed on ice for 5 minutes. The reverse transcription reaction mix (15

μl), containing 4 μl of 5x ImProm-II™ reaction buffer, 1.8 μl of 25 mM $MgCl_2$, 1 μl of dNTP Mix (10 mM each), 20 units of Recombinant Rnasin® Ribonuclease inhibitor and 1 μl of ImProm-II™ reverse transcriptase, was added and mixed gently. The reaction mixture was incubated at 25 ºC for 5 minutes and at 42 ºC for 90 minutes. The reaction was terminated by incubated at 70 ºC for 15 minutes. Concentration and rough quality of the first stranded cDNA was spectrophotometrically examined (OD_{260}).

2.5.2.3 RT-PCR analysis

One microlitter of the first strand cDNA from testis was used as the template for RT-PCR in a 25 μl reaction composing of 10 mM Tris-HCl, pH 8.8, 50 mM KCl, 0.1% Triton X–100, 2 mM $MgCl_2$, 100 mM of each dNTP, 0.4 μM of DMF1 and DMR primers and 1 unit of of DyNAzyme™ II DNA Polymerase. The amplification reaction was composed of denaturation at 92 ºC for 1 minute followed by 5 cycles of denaturation at 92 ºC for 1 minute, annealing at 52 ºC for 1 minute and extension at 72 ºC for 1 minute and additional 35 cycles of 92 ºC for 1 minute, 62 ºC for 1 minute and 72 ºC for 1 minute. The final extension was carried out at 72 ºC for 10 minutes.

Table 2.5 Degenerated primer, primer sequences, length and melting temperature of primers designed from the conserved DM domain of various species

Gene/ Primer	Nucleotide sequence	Length	Tm (°C)
DMF1	5′-CCC AAG TGC KCC CGC TGC CGG AAY CA-3′	26	74
DMF2	5′-CTT AAG GGM CAC AAG CGC TWC TG-3′	23	57
DMR	5′-CGC CAG CTC CTC CTC YTG AGC CTG-3′	24	66

Abbreviations: K = G or T; M = A or C ; Y = C or T and W = K or T

Electrophoresed PCR product was eluted from the gel using a NucleoSpin® Extract Kit acording to protocol recommended by the manufacturer. Semi- nested PCR was carried out with the same condition using DMF2 and DMR primers against eluted DNA template (175 bp). The resulting product was gel eluted and ligated to pGEM®-T Easy and electrotransformed to *E. coli* JM109. Recombinant clones were sequenced using an automated DNA sequencer.

2.5.3 Identification and characterization of sex-involving genes from cDNA libraries of ovaries and testes of *Haliotis asinina*

2.5.3.1 Primer design

Primers were designed from nucleotide sequences of sex-related gene homologue obtained from normal and subtractive cDNA libraries of testes and ovaries of *H. asinina* using the Primer Premier 5 program (Primer Premier, PREMIER Biosoft international). Sequences of primers are shown in Table 2.6.

2.5.3.2 Analysis of sex-specific/differential expression transcripts by RT-PCR

The first standed cDNA was generated from 1.5 μg of total RNA with an ImProm-II™ Reverse Transcription system kit. One microlitre of the first strand cDNA synthesized from total RNA of ovaries, testes or hemocytes of *H. asinina* was used as the template for amplification of sex-related (axonemal protein 66.0, tektin A1, fertilization protein, sperm lysin, DMRT1, VCP1, VCP3, VCP7, VCP49, VCP75 and VTG-1) gene homologues in a 25 μl reaction volume containing 10 mM Tris-HCl, pH 8.8 at 25 °C, 50 mM KCl and 0.1% Triton X–100, 2 mM $MgCl_2$, 100 μM each of dATP, dCTP, dGTP and dTTP, 0.2 μM of each primer and 1 unit of Dynazyme™ DNA Polymerase. PCR was carried oyt by predenaturation at 94 °C for 3 minutes, followed by 35 cycles of a 94 °C denaturation for 30 seconds, a 60 °C annealing for 45 seconds, and a 72 °C extension for 45 seconds. The final extension was carried out at 72 °C for 7 minutes. Five microlitres of each amplification product were electrophoretically analyzed through 1.5% agarose gel and visualized under a UV-transilluminator after ethidium bromide staining.

2.5.4 Tissue distribution analysis using RT-PCR

Expression of 11 transcripts including axonemal protein, tektin A1, fertilization protein, sperm lysin, DMRT1, VCP1, VCP3, VCP7, VCP49, VCP75 and VTG-1 gene homologues were examined in different tissue (foot muscle, mantle, epipodial tentacles, oral tentacles, pairs of eyes, gill, hepatopancreas, digestive gland, left hypobranchial gland, hemocytes, gonads; ovary or testis of both a female and male abalone) using TRI-REAGENT.

The first stranded cDNA was generated and 1 μg of the synthesized first strand cDNA was used as the template for typical RT-PCR (25 ul). Compositions and amplification conditions were identical to those for sex-specific RT-PCR analysis. Eight microlitres of the amplification product was electrophoretically analyzed though 1.8% agarose gels.

2.5.5 Isolation of the full length cDNA of Axonemal p66.0, Tektin A1 and DMRT 1 homologue by Rapid Amplification of cDNA ends-polymerase chain reaction (RACE - PCR)

The SMART ^{TM} RACE cDNA Amplification Kit (BD Bioscience Clontech) provides a flexible method for performing both 5′ and 3′ RACE reaction to be served as the template for generating full-length cDNAs in the subsequent reverse transcription reactions.

2.5.5.1 Primer design

The full length cDNA of 3 sex-involving genes, axonemal p66.0, tektinA1 and DMRT1 were isolated througth RACE-PCR. The 5′ and 3′ gene specific primer (GSP) and 5′ and 3′ nested GSP were designed from EST sequences of axonemal p66.0 and tektinA1 homologues from the testis cDNA library whereas that of the DMRT homologue amplified and cloned from conventional RT-PCR, respectively (Table 2.7).

Table 2.6 Name of primers, nucleotide sequences, length and melting temperature of primers designed from sequences of candidate sex-related gene homologues found in cDNA libraries of *H. asinina.*

Gene /Primer	Sequence	Length (bp)	Tm (°C)
Vitelline coat protein (VCP)			
VCP1-F	5′GGC TGC ACC AGA CCG ATG AAC GAT ACA C3′	28	68
VCP1-R	5′ TGC TTC AAC ACC ATA CCG TCT CCA CAA C 3′	28	65
VCP2-F	5′ ACT GGG CTT TCT ACC ATC AAC GTC CTG T 3′	28	64
VCP2-R	5′ CGA CGA CCC CTT GTT CTG GAT AAT CTC A 3′	28	65
VCP3-F	5′ ATA CAC AAT CTC GTG CTC GTT CGG TTC A 3′	28	64
VCP3-R	5′ CAC TAT TCC GTC TCC GCA ACC TGC TCT T 3′	28	66
VCP7-F	5′ TGC TCA GTG CTG GTG GTG CTC TGC TGG G 3′	28	71
VCP7-R	5′ TGG CTG GGC GGT GCC TTG TTA CCT TCT A 3′	28	70
VCP49-F	5′ ACC CCA CAG GAG GAA CAA ACC 3′	21	66
VCP49-R	5′ CCG CAG CCA GCT CTA AGG ATA 3′	21	66
VCP75-F	5′ AGG ATG TGG TGA CGG TAT TGT G 3′	22	66
VCP75-R	5′ CTT TAT AGT GCG AGC GTT TGG T 3′	22	64
Axonemal p66.0 protein			
Axonemal-F	5′ GCA CGA CGA AAC TTC GCC CTG 3′	21	68
Axonemal-R	5′ AAC GCC CGC TTC GCT CCC CA 3′	20	68
Tektin A1			
Tektin-F	5′ TGT TGA CAG GGA TGG TGC GG 3′	20	64
Tektin-R	5′ CTC CTT TGC CTA CAG TTG AGA TTG 3′	24	70
Vitellogenin-1			
VTG1-F	5′ ACA TCA GAA CCG ACG GCA AC 3′	20	62
VTG1-R	5′ TGA GGC AAG GTA GGC GAG G 3′	19	62
Small androgen receptor-interacting protein (SARIP)			
SARIP-F	5′ GGC TTA GTG ACT GAA CGC CTC TA 3′	23	70
SARIP-R	5′ GCT GCT CTA CTA CGC ACA ACA C 3′	22	68
Gonadotropin inducible ovarian transcription factor 1 (GIOT1)			
GIOT1-F	5′ GAC CAC CCA CGC ACA GGA C 3′	19	64
GIOT1-R	5′ TAG CAG CAC TAA TAA AGC CCC G 3′	22	66
Hydroxysteroid dehydrogenase-like protein (HSD)			
HaHSD-F	5′ GCC GTG GAA GAA GCA GTT GGA 3′	21	68
HaHSD-R	5′ CAG GGG GTT GAG GTT GAG TGG 3′	21	68
Fertilization protein (FP)			
HaFP – F	5′ CGA CCC ATA GCG GCG TAG TT 3′	20	64
HaFP – R	5′ AAG GTC CCA AAG AAA AGC CAG TA 3′	23	66
Sperm lysine			
Halysin–F	5′ CGA GAA GGC AGA CAG CCA GAC 3′	21	68
Halysin–R	5′ ATC CAG TGC TTG ATG CTT GAC G 3′	22	66
DMRT1			
DMRT1Ha-F	5′ GTC TGT TCG CCA ACT GCA GCT G 3′	22	70
DMRT1Ha-R	5′ GTC GAC GAC GGA GGG CAA TCT G 3′	22	72
DMHa-F	5′ CAG CTA CGT TCC TGT GCG CGA TGT 3′	24	77
DMHa-R	5′ GGT ATC GAC CGT CCC GCC TCC TCA 3′	24	80

2.5.5.2 First strand RACE-cDNA synthesis

Total RNA was extracted from testis of *H. asinina* using TRI-REAGENT. Messenger (m) RNA was purified using a QuickPrep *micro* mRNA Purification Kit. The 5′ and 3′ RACE – Ready cDNA was prepared by combining 1 µg of testis mRNA with 1 µl of 5′-RACE CDS primer and 1 µl of 10 µM SMART II^{TM} A oligonucleotide and 1 µg of testis mRNA with 1 µl of 3′-RACE CDS Primer A, respectivly.

The components were mixed and spun briefly. The reaction was incubated at 70 ° C for 2 minutes and cooled on ice for 2 minutes. The reaction tube was spun briefly. After that, 2 µl of 5x First–Strand buffer (250 mM Tris-HCl, pH8.3, 375 mM KCl and 30 mM $MgCl_2$), 1 µl of 20 mM DTT, 1 µl of dNTP Mix (10 mM) and 1 µl of BD PowerScript ^{TM} Reverse Transcriptase were added. The reactions were mixed by gently pipetting and centrifuged briefly to collect the contents at the bottom. The tubes were incubated at 42 ° C for 1.5 hours in a hot-lid thermal cycler. The first strand reaction products were diluted with 250 µl of Tricine–EDTA Buffer and heated at 72 °C for 7 minutes. The 5′- and 3′-RACE first strand cDNA template were stored at –30 °C until further needed.

2.5.5.3 Rapid Amplification of cDNA Ends (RACE)-PCR

The master mix of 5′- and 3′-RACE reactions was prepared. Each amplification reaction included 34.5 µl of PCR-Grade Water, 5 µl of 10x BD Advantage 2 PCR buffer, 1µl of 10 µM dNTP mix and 1 µl of 50x BD Advantage 2 polymerase mix.

5′- RACE – PCR of each gene was performed in 50 µl reaction volume containing 5 µl of 10x UPM, 1 µl of 10µM 5′ RACE specific primer, 41.5 µl of PCR Master Mix and 2.5 µl of 5′-RACE-Ready cDNA. 3′- RACE – PCR was also performed in 50 µl reaction volume compose of 5 µl of 10x UPM, 1 µl of 10µM 3′ RACE specific primer, 41.5 µl of PCR Master Mix and 1.25 µl of 3′-RACE-Ready cDNA. Amplification using the UPM permer and the single GSP of each gene was also carriied out. The reaction was carried out for 20 cycles composing of a 94 °C for 30 seconds, 68 °C for 30 seconds and 72 °C for 3 minutes.

Table 2.7 Name of primers, nucleotide sequences, length and melting temperature of gene-specific primers using for isolation of the full length of axonemall p66.0, tektin A1 and DMRT 1 by RACE-PCR

Gene /Primer	Sequence	Length (bp)	Tm (°C)
5′-RACE-PCR :			
Axonemal p66.0			
Axonemal-R	5′ AAC GCC CGC TTC GCT CCC CA 3′	20	68
NestedAxonemal RACE5-R	5′ GGA GCT TGT CGA GGT GGT CGT TCA G 3′	25	64
Tektin A1			
Tektin-R	5′ CTC CTT TGC CTA CAG TTG AGA TTG 3′	24	70
NestedTektinA1 RACE5-R	5′ TAC TTG GCC GCC CGA AAT CCC TCT G 3′	25	68
DMRT1			
DMRT1Ha-R	5′ GTC GAC GAC GGA GGG CAA TCT G 3′	22	72
NestedDMRT1Ha-R	5′ TTC CCG GGA CAC CTT CTG TC 3′	20	64
3′-RACE-PCR:			
Axonemal p66.0			
Axonemal-F	5′ GCA CGA CGA AAC TTC GCC CTG 3′	21	68
NestedAxonemal RACE3-F	5′ CTG TTC CGG TGT CTG GGC TGC AAT A 3′	25	64
Tektin A1			
Tektin-F	5′ TGT TGA CAG GGA TGG TGC GG 3′	20	64
NestedTektinA1 RACE3-F	5′ CCA GAA ATC CTA TCC ACC CCC GAG T 3′	25	68
DMRT1			
DMRT1Ha-F	5′ GTC TGT TCG CCA ACT GCA GCT G 3′	22	70
NestedDMRT1Ha-F	5′ AGC AGC TGT CGC CTG AGT G 3′	20	64

Table 2.8 Primers used for the first strand cDNA synthesis and RACE – PCR

Primers	Sequence
SMART II^{TM} A Oligonucleotide	5′AAG CAG TGG TAT CAA CGC AGA GTA CGC GGG 3′
3′-RACE CDS Primer A	5′AAG CAG TGG TAT CAA CGC AGA GTA C $(T)_{30}$ 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 Primer A Mix (UPM)	Long : 5′CTA ATA CGA CTC ACT ATA GGG CAA GCA GTG GTA TCA ACG CAG AGT 3′ Short : 5′CTA ATA CGA CTC ACT ATA GGG C 3′
Nested Universal Primer A (NUP)	5′AAG CAG TGG TAT CAA CGC AGA GT 3′

Nested PCR was performed using primers described in Table 2.5. The primary PCR product was 50-fold-diluted (1µl of the product + 49 µl of Tricine-EDTA buffer). The amplification reaction was performed using 5 µl of the diluted PCR product as the template using the same condition for the first PCR for 20 cycles. Products from the first and nested PCR were electrophoretically analyzed.

2.5.5.4 Cloning and sequencing of 5′ and 3′ RACE-PCR products

The expected DNA fragments of each 5′ and 3′ RACE-PCR product were eluted from agarose gel by QIAquick® Gel Extraction and ligated to pGEM®-T Easy vector. The ligation product was transformed to *E. coli* JM109. Recombinant clone was identified by colony PCR. Plasmid DNA was extracted by a QIAprep® Spin Miniprep Kit and sequenced using an automatic sequencer (ABI310) at the Bioservice unit (BSU), National Science and Technology Development Agency (NSTDA).

2.5.5.5 Amplification of a single fragment representing the full-length cDNAs of axonemal p66.0, tektin A1 and DMRT1

Primer primed at the 5′ untranslated region (UTR) and 3′ UTR of axonemal p66.0, tektin and DMRT1 were designed (Table 2.9). The amplification reaction for a single fragment of the full length cDNA was performed in a 50 µl reaction volume composing of 5µl of 10x BD Advantage 2 PCR buffer, 1 µl of dNTP mix (10 mM), 0.2 µM of each primer, 500 ng of the cDNA template synthesized from total RNA of testis and 1 µl of 50x BD Advantage 2 Polymerase Mix.

The reaction was predenaturated at 94 °C for 3 minutes followed by 30 cycles of a 94 °C denaturation step for 1 minute, a 65 °C annealing step for 1 minute and a 72 °C extension step for 2 minutes. The final extension was carried out at 72 °C for 10 minutes. The PCR products were electrophoretically analyzed through 1.6% agarose gels and visualized under a UV transilluminator after ethidium bromide staining. The resulting PCR product was cloned and sequenced for both directions.

Table 2.9 Name of primers, nucleotide sequences, length and melting temperature of gene-specific primers using for amplification of the full length cDNAs of axonemal p66.0, tektin A1 and DMRT1

Primer	Sequence	Length (bp)	Tm (°C)
Axonemal p66.0 protein			
Axonemal 5UTR-F	5′GAC GTC TTG TTA CCA TAA CTT CCC A 3′	25	72
Axonemal 3UTR-R	5′CAT ATT CAG ACC AAC AAG TAA CCT A 3′	25	69
Tektin A1			
TektinA1 5UTR-F	5′CTT CCG CTC ATG GTT GCC GGG GAT C 3′	25	80
TektinA1 3UTR-F	5′TAA GAG ATC TTT GTG TGT GGC TTC A 3′	25	68
DMRT1			
DMRT1 5UTR-F	5′GAC TAC AAG GTG CCT GCC TTT AGT G 3′	25	75
DMRT1 3UTR-F	5′CAA CGC TGA ACA TAA CAT TGA AAGC 3′	25	70

DNA sequences derived from both strands were assembled and analyzed with GENETYX and blasted against data in the GenBank using BlastN and BlastX (www.ncbi.nlm.nih.gov).

2.5.6 Semi-quantitative Reverse Transcription–Polymerase Chain Reaction

2.5.6.1 Samples and the first strand cDNA synthesis

Adults (approximately 8 cm shell length, 18-month-old) of male and female *H. asinina* were sampled from the hatchery stock established from the Samet Island sample (HASAMHE). Ovaries or testes were dissected out from each abalone. Testicular and ovarian development were classified to stages 1, 2, 3 and 4 and stages 1 and 3 according to Sighagaiwan (19XX). Dissceted ovaries and testes were immediately dropped into the liquid N_2 and stored at –80 ^{o}C freezer until required.

Juvenile *H. asinina* was also collected from the hatchery population. Gonad of each abalone were dissected out and performed as for those of adults. Total RNA was extracted from ovaries and testes of *H. asinina* using TRI-REAGENT® (Molecular Research Center). The first stand cDNA was synthesized from 1.5 μg of total RNA using an ImProm-II™ Reverse Transcription system Kit (Promega).

2.5.6.2 Primer design

Expression patterns of five male-specific transcripts; axonemal p66.0, tektinA1, sperm lysin, FP, DMRT1 (Table 2.6) and six female-specific transcripts; VCP1, VCP3, VCP7, VCP49, VCP75 and VTG-1 (Table 2.6) were determined. Elongation factor 1 α (EF1α-F; 5′-CGA GAA GGA GGA AGC CCA GGA GGA AAT G-3′ and EF1 α -R; 5′-GCA TCA ATG ATG GTG AAT GTA GAA C-3′) was used as an internal control for semi-quantitative PCR.

2.5.6.3 Determination of the optimal PCR conditions

The standard amplification was performed in a 25 μl reaction volume containing 500 ng of the first strand cDNA template, 1 x PCR buffer (10 mM Tris-HCl pH 8.8, 50 mM KCl and 0.1% Triton X – 100) , 100 μM each of dATP, dCTP, dGTP and dTTP and 1 unit of Dynazyme™ DNA Polymerase (Finnzymes, Finland). The reaction was carried

out for 35 cycles composing of a 92 ^{o}C denaturation step for 1 minute, a 55 ^{o}C annealing step for 1 minute and a 72 ^{o}C extension step for 1 minute. Ten microlitres of the amplification products are electrophoresed through 1.8% agarose gels and visualized by ethidium bromide staining PCR parameters were further adjusted to measure the amplification product quantitatively.

2.5.6.4 Determination of the optimal primer concentration

The optimal primer concentration for each pair transcript (0.05, 0.1, 0.2, 0.3 and 0.4 μM) was examined using the standard PCR condition as described previously. The resulting product was electrophoretically analyzed. The primer concentration that gave product specificity and clear results were selected for further optimization of PCR conditions.

2.5.6.5 Determination of the optimal $MgCl_2$ concentration

The optimal $MgCl_2$ concentration for amplification of each transcript (1.0, 1.5, 2.0, 2.5, 3.0, 4.0 and 5.0 mM $MgCl_2$) was examined using the standard PCR conditions and the optimal primer concentration obtained. The concentration that gave the highest specificity was chosen.

2.5.6.6 Determination of the optimal number of amplification cycle

The PCR amplifications were carried out at 18, 21, 24, 27, 30, 33 and 36 cycles using the optimal concentration of primers and $MgCl_2$. The amplification product was analyzed by agarose gel electrophoresis. Density of the electrophoresed bands was plotted against the number of cycles. The number of cycles that still provided the PCR product in the exponential rage (did not reach a plateau level of amplification) was chosen.

2.5.6.7 Competitive effect between primer sets

Effects from competition between primers of the target transcripts and EF1-α, were examined by amplification of those respective genes separately and simultaneously (a single tube RT-PCR). The PCR products were analyzed by agarose gel electrophoresis. The band intensity of the target and the control in a single tube reaction was compared with that from separated amplification was compared. Decreasing of band intensity

indicated competition between two primer sets. When competition occurred further adjustment of $MgCl_2$ was carried out.

2.5.6.8 Gel electrophoresis and quantitative analysis

The ratio between the target transcript and EF1-α in different developmental stages of ovaries and testes from broodstock-sized *H. asinina* was determined. Eight microlitres of the PCR product were analyzed by 1.8% agarose gel which were electrophoresed at 6 volts/cm for 1 hours. The gel was stained with 2.5 μg/ml EtBr for 5 minutes and destained in the running tap water for 10 minutes. The intensity of target and control bands was quantifiedly from glossy prints of the gels using the Gel Pro program.

2.5.6.9 Data analysis

The expression level of each transcript in different developmental stages of ovaries and testes from broodstock-sized *H. asinina* was normalized by that of EF1 -α (100%). Significantly different expression levels between different groups of *H. asinina* were tested using one way analysis of variance (ANOVA) and Duncan's new multiple range test .

2.5.7 Analysis of fixed SNPs of sex-related genes in genomic DNA of male and female *H. asinina* by restriction enzyme digestion and SSCP analysis

PCR was carried out using primers designed from cDNA sequences of sex-related gene homologues (Table 2.6). The amplification reaction was performed in a 25 μl reaction volume containing 10 mM Tris-HCl, pH 8.8, 50 mM KCl, 0.1% Triton X-100, 100 μM each of dATP, dCTP, dGTP and dTTP, 2 mM $MgCl_2$, 0.2 μM of each forward and reverse primers, 50 ng of genomic DNA of male or female *H. asinina* and 1.0 unit of DyNAzyme™ II DNA Polymerase (Finnzymes).

The amplification profiles were composed of predenaturation at 94°C for 3 minutes followed 35 cycles of denaturation at 94°C for 1 minute, annealing at 55 °C for 1 minute, and extension at 72°C for 1.5 minutes. The final extension was carried out at the same temperature for 7 minutes. The amplification products were electrophoretically analyzed through 1.8% agarose gels and visualized under a UV transilluminator after ethidium bromide staining. Successful amplification products were further characterized

following a standard single strand conformational polymorphism (SSCP) approach to examine whether fixed single nucleotide polymorphism (SNP) was existent for each gender.

SSCP analysis requires small fragment (usually < 300 bp). Nevertheless, the PCR product of axonemal protein was approximately 940 bp. This amplified gene segment was digested with *Hae* III. The digestion was carried out in a 15 μl reaction volume containing 1x restriction buffer C (10 mM Tris-HCl; pH 7.9, 50 mM NaCl, 10 mM $MgCl_2$ and 1 mM DTT), 6 μl of the PCR product and 3 unit of *Hae* III. The reaction mixture was incubated at 37 °C for 3 hours before electrophoretically analyzed by agarose gel electrophoresis.

2.5.7.1 Single strand conformational polymorphism (SSCP) analysis

The PCR product of sex-related transcript (tektinA1, DMRT1, GIOT1, SARIP, HSD, ADAMTS-9, and α-tubulin2) and *Hae* III-digested axonemal p66.0 was mixed with 4 vol. of the SSCP loading dye (95% formamide, 0.25% bromophenol blue, 0.25% xylene cyanol and 10 mM NaOH) and denatured at 95 °C for 5 minutes. The denatured PCR product were immediately cooled on ice for 10 minutes and electrophoretically analyzed in low crosslink native polyacrylamide gels (37.5:1, acrylamind : bisacrylamind) at 250 volts for 16 – 24 hours at 4°C. Bands were visualized by silver staining

CHAPTER III

RESULTS

3.1 Isolation and characterization of species-specific markers in tropical abalone (*H. asinina*, *H. ovina*, and *H. varia*) in Thailand

3.1.1 DNA extraction

High molecular weight DNA at approximately 23.1 kb was obtained from DNA extraction using a proteinase K-phenol-chloroform method (Figure 3.1). DNA concentrations were spectrophotometrically determined by measuring the optical density at 260 nm (1 OD_{260} unit was equivalent to 50 μg DNA/ml). The ratio of OD_{260} / OD_{280} was 1.8 – 2.0 suggesting that the quality of extracted genomic DNA was acceptable for further used.

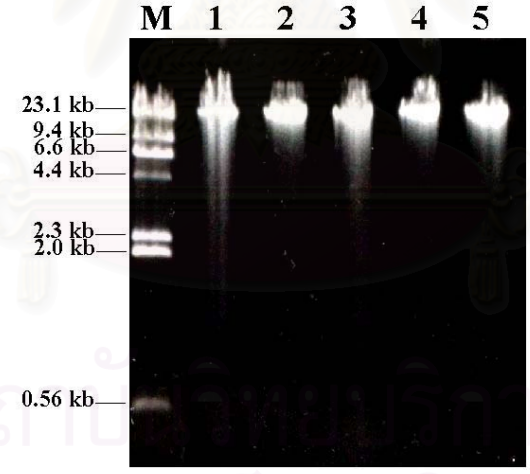


Figure 3.1 A 0.8 % agarose gel showing the quality of total DNA extracted from the foot muscle of abalone. Lanes 1-5 are genomic DNA from five individuals of abalone. Lane M is λ/*Hin*dIII DNA marker.

3.1.2 Screening of RAPD primers

Among the 113 RAPD primers screened against 2-3 representatives of each abalone species, 27 RAPD primers (OPA1, 2, 10, and 20, OPB11, 16 and 17, UBC101, 119, 160, 168, 174, 193, 195, 197, 200, 210, 220, 264, 267, 271, 272, 456, 457 and 459) yielded successful amplification results in three abalone species. Seven RAPD primers (OPB11 and UBC101, 193, 195, 197, 220 and 271) were selected and further tested against larger number of *H. asinina* (*N* = 12), *H. ovina* (*N* = 3) and *H. varia* (*N* = 3) individuals.

Finally, 5 RAPD primers (OPB11 and UBC101, UBC195, UBC197 and UBC271) that gave reproducible RAPD patterns were selected for identification of species- and population- specific markers of *H. asinina*, *H. ovina* and *H. varia* in Thai waters. These primers were tested against different geographic sample of *H. asinina* (Samet Island, Rayong, *N* = 4; Talibong Island, Trang, *N* = 4; Cambodia, *N* = 5; P_0, Cambodia, *N* = 7; and F_1, Philippines, *N* = 4), *H. ovina* (Sichang Island, Chon Buri, *N* = 3; Samet Island, Rayong, *N* = 3; Churk Island, Trang, *N* = 3; and Similan Island, Phangnga, *N* = 4; and *H. varia* (L-Island, Phuket, *N* = 4; and Similan Island, Phangnga, *N* = 3).

Sizes of RAPD fragments of primer OPB11, UBC101, UBC195, UBC197 and UBC271 across three abalone species ranged from 280 bp to 2300 bp, 300 bp to 2200 bp, 420 bp to 2600 bp, 200 bp to 2400 bp and 150 bp to 2400 bp, respectively.

3.1.3 Species-, population- and region- specific markers of three abalone (*H. asinina*, *H. ovina* and *H. varia*)

All primers generated RAPD fragments exhibiting fixed frequencies in at least one species or population (Figure 3.2 and Appendix A1.1-1.5). In total, 9, 2 and 3 fixed RAPD markers were found in limited sample sizes of *H. asinina*, *H. ovina* and *H. varia*. Ten candidate population-specific RAPD fragments were observed in *H. asinina* originating from Rayong and Cambodia (1 marker), Rayong, Cambodia and Philippines (2 markers), Trang (4 markers) and Philippines (3 markers). In addition, an RAPD marker specifically found in *H. ovina* originating from the Gulf of Thailand were also identified (Table 3.2).

Primers OPB11 and UBC195 yielded species-specific RAPD markers in three abalone in Thailand (1400 bp and 800 bp in *H. asinina*, 475 bp in *H. ovina* and 690 bp in

H. varia from primer OPB11 and 760 bp in *H. asinina*, 950 bp in *H. ovina* and 700 bp and 550 bp in *H. varia* from primer UBC195). In addition, candidate population-specific RAPD markers were found in *H. asinina* originated from Talibong Island, Trang (650 bp) and Philippines (1000 bp) using primer UBC195. Moreover, a 1650 bp RAPD marker that was generated from UBC195 was specifically found in *H. ovina*, originated from the Gulf of Thailand (east) but not found in the Andaman Sea sample (west).

Primers UBC101, UBC197 and UBC271 specifically provided RAPD markers in *H. asinina* (1,700 bp, 1325 bp, and 590 bp from primer UBC101, 1400 bp and 710 bp from UBC197 and 1,000 bp and 650 bp from primer UBC271). In addition, the primer UBC197 and UBC271 showed candidate population-specific RAPD markers in *H. asinina* from Philippines (415 bp) and Talibong Island, Trang (850 bp and 450 bp), respectively.

3.1.4 Cloning and sequencing of species-, population- and region- specific RAPD fragments

RAPD-PCR is sensitive to reaction conditions, including the requirement of a good quality DNA template for consistent results. Therefore, false negatives may possibly occur. Candidate species- and population- specific RAPD fragments were converted to sequence-characterized amplified region (SCAR) markers.

RAPD fragments were cloned using a TA cloning approach. The recombinant plasmids were isolated from 2-3 recombinant clones of each RAPD fragment and sequenced for both directions.

In a total, 72 recombinant clones were sequenced. Results showed that 39 different sequences were found. These indicated that the RAPD fragments represented co-migrating fragments that had different nucleotide sequences but similar sizes. A summary from DNA sequencing is shown in Appendix 2.1-2.42. Most of nucleotide sequences (60 clones from 72 clones, accounting for 83.3%) did not reveal significant similarity (E values > 10^{-4}) with any sequence in the GenBank.

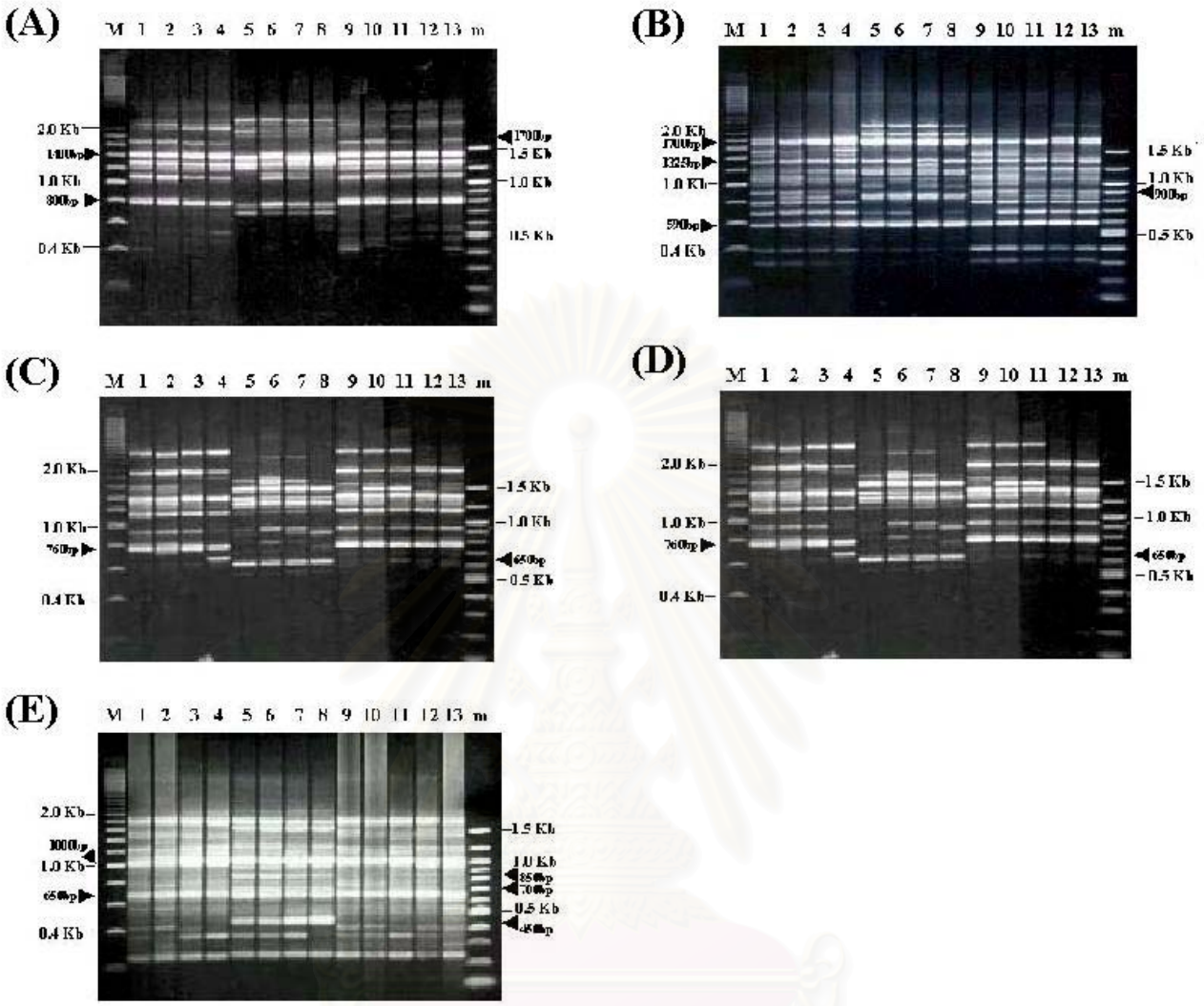


Figure 3.2 RAPD patterns resulted from analysis of *H. asinina* originating from Samet Island, Rayong (HASAME, lanes 1-4, A), Talibong Island, Trang (HATRAW, lanes 5-8, A), Cambodia (HACAME, lanes 9-13,A), P_0, with the primer OPB11 (A), UBC101 (B), UBC195 (C), UBC197 (D), and UBC271 (E). Lanes M and m are 200 bp and 100 bp DNA ladders, respectively. Arrowheads (◂) indicates species – specific markers.

Table 3.1. Species-specific and population-specific markers of *H. asinina, H. ovina* and *H. varia* based on RAPD analysis using primers OPB11, UBC101, UBC195, UBC197 and UBC271

RAPD primer	RAPD marker (bp)	No. of colonies sequenced	Name of clones	Length of clones (bp)	Blast *N* results (P-value)	Blast *X* results (P-value)	Primer designation/ name of primers	Expected size of amplicons (bp)
H. asinina								
OPB11	1400	3	pCUHA14	500	Unknown	Unknown	+/CUHA14	473
			pCUHA16	1243	Microsatellites, *H. kamtchatkana* ($3x\ 10^{-20}$)	Unknown	-	-
	800	4	pCUHA17	816	Lysin precursor, *H. rufescens* (3×10^{-7})	Unknown	-	-
			pCUHA18	832	Unknown	Unknown	-	-
UBC101	1700	3	pCUHA19	1736	Unknown	Unknown	-	-
	1325	3	pCUHA20	1324	Unknown	Unknown	-	-
	590	1	**pCUHA13***	**575**	**Unknown**	**Unknown**	**+/CUHA13**	**296**
UBC197	1400	3	pCUHA23	1388	Unknown	Chitin synthase, *D. melanogaster* (8×10^{-72})	-	-
	710	3	pCUHA1	731		Phosphoinositide-3-kinase, *Homo sapiens* (2×10^{-29})	+/CUHA1	292
UBC271	1000	3	pCUHA24	1033	Unknown	Hsp40 homologue, *Homo sapiens* (2×10^{-9})	-	-
	650	3	**pCUHA12**	644	**Unknown**	**Unknown**	**+/CUHA12**	**312**
			pCUHA5	654	Unknown	Hypothetical protein, *B. glabrata*, (8×10^{-7})	+/CUHA5	264
			pCUHA6	649	Unknown	Hemocyanin, *H. tuberculata* (4×10^{-18})	+/CUHA6	103

Table 3.1. (cont.)

RAPD primer	RAPD marker (bp)	No. of colonies sequenced	Name of clones	Length of clones (bp)	Blast *N* results (P-value)	Blast *X* results (P-value)	Primer designation/ name of primers	Expected size of amplicons (bp)
***H. asinina* (HASAME and HACAME)**								
OPB11	1700	3	pCUHA28	1667	Unknown	Unknown	-	-
			pCUHA29	1713	Hemocyanin, *H. tuberculata* (2×10^{-7})	**KIAA0794(3×10^{-8})**	-	-
UBC101	900	1	pCUHA30	548	Unknown	Unknown	-	-
***H. asinina* (HASAME, HASAMHE and HAPHIE)**								
UBC195	760	5	pCUHA21	725	Unknown	Unknown	-	-
			pCUHA22	807	Unknown	Unknown	-	-
			pCUHA4	664	Unknown	Unknown	+/CUHA4	290
***H. asinina* (HATRAW)**								
UBC195	650	2	pCUHA10	586	Unknown	Unknown	+/CUHA10	472
			pCUHA11	640	IDO-like myoglobin, *H. diversicolor* (1×10^{-70})	Unknown	+/CUHA11	417
UBC271	850	4	pCUHA7	858	Unknown	Unknown	+/CUHA7	554
			PCUHA8	837	Unknown	Unknown	+/CUHA8	142
	450	2	pCUHA9	450	Unknown	Unknown	+/CUHA9	114
	700	3	**pCUHA2***	**710**	**Unknown**	**Unknown**	**+/CUHA2**	**168**
			pCUHA3	719	Microsatellites, *H. kamtschatkana* (7×10^{-14})	Unknown	+/CUHA3	368
			pCUHA25	709	Unknown	Unknown	-	-
***H. asinina* (HAPHIE)**								
UBC195	680	1	pCUHA26	664	Unknown	Unknown	-	261
	1000	3	pCUHA27	1003	Unknown	Unknown	-	-
UBC197	415	3	pCUHA15	454	Unknown	Unknown	+/CUHA15	171

Table 3.1. (cont.)

RAPD primer	RAPD marker (bp)	No. of colonies sequenced	Name of clones	Length of clones (bp)	Blast *N* results (P-value)	Blast *X* results (P-value)	Primer designation/ name of primers	Expected size of amplicons (bp)
H. ovina								
OPB11	475	3	pCUHO1	512	Unknown	Retrotransposable element, *Oryzias latipes* (1×10^{-6})	+/CUHO1	414
			pCUHO2	525	Unknown	Unknown	+/CUHO2	146
			pCUHO3*	**502**	**Unknown**	**Unknown**	**+/CUHO3**	**328**
UBC195	950	2	pCUHO4	817	Unknown	Unknown	+/CUHO4	619
***H. ovina* (Gulf of Thailand)**								
UBC195	1650	3	pCUHO5	1636	Microsatellites, *H. rubra* (2×10^{-6})	Unknown	+/CUHO5	228
H. varia								
OPB11	690	5	pCUHV4	550	Unknown	Unknown	-	-
			pCUHV5	674	Unknown	Unknown	-	-
			pCUHV6	695	Unknown	Unknown	-	-
			pCUHV7	550	Unknown	Unknown	-	-
UBC195	**700**	**4**	**pCUHV1***	**659**	**Unknown**	**Unknown**	**+/CUHV1**	**229**
			pCUHV2	689	Unknown	Unknown	-	-
	550	4	pCUHV3	561	Unknown	Unknown	-	-

*Further characterization with a large number of specimens.

The pCUHA16 and pCUHA3 and pCUHO5 were significantly similar with microsatellites of *H. kamtchatkana* (accession AYO3578 and AY013579) and of *H. rubra* (AF194955), respectively. Homologues of sperm lysin precursor of *H. rufescens* (accession AF076824), chitin synthase of *D. Melanogaster* (AF227729) phosphoinositide-3-kinase of *Homo sapiens* (N_M006218), heat shock protein 40 of *Homo sapiens* (NM_014787), hypothetical protein of *Biomphalaria glabrata* (PC1123) , hemocyanin of *H. tuberculata* (AJ252741), IDO-like myoglobin of *H. diversicolor* (D83984) and retrotransposable element of *Oryzias latipes* (AB073376) were represented by pCUHA17, pCUHA23, pCUHA5, pCUHA24, pCUHA6, pCUHA29, pCUHA1, pCUHA11, and pCUHO1, respectively (Table 3.1).

3.1.5 Development of species-specific SCAR markers of *H. asinina*

Twenty one pairs of primers were designed (15, 5 and 1 for *H. asinina*, *H. ovina* and *H. varia*, respectively) from nucleotide sequence of RAPD markers and subjected to specificity test. Nine primer pairs revealed specificity with small sample sizes. Seven pairs of primers including CUHA1 (Figure 3.3 A), CHHA2, CUHA4, CUHA11, CUHA12, CUHA13 and CUHA14 revealed species-specificity in *H. asinina* while primers CUHO3 and CUHV1 exhibited the expected product restrictively found in *H. ovina* and *H. varia*, respectively. However, CUHA14 provided different sizes of the amplification products. The expected 473 bp was found in HACAMHE and HACAME whereas a larger 515 bp was found in HATRAW and HAPHIE, respectively (Figure 3.3 B). Non-specific amplification products were observed at CUHO4 primer (Figure 3.3 C). CUHO2 provided strong herterozygotic band (the expected 146 bp band and an extra 215 bp band) in *H. ovina* and non-specific bands in *H. asinina* (HACAMHE and HACAME) and *H. varia*.

Additionally, CUHA5 (Figure 3.3D), CUHA6, and CUHO1 showed positive amplification bands in all investigated abalone. The specificity of CUHA5 (264bp) and CUHA6 (103bp) were further tested against 6 gastropod species (the giant African snails, *Achatina fulica* and the apple snail, *Pomacea canaliculata*, *Pila ampullacea*, *P. angelica*, and *P.polita*), 6 bivalve species (oysters; *Crassostrea belcheri*, *C. iredalei*, *Saccostrea cucullata*, *S. forskali* and *Striostrea (Parastriostrea) mytiloides*; the mussel, *Perna viridis*) and 2 crustacean species (the giant tiger shrimp, *Penaeus monodon* and the giant freshwater prawn, *Macrobrachium rosenbergii*).

These primers provided no positive amplification fragments in the non-abalone species and may be used as genus-diagnostic markers of abalone in this study.

Originally, nine primer sets were designed for the development of population-specific markers in *H. asinina*. Four of these (CUHA3, CUHA7, CUHA8 and CUHA10) provided non-specific amplification results; CUHA9 and CUHA15 provided positive amplification in all of Thai abalone species and CUHA2, CUHA4, and CUHA11 yielded a *H. asinina*-specific rather than a population-specific nature (Table3.3). Therefore, population-specific SCAR markers were not found in *H. asinina*. CUHO5 derived from a region-specific RAPD fragment did not retain the original specificity. Cross-species amplification was also observed in *H. varia*.

CUHA2

GCCATCAAGATGTGACATGTTCATGGGCAACTAAAATGGATTGAAAAAACGTTACAA
GTATTTCACAAACCTTTCCCTTTATATATTTATCCTGAATTAAATTGAACTCTGCTG
AAAAGTGGCATATCATTAATCAAAATATAGACTGGTCAAACGTAATGACTTCTTCAT
GACTAAATATAACTGAAGATCAATGAACAAGGTAAATATTGCTGTAAACAATAAACA
GGCACATGTTTCAAAACAATTCCACACAAAGGTAACTTTTGCTATTTGGTTATGTAA
ACAAGCTGCAGAGCCCTCTTATGAAAGTTCACCTCCCAATGACTTCATTTATCAGTA
CUHA2-F
ATGTGTGATCATATCCAGGATGC**TTGTTCAGCATTCTGTGGCAGTTCT**TGACAACTT
GATTTCAGTACTGAATGAACAAGAAAGACTGACAATTTTGTTCCATGTTGCCCCCAA
GTTATTCTTGACAGATGTAGATTCTGCATTAGTGACAAAGATTTTTATATAT**CTCCA**
CUHA2-R
AAGGGTCAGCAAAAAAGAAGTACAGTATACCCTGGTGATATCTTCCACTGTTTGTCA
GAGGGAACATATGGTACTGTATCCAGACACACGTTATAACCAGGGTGCACTGTACAT
GCCCACATGGAAGCATGAATCATGACAGTCAATACCATAACTGTATTTAATTTACAA
AAAAAAATATATATGATCTTGATGGC

Figure 3.4. Sequences of pCUHA2. The locations and sequences of *H.asinina*-specific forward primer (CUHA2-F) and those complementary to reverse primers (CUHA2-R) is labeled in boldface and underlined.

A: CUHA12

GCCATCAAGAATAAGCGTTATGAAAACACTAACTCAAATCAGCCAAATAATCCCACT
AGTGATATCCTCTTTGGACTGAAGTTGTTTGTTTTAAGAGTAAGAGGATTATAAGGC
CUHA12-F
ATGCTCAACTGGGTTTGGTCGTTTAA**CTAATCCCACACAGCCATCACCAG**CAGAGGG
AATGATTTACACATCTCATTTTTTGTGCTTAATTATTGTGTTTGTTTTTTATTGAAA
AAAAAATGAAATGGGAGTGTGAACAGCTTTACAGTTTATATCGACACTGGTACCATA
CAACATACAATACAAAATTTAAATGCAGCATGTGTTGTCACATATGTATATAAATTT
GAATTACAAACATGGGTAAAGCTTGAAGACAATCTTATATGGAAATAGTGAATTATC
CUHA12-R
TGATCTTGATAATCTGTCCATCTTAAGAT**CTGCCTACCTCTTCGTCACTTCTT**CTTT
TAGATTTGATACTAAACATGATTGAAGAATTAACCTCGTGGTTTGTCATATGACAAT
CTTGTATTAGAACTAGGTTAGGTTTATTTACCTTTGTAAAGTGCATTACACACATAC
AAACTAACTGTATTAATATAGAGCATTTATGAGAATGGTATTGGTAAGANAATTATT
TTTGGTGTCTTGATGGC

B: CUHA13

GCGCCTGGAGACAAGGTTACGTGTAATTATATACTGACACATATCATTGTGAGATGA
TAAGAGTTAACACTGCATTCTGGATTATTCAGTGTGTGTGAACAGTGTCAAACTTAA
CUHA13-F
TGTTAGCGCAGTACAGCACATCAC**TGACCTGTGTTGAGACTCTACGGA**AAAGTCTAA
TATCGTAATGCACAACAATAAAAACCATGGTTACATTTCTATTACCAAAAAACTAGA
AAAACAACATCTACCAGAATGAGTAAGTGTGTGAAGTTTAGTTCTACACCACTTCTA
GCAATAATCCAGCAATATGACGGGAGTGGGGGAGATAATAGAAACGCTAGAAATGGG
CTTTACACATTGCATCAATGAGATGGTTTGCACTTTGGTCTTGGACGTGACGTGCAA
CUHA13-R
ATGCTTTATCCATG**CGGCTACTCCATCTCCCCTCA**TCTAACTGACACAAACACAGGA
ATTGTCAGCTGTCAAAGCTNCAGTACAGTGACTCTCTGCAACCATTGTTGTTTTGTG
TTGTCACAGGACCAGGAGCCCTATGGGTTAAAAAGTTGCCAAGTCCCACCCACCTTC
AGCGC

Figure 3.5. Sequences of pCUHA12 (A) and pCUHA13 (B). The locations and sequences of *H.asinina*-specific forward primers (CUHA12-F and CUHA13-R) and those complementary to reverse primers (CUHA12-R and CUHA13-R) are labeled in boldface and underlined.

A: CUHO3

CUHO3-F

GTAGACCCGTCATT**GGGTATCTTCCCACAACAGC**TGTCTCGTGCTGTAGGTATCCTA
CAAATGAGAAAAACATCTGCTAGTTGTTATAATTATTTGGGTGATAAAGTTATGTAT
TACCCCCCCTGATGTTAGTTATTGGAATGTATACACACACATATACCATATACAAAG
AAAATACATACACACCAAACTGAACATGCTTAATCATAAAGTTTGTGGTGTCATATA
CTGAATGGCAAACAAGGATAAGTTGGACTGGCACAGAGAACATGTTCATGGGAAGTA
CUHO3-R
GGCAGGGCACAGCAGTTAGTAAAACCATTTGTTTT**GTGAAAGGATGTAGGCAAGTGC**
AGATGCAGTGCATGGCGATGAAATTGATTCTGTGCATAGCTGCACATGCCTTACTCA
TAAGATATTCACCTGGTTGGTAACTACACTCTATCATGGATGACTTGGGACCTCTTC
ACTCACAATAGCAAGAACCAGTTCATATGTGCTAATACGGGTCTAC

B: CUHV1

GATCTCAGCGCATACATAATCATATTCATTACAAAGGATATAACACCAGAAATGTCG
CGCCATAATTCTTATGACATACAGAAACCGGGGTAATTTTATCGGCATCGCTGCTAT
ATAATTTCCCCTCCTGAAACTAATAGAATTTCCTAATTGGGGTTGTCATGGTAGGTT
CUHV1-F
CTTTGCCTAA**CCCCTTGTTTCTCCTTCTTG**TCGACATGAAAAGACCACTTTATCAAA
TATTTCATCCTCACTTCATTCACAATAATCTCTTGATGAGCCCAACATTCGTCTGAA
ACATCCGCACGAGGCAGCCGGAACCACCAGGCTGTTGGCAACTTATTCAAAGTCTCT
TCTCTTCACCCACAGAGATTTTTTCACAATACCAACAAACAATATCCA**TCAAACCGC**
CUHV1-R
CTGCGTCATCGTTTGTACTCAAGCCTCAATCGAACTTCGCCAGTGCACTATGGACGT
GCTTCGACGGTCGGCTGCAAACCTTTGCAACAAAACCACCGATTCTAGAAGTTTCGT
TCATTGCATTCCCCCGACAACTGCTATTATAAAACATTTAGTTAGTGTAATTGTAAC
TGGTCTATCTCGAGTCTTCGTCTTCATTACGGATAATTAGCCCGTCGCTGAGGGGGG
CACTTTCTACGCTCATCCATCCCGCTGAGATC

Figure 3.6. Sequences of pCUHO3 (A) and pCUHV1 (B). The locations and sequences of *H. ovina*- and *H. varia*-specific forward primers (CUHO3-F and CUHV1-F) and those complementary to reverse primers (CUHO3-R and CUHV1-R) are labeled in boldface and underlined.

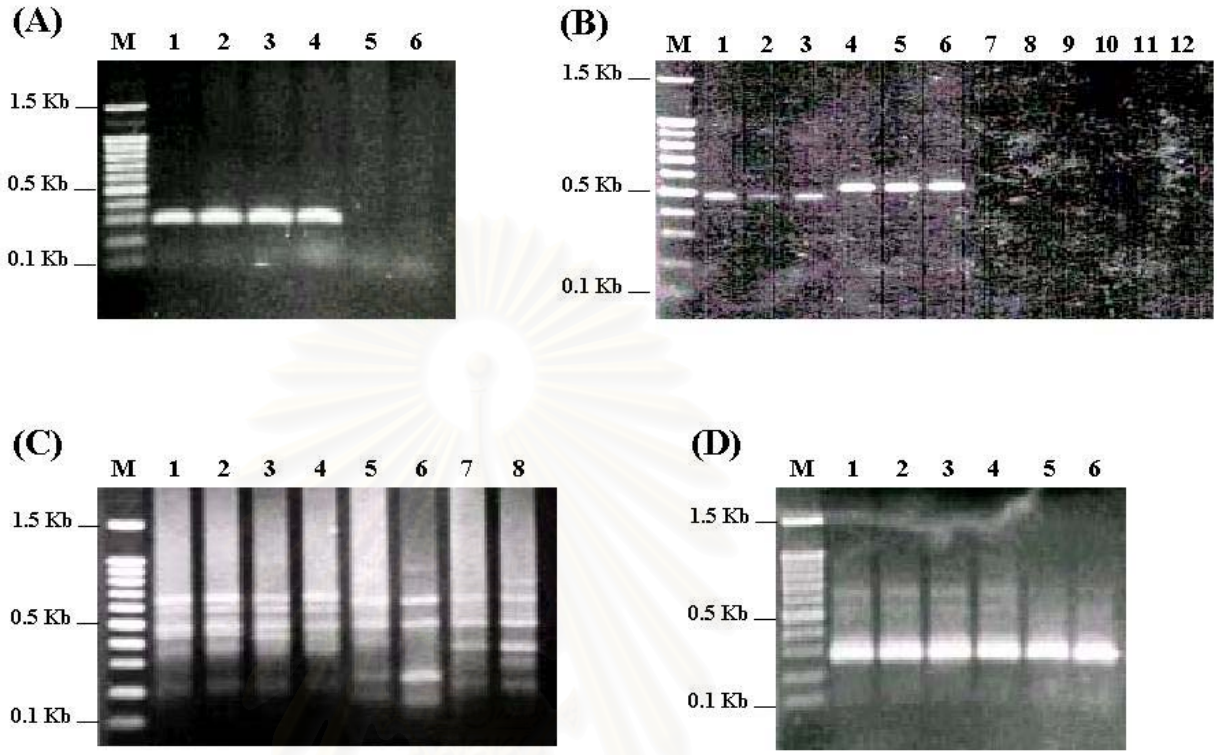


Figure 3.3 Agarose gel electrophoresis showing results from amplification of genomic DNA of *H. asinina* (lanes 1-4, A, C, and D), *H. ovina*, (lane 5, A and D and lanes 5-6, C) and *H. varia* (lane 6, A and D and lanes 7-8, C) from CUHA01 (A), CUHO04 (C), and CUHA05 (D) and that of HASAME (lanes 1-3), HATRAW (lanes 4-6) of *H. asinina*, *H. ovina*, (lanes 7-9) and *H. varia* (lane 10-12) using CUHA14 (B). A 100 bp ladder is used as a DNA marker.

Table 3.2. A specificity test of SCAR markers developed from population-specific or species-specific RAPD markers of *H. asinina*, *H. ovina* and *H. varia*

Primers	Expected amplicon (bp)	*H. asinina*				*H. ovina*	*H. varia*
		HACAMHE	HACAME	HATRAW	HAPHIE		
CUHA1	292	+	+	+	+	-	-
CUHA2	**168**	**+**	**+**	**+**	**+**	**-**	**-**
CUHA3	368	NS	NS	NS	NS	NS	NS
CUHA4	290	+	+	+	+	-	-
CUHA5	264	+	+	+	+	+	+
CUHA6	103	+	+	+	+	+	+
CUHA7	554	NS	NS	NS	NS	NS	NS
CUHA8	114	NS	NS	NS	NS	NS	NS
CUHA9	142	+	+	+	+	+	+
CUHA10	472	NS	NS	NS	NS	NS	NS
CUHA11	417	+	+	+	+	-	-
CUHA12	**312**	**+**	**+**	**+**	**+**	**-**	**-**
CUHA13	**296**	+	+	+	+	-	-
CUHA14	473	+	+	515 bp	515 bp	-	-
CUHA15	171	+	+	+	+	Faint band	Faint band
CUHO1	414	+	+	+	+	+	+
CUHO2	146	Faint band	Faint band	-	-	+/215	-
CUHO3	**328**	**-**	**-**	**-**	**-**	**+**	**-**
CUHO4	619	NS	NS	NS	NS	NS	NS
CUHO5	228	-	-	-	-	+	Faint band
CUHV1	**229**	**-**	**-**	**-**	**-**	**-**	**+**

Abbreviations: +, successful amplification with the expected product; -, no amplification product; NS, non-specific amplification fragments obtained.

Five SCAR markers (CUHA2, CUHA12, CUHA13, CUHO3 and CUHV1) were eventually selected and further examined against a large sample size of abalone (*N = 216*). Nucleotide sequences of these clones and location of primers are shown in Figure 3.4 -3.6.

Species-specificity was observed from CUHA2, CUHA12, CUHA13, and CUHV1 (100% without false positive/negative results, Figure 3.7 - 3.8). No heterozygotes that exhibited two different sizes (alleles) of the amplification products were observed across overall specimens that were analyzed by these primers. This implies the retention of a dominant segregated fashion of the original RAPD markers.

Although CUHO3 yielded a strong amplification products in the target species (100%), a very faint product was also observed in some individuals of the non-target species ; *H. asinina and H. varia*, which suggested that problems may be arisen from this primer pair

Sensitivity of *H. asinina*-specific primers (CUHA2, CUHA12, and CUHA13) and *H. varia*-specific primer (CUHV1) was examined against varying concentrations of the target DNA template between 10 pg – 25 ng. All primer pairs revealed good correlation between the amount of DNA template and intensity of the PCR product. The sensitivity of detection was approximately 100 pg and 500 pg of the target DNA template for primers CUHA2 and CUHA12 and CUHA13 but greater sensitivity was observed with CUHV1 (20 pg) (Figure 3.9). The sensitivity levels of the species-specific PCR developed in this study were sufficient for the identification of the species-origins of abalone, beginning with the early development stages.

For rapid species-identification of *H. asinina*, the tedious and time-consuming phenol/chloroform extraction method was omitted by simplification of the extraction method to a rapid 5% chelex-based method. Species-specific PCR (CUHA2, CUHA12, and CUHA13) was test against genomic DNA of frozen specimens (up to 3 years at -30 °C), ethanol-preserved larvae (several months at 4 °C), dried (80 °C for 72 hours and kept at room temperature for 2 weeks) and boiled (10 minutes) broodstock of *H. asinina* using primers CUHA2, CUHA12, and CUHA13 (Figure 3.10). Positive amplification products were consistenly observed indicating reliability of species-specific SCAR markers developed in this thesis.

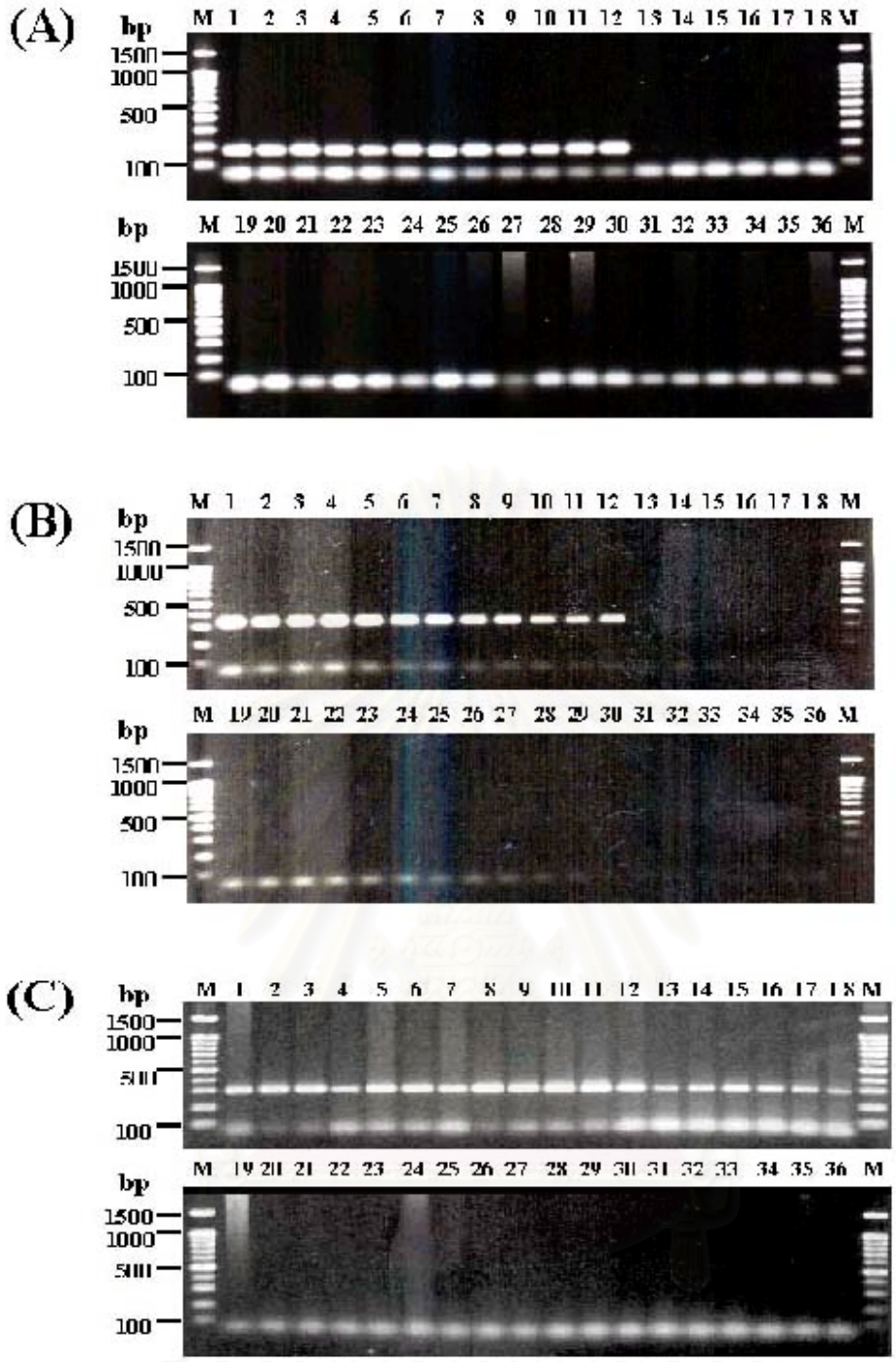


Figure 3.7. Agarose gel electrophoresis illustrating species-specificity of CUHA2 (A), CUHA12 (B), and CUHA13 (C) against genomic DNA of *H. asinina* (lanes 1-12 for CUHA02 and CUHA12 and lanes 1-18 for CUHA13), *H. ovina* (lane 13-24 for CUHA02 and CUHA12 and lanes 19-27 for CUHA13) and *H. varia* (lanes 25-36 for CUHA02 and CUHA12 and lanes 28-36 for CUHA13). A 100 bp ladder (lanes M) was used as a DNA marker.

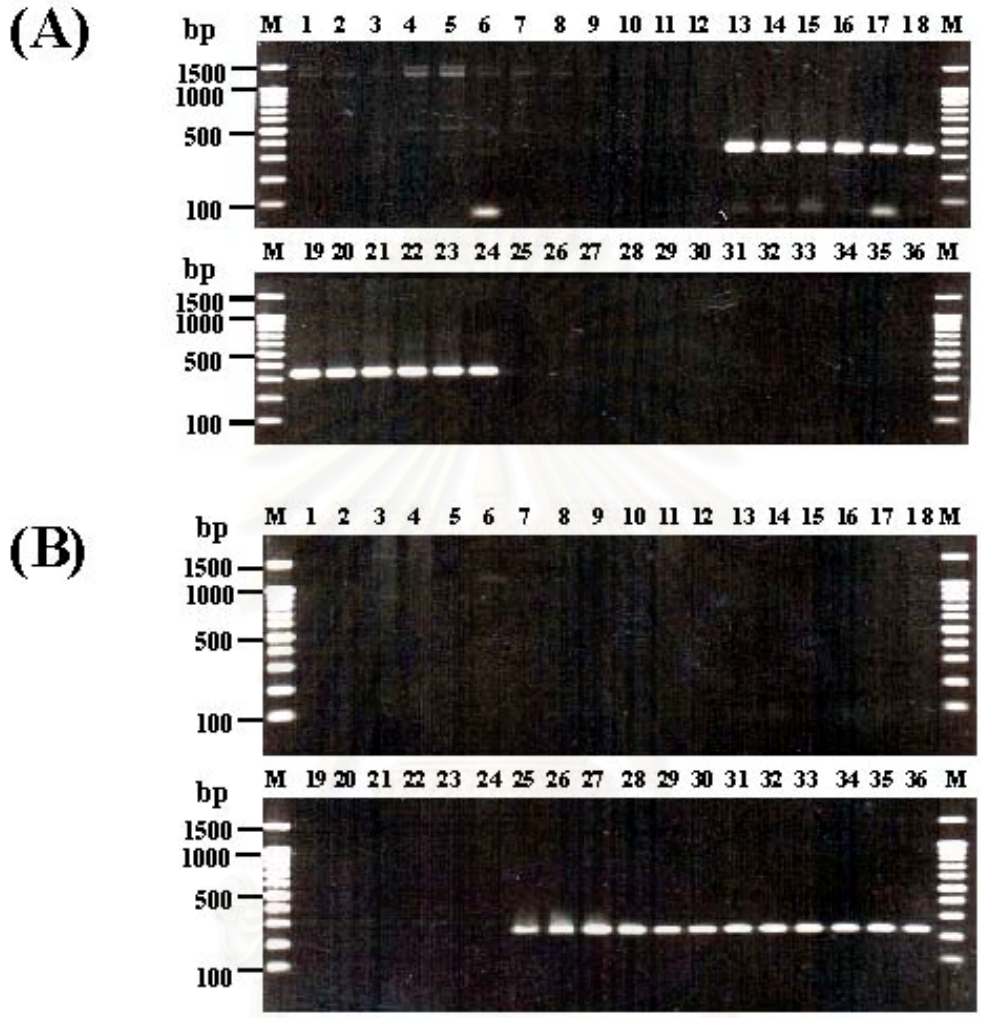


Figure 3.8. Agarose gel electrophoresis illustrating amplification results of CUHO03 (A) and CUHV01 (B) against genomic DNA of *H. asinina* (lanes 1-12), *H. ovina* (lane 13-24) and *H. varia* (lanes 25-36). A 100 bp ladder (lanes M) was used as a DNA marker.

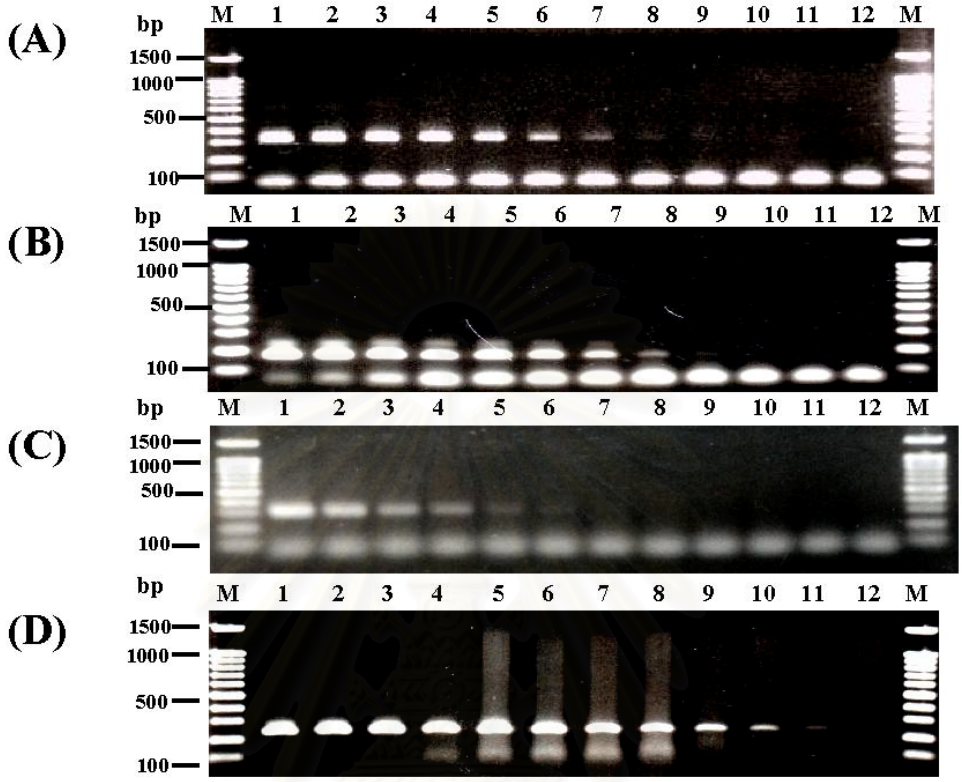


Figure 3.9. Sensitivity of CUHA2 (A), CUHA12 (B), CUHA13 (C) and CUHV1 (D) was examined against varying concentrations of *H. asinina* (A and B) and *H. varia* (C) DNA template (25 ng, 10 ng, 5 ng, 2.5 ng, 1 ng, 500 pg, 250 pg, 100 pg, 60 pg, 30 pg 20 pg and 10 pg corresponding to lanes 1-12, respectively). A 100 bp ladder (lanes M) was used as a marker.

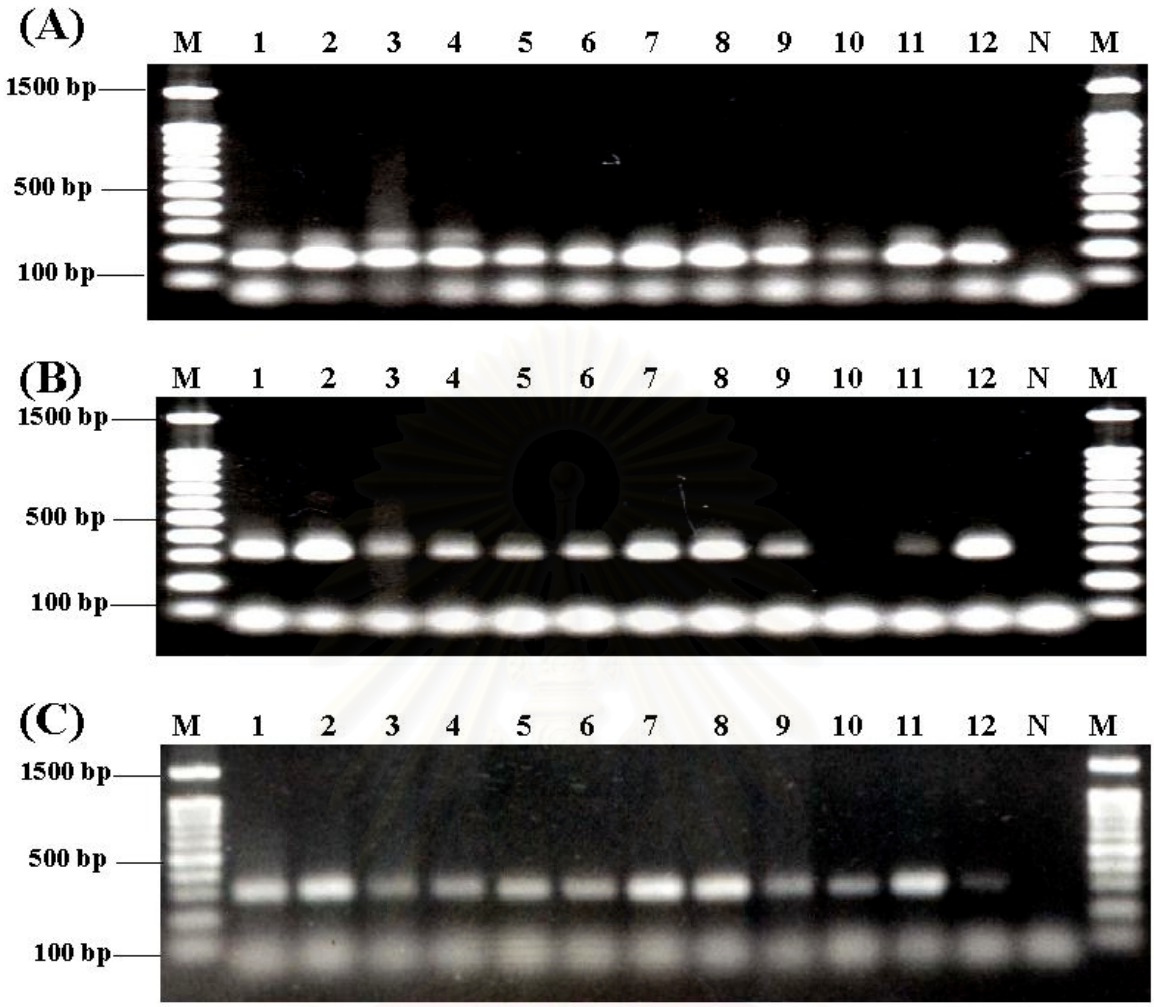


Figure 3.10 Agarose gel electrophoresis showing results from amplification of total DNA extracted with phenol/chloroform (lanes 1 - 2) and 5% Chelex extraction methods (lanes 3 - 12) of frozen (lanes 1 – 4), ethanol-preserved larvae (approximately 1 mM in size, lanes 5 – 8) and dried (lanes 9 – 10) and boiled (11 – 12) broodstock of *H. asinina* with primers CUHA2 (panel A), CUHA12 (panel B) and CUHA13 (panel C). Lanes 13 are negative controls (without DNA template). A 100 bp ladder (lanes M) was used as a DNA marker.

3.2 Identification and characterization of sex-related markers in genomic DNA of *H. asinina* using AFLP analysis

3.2.1 Screening of candidate sex-specific AFLP

A combination of bulked segregant analysis (BSA) and AFLP methodology was used to identify sex determination markers in *H. asinina*. Bulked genomic DNAs of male (HAM1, N = 10 and HAM2, N = 8) and female (HAF1, N = 10 and HAF2, N = 8) were generated by pooling genomic DNA of different male and female individuals.

Pre-amplification was carried out using a single pre-selective nucleotide of *Eco*RI primer (E_{+A}) and *Mbo*I primer (M_{+C}). The preamplification products indicated successful digestion/ligation of *H. asinina* genomic DNA with *Eco*RI and *Mbo*I adaptors. A total of 214 primer combinations were tested against two pooled genomic DNA of males and females of *H. asinina*, respectively.

The pre-amplification products of HAM1 and HAF1 were subjected to selective amplification for the primary screening. The selective amplification products was examined for the amplification success by 1.8% agarose gel electrophoresis (Figure 3.11).

To characterize the selective amplification fragments, products were size-fractionated through 6 % denaturing polyacylamide gels (19:1 crosslink). A total of 28 candidate male- and 22 female-related AFLP fragments from 28 polymorphic primer combinations were found from screening of HAM1 and HAF1 of *H. asinina* (Table 3.3).

To eliminate the false positive results, primer combinations that gave candidate AFLP fragment from the primary screening were re-screened with bulk HAM1 and HAF1 and second sample set (HAM2 and HAF2). Only 3 fragments provided reproducible results suggested that the use of only 2 bulked genomic DNA (one from females and the other from males) provided a large number of false positive AFLP fragments.

In addition, 150 combination primers (E_{+3}-1-8 + M_{+3}-9-16 and E_{+3}-9-14 + M_{+3}-1-16) were simultaneously screened against 4-bulked genomic DNA of HAM1 and HAF1, HAM2 and HAF2. A total of 5 and 6 promising AFLP bands exhibiting male and female specificity were identified, respectively (Figure 3.12- 3.13, Appendix B1.1-1.10). Finally, 7 candidate female-specific and 7 male-specific AFLP markers were observed (Table 3.4)

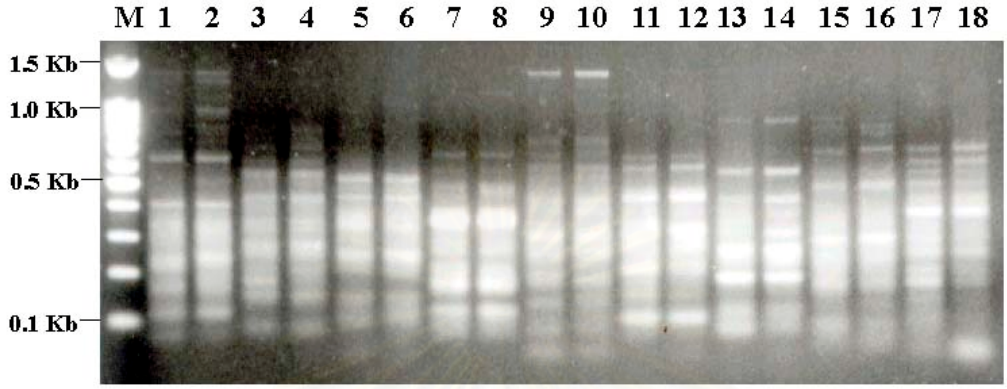


Figure 3.11 Agarose gel electrophoresis showing selective amplification products of bulked HaM1 (lanes 1, 3, 5, 7, 9, 11, 15, and 17) and HaF1 (lanes 2, 4, 6, 8, 10, 12, 14, 16, and 18) using primers E_{+3}-1/M_{+3}-6 (lanes 1 - 2) E_{+3}-1/M_{+3}-7 (lanes 3 – 4), E_{+3}-1/M_{+3}-8 (lanes 5 – 6), E_{+3}-2/M_{+3}-6 (lanes 7 – 8), E_{+3}-2/M_{+3}-7 (lanes 9 –10), E_{+3}-2/M_{+3}-8 (lanes 11 – 12), E_{+3}-3/M_{+3}-6 (lanes 13 –14), E_{+3}-3/M_{+3}-7 (lanes 15 – 16) and E_{+3}-3/M_{+3}-8 (lanes 17 – 18). Lane M is a 100 bp DNA marker.

Table 3.3 Candidate sex-specific AFLP markers of *H. asinina* generated by primer combinations E_{+3}-1-8/M_{+3}-1-8

Primer combinations	AFLP bands from 4 DNA bulks (bp)			
	Male (HAM1, *N* =10)	**Female (HAF1, *N* =10)**	**Male (HAM2, *N* =8)**	**Female (HAF2, *N* =8)**
E_{+3}-1/M_{+3}-1	320 and 550		-	-
E_{+3}-1/M_{+3}-2	350 and 560	800	-	-
E_{+3}-1/M_{+3}-3	-	1500	-	-
E_{+3}-1/M_{+3}-6	280	270	-	-
E_{+3}-1/M_{+3}-7	**300***	-	**300***	-
E_{+3}-1/M_{+3}-8	1500	280	-	-
E_{+3}-2M_{+3}-4	600	-	-	-
E_{+3}-2/M_{+3}-5	-	1100 and 355	-	-
E_{+3}-2/M_{+3}-6	480	350 and 450	-	-
E_{+3}-2/M_{+3}-7	280 and 1000	-	-	-
E_{+3}-2/M_{+3}-8	180 and 425	-	-	-
E_{+3}-3/M_{+3}-3	425	-	-	-
E_{+3}-3/M_{+3}-6	350 and 500	-	-	-
E_{+3}-4/M_{+3}-6	-	450	-	-
E_{+3}-4/M_{+3}-8	-	310	-	-
E_{+3}-5/M_{+3}-2	-	170	-	-
E_{+3}-5/M_{+3}-3	-	225	-	-
E_{+3}-5/M_{+3}-6	325		-	-
E_{+3}-5/M_{+3}-8	250 and 350	**400***	-	**400***
E_{+3}-6/M_{+3}-1	-	280	-	-
E_{+3}-6/M_{+3}-2	210 and 255	-	-	-
E_{+3}-6/M_{+3}-5	245 and 400	700 and 1300	-	-
E_{+3}-6/M_{+3}-6	950	315 and 750	-	-
E_{+3}-6/M_{+3}-7	-	315 and 325	-	-
E_{+3}-6/M_{+3}-8	360	-	-	-
E_{+3}-8/M_{+3}-6	265	230	-	-
E_{+3}-8/M_{+3}-7	**280***	-	**280***	-
E_{+3}-8/M_{+3}-8	245	550	-	-
Total	**28**	**22**	**2**	**1**

*AFLP fragments showing consistent results against 2 bulks of female and male genomic DNA are illustrated in bolfface.

Table 3.4 Candidate male- and female specific AFLP markers of *H. asinina* across primer combinations E_{+3}-1-14/M_{+3}-1-16

Candidate AFLP marker	Primer combination	Size (bp)
Male-specific AFLP fragment	*Eco* RI_{+3}-1/*Mse* I_{+3}-7	300
	Eco RI_{+3}-8/*Mse* I_{+3}-7	280
	Eco RI_{+3}-5/*Mse* I_{+3}-16	325
	Eco RI_{+3}-7/*Mse* I_{+3}-16	520
	Eco RI_{+3}-11/*Mse* I_{+3}-11	280
	Eco RI_{+3}-6/*Mse* I_{+3}-9	517
	Eco RI_{+3}-9/*Mse* I_{+3}-15	320
Female-specific AFLP fragment	*Eco* RI_{+3}-4/*Mse* I_{+3}-13	235
	Eco RI_{+3}-4/*Mse* I_{+3}-13	260
	Eco RI_{+3}-7/*Mse* I_{+3}-15	275
	Eco RI_{+3}-5/*Mse* I_{+3}-15	375
	Eco RI_{+3}-5/*Mse* I_{+3}-8	400
	Eco RI_{+3}-6/*Mse* I_{+3}-15	460
	Eco RI_{+3}-11/*Mse* I_{+3}-15	490

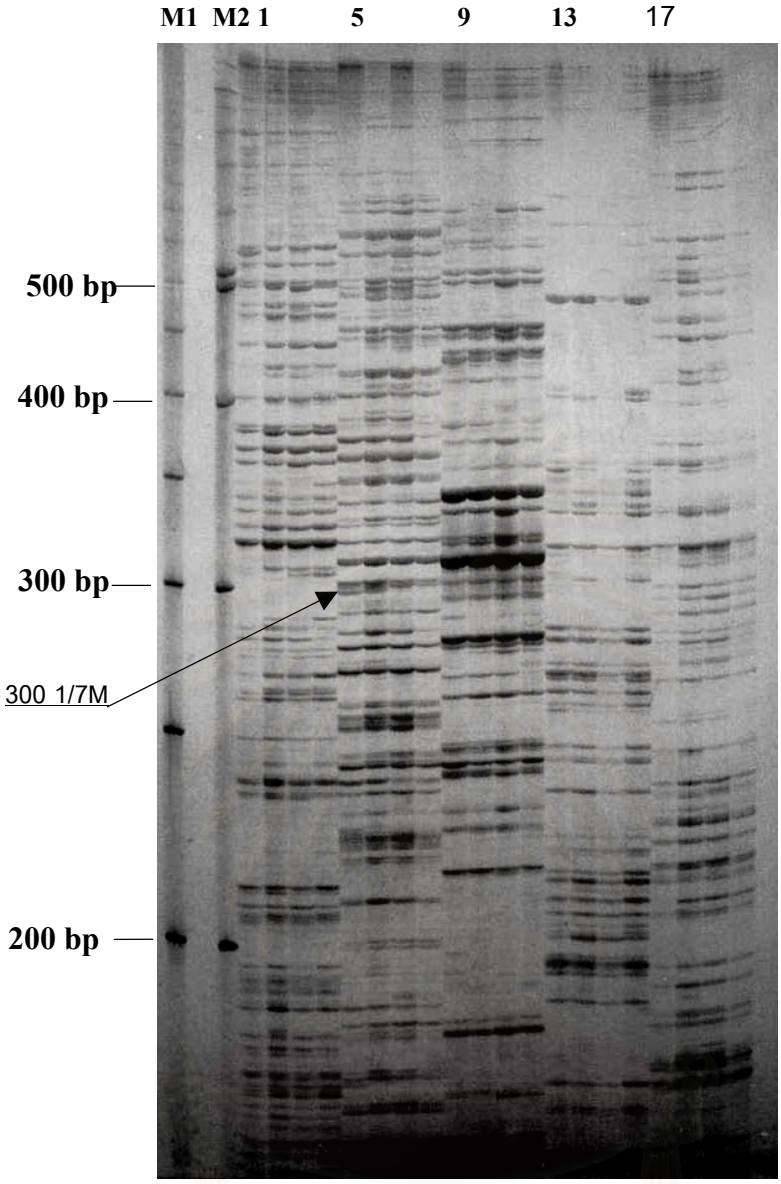


Figure 3.12 A 6% denaturing polyacrylamide gel electrophoresis showing AFLP products of 4 bulked ABM1 (lanes 1, 5, 9, 13 and 17), ABM2 (lanes 2, 6, 10, 14 and 18), ABF1 (lanes 3, 7, 11, 15 and 19) and ABF2 (lanes 4, 8, 12, 16 and 20) using primers E_{+3}-1/M_{+3}-1 (lanes 1 – 4), E_{+3}-1/M_{+3}-7 (lanes 5 – 8): E_{+3}-2/M_{+3}-5 (lanes 9 – 12) E_{+3}-2/M_{+3}-6 (lanes 13 – 16), E_{+3}-2/M_{+3}-8 (lanes 17 – 20). An arrowhead indicates a candidate sex-specific AFLP marker from E_{+3}-1/M_{+3}-7.Lanes M1 and M2 were 100 bp and 50 bp DNA markers, respectively.

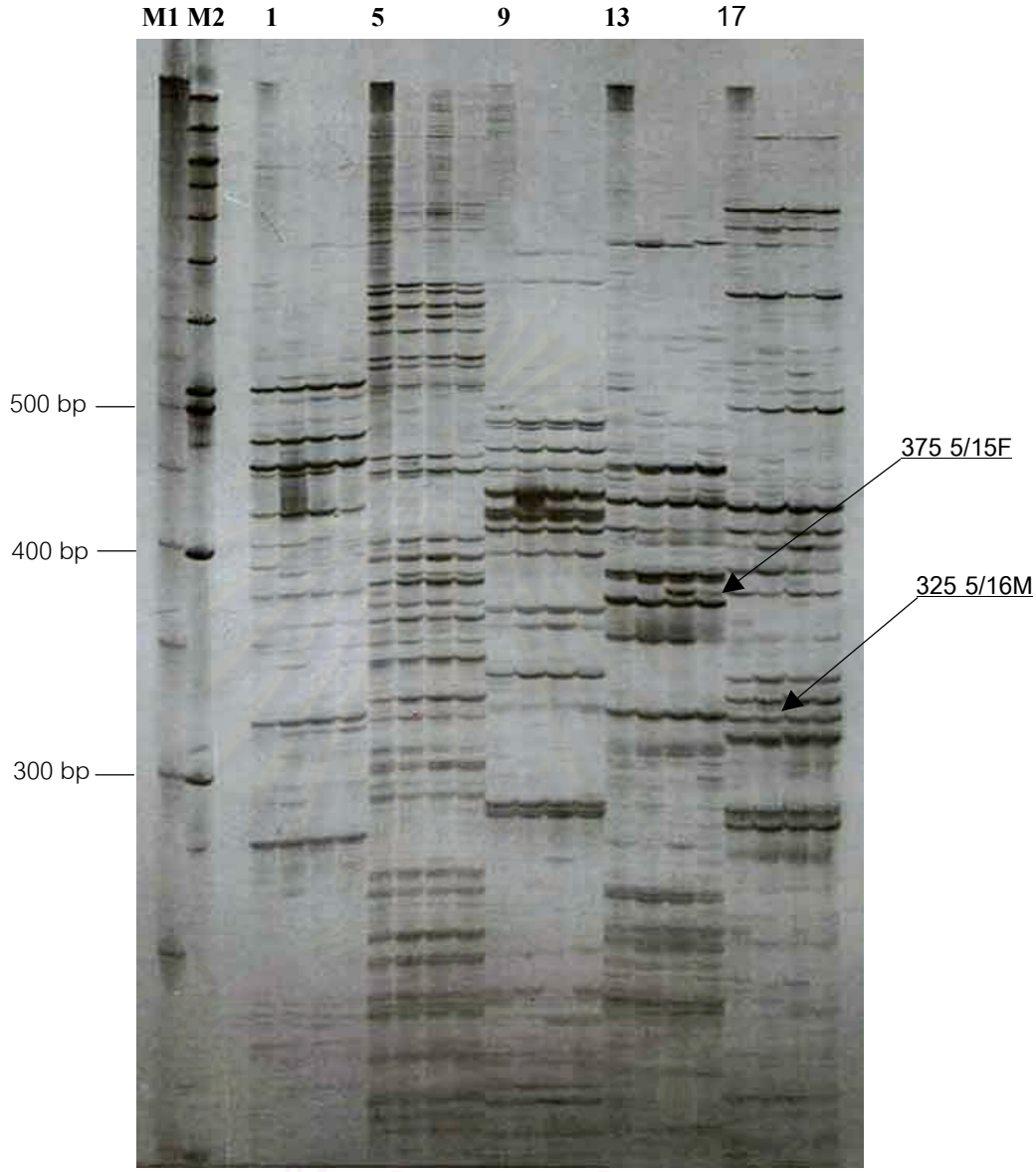


Figure 3.13 A 6% denaturing polyacrylamide gel electrophoresis showing AFLP products of 4 bulked ABM1 (lanes 1, 5, 9, 13 and 17), ABM2 (lanes 2, 6, 10, 14 and 18), ABF1 (lanes 3, 7, 11, 15 and 19) and ABF2 (lanes 4, 8, 12, 16 and 20) using primers E_{+3}-5/M_{+3}-12 (lanes 1 – 4), E_{+3}-5/M_{+3}-13 (lanes 5 – 8): E_{+3}-5/M_{+3}-14 (lanes 9 – 12) E_{+3}-5/M_{+3}-15 (lanes 13 – 16), E_{+3}-5/M_{+3}-16 (lanes 17 – 20). An arrowhead indicates a candidate sex-related AFLP fragment from E_{+3}-5/M_{+3}-15 and E_{+3}-5/M_{+3}-16. Lanes M1 and M2 were 100 bp and 50 bp DNA markers, respectively.

3.2.2 Cloning and characterization of candidate sex-specific AFLP fragments in H. asinina

Candidate sex-specific AFLP fragments were converted to sequence-characterized amplified region (SCAR) markers for more reliable amplification by PCR. All fourteen candidate sex-related AFLP fragments (7 male- and 7 female-related fragments) were cloned (Table 3.4). Recombinant plasmids were extracted and 1-5 positive clones of each AFLP fragment and unidirectional sequenced.

Fifty recombinant clones were sequenced. Twenty-nine different sequences were found indicating that an AFLP fragment represented co-migrating fragments that had different nucleotide sequences but similar size. Nucleotide sequences of characterized clones are shown in APPENDIX B 2.1-2.14. Original nucleotide sequences of AFLP primer combinations were found in all sequences indicating that each insert was not generated from non-specific amplification or from product of only a single primer.

Name of informative primers, sizes of AFLP markers and the original specificity are illustrated by Table 3.5. Nucleotide sequence of each clone was blasted against data in the GenBank. Almost all of the clones were unknown transcript ($E > 10^{-4}$). Only 4005/8F2 showed significant similarity with LTR retrotransposable element in the Japanese medaka (*Oryzias latipes*).

Table 3.5: Candidate sex-specific AFLP fragments of males and females of *H. asinina* generated by AFLP analysis, nuber of recombinant clones sequenced length of each insert and similarity analysis using BLAST*N* and BLAST*X*

AFLP primer	Male-specific marker	No. of clone	Clone	Length	BLASTN	BLASTX	SCAR primer
EcoRI01/MseI07	300 1/7M	3	3001/7M1	309 bp	Unknown	Unknown	HaMale3F/R
			3001/7M2	297 bp	Unknown	Unknown	HaMale4F/R
			3001/7M3	297 bp	Unknown	Unknown	HaMale5F/R
EcoRI08/MseI07	280 8/7 M	3	2808/7 M1	237 bp	Unknown	Unknown	
			2808/7 M2	238 bp	Unknown	Unknown	
			2808/7 M3	242 bp	Unknown	Unknown	
EcoRI05/MseI16	325 5/16M	4	3255/16M1	322 bp	Unknown	Unknown	
			3255/16M2	322 bp	Unknown	Unknown	
			3255/16M3	312 bp	Unknown	Unknown	
			3255/16M4	325 bp	Unknown	Unknown	
EcoRI07/MseI16	520 7/16M	5	5207/16M	522 bp	Unknown	Unknown	HaMale2F/R
EcoRI11/MseI11	280 11/11M	3	28011/11M1	296 bp	Unknown	Unknown	
			28011/11M2	285 bp	Unknown	Unknown	HaMale6F/R
			28011/11M3	290 bp	Unknown	Unknown	
EcoRI06/MseI09	517 6/9M	4	5176/9M1	519 bp	Unknown	Unknown	
			5176/9M2	520 bp	Unknown	Unknown	
EcoRI09/MseI15	320 9/15M	2	3209/15M	318 bp	Unknown	Unknown	HaMale1F/R

Table 3.5 (cont.)

AFLP primer	Female-specific marker	No. of clone	Clone	Length	BLASTN	BLASTX	SCAR primer
EcoRI04/MseI13	235 4/13F	2	2354/13F	220 bp	Unknown	Unknown	
EcoRI04/MseI13	260 4/13F	2	2604/13F1	269 bp	Unknown	Unknown	HaFemale5F/R
			2604/13F2	263 bp	Unknown	Unknown	
EcoRI07/MseI15	275 7/15F	5	2757/15F1	276 bp	Unknown	Unknown	HaFemale1F/R
			2757/15F2	271 bp	Unknown	Unknown	HaFemale2F/R
EcoRI05/MseI15	375 5/15F	3	375 5/15F	372 bp	Unknown	Unknown	HaFemale3F/R
EcoRI05/MseI08	400 5/8F	5	4005/8F1	372 bp	Unknown	Unknown	
			4005/8F2	434 bp	Unknown	LTR retrotransposable element , *Oryzias latipes* ($6X10^{-22}$)	HaFemale4F/R
			4005/8F3	410 bp	Unknown	Unknown	
EcoRI06/MseI15	460 6/15F	2	4606/15F	472 bp	Unknown	Unknown	
EcoRI11/MseI15	490 11/15F	3	49011/15F1	501 bp	Unknown	Unknown	
			49011/15F2	474 bp	Unknown	Unknown	

3.2.3 Development of sex-specific SCAR marker in *H. asinina*

Six (HaMale1 -6) and five (HaFemale1-5) sequence specific primer pairs were designed from male and female AFLP fragments, respectively (Table 3.5). Specificity of each pair of primers was tested against genomic DNA of male (N = 3) and female (N = 3) *H. asinina* individuals. The amplification product was analyzed by agarose gel electrophoresis.

Sex-specific SCAR markers were only considered from those were clearly present in males but were absent from female and *vice versa*. Among 11 primer pairs screened, the expected amplification product was observed from each marker except HaFemale1, which provided non-specific amplification products (Figure 3.14B). Additional minor amplification products were also observed along with the expected product when genomic DNA of *H. asinina* was amplified with HaMale6 (the expected 148 bp along with 425, 780 and 1400 bp fragments), HaFemale3 (the expected 175 bp along with 300 and 550 fragments, Figure 3.14C) and HaFemale4 (an expected 284 bp along with 300 and 350 bp fragments). Nevertheless, the amplification product was observed in both male and female *H. asinina* indicating the loss of specificity of original AFLP markers.

3.2.4 Characterization of SCAR markers in *H. asinina* by PCR followed by agarose gel electrophoresis and SSCP analysis

In vertebrates, sex determination markers usually represent single nucleotide polymorphism (SNP) between the same DNA fragment of males and females. To verify that positive amplification fragments of SCAR markers after PCR may still provide sex-specific SNP, SSCP was then applied.

The PCR products of HaM1-5 and HaF2-6 primers were denatured and electrophoresed through low crosslink non-denaturing polyacrylamide gels (37.5:1 or 75:1) (Figure 3.15). Results indicated that most of the markers (HaM1, HaM5, HaM6, HaF2, HaF3, HaF4, and HaF5) were polymorphic but their SSCP patterns/fragments were not related with sex determination in *H. asinina* (Table 3.6). Therefore, sex-specific SCAR markers derived from AFLP markers of *H. asinina* were not successfully developed at the genomic DNA level

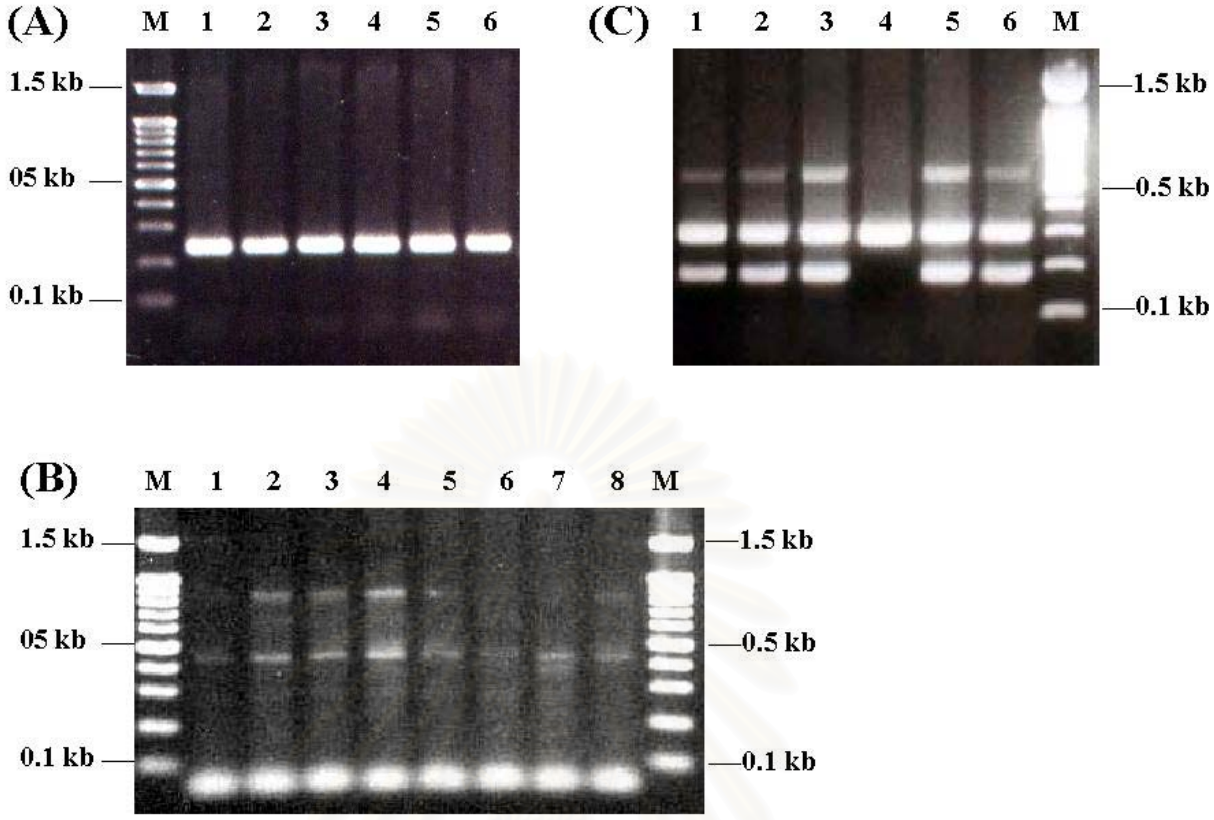


Figure 3.14 Agarose gel electrophoresis showing the PCR product of HaMale3 (A), HaFemale1 (B) and HaFemale3 (C) primers against male (lanes 1 – 3, A and C and lanes 1-4, B) and female (lanes 4 – 6, A and C and lanes 5-8, B) genomic DNA of *H. asinina*. A 100 bp ladder is used as the DNA markers.

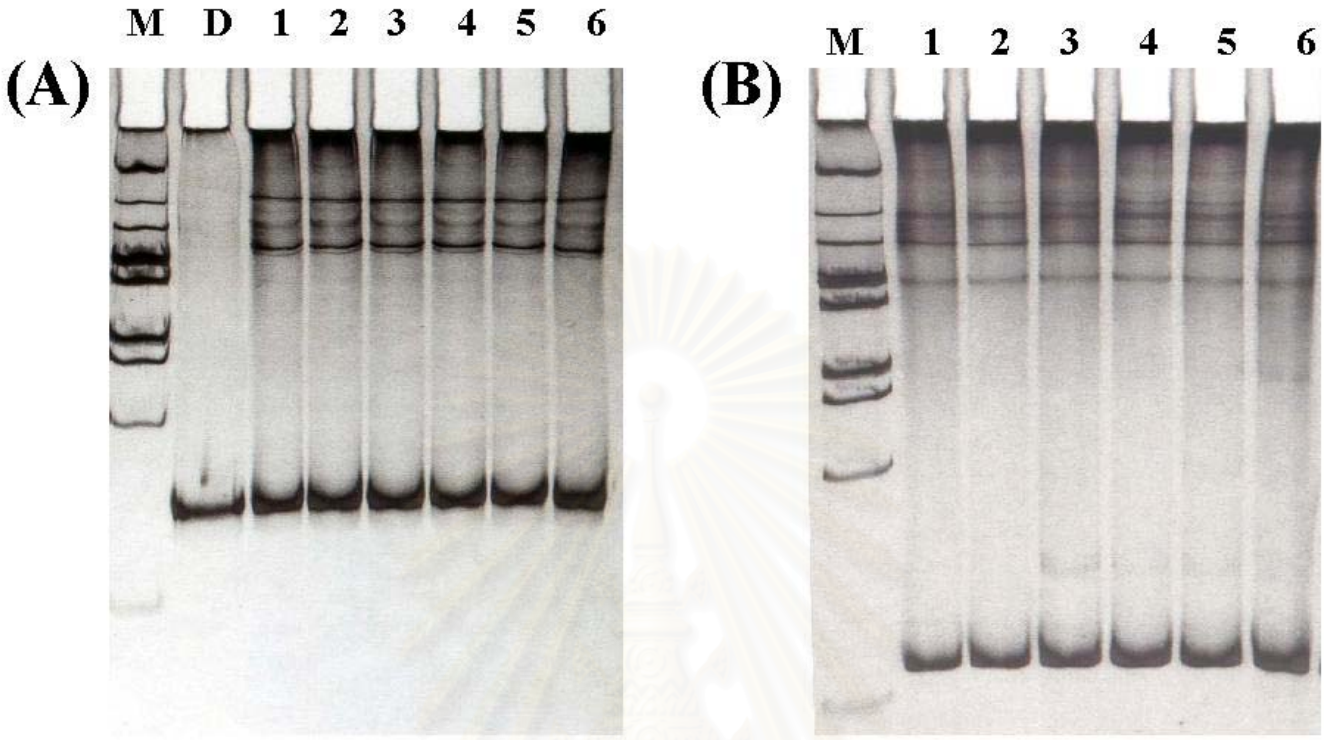


Figure 3.15 SSCP patterns of a SCAR marker derived from HaMale3 (A) and HaFemale5 (B) of male (lanes 1– 3) and female (lanes 4 – 6) genomic DNA of *H. asinina*. The PCR product was size-fractionated through native 15% PAGE (37:5:1) at xxx volt for xx hours. Lanes M and D are a 100 bp DNA marker and the non-denatured PCR product, respectively.

Table 3.6 Amplification results of SCAR markers in *H. asinina* followed by agarose gel electrophoresis and SSCP analysis

Primer	Expected product (bp)	PCR		SSCP	
		Male (*N*=3)	Female (*N*=3)	Female	Male
HaMale1-F/R	169 bp	169 bp	169 bp	Polymorphism	Polymorphism
HaMale2-F/R	187 bp	187 bp	187 bp	Monomorphism	Monomorphism
HaMale3-F/R	227 bp	227 bp	227 bp	Monomorphism	Monomorphism
HaMale4-F/R	166 bp	166 bp	166 bp	Monomorphism	Monomorphism
HaMale5-F/R	209 bp	209 bp	209 bp	Polymorphism	Polymorphism
HaMale6-F/R	148 bp	1400, 780, 425 and 148 bp	1400, 780, 425 and 148 bp	Polymorphism	Polymorphism
HaFemale1-F/R	200 bp	NS	NS	ND	ND
HaFemale2-F/R	230 bp	230 bp	230 bp	Polymorphism	Polymorphism
HaFemale3-F/R	175 bp	550, 300 and 166 bp	550, 300 and 166 bp	Polymorphism	Polymorphism
HaFemale4-F/R	284 bp	350, 300 and 284 bp	350, 300 and 280 bp	Polymorphism	Polymorphism
HaFemale5-F/R	213 bp	213 bp	213 bp	Polymorphism	Polymorphism

Abbreviations:, - = no amplification product; NS = non-specific amplification; ND = not determined.

3.3 Expressed sequence taq analysis of ovaries and testes from tropical abalone (*Haliotis asinina*)

3.3.1Total RNA and mRNA isolation

An average amount of total RNA isolated from testes, ovaries, and hemocytes was approximately around 40, 100 and 22 μg per individual The A_{260}/A_{280} ratio of total RNA was 1.7-2.0 indicating accepted quality of extracted total RNA. An ethidium bromide stained 1% agarose gel of total RNA from testes, ovaries and hemocytes revealed predominant; 28S and 18S ribosomal RNAs bands and low molecular weight RNAs along with smeared RNA with molecular sizes up to approximately 12 kb (Figure 3.16).

After purification, 2.2, 5.1 and 0.13 μg of mRNA were obtained from testes, ovaries and hemocytes of each abalone, respectively.

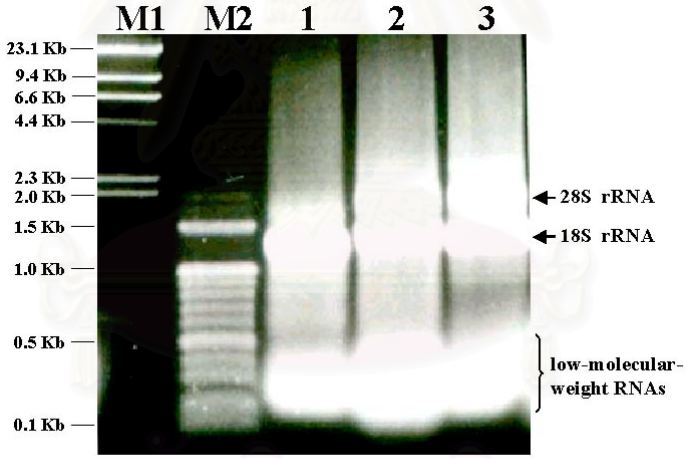


Figure 3.16 A 1.0% agarose gel electrophoresis of total RNA extracted from hemocyte (lane 1), testis (lane 2) and ovary (lane3) of *H. asinine*. Lanes M1 and M2 are a λ-*Hind* III DNA marker and a 100 bp DNA ladder, respectively.

3.3.2 Construction of normal cDNA library from H. asinina ovaries and testes

Three micrograms of purified mRNA from testes and ovaries were used to construct the typical cDNA library. Synthesized double strand cDNA (>500 bp) was sized fractionated by column chromatography (Figure 3.17). The 8^{th}- 13^{th} and the 7^{th}- 12^{th} fractions of cDNA from testes and ovaries were pooled and cloned into pSPORT1.

The insert sizes of recombinant clones were usually greater than 500 bp (Figure 3.18 and 3.19) as analyzed by 1.0% agarose gel electrophoresis. The percentage of recombinant clones having insert sizes > 500 bp was 65% and 80% for normal cDNA libraries of testes and ovaries, respectively.

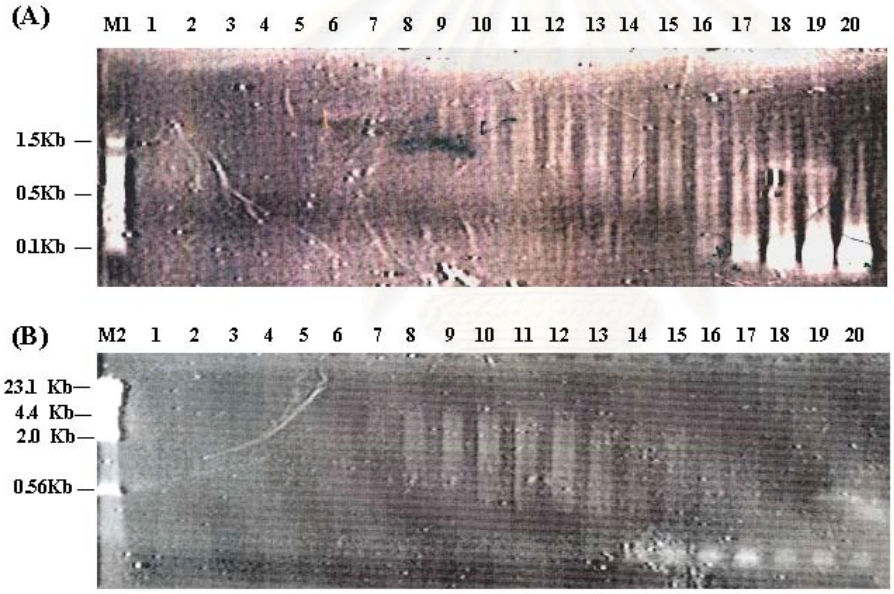


Figure 3.17 Size-fractionation of *Not* I-digested cDNA of testes (A) and ovaries (B) using column chromatography. Lanes 1- 20 are the effluent fraction no. 1 - 20. Lane M1 and M2 are a 100 bp DNA marker and a λ–*Hind* III DNA marker, respectively.

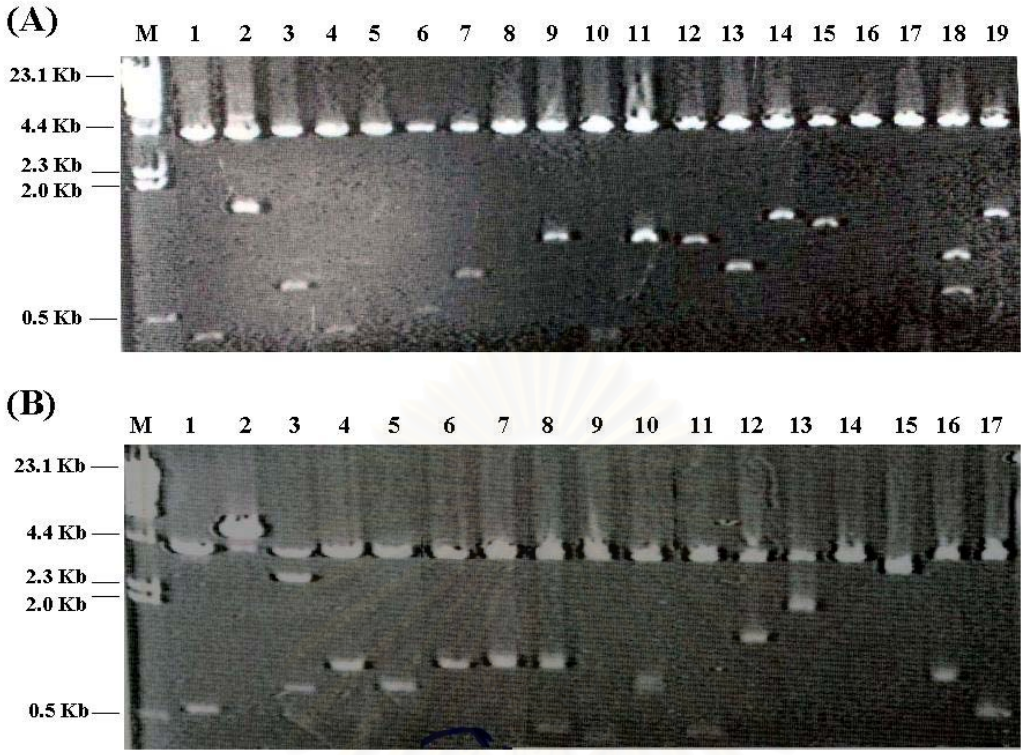


Figure 3.18 Determination of insert sizes of recombinant clones from the ovarian cDNA library by digestion of recombinant plasmids with *Not* I and *Sal* I (lanes 1-19, A) and *Sph* I and *Sal* I (lanes 1-17, B) and electrophoressed through 1.0% agarose gel electrophoresis. Lanes M are a λ-*Hind* III marker

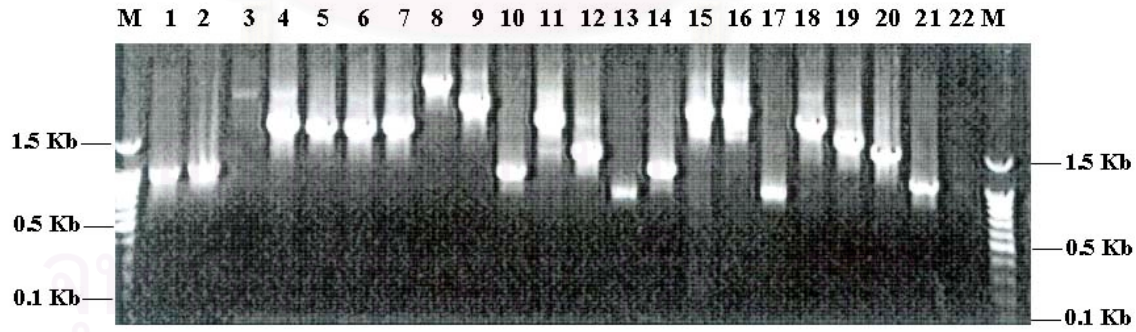


Figure 3.19 Determination of insert sizes of recombinant clones from the testis cDNA library using colony PCR analysis. The PCR product (insert + 326 bp) from each clone was electrophoresed through 1.0% agarose gel electrophoresis (lanes 1-22). Lanes M are a 100 bp DNA marker.

3.3.4 Construction of subtractive cDNA library from H. asinina ovaries and testes

Subtractive cDNA libraries were constructed using cDNA from ovaries as the tester and cDNA from testes as the driver and *vice versa*. As can be seen from Figure 3.20, the first and second PCR products indicated that forward and subtraction were only partially successful because the major products (bands) were still observed in both reactions.

The resulting forward (cDNA of ovaries as the tester) and reverse (cDNA of testes as the tester) subtracted cDNA fragments were ligated to pGEM-T Easy vector and transformed to *E. coli* JM109. The insert sizes of recombinant clones were determined by colony PCR (Figure 3.20). The percentage of recombinant clones of both libraries with insert sizes > 200 bp was approximately 70%.

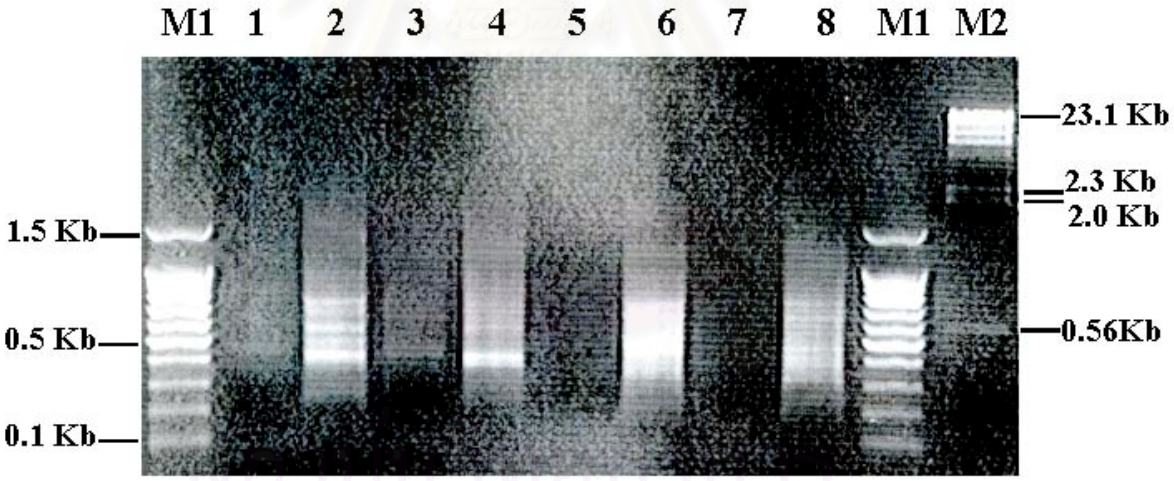


Figure 3.20 Electrophoresis of the first and the second PCR products of subtracted cDNA of testes (lane 1 and 2) and ovaries (lane 5 and 6) and unsubtracted cDNA of testes (lane 3 and 4) and ovaries (lane 7 and 8). Lane M1 and M2 are a 100 bp DNA ladder and a λ-*Hind* III, respectively.

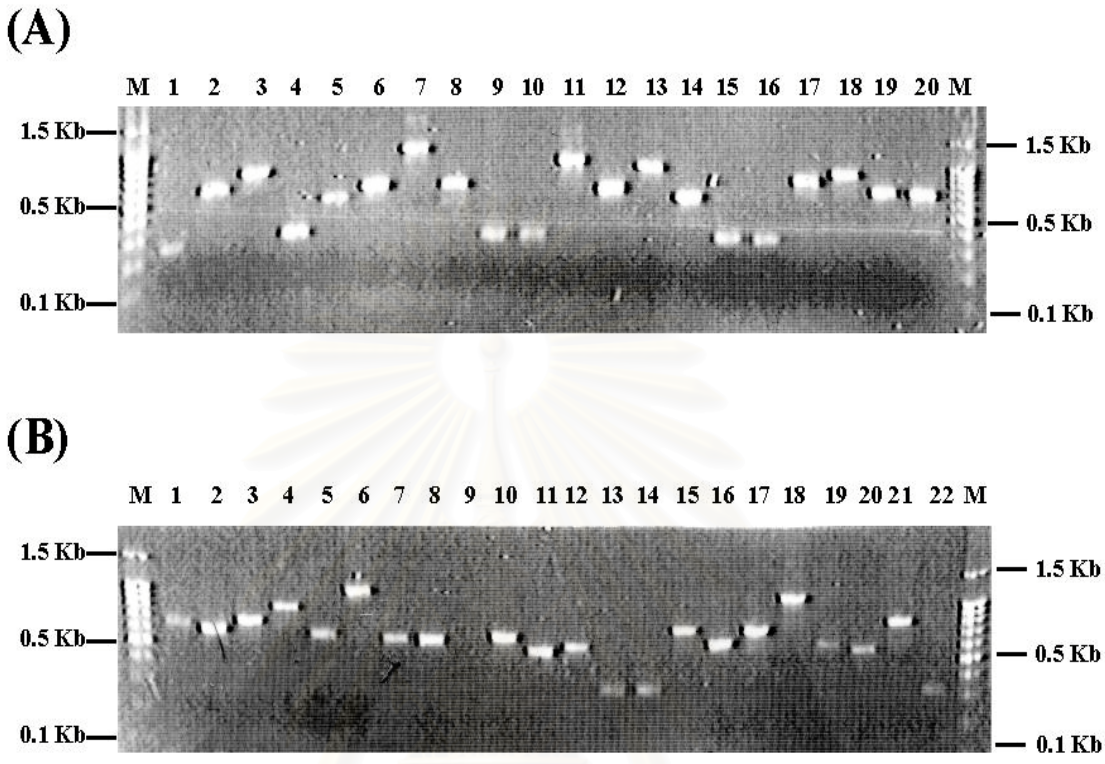


Figure 3.21 Determination of insert sizes of recombinant clones from subtractive cDNA library of ovaries (lanes 1-20,A) and testes (lanes 1 - 22, B) by colony PCR followed by agarose gel electrophoresis. Lanes M are a 100 bp DNA marker.

3.3.5 Nucleotide sequence analysis of EST clones

A total of 588 cDNAs from normal and subtractive libraries of ovaries (200 and 110 clones) and testes (118 and 160 clones) of *H. asinina* were unidirectional sequenced from 5′ direction of each clone for normal libraries and either 5′ or 3′ of each clone for subtractive libraries (Table 3.6).

Disregarding redundancy of sequenced clones, the total nucleotide length examined for normal and subtractive ovarian and testis cDNA libraries was 317 kb (134 for testis cDNAs and 183 kb for ovarian cDNAs). The average nucleotide sequences of cDNA insert are 669 and 707 bp from normal and 377 and 383 bp from subtractive cDNA of testes and ovaries, respectively.

The nucleotide sequences of each clone were analyzed by homology searches against data in the GenBank based on nucleotide similarity (BLAST*N*) and similarity of translated protein sequence (BLAST*X*).

For overall libraries, 309 ESTs (accounting for 52.6%) showed significant similarity with deposited sequences ($E<10^{-4}$) whereas 279 ESTs (accounting for 47.4%) did not match with any previously deposited gene in the GenBank ($E>10^{-4}$) and were regarded as unknown transcripts.

For normal cDNA libraries, 109 ESTs (59.5%) of ovaries and 73 ESTs (61.9%) of testes represented gene homologues whereas the remaining clones (40.5% for ovaries and 38.1% for testes) did not match to any sequence in the GenBank.

For subtractive cDNA libraries, 71 ESTs (64.5%) and 56 ESTs (35%) significant matched with deposited sequences in the GenBank whereas 39 ESTs (35.5%) and 104 ESTs (65%) of subtractive cDNA of ovaries and testes were unknown.

Table 3.6 Number of investigated EST clones from normal and subtractive cDNA libraries established from ovaries and testes of the tropical abalone (*H. asinina*)

cDNA library	No. of Clone	Homologues	Unknown
Normal ovary	200 clones	109 (59.5%)	91 (45.5%)
Normal testis	118 clones	73 (40.5%)	45 (38.1%)
Subtractive ovary	110 clones	71 (64.5%)	39 (35.5%)
Subtractive testis	160 clones	56 (35.5%)	104 (65.0%)
Total	**588 clones**	**309 (52.6%)**	**279 (47.4%)**

Matched ESTs found in this thesis were classified into seven broad categories: (1) sex-related genes, (2) stress response, detoxification, and cell defense proteins, (3) protein synthesis and ribosomal protein, (4) cell division and DNA replication, (5) metabolism, (6) unidentified function, and (7) unknown on the basis of their putative functions of the predicted proteins. Distribution of identified ESTs for different categories in normal and subtractive cDNA libraries of ovaries and testes is summarized in Table 3.7.

Table 3.7 Numbers and functional categories of ESTs from normal and subtractive cDNA libraries of the tropical abalone (*Haliotis asinina*)

Category	**No. of clone (%)**			
	Normal ovarian cDNA library	**Normal testis cDNA library**	**Subtractive ovarian cDNA library**	**Subtractive testis cDNA library**
Sex-related genes	43 (21.5)	11 (9.3)	40 (36.4)	20 (12.5)
Stress response, detoxification and cell defense proteins	18 (9.0)	8 (6.8)	1 (0.9)	0 (0.0)
Protein synthesis and ribosomal proteins	31 (16.5)	16 (13.6)	5 (4.5)	10 (6.25)
Cell division and DNA replication	0 (0)	16 (13.6)	8 (7.3)	11 (6.9)
Metabolisms	9 (4.5)	10 (8.5)	11 (10.0)	11 (6.9)
Unidentified function	8 (4.0)	1 (0.85)	26 (23.6)	4 (2.5)
Unknown	91 (40.5)	45 (38.1)	39 (35.5)	104 (65.0)
Total	**200**	**118**	**110**	**160**

(1) Sex-related genes

In normal cDNA library, 43 (21.5%) and 11 (9.3%) analyzed ESTs of ovary and testis libraries were classified as cDNA homologues of sex-related transcripts. Nevertheless, they represented only 4 different genes; vitelline coat protein (VCP) 41, VCP42; vitelline envelope sperm for lysin receptor (VERL) and hydroxysteriod

dehydrogenase -like protein (HSD) for cDNA from ovaries and 4 additional genes; sperm lysin, gonadotropin inducible ovarian transcription factor 1 (GIOT1), axonemal p66.0 and tektinA1 for cDNA from testes. The major EST encoding putative proteins in this category were vitelline coat protein 41 (38 EST clones, 19.0%) and sperm lysin (9 EST clones, 7.6%), respectively (Table 3.8).

In subtractive cDNA library, 40 (36.4%) and 20 (12.5%) ESTs of ovary and testis cDNA represented 5 different sex-related gene homologues including VCP 41, VCP42, VERL, vitellogenin-1 (VTG-1) and a disintegrin and metalloproteinase with thrombospondin motifs 9 (ADAMTS-9) and 3 different sex-related gene homologues including sperm lysin, fertilization protein (FP) and small androgen receptor-interacting protein (SARIP), respectively (Table 3.7). VCP41 (21 EST clones, 19%) and VERL (15 EST clones, 13.6%) were the most and the second abundant transcripts for a subtractive cDNA library of ovaries. On the other hand, sperm lysin (7 EST clones, 4.4%) was the most abundant transcript in the subtractive cDNA libraries of testes. Two EST representing a FP homologue were also found.

The full-length cDNA of sperm lysin was obtained in this study. Two ESTs (ST162 and ST124) represent a homologue of FP of *H. rufescens* were found in the subtractive cDNA library of testes. The abalone spermatozoan has a large acrosomal vesicle containing two major proteins (a 16 kDa sperm lysin) and (a 18 kDa fertilization protein). Abalone sperm lysin creates a hole in the egg vitelline envelope species-specifically and binds tightly to VERL which the sperm passes to reach the egg cell membrane. FP coats the sperm acrosomal process and probably mediate fusion of the two gametes.

A homologues of axonemal p66.0 (PT004; E-value 1 $x10^{-21}$) and tektin A1 (PT108; E-value 7 x 10^{-81}) proteins of *Ciona intestinalis* was obtained. Axonemal protein is associated with axonemal structures (microtubule-based structures) of sperm tail. Tektin is a group of proteins that form filamentous polymers in the wall of ciliary and fragellar microtubules and has biochemical and immunological properties similar to those of intermediate-filament proteins.

PT090 and ST043 clones from normal and subtractive testis cDNA libraries showed significant with GIOT1. The nucleotide sequences of these clones were highly similar with GIOT1 (E-value 7 x 10^{-30}) and SARIP (E-value 2 x 10^{-19}) of Norway rat

(*Rattus norregicus*). Gonadotropins are essential for ovarian follicular development and differentiation. In rat, (Cys)(2)-(His)(2)-type zinc finger protein family GIOT1 restrictively expressed in the pituitary gland, adrenal, testis and ovary may function as a novel transcriptional repressor. SARIP is a protein that interacting with the receptor of androgen and may play the significant role in spermatogenesis and testicular development.

A large number of the VCP ESTs and VERL ESTs from normal (40 clones and 2 clones) and subtractive (23 clones and 15 clones) cDNA libraries of ovaries were found in this study could be gathered into several VCP clusters. The vitelline coat proteins are extracellular glycoprotein coats that are deposited around egg and perform several functions that relate to fertilization and interaction of sperm and egg. VERL is a rod-like molecule, unbranched glycoprotein of vitelline envelope. To create the hole of egg, sperm lysin binds to VERL species-specifically, causes the VERL molecules to lose cohesion and splay apart creating the hole.

SO079 from the subtractive cDNA library of ovaries was homologue of vitellogenin-1 (VTG-1) of brown-winged green bug (*Plautia stali*, E-value 4 x 10^{-05}). VTG is a large phospholipoglycoprotein (200-700 kDa) and is the principal precursor of egg-yolk proteins crucial for successful embryonic and larval development. In numerous egg-laying animals, VTG is secreted into the bloodstream by the intestines of nematodes, fat bodies of insects, and livers of vertebrate.

SO006 and PO002 were significant sequence similarity to the ADAMTS-9 precursor (E-value 4 x 10^{-05}) and hydroxysteriod dehydrogenase-like protein (HSD, E-value 1 x 10^{-83}) of *Homo sapiens*. ADAMTS-9 is a member of ADAM-TS/metallospondin genes encodes a new family of proteins with structural homology to ADAM metalloprotease-disintegrin family. ADAMTS9 has a metalloprotease domain, a disintegrin-like domain, one internal thrombospondin 1 (TSP1) motif, and three carboxyl-terminal TSP1-like submotifs. Human ADAMTS-9 is orthologous of abnormal gonad development protein (GON-1), an ADAMTs protease required for gonadal morphogenesis in *Caenorhabditis elegans.*

HSDs are responsible for the biosynthesis of steroid hormones and play a crucial role in mammalian physiology and development. 17β-hydroxysteriod dehydrogenase is a steroidogenic enzyme which is essential for invertebrate spermatogenesis (Kho, 2004).

(2) Stress response, detoxification, and cell defense proteins

Twenty-seven ESTs; 18 clones (9.0%), 8 clones (6.8%), and 1 clone (0.9%) from the normal ovarian and testis libraries and the subtractive ovarian cDNA libraries, were classified as homologues putative stress response, detoxification, and cell defense proteins. Transcripts representing of the gene categories were not found in the subtractive cDNA library of testis.

Fifteen different gene homologues were found and categorized in 3 different subgroups (Table 3.9). The first subgroup was those representing stress response genes. Transcripts in this group were composed of stress-inducible chaperone mitochondria-GrpE#1 (GrpE-1), chaperonin containing t-complex polypeptide 1 (TCP-1), importin 9 and heat shock protein 70 (HSP70). The second subgroup was composed of thiosulfate sulfurtransferase, which plays a central role in detoxification of cyanide. The last subgroup was transcripts involving with cell defense proteins. They were tumor-specific transplantation antigen P198 homolog p23 (Tum-P198), IgE-dependent histamine release factor (IgE-HRF), translationally controlled tumor protein homolog (TCTP), variable region-containing chitin-binding protein 5, cysteine proteinase 1, T-cell activation WD repeat protein, proteasome beta 6, proteasome subunit LMPX, ubiquitin, high mobility group protein 1 (HMG1), FK506-binding protein12 (FKBP12) and HLA-B associated transcript 1 (BAT1), respectively.

Nucleotide sequences of PT007 and PT122 clones obtained from normal the cDNA library of testes showed significant similarity with HSP70 of *Danio rerio* (E-value 5 x 10^{-36}) and to GrpE-1 of *Rattus norvegicus* (E-value 7 x 10^{-51}). HSP70 play a vital role in mediating the folding of newly synthesized proteins and the refolding of denatured proteins. The mode of action involves binding to expose hydrophobic regions of nascent proteins, thereby facilitating the folding of the protein to its native tertiary structure. HSP70s are involved in combating cellular stress. Accordingly, they are found in almost all compartments of the cells. GrpE has also been shown to play a role as a nucleotide exchange factor for the stimulation of HSP70-ATPase activity.

Table 3.8 Sex-related transcripts found in normal and subtractive cDNA libraries of ovaries and testes of *H. asinina*

Transcript	Closest species	Accession	E-value	Length (bp)	No. of clones
Normal cDNA from ovaries					
Vitelline coat protein (VCP) 41	*Tegula pfeifferi*	BAB15843	4×10^{-04} - 1×10^{-60}	531-1090	38
VCP42	*Tegula pfeifferi*	BAB15930	7×10^{-24} – 2×10^{-30}	711-715	2
Vitelline envelope receptor for sperm lysin (VERL)	*Haliotis rufescens*	AAL50827	2×10^{-16} – 4×10^{-20}	547-747	2
Hydroxysteriod dehydrogenase –like protein (HSD)	*Homo sapiens*	AAM14670	1×10^{-83}	925	1
Normal cDNA from testes					
Sperm lysine	*Haliotis ovina*	L26276	3×10^{-52} – 3×10^{-81}	645-745	9
Gonadotropin inducible ovarian transcription factor 1 (GIOT1)	*Rattus norvegicus*	XP_230906	7×10^{-30}	745	1
Axonemal p66.0	*Ciona intestinalis*	BAB88833	1×10^{-21}	503	1
Tektin A1	*Ciona intestinalis*	BAB86298	7×10^{-81}	782	1
Subtractive cDNA from ovaries					
VCP41	*Tegula pfeifferi*	BAB15843	4×10^{-07} – 5×10^{-22}	223-665	21
VCP42	*Tegula pfeifferi*	BAB15930	5×10^{-07} – 2×10^{-17}	381-403	2
VERL	*Haliotis rufescens*	AAL50827	$1x\ 10^{-11}$- 2×10^{-23}	298-465	15
Vitellogenin-1 (VTG-1)	*Plautia stali*	BAA88075	4×10^{-05}	350	1
ADAMTS-9	*Homo sapiens*	Q9P2N4	4×10^{-05}	284	1
Subtractive cDNA from testes					
Sperm lysine	*Haliotis ovina*	L26276	1×10^{-12} – 6×10^{-51}	194-386	17
Fertilization protein (FP)	*Haliotis rufescens*	AAC37229	7×10^{-06}	438	2
Small androgen receptor-interacting protein (SARIP)	*Rattus norvegicus*	NM147146	2×10^{-19}	343	1

Table 3.9 Transcripts representing homologues of stress response, detoxification, and cell defense proteins found in normal and subtractive cDNA libraries of ovaries and testes of *H. asinina*

Transcript	Closest species	Accession	E-value	Length (bp)	No. of clones
Normal cDNA from ovaries					
Stress response					
Stress-inducible chaperone mitochondria-GrpE-1	*Rattus norvegicus*	NP_077813	7×10^{-51}	752	1
T-complex protein 1 theta subunit (TCT-1)	*Homo sapiens*	CAB90433	2×10^{-10}	672	1
Detoxification					
Thiosulfate sulfurtransferase	*Gallus gallus*	A37209	5×10^{-26}	954	1
Cell defense proteins					
Tumor-specific transplantation antigen P198 homolog p23 (Tum-P198)	*Bos taurus*	A44367	1×10^{-35}-3×10^{-58}	606-716	4
IgE-dependent histamine release factor	*Dermacentor variabilis*	AAL75585	4×10^{-11}-4×10^{-13}	662-729	2
Translationally controlled tumor protein homolog (TCTP)	*Drosophila melanogaster*	Q9VGS2	1×10^{-08}	861	1
Cysteine proteinase 1	*Dictyostelium discoideum*	P04988	2×10^{-14}	555	1
Variable region-containing chitin-binding protein 5	*Branchiostoma floridae*	AAN62911	3×10^{-07}	492	1
Proteasome beta 6 subunit	*Homo sapiens*	NP_002789	5×10^{-85}	869	1
LMPX	*Myxine glutinosa*	BAA10931	3×10^{-64}-2×10^{-80}	748-780	2
FKBP12	*Schizosaccharomyces pombe*	CAB46710	2×10^{-35}	700	1
Ubiquitin	*Biomphalaria glabrata*	AAG49553	3×10^{-65}	485	1
T-cell activation WD repeat protein	*Homo sapiens*	AAM43838	9×10^{-40}	725	1

Table 3.9 (cont.)

Transcript	Closest species	Accession	E-value	Length (bp)	No. of clones
Normal cDNA from testes					
Stress response					
Heat shock protein 70 (HSP70)	*Danio rerio*	AAH48063	5×10^{-36}	550	1
Importin 9	*Homo sapiens*	NP_060555	1×10^{-04}	763	1
Cell defense proteins					
HLA-B associated transcript 1	*Homo sapiens*	NP_004631	1×10^{-75}- 1×10^{-106}	759-789	2
Tum-P198	*Bos taurus*	A44367	3×10^{-64}- 4×10^{-69}	670-671	2
FKBP12	*Schizosaccharomyces pombe*	CAB46710	6×10^{-32}	668	1
High mobility group protein 1	*Biomphalaria glabrata*	AAN31640	3×10^{-48}	702	1
Subtractive cDNA from ovaries					
Stress response					
TCT-1	*Danio rerio*	AAD48818	4×10^{-50}	392	1

(3) Protein synthesis and ribosomal proteins

A total of 31 (16.5%), 16 (13.6%), 5 (4.5%), and 10 (6.25%) ESTs of normal and subtractive cDNA libraries cDNA libraries of ovaries and testes were identified. These EST represented 36 gene homologues categorized to protein synthesis and ribosomal proteins (Table 3.10).

Major transcripts representing putative proteins of this function category were those of ribosomal protein (55 ESTs). These included ESTs of 60S large subunit ribosomal proteins (33 clones), 40S small subunit ribosomal proteins (19 clones), 16S mitochondrial rRNA (2 clones) and 12S mitochondrial rRNA (1 clone). Other transcripts were ESTs of of eukaryotic translation initiation factor 2B (eIF-2B), EF-1β and EF-1δ, 34/67 kD laminin binding protein, cyclin B, RNA-binding protein and alanyl tRNA synthetase, respectively.

Table 3.10 Transcripts representing homologues of protein synthesis and ribosomal proteins found in normal and subtractive cDNA libraries of ovaries and testes of *H. asinina*

Transcript	Closest species	Accession	E-value	Length (bp)	No. of clones
Normal cDNA from ovaries					
Ribosomal protein L4	*Urechis caupo*	T12048	4×10^{-60}	675	1
Ribosomal protein L7	*Crassostrea gigas*	CAD89885	6×10^{-70}	698	1
Ribosomal protein L7	*Argopecten irradians*	AAN05591	6×10^{-70}	633	1
Ribosomal protein L7a	*Takifugu rubripes*	CAA75444	9×10^{-77}	974	1
Ribosomal protein L9	*Argopecten irradians*	AAN05606	6×10^{-67}	668	1
Ribosomal protein L10A	*Spodoptera frugiperda*	AAK76990	5×10^{-19}-5×10^{-79}	533-734	2
Ribosomal protein L10a	*Xenopus laevis*	AAH41308	9×10^{-75}	732	1
Ribosomal protein L11	*Argopecten irradians*	AAN05587	3×10^{-82}	638	1
Ribosomal protein L13	*Scyliorhinus canicula*	AAN73374	2×10^{-39}	678	1
Ribosomal protein L21	*Argopecten irradians*	AAN05604	1×10^{-71}	565	1
Ribosomal protein L22	*Argopecten irradians*	AAN05585	1×10^{-40}	524	1
Ribosomal protein S3	*Homo sapiens*	AAB19349	9×10^{-95}	783	1
Ribosomal protein S19	*Mya arenaria*	AAB09536	4×10^{-54}	538	1
Ribosomal protein S21	*Homo sapiens*	BAB79481	5×10^{-04}	411	1
Ribosomal protein S28	*Branchiostoma belcheri tsingtaunese*	AAP21778	5×10^{-10}	318	1
Ribosomal protein S28	*Branchiostoma belcheri tsingtaunese*	AAP21778	2×10^{-14}	363	1
40S ribosomal protein	*Perinereis aibuhitensis*	AAO43049	1×10^{-60}	508	1
40S ribosomal protein S2	*Urechis caupo*	P49154	1×10^{-97}	878	1
40S ribosomal protein S11	*Danio rerio*	AAH46054	2×10^{-63}	563	1
40S ribosomal protein S15 (RIG PROTEIN)	*African clawed frog*	P20342	1×10^{-12}	527	1
16S ribosomal RNA gene	*Haliotis diversicolor*	U51989.1	6×10^{-42}	603	1
16S ribosomal RNA gene	*Haliotis asinina*	AY163259	0.0	385	2
RNA-binding protein KIAA0117 (HAL845)	*Rattus norvegicus*	XP_234003	3×10^{-09}	925	1
Eukaryotic translation initiation factor 2B, subunit 5 epsilon	*Mus musculus*	AAH38620	4×10^{-89}	930	1
Eukaryotic translation elongation factor 1 beta 2	*Mus musculus*	AAH39635	7×10^{-59}	774	1
Laminin binding protein 34/67 kD		AAA90978	2×10^{-51}	602	1
Cyclin B	*Patella vulgata*	CAA41255	1×10^{-99}	965	1
Cyclin B	*Marthasterias glacialis*	CAA34624.1	9×10^{-25}	508	1
Elongation factor 1 delta	*Bombyx mori*	BAB21109.1	5×10^{-34}	822	1

Table 3.10 (cont.)

Transcript	Closest species	Accession	E-value	Length (bp)	No. of clones
Normal cDNA from testes					
Ribosomal protein L41	*Cyprinus carpio*	AY117540	9×10^{-05}	766	1
Ribosomal protein L7	*Crassostrea gigas*	CAD89885	9×10^{-49}	620	1
Ribosomal protein L36A	*Spodoptera frugiperda*	AAK92170	1×10^{-17}	420	1
Ribosomal protein S12	*Branchiostoma belcheri*	AF548331	1×10^{-36}	552	1
Ribosomal protein L36	*Drosophila melanogaster*	NP_476629	3×10^{-21}	421	1
Ribosomal protein S10	*Branchiostoma belcheri*	AAO31776	1×10^{-61}	622	1
Ribosomal protein S24	*Rattus norvegicus*	XP_235376	1×10^{-26}	529	1
Ribosomal protein S25	*Branchiostoma belcheri*	AAN52391	3×10^{-32}	423	1
60S acidic ribosomal protein P0 (L10E)		P47826	3×10^{-76}	814	1
60S Ribosomal protein L17 (L23)	*Mus musculus*	XP_127004	2×10^{-34}	562	1
28S large subunit ribosomal RNA gene	*Aplysia californica*	AY026366	4×10^{-96}	694	1
16S ribosomal RNA gene	*Haliotis asinina*	AY163259	0.0	719	1
16S ribosomal RNA gene	*Haliotis diversicolor*	U51989	1×10^{-37}-1×10^{-40}	635-756	3
12S ribosomal RNA gene	*Ocinebrellus inornatus*	AY148653	3×10^{-04}	716	1
Subtractive cDNA from ovaries					
Ribosomal protein L18	*Branchiostoma lanceolatum*	AAN73381	2×10^{-56}	410	2
Ribosomal protein L19	*Argopecten irradians*	AF526203	8×10^{-21}	238	1
40S ribosomal protein S9	*Ictalurus punctatus*	AAK95191	7×10^{-64}	442	1
Alanyl tRNAsynthetase	*Mesocricetus auratus*	BAC44844	2×10^{-21}	531	1
Subtractive cDNA from testes					
40S ribosomal proteinS26-2	*Ictalurus punctatus*	AF402835	3×10^{-21}	239	1
23S and 5S rRNA	*Calanus finmarchicus*	X06056	2×10^{-05}-3×10^{-06}	194-200	2
28S rRNA gene	*Salvelinus namaycush*	U1796	6×10^{-26}	344	1
28S rRNA gene	*Hydrolagus colliei*	AF061799	9×10^{-19}-1×10^{-36}	351-378	3
28S rRNA gene	*Episiphon yamakawai*	AB103133	1×10^{-112}-1×10^{-116}	328	3

(4) Cell division/DNA synthesis, repair and replication

Sixteen (13.6%), eight (7.3%) and eleven (6.9%) ESTs of normal cDNA libraries of testis and subtractive cDNA libraries of ovary and testis were classified as homologues of genes involving with cell division/DNA synthesis, repair and replication (Table 3.11).

The major ESTs of putative protein in this category were groups of tubulin genes (11 clones; α-tubulin 1, α-tubulin 2 and β-tubulin) and histone proteins (10 clones; histone H1 and H2A). Other were ESTs of nucleotide translocator, meiosis specific nuclear structure protein, nucleolar protein family A member 2, neural wiskott-aldrich syndrome protein (N-WASP), transport protein sec61 gamma subunit, alternative splicing factor ASF-2, G2/Mitotic-specific cyclin A, mitotic apparatus protein P62, Myc homolog and polyadenylate binding protein, respectively.

The α-tubulin 2 was found in the subtractive cDNA library of ovaries whereas α-tubulin 1 and β-tubulin were found in the normal and subtractive cDNA libraries of testes.

Table 3.11 Transcripts representing homologues of proteins involving with cell division/DNA synthesis, repair and replication found in normal and subtractive cDNA libraries of ovaries and testes of *H. asinina*

Transcript	Closest species	Accession	E-value	Length (bp)	No. of clones
Normal cDNA from testis					
Histone 1	*Mytilus edulis*	CAA11816	3×10^{-19}	527	1
Histone H1	*Mytilus chilensis*	CAC94905	2×10^{-19}	760	1
Histone H1	*Mytilus chilensis*	CAC94905	1×10^{-14}	758	1
Histone H1	*Mytilus californianus*	CAC94904	8×10^{-24}	743	1
Alternative splicing factor ASF-2	*Homo sapiens*	B40040	8×10^{-06}	307	1
Nucleotide translocator	*Arabidopsis thaliana*	1908224A	3×10^{-57}	727	1
Meiosis specific nuclear structure protein	*Astropecten brasiliensis*	CAC18535	5×10^{-30}	806	1
Nucleolar protein family A member 2	*Branchiostoma belcheri tsingtaunese*	AAN86977	5×10^{-39}	781	1
α-tubulin 1	*Meriones unguiculatus*	AF052694	7×10^{-91}	555	1
α-tubulin 1	*Urechis caupo*	U30467	1×10^{-133}	725	1

Table 3.11 (cont.)

Transcript	Closest species	Accession	E-value	Length (bp)	No. of clones
β-tubulin	*Paracentrotus lividus*	X15389	4 x 10^{-56}- 1 x 10^{-102}	707-754	2
β-tubulin	*Halocynthia roretzi*	BAA22382	7 x 10^{-77}	557	1
Neural Wiskott-Aldrich syndrome protein	*Rattus norvegicus*	O08816	2 x 10^{-27}	733	1
Transport protein sec61 gamma subunit	*Ciona intestinalis*	CAC82549	1 x 10^{-22}	560	1
Subtractive cDNA from ovaries					
G2/ Mitotic-specific clclin A	*Chlorohydra viridissima*	P51986	1 x 10^{-24}	491	1
Eukaryotic translation elongation factor 1 beta 2	*Homo sapiens*	AAP36790	4 x 10^{-37}	428	1
Peptide elongation factor1-beta	*Gallus gallus*	AAD16874	1 x 10^{-33}	431	2
Mitotic apparatus protein P62	*Lytechinus pictus*	P91753	4 x 10^{-04}	216	2
Myc homolog	*Crassostrea virginica*	AAB34577	2 x 10^{-04}	616	2
α-tubulin 2	*Aplysia californica*	AF481056	1 x 10^{-66}-8 x 10^{-71}	321	2
Subtractive cDNA from testes					
Histone H1	*Mytilus californianus*	CAC94904	1 x 10^{-07}	636	1
Histone H1	*Mytilus chilensis*	CAC9490	1 x 10^{-07}-5 x 10^{-11}	635-637	2
Histone 1	*Mytilus edulis*	CAA11816	7 x 10^{-11}	636	1
Histone H2A	*Mytilus edulis*	CAD37821	4 x 10^{-35}	626	1
Histone H2A	*Sea urchin*	X05547	3 x 10^{-20}	247	1
Polyadenylate binding protein	*Petromyzon marinus*	AAB88449	3 x 10^{-31}	558	1
α-tubulin 1	*Aplysia californica*	AF481055	2 x 10^{-37}-5 x 10^{-40}	171-200	2
α-tubulin 1	*Urechis caupo*	U30467	1 x 10^{-104}- 1 x 10^{-165}	530-697	1

(5) Metabolism

Forty-one ESTs; 9 clones (4.5%) of the normal cDNA library of ovaries, 10 clones (8.5%) of the normal cDNA library of testis, 11 clones (10.0%) of the subtractive cDNA library of ovaries and 11 clones (6.9%) of the subtractive cDNA library, which are homologous with gene involving with metabolism were identified. They represented 18 different putative genes in this functional category (Table 3.12).

These EST included transcripts of ATP synthase, ADP/ATP carrier, ATP-binding protein, fatty acid synthase, mitochondrial ATPase inhibitor, ligase family member 60.3 kD, solute carrier family 25, arginine kinase, protein tyrosine phosphatase, NADH dehydrogenase subunit II, isocitrate dehydrogenase, fructosamine 3 kinase, α-methylacyl-CoA racemase, carnitine palmitoyltransferase 1 alpha, mitochondrial long-chain enoyl-CoA hydratase/3-hydroxycyl-CoA dehydrogenase alpha-subunit, UDP-Gal:betaGlcNAc beta 1,3-galactosyltransferase, cytochrome c oxidase subunit IV, cytochrome c oxidase subunit III and cytochrome b. The major transcripts of ESTs in this category were arginine kinase (16 clones), and mitochondrial NADH dehydrogenase subunit II (4 clones).

Table 3.12 Transcripts representing homologues of proteins involving with metabolism found in normal and subtractive cDNA libraries of ovaries and testes of *H. asinina*

Transcript	Closest species	Accession	E-value	Length (bp)	No. of clones
Normal cDNA from ovaries					
ADP/ATP carrier	*Trypanosoma brucei brucei*	AAC23561	4×10^{-44}	923	1
Solute carrier family 25	*Homo sapiens*	AAH01328	1×10^{-119}	955	1
ATP-binding protein	*Ralstonia solanacearum*	CAD14642	7×10^{-25}	715	1
NADH dehydrogenase subunit 2	*Loligo bleekeri*	NP_062838	3×10^{-31}	877	1
Mitochondrial ATPase inhibitor	*Caenorhabditis elegans*	AAG49389	2×10^{-13}	669	1
Cytochrome c oxidase subunit III	*Melanocetus murrayi*	BAC23352	1×10^{-44}	665	1
Arginine kinase	*Nordotis madaka*	BAA05100	1×10^{-68}	650	1
Mitochondrial long-chain enoyl-CoA hydratase/3-hydroxycyl-CoA dehydrogenase alpha-subunit	*Rattus norvegicus*	BAA03939	9×10^{-75}	785	1
Tyrosine phosphatase	*Rattus norvegicus*	I49365	8×10^{-45}	666	1

Table 3.12 (cont.)

Transcript	Closest species	Accession	E-value	Length (bp)	No. of clones
Normal cDNA from testes					
ATP synthase	*Mus musculus*	NP_663402	2×10^{-71}	802	1
ADP/ATP carrier	*Neocallimastix patriciarum*	AAL79525	3×10^{-48}	818	1
NADH dehydrogenase subunit 2	*Loligo bleekeri*	NP_062838	2×10^{-27}	729	2
NADH dehydrogenase subunit 2	*Heloderma suspectum*	AAD51498	4×10^{-12}	548	1
Cytochrome b	*Phlebotomus papatasi*	AAD45005	8×10^{-60}	628	1
Arginine kinase	*N. madaka*	D26104	1×10^{-130}	800	1
Fructosamine 3 kinase	*Mus musculus*	NP_071297	5×10^{-09}	850	1
α-methylacyl-CoA racemase	*Homo sapiens*	Q9UHK6	2×10^{-52}	726	1
Carnitine palmitoyltransferase 1	*Rattus norvegicus*	NP_113747	3×10^{-68}	729	1
Subtractive cDNA from ovaries					
ATP synthase	*Mus musculus*	BC003854	2×10^{-29}	446	1
Arginine kinase	*Corbicula japonica*	BAB91357	2×10^{-22}	698	1
Arginine kinase	*Haliotis madaka*	P51544	4×10^{-64}-4×10^{-68}	705	3
Isocitrate dehydrogenase (NADP)	*Homo sapiens*	S57499	4×10^{-76}	489	2
UDP-Gal:betaGalNAc beta 1,3-galactosyltransferase, polypeptide 4	*Mus musculus*	NP_062293	3×10^{-17}	788	1
Fatty acid synthase	*Gallus gallus (chicken)*	AAA48767	8×10^{-23}	202	1
Ligase family member	*Caenorhabditis elegans*	NP_505451	1×10^{-21}-3×10^{-22}	378	2
Subtractive cDNA from testes					
Arginine kinase	*Haliotis madaka*	D26104	2×10^{-53}-1×10^{-106}	367-522	10
Cytochrome C oxidase subunit IV	*Urechis caupo*	AAA7439	9×10^{-09}	210	1

(6) Unidentified function

A total of 8 (4.0%) and 12 (10.2%) ESTs of normal cDNA libraries of ovaries and testes and 6 (5.5%) and 4 (2.5%) ESTs of subtractive cDNA libraries of ovaries and testes were regarded as homologous genes which non-identified functions (Table 3.13).

Hypothetical proteins (or unnamed proteins) are the largest group in this category. In a total, 15 ESTs of hypothetical proteins are found. Other major were gene products of *Anopheles gambiae* (ENSANGP, 6 clones) and *Drosophila melanogaster* (agCP, 2 clones). Four microsatellite (0.68%) containing clones were found across libraries.

(7) Unknown genes

Unmatched cDNA sequences of the normal and subtractive cDNA libraries of ovaries and testes were 91 clones (45.5%), 39 clones (35.5%), 45 clones (38.1%), and 104 clones (65.0%), respectively. The ORFs of unmatched ESTs having sizes greater than 300 bp in length were searched using the Genetyx-Win software and revealed that 50 and 15 ESTs of the normal and subtractive cDNA of ovaries and 20 and 20 ESTs of the normal and subtractive cDNA of testes containing ORFs of the genes.

Table 3.13 ESTs with unidentified function found in normal and subtractive cDNA libraries of ovaries and testes of *H. asinina*

Transcript	Closest species	Accession	E-value	Length (bp)	No. of clones
Normal cDNA from ovaries					
Hypothetical protein	*Ralstonia metallidurans*	ZP_00021518	7×10^{-38}	603	1
agCP3870	*Anopheles gambiae*	EAA04480	7×10^{-11}	528	1
agCP12540	*Anopheles gambiae*	EAA01351	1×10^{-21}	330	1
ENSANGP00000014551	*Anopheles gambiae*	EAA03222	2×10^{-09}	942	1
ENSANGP00000012579	*Anopheles gambiae*	XP_309517	1×10^{-08}	594	1
Microsatellite VNTR	*Haliotis rubra*	AF302835	1×10^{-09}	536	1
RIKEN cDNA 1110037D14	*Rattus norvegicus*	XP_214003	3×10^{-19}	715	1
1190005P08Rik protein	*Mus musculus*	AAH16563	3×10^{-14}	885	1
Normal cDNA from testes					
Hypothetical protein MGC2655	*Homo sapiens*	NP_077315	8×10^{-27}	623	1
Hypothetical protein FLJ32884	*Rattus norvegicus*	XP_227479	3×10^{-09}	666	1
ENSANGP00000014099	*Anopheles gambiae*	XP_316497	2×10^{-75}	761	1
Unknown protein	*Homo sapiens*	AAH14022	2×10^{-77}	810	1
PCO064652	*Zea mays*	AY109187	6×10^{-05}	538	1
Unnamed protein product	*Mus musculus*	BAC37272	1×10^{-74}	767	1
DKFZP434C245 protein	*Rattus norvegicus*	XP_218509	1×10^{-31}	610	1
CG33188-PA	*Drosophila melanogaster*	NP_788605	7×10^{-08}	710	1
Hypothetical protein FLJ20457	*Homo sapiens*	NP_060302	2×10^{-35}	750	1
Unknown for MGC:64164	*Danio rerio*	AAH5329	2×10^{-33}	697	1
RIKEN cDNA 4930434E21	*Mus musculus*	XP_132523	2×10^{-07}	722	1
ENSANGP00000014099	*Anopheles gambiae*	XP_316497	2×10^{-82}	718	1
Subtractive cDNA from ovaries					
Hypothetical protein FLJ11196	*Homo sapiens*	AAH06082	2×10^{-10}	279	1
ENSANGP00000021006	*Anopheles gambiae*	XP_312251	5×10^{-09}	432	1
Unnamed protein product	*Mus musculus*	BAC29193	6×10^{-06}	282	1
KIAA1412 protein	*Homo sapiens*	BAA92650	2×10^{-13}	419	1
CG10211-PA	*Drosophila melanogaster*	NP_609883	2×10^{-21}	389	1
Microsatellite DNA	*Entandrophragma cylindricum*	AJ420893	1×10^{-143}	467	1
Subtractive cDNA from testes					
Hypothetical protein MGC26778	*Danio rerio*	AAH49447	2×10^{-21}	402	1
ENSANGP00000014144	*Anopheles gambiae*	EAA10186	2×10^{-26}	667	1
Microsatellite VNTR	*Haliotis rubra*	AF302830.1	2×10^{-19}	365	1
Microsatellite DNA	*Cocos nucifera*	CNU458311	2×10^{-63}	283	1

3.3.6 Clustering analysis ESTs coding for vitelline coat protein and unknown proteins from cDNA libraries of ovaries and testes of *H. asinina*

3.3.6.1 Vitelline coat proteins

Sex-related transcripts predominated among known transcripts in all cDNA libraries, but the most abundant transcripts were VCP41 in ovaries of *H. asinina*. Clustering analysis of each EST indicated that 90 clones (Table 3.14) representing the vitelline envelope proteins (VCP41, VCP42 and VERL) found in this study could be gathered into 18 clusters (HAVCP 1-16, HAVERL-1 and HAVERL-2). These clusters were composed of both ESTs from normal and subtractive cDNA libraries. ESTs originally regarded as homologues of VERL and unknown transcripts were clusters to VCP homologues.

The nucleotide sequence of each cluster of VCP homologues was re-blasted against data in the GenBank. Results showed that clusters HAVCP1-10 exhibited significant matches with VCP41 of *T. pfeifferi* whereas cluster 16 showed significant similarity with VCP42 of the same species, The complete ORF of VCP were represented by HAVCP1 and HAVCP7 clusters. A homologue of VERL (*H. rufescens*) was represented by HAVERL-1 and HAVERL-2.

3.3.6.1 Unknown genes

Clustering of unknown transcripts was also carried out. Unknown transcripts in the normal cDNA libraries of ovaries (91 clones) and testes, (45 clones) were classified to 5 and 8 clusters and 78 and 30 singletons, respectively (Table 3.15). In contrast, unknown transcripts (39 and 104 clones) obtained from the subtractive cDNA library of ovary and testis were allocated to 9 and 7 different clusters and 9 and 74 singletons, respectively. After re-blasted analysis, clusters of unknownO8 and unknown14 exhibited significant matches with VCP41 of *T. pfeifferi*. These clusters were classified into clusters HAVCP3 and HAVCP7, respectively.

Table 3.14 Clusters of EST exhibiting significant similarity with VCP and VERL genes from normal and subtractive ovary cDNA libraries of *H. asinina*

Name	No. of clone	Normal ovary	No. of clone	Subtractive ovary	Length (bp)	Gene homologue	Closed species	E-value
HAVCP1	6	NO012, NO016[B], NO070, NO147, NO187, and NO196	4 clones	SO009[B], SO018[C], SO082 [B], and SO086	1094	VCP41	*Tegula pfeifferi*	3×10^{-48}
HAVCP2	4	NO 007, NO 133, NO161, and NO164	-	-	941	VCP41	*Tegula pfeifferi*	8×10^{-25}
HAVCP3	8	NO017, NO028, NO029, NO031, NO111, NO136, NO169 [A], and NO178	5 clones	SO005, SO015, SO021, SO043, and SO091[A]	1027	VCP41	*Tegula pfeifferi*	1×10^{-51}
HAVCP4	3	NO075, NO140, and NO174	-	-	1001	VCP41	*Tegula pfeifferi*	2×10^{-61}
HAVCP5	2	NO066 and NO073	5 clones	SO029[C], SO045, SO074[C], SO097[B], and SO107[A]	1026	VCP41	*Tegula pfeifferi*	3×10^{-39}
HAVCP6	2	NO023 and NO120	4 clones	SO004, SO012, SO078 [C], and SO085 [B]	1017	VCP41	*Tegula pfeifferi*	1×10^{-58}
HAVCP7	6	NO054, NO089, NO102, NO183[B], NO194, and NO217	14 clones	SO024[C], SO034[B], SO038[B], SO050, SO052[B], SO054[B], SO062[B], SO071[C], SO077[B], SO087[B], SO090[B], SO100, SO104, and SO108	1130	VCP41	*Tegula pfeifferi*	7×10^{-52}
HAVCP8	2	NO055 and NO127	-	-	909	VCP41	*Tegula pfeifferi*	1×10^{-49}
HAVCP9	1	NO049	3 clones	SO044, SO049, and SO064	783	VCP41	*Tegula pfeifferi*	4×10^{-41}
HAVCP10	3	NO038[C], NO123 and NO193 [C]	2 clones	SO025 [C] and SO048 [C]	1631	VCP41	*Tegula pfeifferi*	6×10^{-37}
HAVCP11	1	NO036	2 clones	SO019 and SO028	1200	VCP41	*Tegula pfeifferi*	5×10^{-33}
HAVCP12	-	NO064	-	-	828	VCP41	*Tegula pfeifferi*	3×10^{-36}
HAVCP13	1	NO095		SO059[B]	1107	VCP41	*Tegula pfeifferi*	6×10^{-16}
HAVCP14	-	-	2 clones	SO096 and SO106	412	VCP41	*Tegula pfeifferi*	4×10^{-14}
HAVCP15	-	-	1 clone	SO008	485	VCP41	*Tegula pfeifferi*	2×10^{-22}
HAVCP16	2	NO003 and NO156[A]	2 clones	SO002[C] and SO053	1149	VCP42	*Tegula pfeifferi*	1×10^{-27}
HAVERL-1	2	NO078 and NO128	-	-	1102	VERL	*Haliotis rufescens*	2×10^{-30}
HAVERL-2	1	NO096[C]	2 clones	SO035[B] and SO098[B]	552	VERL	*Haliotis rufescens*	7×10^{-17}

Blast results of the original ESTs were classified to be homologues of VCP42 (A), VERL (B) and unknown (C)

Table 3.15 Clusters of unknown transcripts found in normal and subtractive cDNA libraries of ovaries and testes of *H. asinina*

Transcript	Name of clones	No. of clones	Length (bp)	Gene homologues (E-value)
Normal cDNA from ovaries				
UnknownO1	PO038 PO193	2	864	Unknown
UnknownO2	PO018 PO027 PO030	3	782	Unknown
UnknownO3	PO022 PO114	2	509	Unknown
UnknownO4	PO057 PO129	2	893	Unknown
UnknownO5	PO160 PO088	2	776	Unknown
Subtractive cDNA from ovaries				
UnknownO6	SO004 SO012	2	309	Unknown
UnknownO7	SO019 SO028	2	395	Unknown
UnknownO8	SO043 SO005 SO021 SO015 SO088	5	494	VCP41 of *Tegula pfeiffer* (1×10^{-49})
UnknownO9	SO044 SO049 SO064	3	224	Unknown
UnknownO10	SO050 SO108 SO100	3	680	Unknown
UnknownO11	SO096 SO106	2	412	Unknown
UnknownO12	SO107 SO097	2	404	Unknown
UnknownO13	SO009 SO082	2	402	Unknown
UnknownO14	SO034 SO054 SO062 SO090 SO052 SO038 SO077 SO087 SO104	9	465	VCP41 of *Tegula pfeiffer* (1×10^{-490})
Normal cDNA from testes				
UnknownT1	PT006 PT117	2	1034	Unknown
UnknownT2	PT028 ST011 ST049	3	708	Unknown
UnknownT3	PT048 ST131 ST153 ST037 ST038 ST057 ST062 ST018 ST112	9	805	Unknown
UnknownT4	PT049 PT051	2	503	Unknown
UnknownT5	PT067 PT106	2	799	Unknown
UnknownT6	PT078 PT115	2	799	Unknown
UnknownT7	PT095 PT103	2	791	Unknown
UnknownT8	PT110 PT114	2	762	Unknown
Subtractive cDNA from testes				
UnknownT9	ST015 ST001 ST125 ST073 ST075	5	378	Unknown
UnknownT10	ST079 ST078	2	197	Unknown
UnknownT11	ST099 ST119	2	421	Unknown
UnknownT12	ST103 ST134	2	497	Unknown
UnknownT13	ST108 ST140	2	630	Unknown
UnknownT14	ST110 PT008	2	599	Unknown
UnknownT15	ST116 ST149 ST065 PT027 ST109 ST028 ST123	7	361	Unknown

3.3.7 Full-length cDNA obtained from EST analysis

Nucleotide sequences of each EST homologous with known genes were searched for the open reading frame (ORF). Five different genes contained complete ORF. They were VCP1, VCP7, sperm lysin, TCTP and ATPase inhibitor.

(1) VCP1 and VCP7 of *H. asinina*

The coats surrounding non-mammalian and invertebrate eggs and mammalian eggs have been called vitelline envelope (VE) and zona pellucida (ZP), respectively. These extracellular coats apparently function to restrict fertilization of eggs and sperms from the same species, to prevent polyspermy, and to protect development of embryos either outside or within the female reproductive tract.

The full-length nucleotide sequence of HAVCP1 and HAVCP7 cDNA clusters containing 1094 bp (Figure 3.22) and 1130 bp (Figure 3.23), respectively and could be deduced from two recombinant clones (PO187 and PO196 for HAVCP1 and PO102 and PO194 for HAVCP7). ORF was examined by the presence of the in-frame methionine initiation (ATG) compared with the potential Kozak consensus sequence, A/GNNATGA/G, (Kozak, 1986) and the presence of in-frame stop codon (TAG, TGA and TAA). The ORF of HAVCP1 and HAVCP7 was 942 and 1002 nucleotides encoded for a protein of 313 and 333 amino acid residues, respectively. The putative polyadenylation signal, AATAAA, is present 10 and 12 nucleotides uptstrem from the poly(A) tail of VCP1 and VCP7 cDNAs, respectively.

Sequence analysis of deduced HAVCP1 and HAVCP7 revealed putative cleavage site of signal peptidase located between $A_{16} - V_{17}$.and $S_{20} - A_{21}$, respectively. The COOH-terminus of HAVCP1 and HAVCP7 contain the furin cleavage sequence (RKRR and RRRR where a conserved consensus amino acid sequence of a furin-like protease R-X-K/R-R). The predicted *N*-linked glycosylation site (NXS/T) was observed in the deduced HAVCP1 (NFT, 258th amino acid residues) and HAVCP7 (NFT, 261st amino acid residues). Deduced amino acid of these EST using the Blast*X* program showed 37% and 39% identy to VCP41 of *Tegula pfeifferi* with E-value 3 x 10^{-48} and 7 x 10^{-52}, respectively. Nucleotide sequences and deduced amino acid of HAVCP1 and HAVCP7 are highly similar (Figure 3.24).

```
        10        20        30        40        50        60
CCACGCGTCCGGGTTTCATTTGTTGCGCCAAAAGGTGAAAATGGCAGCAATTTTCCTTGT
                                          M  A  A  I  F  L  V
        70        80        90       100       110       120
TCTTGGCTTCATAGCTACCGTCCAGGCTGTTGTACCCCCGGGCTATGTATTGGACATCGA
  L  G  F  I  A  T  V  Q  A  V  V  P  P  G  Y  V  L  D  I  E
       130       140       150       160       170       180
ACCAGATTGTGGAAAAAATGGTATCGCCGACGCAACAATAAATCTGTTATCCGATTATGA
  P  D  C  G  K  N  G  I  A  D  A  T  I  N  L  L  S  D  Y  E
       190       200       210       220       230       240
AGCCGAGGCTAAAGCCTTTTGTGCAGGCAATAAGGAAGTCAGCTTCACCTCCAATGATGG
  A  E  A  K  A  F  C  A  G  N  K  E  V  S  F  T  S  N  D  G
       250       260       270       280       290       300
CGTGTACTTCACCTTACCCGTGTCCTATCCAGTCACTGGTGGGGGATCTTCTACCTGCAG
  V  Y  F  T  L  P  V  S  Y  P  V  T  G  G  G  S  S  T  C  R
       310       320       330       340       350       360
ATTTGTGAAAAAGAAAGATGCATTTGTTTACACAATTCTGGTGATTGTGGCATTCGGCAG
  F  V  K  K  K  D  A  F  V  Y  T  I  L  V  I  V  A  F  G  S
       370       380       390       400       410       420
CCCGGGTAGCAGGCTGCACCAGACCGATGAACGATACACAATCACCTGCTCCTACCAACC
  P  G  S  R  L  H  Q  T  D  E  R  Y  T  I  T  C  S  Y  Q  P
       430       440       450       460       470       480
TGGTGCCAAGAAGGAGAGCTCGCCCATTAAGGTTAAGCCTGGTGCTGTTGCTCCAAAGGT
  G  A  K  K  E  S  S  P  I  K  V  K  P  G  A  V  A  P  K  V
       490       500       510       520       530       540
AATAGAGGGAAATACGCCCCCGAAGAGCCCATCTGTCATCCTCATGTACCTTGTTGACGT
  I  E  G  N  T  P  P  K  S  P  S  V  I  L  M  Y  L  V  D  V
       550       560       570       580       590       600
CATCGGCAGAAGGATTTCAGGATCCGTCAATGCCGGCAAGAGCGTCCGTCTGAAGGCCGT
  I  G  R  R  I  S  G  S  V  N  A  G  K  S  V  R  L  K  A  V
       610       620       630       640       650       660
TACCCTTGGACCTGCTGACAAAGGTCTAAGGCCGGAATCTTGCGATGCTTTGAATTCAAA
  T  L  G  P  A  D  K  G  L  R  P  E  S  C  D  A  L  N  S  K
       670       680       690       700       710       720
GGGTGGTAGACTTTCGATTCTGAGATCAGGTTGTGGAGACGGTATGGTGTTGAAGCAGAC
  G  G  R  L  S  I  L  R  S  G  C  G  D  G  M  V  L  K  Q  T
       730       740       750       760       770       780
ACAAGGCTTTAAAACCCTTGGCAAGAGAACCTACAGCGACTTCTTCAAAGTTTTTACTGT
  Q  G  F  K  T  L  G  K  R  T  Y  S  D  F  F  K  V  F  T  V
       790       800       810       820       830       840
TAACGGAGATCCACAACTCAAGTTTGAATGCAACTTCACGGTCTGCTCCAAAAAGTGCAA
  N  G  D  P  Q  L  K  F  E  C  N  F  T  V  C  S  K  K  C  N
       850       860       870       880       890       900
CGGCCCGTCCTGTTTTCCTAAAGGCCGAAAACGCAGAGAACAACCAGTGGGAAGTCTAGC
  G  P  S  C  F  P  K  G  R  K  R  R  E  Q  P  V  G  S  L  A
       910       920       930       940       950       960
AAGCTACCTGTGGAACTCCCGTAGTACTGCCATGACGGACGTCTACACATTGAACGGTGC
  S  Y  L  W  N  S  R  S  T  A  M  T  D  V  Y  T  L  N  G  A
       970       980       990      1000      1010      1020
TGCGGTAGACTTGCCAGTTTAACAACGCGAACTCCTCGGTCACGTTTCCCACGTAAACAG
  A  V  D  L  P  V  *
      1030      1040      1050      1060      1070      1080
TAATTCGTGTAACCGGGTTTTATTTGGTGACTGGTAAATAAAAACAGAAACTAAAAAAAA
      1090      1100
AAAAAAAAAAAAAA
```

Figure 3.22 The full length of nucleotide and deduced amino acid sequences of HAVCP1 of *H. asinina*. The putative initiation codon, stop codon, signal peptide, *N*-linked glycosylation site (NXS/T), and polyadenylation signal (AATAAA) are bold-underlined. The furin cleavage sequence are shaded and bold-italicized.

```
        10        20        30        40        50        60
CCACGCGTCCGAACCGCCGTGGCGAGACAGAGCTGTGAGACACACATCAAAGGATTAAAA
        70        80        90       100       110       120
ATGAATATTGTCTGCTCAGTGCTGGTGGTGCTCTGCTTGGCGGTGAGCTACGTCAGATCG
 M  N  I  V  C  S  V  L  V  V  L  C  L  A  V  S  Y  V  R  S
       130       140       150       160       170       180
GCAATTCCCCCCGGTTATATCCTTGACATTACTCCAGACTGTGGTTCTAATGGAGTCGAG
 A  I  P  P  G  Y  I  L  D  I  T  P  D  C  G  S  N  G  V  E
       190       200       210       220       230       240
GACGGAGTCGTGAACCTGATGACGGACTTCAGTGCAGAAGCTAAAGCCGTATGTGCAGGG
 D  G  V  V  N  L  M  T  D  F  S  A  E  A  K  A  V  C  A  G
       250       260       270       280       290       300
GGAGTCAAGGTCAAGTTCACGTCTAAAGATGGCGTCAATTATGCACTGCCCGTATCCTAC
 G  V  K  V  K  F  T  S  K  D  G  V  N  Y  A  L  P  V  S  Y
       310       320       330       340       350       360
CCAGGAAAGGGAGGAAAGTCAGTCTGTAAATACAAGAAAAAGCGAGATTCTTTGGTCTAC
 P  G  K  G  G  K  S  V  C  K  Y  K  K  K  R  D  S  L  V  Y
       370       380       390       400       410       420
ACGGTCCTAGTGATCGTGGCCTTCGGCGAGCCTGGCAGTCGTCTTCACCAGTCTGATGAG
 T  V  L  V  I  V  A  F  G  E  P  G  S  R  L  H  Q  S  D  E
       430       440       450       460       470       480
GAATACACCATTACATGTACCTTCCAACCCAAGGGCAAGAAAGACAGTGGTCCTCTGACA
 E  Y  T  I  T  C  T  F  Q  P  K  G  K  K  D  S  G  P  L  T
       490       500       510       520       530       540
ATCGTTCCAGGTGTAAGTGCTCCAAAGGTCATAGAAGGTAACAAGGCACCGCCCAGCCAG
 I  V  P  G  V  S  A  P  K  V  I  E  G  N  K  A  P  P  S  Q
       550       560       570       580       590       600
TCCGTGGTCAAAGTATACTTAGTCGACGTGAAGGGTGCCGCACTCAAGGGAAACATACCA
 S  V  V  K  V  Y  L  V  D  V  K  G  A  A  L  K  G  N  I  P
       610       620       630       640       650       660
ACTGGCAAGAGCGTCCGACTCATGGCACGGACCTACACCAAGACGGACATTGGCATACAA
 T  G  K  S  V  R  L  M  A  R  T  Y  T  K  T  D  I  G  I  Q
       670       680       690       700       710       720
GTAGAGTCGTGTGACGCAATGGATTCTAAAGCAAACAGATACTCCATACTGAGAGCGGGC
 V  E  S  C  D  A  M  D  S  K  A  N  R  Y  S  I  L  R  A  G
       730       740       750       760       770       780
TGCGGGGACGGCATGATCATCGATAAAAACGAGGGTTTCATCACCGTCAGTAGGCGTTCC
 C  G  D  G  M  I  I  D  K  N  E  G  F  I  T  V  S  R  R  S
       790       800       810       820       830       840
TTCAGCCCCTTCTTCAAAGTATTCACCGTCAATGGTGACTGGACGCTCCGCTTCGAGTGT
 F  S  P  F  F  K  V  F  T  V  N  G  D  W  T  L  R  F  E  C
       850       860       870       880       890       900
AACTTTACTATCTGTGCTGCGAAATGCAATGGTCCCTCGTGTTCATCACAAAAGGTGGTA
 N  F  T  I  C  A  A  K  C  N  G  P  S  C  S  S  Q  K  V  V
       910       920       930       940       950       960
CGAAGACGCCGAGACACAAGTTTAGTTGGAAGACTCCAGTCCTTCATAACGGGTTGGAAG
 R  R  R  R  D  T  S  L  V  G  R  L  Q  S  F  I  T  G  W  K
       970       980       990      1000      1010      1020
ACCAATGTTCTCACAGAACCTTATACTTTGGACAGCGGAATTGCTCTCCCGCAGAACGTC
 T  N  V  L  T  E  P  Y  T  L  D  S  G  I  A  L  P  Q  N  V
      1030      1040      1050      1060      1070      1080
AACACAGACCACGTTCGTTCTTCTAGGCTGTCCGTGACCTAAACAACCTATAAATGTTAG
 N  T  D  H  V  R  S  S  R  L  S  V  T  *
      1090      1100      1110      1120      1130
AATGTACAAGGTGGTTGGTTTGCAATAAACTTTGAAAATTTAAAAAAAAA
```

Figure 3.23 The full length nucleotide and deduced amino acid sequences of HAVCP7 of *H. asinina*. The putative initiation codon, stop codon, signal peptide, *N*-linked glycosylation site (NXS/T), and polyadenylation signal (AATAAA) are bold-underlined. The furin cleavage sequence are shaded and bold-italicized.

Alignments of deduced amino acid sequences of VCP41 and VCP42 of snail (*Tegula pfeiffei*), COOH-terminus of VERL of abalone (*H. rufessen*) and HAVCP1 and HAVCP7 of *H. asinina* revealed high similarity between *H. asinina* HAVCP1 and HAVCP7 and VCP41 of *Tegula pfeiffei.* In addition, ten conserve cysteine residues (C1-C10) were observed in vitelline envelope proteins of gastropod (VCP41 and VCP42 of *Tegula pfeiffei* , VERL of *H. rufessen* and HAVCP1 and HAVCP7 of *H. asinina* (Figure 3.17).

```
HAVCP1          M----AAIFLVLGFIATVQAVVPPGYVLDIEPDCGKNGIADATINLLSDYEAEAKAFCAG
HAVCP7          MNIVCSVLVVLCLAVSYVRSAIPPGYILDITPDCGSNGVEDGVVNLMTDFSAEAKAVCAG
                *    :.:.::   :: *:::.****:*** ****.**: *..:**::*:.*****.***

HAVCP1          NKEVSFTSNDGVYFTLPVSYPVTGGGSSTCRFVKKKDAFVYTILVIVAFGSPGSRLHQTD
HAVCP7          GVKVKFTSKDGVNYALPVSYPGKGG-KSVCKYKKKRDSLVYTVLVIVAFGEPGSRLHQSD
                . :*.***:*** ::****** .** .*.*:: **:*::***:*******.*******:*

HAVCP1          ERYTITCSYQPGAKKESSPIKVKPGAVAPKVIEGNTPPKSPSVILMYLVDVIGRRISGSV
HAVCP7          EEYTITCTFQPKGKKDSGPLTIVPGVSAPKVIEGNKAPPSQSVVKVYLVDVKGAALKGNI
                *.*****::** .**:*.*:.: **. ********..* * **: :***** *  :.*.:

HAVCP1          NAGKSVRLKAVTLGPADKGLRPESCDALNSKGGRLSILRSGCGDGMVLKQTQGFKTLGKR
HAVCP7          PTGKSVRLMARTYTKTDIGIQVESCDAMDSKANRYSILRAGCGDGMIIDKNEGFITVSRR
                 :****** * *   :* *:: *****::**..* ****:******::.:.:** *:.:*

HAVCP1          TYSDFFKVFTVNGDPQLKFECNFTVCSKKCNGPSCF-PK-GRKRR-EQPVGSLASYLWNS
HAVCP7          SFSPFFKVFTVNGDWTLRFECNFTICAAKCNGPSCSSQKVVRRRRDTSLVGRLQSFITGW
                ::* **********  *:******:*: *******   *  *:**  . ** * *:: .

HAVCP1          RSTAMTDVYTLNGAAVDLPV--------------
HAVCP7          KTNVLTEPYTLDSGIALPQNVNTDHVRSSRLSVT
                ::..:*: ***:...
```

Figure 3.24 Alignments of predicted amino acid sequences of HAVCP1 and HAVCP7 from *H. asinina*. Conserved residues are shown in asterisks (*). The putative signal peptide are bold-underlined. The furin cleavage sequence are shaded and bold-italicized.

```
VCP1Ha          ------------------------------------------------------------
VCP7Ha          ------------------------------------------------------------
VCP41Tp         ------------------------------------------------------------
VCP42Tp         ------------------------------------------------------------
VERLHr          MMLIQYTRNELLDSPGMCVFWGPYSVPKNDTVVLYTVTARLKWSEGPPTNLSIECYMPKS


VCP1Ha          ------------------------------------MAAIFLVLGFIATVQAVVP------
VCP7Ha          -------------------------------MNIVCSVLVVLCLAVSYVRSAIP------
VCP41Tp         -------------------------------MASLSTVVYFIASVAMTVARIP------
VCP42Tp         -------------------------------MATLSTAIVIFALAAAAVAEIP------
VERLHr          PVAPKPEASPTSNAPEPQTYPTSSAPGTSPEGSATAAPGTSPEGNTTAARNAYPRKSNQT
                                                :       : .    *
                                      C1                    C2
VCP1Ha          ----------PGYVLDIEPDCGKNGIADATINLLSDYEAEAKAFCAGNKEVSFTSNDGVY
VCP7Ha          ----------PGYILDITPDCGSNGVEDGVVNLMTDFSAEAKAVCAGGVKVKFTSKDGVN
VCP41Tp         ----------DKYILKVTPFCGSDVSSDARIEIVTDLVIEAQASCAGGVDVNFTSTDKVN
VCP42Tp         ----------ENYTLKVTPYCGKGAEADARIEIVTDLVIKARLECAGGVNVDFTTTDKVN
VERLHr          TSTEDVLDDTSNYIIKVIPHCRTR-GDVALIEIITDVDLSAVAVCSNGSRHHFNSTDFVH
                          * :.: * * .    . :::::*   .*   *:..    *.:.* *
                                    C3
VCP1Ha          FTLPVSYPVTGGGSSTCRFVKKKDAFVYTILVIVAFGSPG-------------------S
VCP7Ha          YALPVSYPGKGG-KSVCKYKKKRDSLVYTVLVIVAFGEPG-------------------S
VCP41Tp         FMLLVSYPGQGS--QPCVFMKQRNALIFYLNVSVAYGEEP-------------------G
VCP42Tp         YVAYVSYPTHST-PHPCVFLKKVNAMVFSVKVLASYGEGPRDLDCPDPDPEGSGDPDVPG
VERLHr          FYLPVSYNFTPS---VCAFTRSKANLFKLHIGVSWKDRLH--------------------
                :   ***          * : :.   :.         .
                          C4
VCP1Ha          RLHQTDERYTITCSYQPGAKKESSPIKVKPGAVAPKVIEGNTPPKSPSVILMYLVDVIGR
VCP7Ha          RLHQSDEEYTITCTFQPKGKKDSGPLTIVPGVSAPKVIEGNKAPPSQSVVKVYLVDVKGA
VCP41Tp         LVRMDEEEYTVTCTFSPHGTDGSPEQSIIEGLIAPPELLTNVGPDSPSTFSLELVDVLGG
VCP42Tp         LVRMEEGQWTVTCSFGLMGN-GKVGKNVIEGCIAPTELMTNAGPASSSSFDLELVDVTGR
VERLHr          DVTTQKKDFLITCTFDPHKSHRGPTSASSEPLIAAKEIQSHQGPQSDAEEVFLKLVDIRN
                 :   .  : :**::    .              *.  :  :  * * :   .  :
                                                C5                  C6
VCP1Ha          RISGS-VNAGKSVRLKAVTLG-PADKGLRPESCDALNSKGG-RLSILRSGCGDGMVLKQT
VCP7Ha          ALKGN-IPTGKSVRLMARTYT-KTDIGIQVESCDAMDSKAN-RYSILRAGCGDGMIIDKN
VCP41Tp         NLANENVHLGRTVQLRGHTAGEGGESVYAQLQCIAMDGSQ--QYAILRGGCGDGIVFPKD
VCP42Tp         SLAGQDVHNGRTVQLRGVATSASGVIGIRPVSCFAIDGVA--AYAILRAGCGDGIVFPKD
VERLHr          ETLAAAVPLSKKVRLVGEVHGSSLESGLKPVACDAVGVQQGQRYTILRDGCGDGIVFAKD
                     :  .:.*:* . .                * *:.        :*** *****::: :
                                              C7   C8  C9   C10
VCP1Ha          QGFKTLGKRTYSDFFKVFTVNGDPQLKFECNFTVCSKKCNGPSCFPK---GRKRREQPVG
VCP7Ha          EGFITVSRRSFSPFFKVFTVNGDWTLRFECNFTICAAKCNGPSCSSQKVVRRRRDTSLVG
VCP41Tp         VGFTSNGTTTLSPYFEAFGINFTPVIQFKCTFVTCTTDCNGSSCESE-----ARKRRDVS
VCP42Tp         RGFITNGLTTLSPYFMAFEINFASTVTYKCHFVTCDENCDGSSCPVE-----RGRRSTSD
VERLHr          IGFITEGNKAFSPVFEVFKLHGNLHLTFMCNFTLCSHSCDGSSCSNQ----RRTRRSMAW
                 ** : .  : *  * .* ::    : : * *. *  .*:*.**  :

VCP1Ha          SLASYLWNSRSTAMTDVYTLNGAAVDLPV------------------
VCP7Ha          RLQSFITGWKTNVLTEPYTLDSGIALPQNVNTDHVRSSRLSVT----
VCP41Tp         DAMQPMGSESKLVRSSIVKFYGAQPLEAQSGFNMK------------
VCP42Tp         CRQTDCGAFRSKSLEAPTILSRLLRRR-------------------
VERLHr          QDIPHVADFDSSATPSTDMATVQVALLVAVALLITQLAGLAIYVNIN
                          .
```

Figure 3.25 Multiple alignments of deduced amino acid sequences of HaVCP1 (VCP1Ha) and HaVCP7 (VCP7Ha) of *H. asinina*, VCP41 (VCP41Tp) and VCP42 (VCP42Tp) of *Tegula pfeifferi*, and VERL (VERLHr) of *H. rufescens.* Conserved residues are show in asterisks (*). Conserved residues are shaded. Blank (-) are inserted to optimize alignment of the sequences.

(2) Sperm lysin of H. asinina

Abalone sperm lysin is released by the sperm acrosome reaction and creates a hole in the egg vitelline envelope by a species-specific, non-enzymatic mechanism. To create the hole, lysin binds tightly to VERL that causes the VERL molecules to lose cohesion and splay apart creating the hole. Sequences of sperm lysin have been reported in 27 abalone species including *H. ovina* and *H. varia* but no information is available for *H. asinina*.

The full length of cDNA (from PT060) representing a sperm lysin homologue was first isolated in testes of *H. asinina*. The full-length cDNA obtained contained the ORF of 453 nucleotides encoding for 150 amino acid residues. The potential Kozak consensus sequence (A/GNNATGG/A) was also found. The putative polyadenylation signal, AATAAA is present 12 nucleotide upstream of the poly (A) tail. The putative *N*-linked glycosylation site (NXS/T) was not found (Figure 3.26).

The sequence analysis showed 71 % identity with that of *H. ovina* sperm lysin. The deduced amino acid sequence of *H. asinina* sperm lysin (this study), *H. ovina* sperm lysin, and *H. varia* sperm lysin cDNAs were aligned using ClustalX. Abalone sperm lysin sequences are highly conserved among different *Haliotis* species. Sperm lysin of *H. ovina* revealed a higher homology to that of *H. varia* more than sperm lysin of *H. asinina* (Figure 3.27).

(3) Translationally controlled tumor proteins (TCTP) of H. asinina

TCTP, an orthologue of the IgE-dependent histamine release factor (IgE-HRF), is a highly conserved protein widely expressed in all eukaryotic organisms. TCTP expression is regulated both at the transcriptional and translation levels by a wide range of extracellular signals. TCTP has been several implicated in important cellular processes, such as cell growth, cell cycle progression, malignant transformation and in the protection of cells against various stress conditions and apoptosis.

```
        10        20        30        40        50        60
CGAACAGATCACAAGATGAAGCTGTTAGTGCTCTGTACGCTGGCCATGATGGTGACGATG
              M  K  L  L  V  L  C  T  L  A  M  M  V  T  M
        70        80        90       100       110       120
GCGATGTCCCGCAGTTGGGGCTTTATTCCTCGCAGAAGTCTTGCCAGGCCTTATGAAGTT
 A  M  S  R  S  W  G  F  I  P  R  R  S  L  A  R  P  Y  E  V
       130       140       150       160       170       180
GCAATGAAGACACAGATCATTGCTGGGTTTGATAGAAAGCTGGTCAGCTGGTTAGCACGT
 A  M  K  T  Q  I  I  A  G  F  D  R  K  L  V  S  W  L  A  R
       190       200       210       220       230       240
CATGGCAGCGGTTTGAATGCTATTCAGAGGAAGACGCTGTACTTCGTAAACAGACGACAT
 H  G  S  G  L  N  A  I  Q  R  K  T  L  Y  F  V  N  R  R  H
       250       260       270       280       290       300
ATGCAGACTTACTGGCCATCCTACACAAGATACGCGAGAAGGCAGACAGCCAGACTTGGT
 M  Q  T  Y  W  P  S  Y  T  R  Y  A  R  R  Q  T  A  R  L  G
       310       320       330       340       350       360
AGGCCAGCCACTGTTAACGACTACAGGCGCATTGGTGCCGAGATCGGAAGACGTATTCCT
 R  P  A  T  V  N  D  Y  R  R  I  G  A  E  I  G  R  R  I  P
       370       380       390       400       410       420
ATGGAATTATCCTACGAAGTGTTGGTTAGGAGAAACATGATTCCAGCATGGCGTCAGTAC
 M  E  L  S  Y  E  V  L  V  R  R  N  M  I  P  A  W  R  Q  Y
       430       440       450       460       470       480
ATGGCAGACCTTATGGCCAAACGCGTGGAAGATATCCCAGTTGGATTAATTGCACGTAGAA
 M  A  D  L  M  A  K  R  V  E  D  I  P  V  G  *
       490       500       510       520       530       540
GTTGTGCCGCACCCGGGACATGACATAAAGAAGACGTCAAGCATCAAGCACTGGATGGAG
       550       560       570       580       590       600
GCATGACCTCAATCTAGCGCAGGGTGGCCAGATTTTGCCCCCGTGATATTTTAACTCGTA
       610       620       630       640       650       660
AACATTGTACAGTAGTAAATAAATAATTTTGCCACAAAAAAAAAAAAAAAAAAAAAAAAA
       670
AAAAA
```

Figure 3.26 The full length of nucleotide and deduced amino acid sequences of sperm lysin of *H. asinina*. The putative initiation codon, stop codon, and polyadenylation signal (AATAAA) are bold-underlined

```
H. ovina     MKLLVLCVLAMMVTVAVSRRWAFIPPTRIPRAHEVALKVQIIAGWNRKLANWLARHGSRL
H. varia     MKLLVLCVLTMTVTVAMSRRWTFIPRKRLPMAHEVALKVEIIAGFNKKLDIWLARHGSRL
H. asinina   MKLLVLCTLAMMVTMAMSRSWGFIPRRSLARPYEVAMKTQIIAGFDRKLVSWLARHGSGL
             *******.*:* **:*:** * ***   :. .:***:*.:****:::**  ******* *

H. ovina     SAIQKKTLYFVNRRYMQTHWPSYMIFVKKQIARLGRPGNTNDYSRLGAEIGRRIPMEVNY
H. varia     SPIQKKTLYFVNRRYMQTHWSNYMLWVKRRINSLGRPGTTADYRNLGAEIGRRVPLELTY
H. asinina   NAIQRKTLYFVNRRHMQTYWPSYTRYARRQTARLGRPATVNDYRRIGAEIGRRIPMELSY
             ..**:*********:***:*..*  :.:::   ****... ** .:*******:*:*:.*

H. ovina     SFLVRRNLIPRWRQYMANLMAKRVEDIPVQ--
H. varia     SFLVRRNLIPQWRQYMADLMARRVADIPVARG
H. asinina   EVLVRRNMIPAWRQYMADLMAKRVEDIPV--G
             ..*****:** ******:***:** ****
```

Figure 3.27 Multiple alignments of deduced amino acid sequences of sperm lysin homologue of *H. asinina*, *H. ovina* and *H. varia*. Blank (-) are inserted to optimize alignment of the sequences. Conserved residues are show in asterisks (*).

Three ESTs (PO172, PO191 and PO157) represented a TCTP homologue were combined to produce the full length of TCTP gene. The putative Kosak sequence of HaTCTP is ACGATGA as compared with the consensus sequence of A/GNNATGG/A. The full-length cDNA of this protein contained an ORF of 546 nucleotides encoding for 181 amino acid residues. Two (48^{th}, NAS) and 162^{nd}, NST) amino acid residues) revealed putative *N*-linked glycosylation sites (NXS/T). The putative polyadenylation signals (AATAAA) were found at 12 nucleotide upstream from polyA tail (Figure 3.28). The HaTCTP of *H. asinina* showed 34% and 32% identities to TCTP and IgE-HRF of *Bombyx mori* (8×10^{-17}) and *Dermacentor variabilis* (3×10^{-16}), respectively.

Deduced amino acid sequences of *H. asinina* TCTP, *Homo sapiens* TCTP/HRF (AAQ01550), *Gallus gallus* TCTP/HRF (A38960), *Bombyx mori* TCTP (AB124799), and *Dermacentor variabilis* HRF (AF467699) cDNAs were multiple aligned. TCTP/HRF sequences of *Homo sapiens* and *Gallus gallus* are highly conserved. TCTP of *H. asinina* revealed low similarity with TCTP and IgE-HRF from other species (Figure 3.29).

(4) ATPase inhibitor of H. asinina

ATPase inhibitor protein exerts a regulatory action of the catalytic properties of mitochondrial ATPase, and plays an active role in the process of energy conservation (Gomez-Puyou et al, 1983). The full-length cDNA of ATPase inhibitor of *H. asinina* (PO046) contained the ORF of 315 nucleotides encoding for 104 amino acid residues (Figure 3.30). The putative polyadenylation signal, AATAAA, is present at 5 nucleotides upstream of the poly A tail. The putative *N*-linked glycosylation site (NXS/T) was not found. Its showed 52% identity to ATPase inhibitor of *C. elegans* (3×10^{-25}).

Deduced amino acid sequence of *H. asinina* ATPase inhibitor , *Homo sapiens* ATPase inhibitor (JC7175), *Mus musculus* ATPase inhibitor (AAH12680), and *C. elegans* ATPase inhibitor (O44441) cDNAs were aligned using ClustalX. ATPase inhibitor sequences are highly conserved among different species. The *H. asinina* ATPase inhibitor revealed the highest similarity to that of *C. elegans* whereas *H. sapiens* and *M. musculus* also exhibited highly conserved sequences (Figure 3.23).

```
        10        20        30        40        50        60
CACGCGTCCGTCGCGACAACAGGCGAGATAGACAGAACCAGTGAGATTTAAAGTCAGTCT
        70        80        90       100       110       120
CCATCTGCAACGATGATCGTCTTTCATGACGCCTTCACAGGTGATGAGTTGTTTATTGAC
            M  I  V  F  H  D  A  F  T  G  D  E  L  F  I  D
       130       140       150       160       170       180
AGTAACAAGTTCAAAAGGTTGCGCGAGAGGGTCTATCGCGTGACATGTACGGAGATAGTT
 S  N  K  F  K  R  L  R  E  R  V  Y  R  V  T  C  T  E  I  V
       190       200       210       220       230       240
AAGAAAGAGTCTGGGGACTTCGACATCGGTGCAAATGCCAGTCAGGAAGAGGCAGCAGAA
 K  K  E  S  G  D  F  D  I  G  A  N  A  S  Q  E  E  A  A  E
       250       260       270       280       290       300
GAATTAGAGGAGAGCGTTGTAAGAGGTGTTGACATTGTCATCAACCAGCACCTGGAAGAA
 E  L  E  E  S  V  V  R  G  V  D  I  V  I  N  Q  H  L  E  E
       310       320       330       340       350       360
TGCACTCAGGAATATACCATGAAAAATTACAATTTATTCTTAAGGGATTACTTCAAAAGA
 C  T  Q  E  Y  T  M  K  N  Y  N  L  F  L  R  D  Y  F  K  R
       370       380       390       400       410       420
ATTATTAATTCGATCAAGGAATCAGACTTGAGTGATGAAGAGAAGTCAGCACAGACTGCA
 I  I  N  S  I  K  E  S  D  L  S  D  E  E  K  S  A  Q  T  A
       430       440       450       460       470       480
GCTTTCAAGGAGGAGGCTGTACAATTTACTAACCTGGTGAAGAAGAATTTTGACGACATC
 A  F  K  E  E  A  V  Q  F  T  N  L  V  K  K  N  F  D  D  I
       490       500       510       520       530       540
CAGTTTTTCAGATTTCCAATAGTTTTTGAAGACCCCAAAGACATCAGTGGCTTTGTTATT
 Q  F  F  R  F  P  I  V  F  E  D  P  K  D  I  S  G  F  V  I
       550       560       570       580       590       600
CCCCTTCATCGTGACAACAGCACAGATTGCATCACCATGTGGTTTTTCCGACGTGGACTC
 P  L  H  R  D  N  S  T  D  C  I  T  M  W  F  F  R  R  G  L
       610       620       630       640       650       660
AGGGAAGAGAAGTGTTAACAACCGACAGTGTTTATTTGGAGCTCATCATTGGCCGCCCTG
 R  E  E  K  C  *
       670       680       690       700       710       720
ACAAAAGAGCTGCCGTCAGACTTTTGAAATGTGAACATTTATTCCCAACAGTTGTCTGCT
       730       740       750       760       770       780
TTTATCTCTTATATCTTTTAGAAATCGCATGTGCCAAAATATAAGTCAGTGAATTTTTGT
       790       800       810       820       830       840
ACAGATAATTTATTGTAAGGGTTTTACATTTTTATGTTTTAAGTATGCATCCTTCGCGGT
       850       860       870       880       890       900
GCATGCATTTTGAACATTTTTACATTGCAATTAGTATAATTCACATAATGCGATTTCCGA
       910       920       930       940       950       960
AAGATGTCACAACGGGCAAGTTTTGAAAAGATCAATAAAGTTTAAAATGATAAAAAAAAA
       970       980
AAAAAAAAAAA
```

Figure 3.28 The full length of nucleotide and deduced amino acid sequences of TCTP homologue of *H. asinina*. The putative initiation codon, stop codon, *N*-linked glycosylation site (NXS/T), and polyadenylation signal (AATAAA) are bold-underlined

```
Homo sapiens(TCTP/HRF)          MIIYRDLISHDEMFSDIYKIREIADGLCLEVEGKMVSRTEGNIDDSLIGG
Gallus gallus(TCTP/HRF)         MIIYRDCISQDEMFSDIYKIREVANGLCLEVEGKMVTRTEGQIDDSLIGG
Bombyx mori(TCTP)               MKIYKDIITGDEMFSDTYKMKLV-DEVIYEVTGRLVTRAQG---DIQIEG
Dermacentor variabilis(HRF)     MLIFKDKISGDEMFTDSSKYKLV-DDCIFKIECHHVTRKQG---EIQLDG
Haliotis asinina                MIVFHDAFTGDELFIDSNKFKRL-RERVYRVTCTEIVKKES---GDFDIG
                                * :::* :: **:* *  * : :      .:    : : :.       *

Homo sapiens(TCTP/HRF)          -NASAEGPEGEGTESTVITGVDIVMNHHLQETSF--TKEAYKKYIKDYMK
Gallus gallus(TCTP/HRF)         -NASAEGPEGEGTEATVITGVDIVINHHLQETSF--TKESYKKYIKDYMK
Bombyx mori(TCTP)               FNPSAEEAD-EGTDSAVESGVDIVLNHRLVETYAFGDKKSYTLYLKDYMK
Dermacentor variabilis(HRF)     ANPSAEELD-EGTDENVESGLDLVLNMRLTETCF--TKADYKNYLKTYTK
Haliotis asinina                ANASQEEAA-EELEESVVRGVDIVINQHLEECTQEYTMKNYNLFLRDYFK
                                 * * *    *  :  *  *:*:*:* :* *         *. ::: * *

Homo sapiens(TCTP/HRF)          SIKGKLEE------QRPERVKPFMTGAAEQIKHILANFKNYQFFIGENMN
Gallus gallus(TCTP/HRF)         AIKARLEE------HKPERVKPFMTGAAEQIKHILANFKNYQFFVGENMN
Bombyx mori(TCTP)               KLVAKLEE------KAPDQVEVFKTNMNKVMKDILGRFKELQFFTGESMD
Dermacentor variabilis(HRF)     ALQEKWKEE----GKSPEEIEDAKSKLTTAVKKVLPKLDDYQFFIGESCN
Haliotis asinina                IRINSIKESDLSDEEKSAQTAAFKEEAVQFTNLVKKNFDDIQFFRFPIVF
                                :    :*      . . .             : :  .:.: ***

Homo sapiens(TCTP/HRF)          PD--GMVALLDYREDG---VTPYMIFFKDGLEMEKC
Gallus gallus(TCTP/HRF)         PD--GMVALLDFREDG---VTPYMIFFKDGLEIEKC
Bombyx mori(TCTP)               CD--GMVAMMEYRDFDGT-QIPIMMFFKHGLEEEKF
Dermacentor variabilis(HRF)     AE--GIVGLLEYREQDGGGEKAVMMFFKHGLDEEKM
Haliotis asinina                EDPKDISGFVIPLHRDNSTDCITMWFFRRGLREEKC
                                :  .: .::   . .       * **: **  **
```

Figure. 3.29 Multiple alignment of deduced amino acid sequences of TCTP and IgE-dependent histamine release factor (IgE-HRF) homologues of *H. asinina*, *Homo sapiens* (AAQ01550), *Gallus gallus* (A38960), *Bombyx mori* (AB124799), and *Dermacentor variabilis* (AF467699). Asterisks (*) indicated conserved amino acids across compared species.

```
        10        20        30        40        50        60
GACATGGCAGCACGAAGACTACCAGGAACATTGCGGCTTTTCGGAGTACGCTACATGAGC
    M  A  A  R  R  L  P  G  T  L  R  L  F  G  V  R  Y  M  S
        70        80        90       100       110       120
GGAGAATGGGGATCCGGGGCAGGCAAGGGCGGTGGATCAGGCGGCAGTGTCAGAGAAGCT
 G  E  W  G  S  G  A  G  K  G  G  G  S  G  G  S  V  R  E  A
       130       140       150       160       170       180
GGTGGTAGTTTTGGCAAGATGGAAGCGGCACAAGAAGAGCAGTATTTTCGCAAGCGGCAA
 G  G  S  F  G  K  M  E  A  A  Q  E  E  Q  Y  F  R  K  R  Q
       190       200       210       220       230       240
GCTGCTCAGCTGAGTGCCCTACATGACCATCTGGAAGAGGAGATCAAACACCACCAGAAG
 A  A  Q  L  S  A  L  H  D  H  L  E  E  E  I  K  H  H  Q  K
       250       260       270       280       290       300
CAGATTAACAATCACCAGGACGCCATTGAAAGGCACAAGAAGAAGATATCTCAGTTAAAA
 Q  I  N  N  H  Q  D  A  I  E  R  H  K  K  K  I  S  Q  L  K
       310       320       330       340       350       360
CATGCAGGGGACAAGTAGTTCCACTGTAGTGATAAATACCAGTAAATCTATACTACAATG
 H  A  G  D  K  *
       370       380       390       400       410       420
CTGGAAGAAACTGTCAAATTAAAGTTCAGTACAACAGAAGCCCTCACATGGGGAATGGAA
       430       440       450       460       470       480
ATCATCTGTCACTAAATGGAGGCATTGTCAGTTTTTTGTAAATTATCAGAAATGATTATT
       490       500       510       520       530       540
TGATAGTTCCGGCAGAACAGTTGGAGTCCCGATTTAAATGGTAGATATACAAGAGTGTGT
       550       560       570       580       590       600
GTATTTGTGTTCTCCTTGAGGATGTGTGTATGACTGCCTTCTGTGGTTGTATGGCATTAC
       610       620       630       640       650       660
AGAGTTCTGTGCCTCAACTAGGAAACAAAAATAAATTTCACTGCATATTTAAAAAAAAAA
          670
AAAAAAAA
```

Figure 3.30 Nucleotide and deduced amino acid sequences of a homologue of ATPase inhibitor of *H. asinina*. The start codon, stop codon, and putative polyadenylation signal (AATAAA) are illustrated in bold face and underlined.

```
H. sapiens    MAVTALAARTWLGVWGVRTMQARGFGSDQSENVDRGAGSIREAGGAFGKREQAEEERYFR
M. musculus   MAGSALAVRARFGVWGMKVLQTRGFVSDSSDSMDTGAGSIREAGGAFGKREKAEEDRYFR
H. asinina    MA--ARRLPGTLRLFGVRYMSG--EWGSGAGKGGGSGGSVREAGGSFGKMEAAQEEQYFR
C. elegans    ML-SVSRAATRMTGMVARFSAG--GHGDGAGRGGGSGGSIRDAGGAFGKMEAAREDEYFY
              *   .     :     :        .. :   . ..**:*:***:*** * *.*:.**

H. sapiens    AQSREQLAALKKHHEEEIVHHKKEIERLQKEIERHKQKIKMLKHDD------
M. musculus   EKTKEQLAALRKHHEDEIDHHSKEIERLQKQIDRHKKKIQQLKNNH------
H. asinina    KRQAAQLSALHDHLEEEIKHHQKQINNHQDAIERHKKKISQLKHAGDK----
C. elegans    KKQKAQLQELREHIQEEVKHHEGQLENHKKVLERHQQRISEIEAQERALGKE
               :   **  *:.* ::*: **. :::. :. ::**:::*. ::
```

Figure. 3.31 Multiple alignments of deduced amino acid sequence of ATPase inhibitor homologues of *H. asinina*, *Homo sapiens, Mus musculus,* and *Caenorhabditis elegans*. Asterisks indicated conserved amino acids across compared species.

3.4 Isolation and characterization of sex-related genes in ovaries and testes of the tropical abalone (*H. asinina*)

3.4.1 *Sex-specific expression analysis*

Primers were designed from nucleotide sequences of sex-related genes consist of vitelline coat protein1 (VCP1F/R), VCP2 (VCP2F/R), VCP3 (VCP3F/R), VCP7 (VCP7F/R), and individual EST clone of VCP49 (VCP49F/R), VCP75 (VCP75F/R), hydroxysteriod dehydrogenase –like protein (HSDF/R), vitellogenin-1 (VTG-1F/R), a disintegrin and metalloproteinase with thrombospondin motifs 9 (ADAMTS-9F/R), sperm lysin (lysinF/R), fertilization protein (FP-F/R), axonemal p66.0 (axonemal F/R), tektinA1 (Tektin F/R), gonadotropin inducible ovarian transcription factor 1 (GIOT1 F/R), small androgen receptor-interacting protein (SARIP-F/R), α-tubulin2 (Tubulin2F/R), and DMRT1 (which was cloned and characterized from the conventional RT-PCR product, DMRT1HaF/R) using Primer Premier 5.0.

RT-PCR was carried out to examine expression levels of interesting gene in ovaries, testes and hemocytes of *H. asinina*. The results show that VCP2, SARIP homologues were all expressed in ovaries, testes and hemocytes of *H. asinina* (Figure 3.32). Homologues of GIOT-1 and HSD were non-differentially expressed in both ovaries and testes of *H. asinina* (Figure 3.32). ADAMTS-9 and α-tubulin2 homologues were abundantly expressed in ovaries and in low testes of *H. asinina* (Figure 3.33).

In contrast, vitelline coat protein 1, VCP3, VCP49, VCP75, and VTG-1 were specifically expressed in *H. asinina* ovaries. Moreover, VCP7 was detected in ovaries and hemocytes but not in testes. (Figure 3.32). In addition, homologues of axonemal p66.0, tektinA1, sperm lysin, fertilization protein, DMRT1 were restrictively expressed in testes of *H. asinina* (Figure 3.32).

Amplification of VCP1, VCP3, VCP7, VCP49, VCP75 and VTG-1 and axonemalp66.0 tektinA1, sperm lysin, fertilization protein and DMRT1 indicated specific expression of transcripts in ovaries and testes of *H. asinina*, respectively. The large sample sizes of sex-specific transcripts in female (N = 16) and male (N = 16) *H. asinina* were examined. Results were consistent when sample sizes of ovary and testis

H. asinina were increased. Therefore, six (VCP1, VCP3, VCP7, VCP49, VCP75 and VTG-1) and five transcripts (axonemalp66.0 tektinA1, sperm lysin, fertilization protein and DMRT1) showed sex-specific expression in female and male *H. asinina*, respectively.

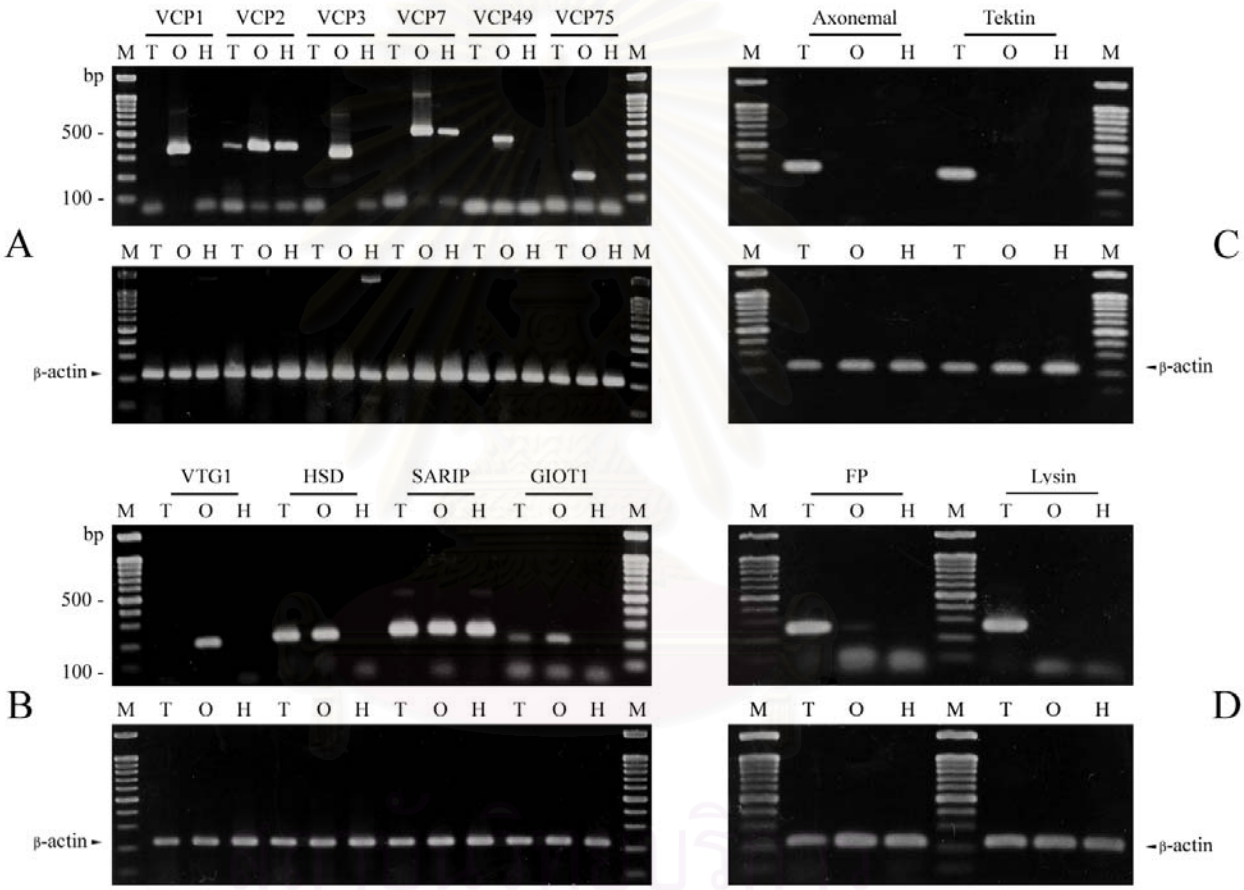


Figure 3.32 RT-PCR of VCP1, VCP2, VCP3, VCP7, VCP49, and VCP75 (A); VTG-1, HSD, SARIP, and GIOT-1 (B); axonemal p66.0 and tektinA1 (C); and FP and sperm lysin (D). Expression of each gene and β-actin (210 bp) was examined in testes, ovaries and hemocytes, respectively. Lanes M are 100-bp DNA markers

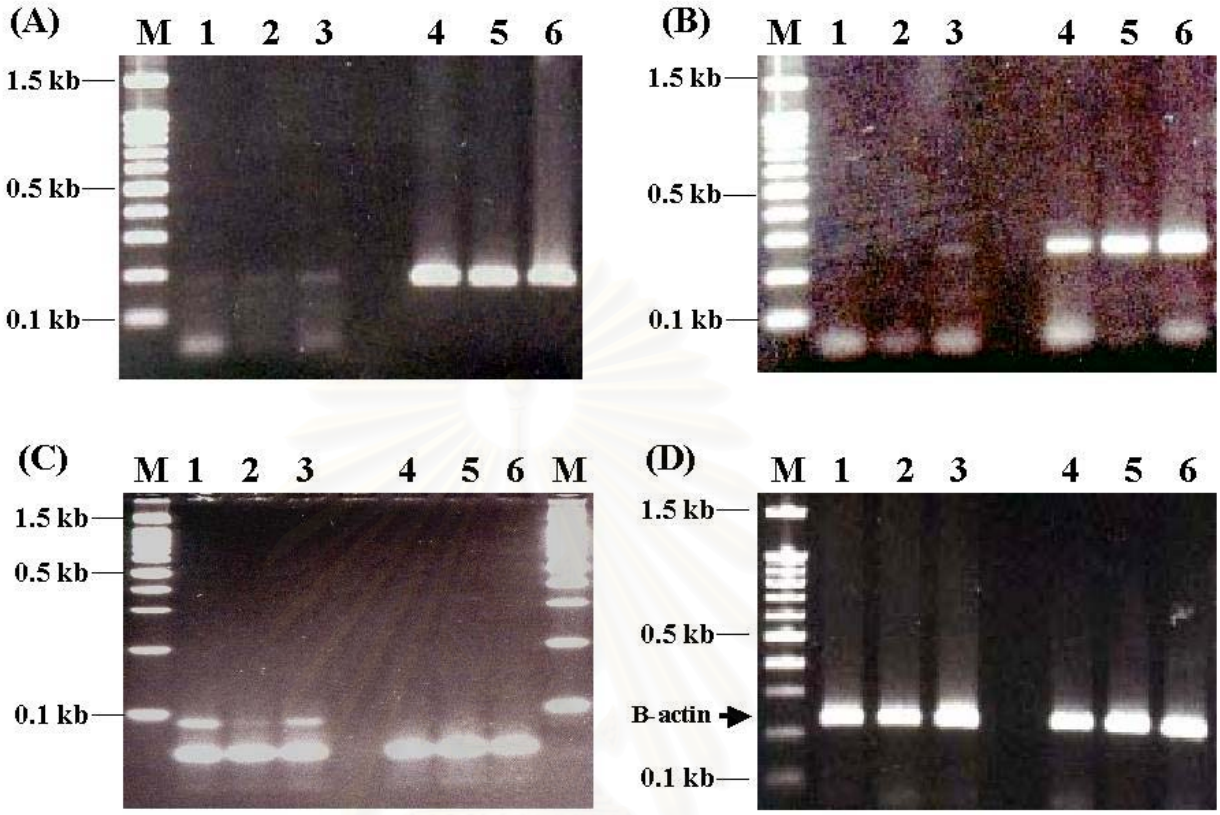


Figure 3.33 RT-PCR of the first strand cDNA synthesized from total RNA of testes (lanes 1-3) and ovaries (lanes 4-6) of *H. asinina* using α-tubulin 2 (A), ADAMTS9 (B), DMRT1 (C, HaDMRT1-F/R primer), and β-actin (D) primers. Lanes M are 100-bp DNA markers.

3.4.2 Tissue distribution analysis

RT-PCR was carried out against the first strand cDNA synthesized from total RNAs of various tissues (ovary, testis. foot, hemocyte, mantle, epipodial tentacles, oral tentacles, pairs of eyes, gill, hepatopancreas, digestive gland and left hypobranchial gland) of male and female *H. asinina* for tissue distribution analysis of VCP1, VCP2, VCP3, VCP3, VCP7, VCP49, VCP75, VTG-1, axonemal p66.0, tektinA1, sperm lysin, fertilization protein and DMRT1 homologues.

For sex-related transcripts from testes of *H. asinina*, RT-PCR further illustrated that sperm lysin and FP was abundantly expressed in testis followed hepatopancreas, and digestive gland of male but these transcripts were not expressed in foot, hemocyte, mantle, epipodial tentacles, oral tentacles, pairs of eyes, gill, and left hypobranchial gland of both sexes (Figure 3.34A).

A tektinA1 homologue was mainly expressed in testes and several other tissues including gill, digestive gland, left hypobranchial gland, and mantle of both males and females but it was not expressed in ovary and foot, hemocyte, mantle, epipodial tentacles, oral tentacles, pairs of eyes, and hepatopancreas of both males and females (Figure 3.34B). The axonemal p66.0 and DMRT1 homologous of *H. asinina* were only detected in testis cDNA but not in other tissues (Figure 3.34C).

For sex-related transcripts from ovaries of *H. asinina*, VCP1, VCP3, VCP49, VCP75, and VTG-1 were expressed in ovary cDNA but not in other tissue (Figure 3.35A). In addition, the VCP2 transcript was detected in hemocyte of males and ovary of females (Figure 3.35C). The VCP7 homologue was expressed ovary, eye, and digestive gland of female but not in any tissue of males (Figure 3.35B).

β-actin was included as a positive control and was positively amplified from all transcripts with no large differences in amounts of PCR products.

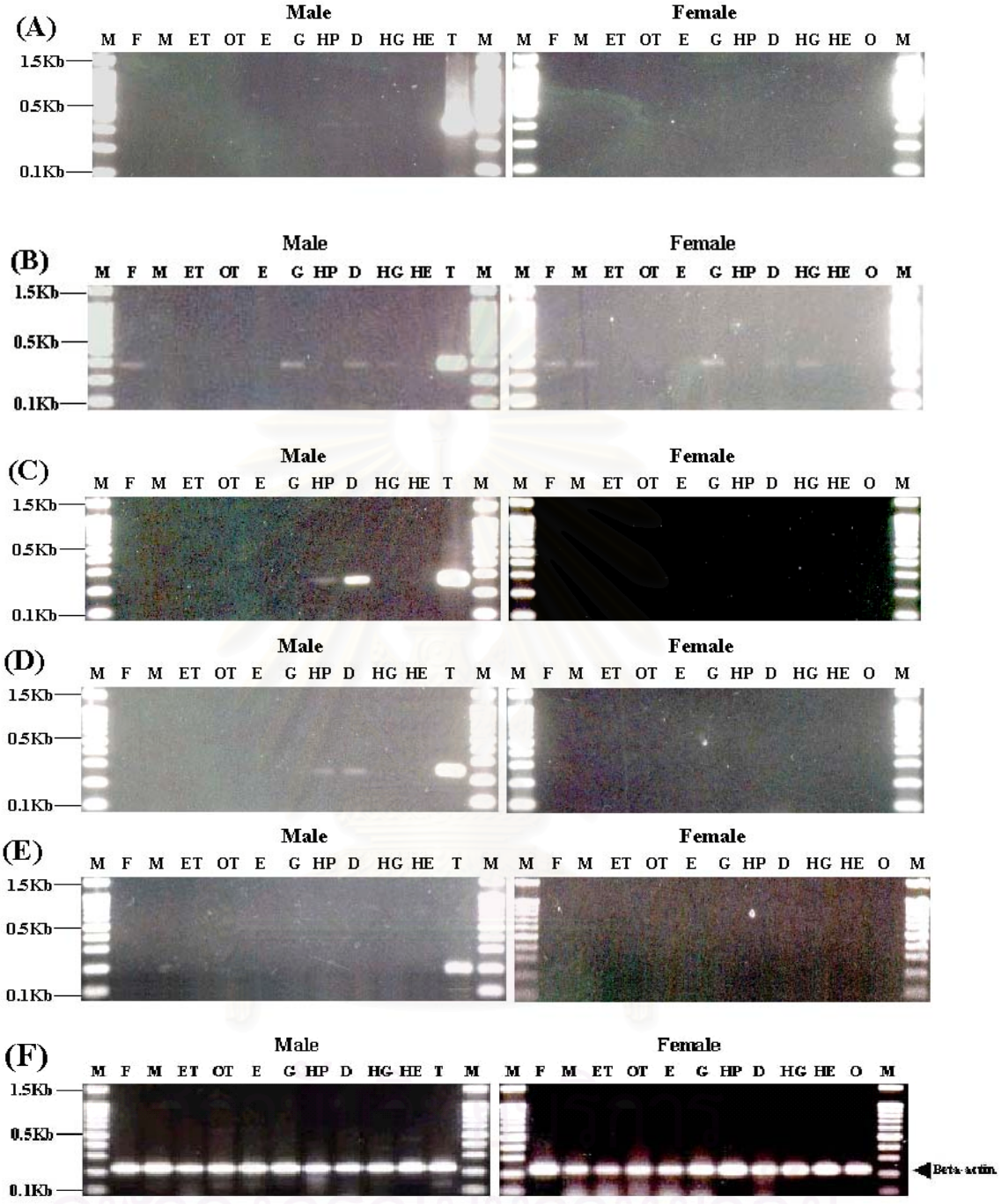


Figure 3.34 Expression of axonemal p66.0 (A), tektinA1 (B), sperm lysin (C), fertilization protein (D), DMRT1 (E, DMHa-F/R primer) homologues across various tissues; F= foot, M= mantle, ET= epipodial tentacles, OT= oral tentacles, E= eye, G= gill, HP= hepatopancreas, D= digestive gland, HG= left hypobranchail gland, HE= hemocyte, T= testis and O= ovary. β-actin (210 bp) was included as the positive control. Lanes M were a 100 bp DNA marker.

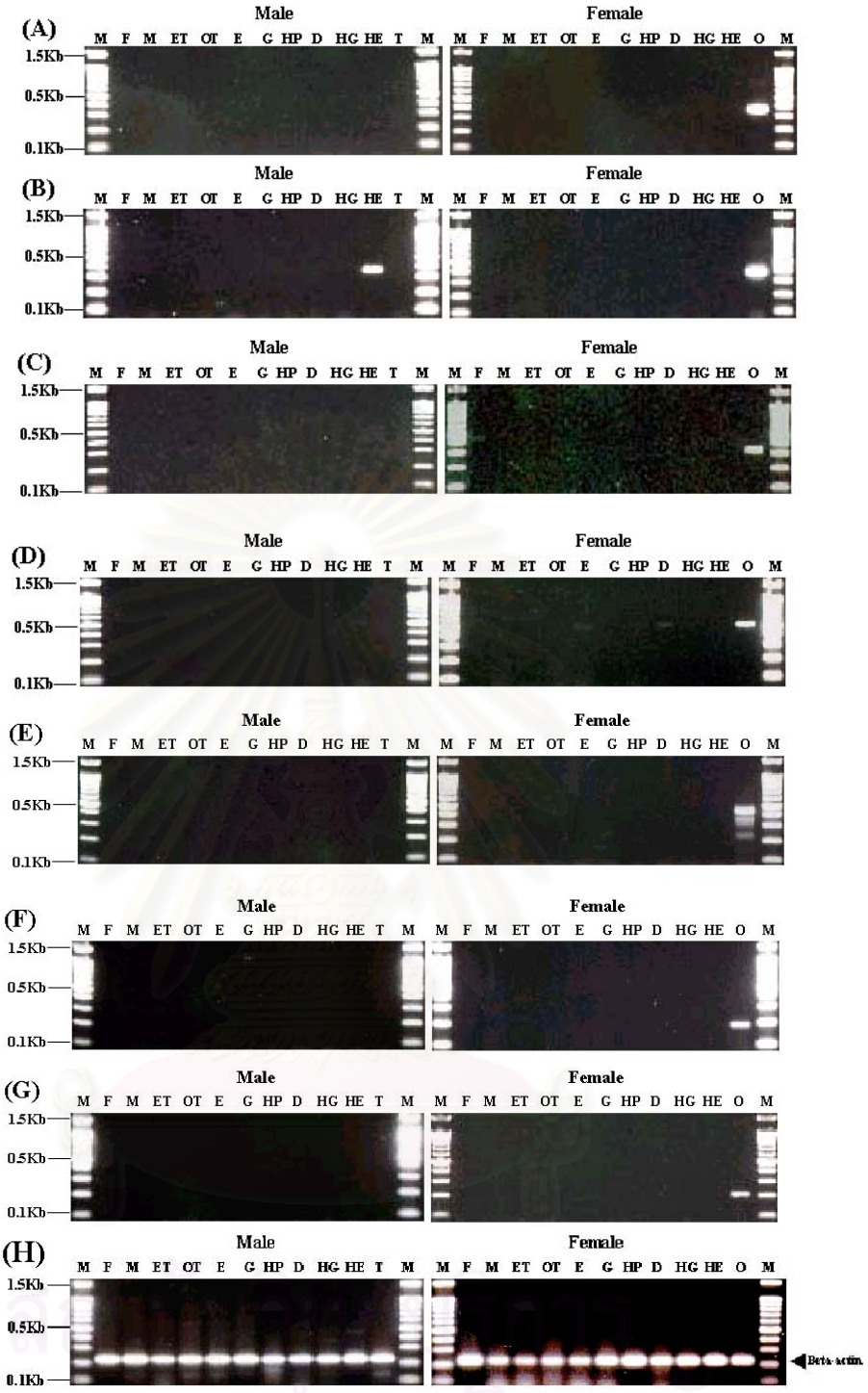


Figure 3.35 Expression of VCP1 (A), VCP2 (B), VCP3 (C), VCP7 (D), VCP49 (E), VCP75 (F), and VTG-1 (G) homologues across various tissues; F= foot, M= mantle, ET= epipodial tentacles, OT= oral tentacles, E= eye, G= gill, HP= hepatopancreas, D= digestive gland, HG= left hypobranchail gland, HE= hemocyte, T= testis and O= ovary. β-actin (210 bp) was included as the positive control. Lanes M were a 100 bp DNA marker.

3.4.3 Molecular cloning and characterization of DMRT1 gene homologue in *Haliotis asinina* by conventional RT-PCR

DMRT1, DSX, and MAB3 genes share a number of properties and contain the conserved DM domain (a zinc finger-like DNA binding motif) that the conserved genes involved with sex differentiation found from invertebrates to vertebrates and human.

Amino acid sequences of these genes in various species previously deposited in the GenBank were retrieved and aligned (Figure3.36). Conserved DM domains were found and used to design degenerated primers (and nested primers) for further characterization of DMRT-1/*dsx*/mab-3 homologues in *H. asinina*.

Initially, the transcript in testis of *H. asinina* was amplified using DMF1 and DMR primers (Figure 3.37, A). The expected 177-bp amplicon was gel-purified. Semi-nested PCR (DMF2 and DMR primers) was carried out using the gel-eluted PCR product (177 bp) from the primary amplification (Figure 3.37, B). The resulting of the expected 135 bp products was ligated to pGEM-Teasy and electrotransformed to *E. coli* XL1-Blue.

Nine recombinant clones were sequenced. Nucleotide sequences were further blasted against data in the GenBank using Blast*N* and Blast*X* programs. Results indicated that DM13561testis (Figure 3.38) and DM13567testis clones showed 63% sequence identity with the DMRT1 protein in *Oryzias curvinotus* (E-value 3 x 10^{-10}). The remaining clones (DM13553testis, DM13556testis, DM13562testis, DM13563testis, DM13565testis, DM13568testis, and DM13569testis) did not show significant similarity with any deposited data in the GenBank (Table 3.16).

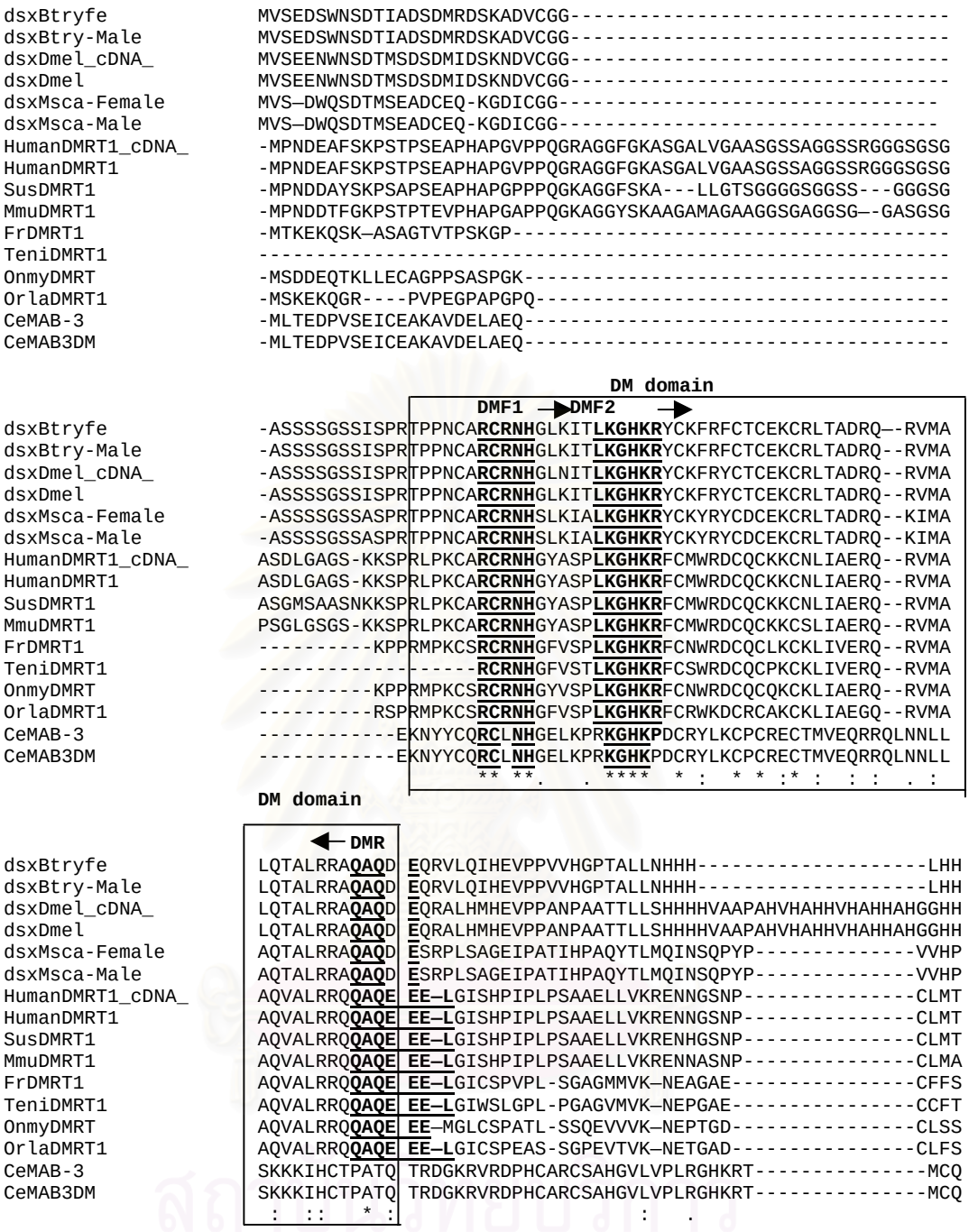


Figure 3.36 Multiple alignments of genes encoding DMRT1, *dsx* and mab-3 from *B. tryoni* (Btry), *D. melanogaster* (Dmel), *M. scalaris* (Msca), *H. sapiens* (Human), *S. scrota* (Sus), *M. musculus* (Mmu), *F. rubribes* (Fr), *T. nigroviridis* (Teni), *O. mykiss* (Onmy), *O. latipes* (Orla) and *C. elegans* (Ce). Conserved amino acid sequences in the DM domain (box) were used for designation of primers and nested primers and illustrated in boldface and underlined. Primer sequences are shown in Table 2.5.

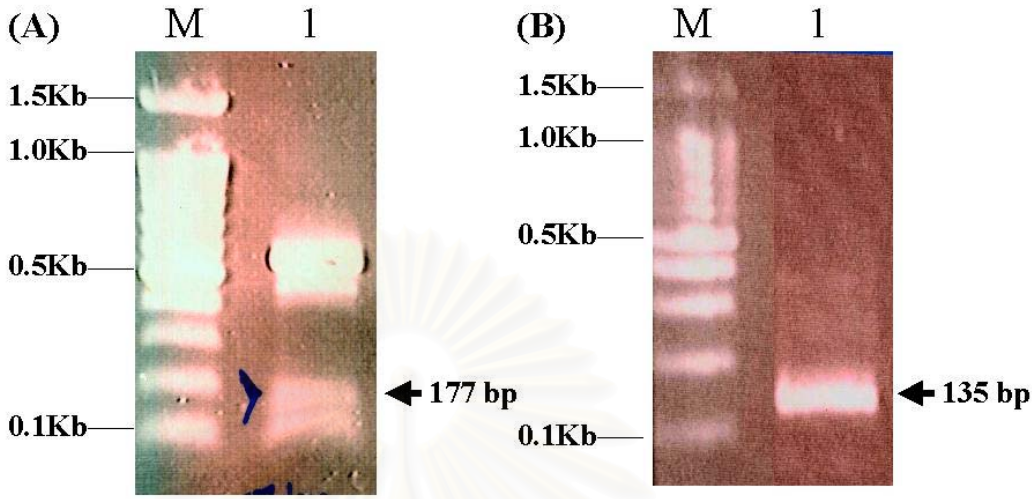


Figure 3.37 RT-PCR of the first strand cDNA synthesized from total RNA of testes (lanes 1) *H. asinina* using forward (DMF1) and reverse (DMR) primers (A) and nested-forward (DMF2) and reverse (DMR) primers. Lanes M are 100-bp DNA markers.

Table 3.16 Blast result of recombinant clones of putative DM-domain of *H. asinina* using RT-PCR

Clone	Length	Result from BlastX analysis
DM13553testis	160 bp	Unknown
DM13556testis	152 bp	Unknown
DM13561testis	135 bp	DMRT1 of *Oryzias curvinotus* (E-value 3 x10^{-10})
DM13562testis	204 bp	Unknown
DM13563testis	190 bp	Unknown
DM13565testis	170 bp	Unknown
DM13567testis	135 bp	DMRT1 of *Oryzias curvinotus* (E-value 3 x10^{-10})
DM13568testis	138bp	Unknown
DM13569testis	139 bp	Unknown

```
        10        20        30        40        50        60
CTAAAGGGACACAAGCGCTACTGTCTGTTCGCCAACTGCAGCTGTAGCAGCTGTCGCCTG
 L  K  G  H  K  R  Y  C  L  F  A  N  C  S  C  S  S  C  R  L

        70        80        90       100       110       120
AGTGCACAAAGACAGAAGGTGTCCCGGGAACAGATTGCCCTCCGTCGTCGACAGGCTCAG
 S  A  Q  R  Q  K  V  S  R  E  Q  I  A  L  R  R  R  Q  A  Q

       130       140
GAGGAGGAGCTGGCG
 E  E  E  L  A
```

Figure 3.38 Nucleotide sequence of putative DMRT1 homologue of *H. asinina* from a DM13561testis clone.

3.4.4 Isolation and characterization of the full length cDNA using RACE-PCR

5′ RACE-PCR and 3′ RACE-PCR were performed for isolation and characterization of the full-length cDNA of DMRT1, tektinA1, and axonemal p66.0.

3.4.4.1 Full length cDNA of DMRT1 homologue of *H. asinina*

5′ end and 3′ end of *H. asinina* DMRT1 gene homologues were isolated and characterized using the RACE-PCR approach. Gene specific primers (DMRT1Ha-R and DMRT1Ha-F) were designed (corresponding to 184 - 205 and 116 - 137 of the full-length cDNA of a DMRT1 homologue of *H. asinina*, respectively.) The nested PCR was performed using nested gene specific primers (nDMRT1Ha-R located at position 164 - 183 for 5′ RACE and nDMRT1Ha-F located of 139 - 158 for 3′ RACE). Results showed two (280 bp and 580 bp) and three (850 bp, 1250, and 1600 bp) discrete PCR products, respectively (Figure 3.39). The expected PCR products (a 280 bp for 5′ RACE and a 1600 bp for 3′ RACE) were cloned and sequenced.

The full-length cDNA of a DMRT1 homologue (Figure 3.40) was 1727 bp in length comprising a 5′ untranslated region of 55 bp, an open reading frame of 732 bp encoding for 243 amino acid residues and a 3′ untranslated region of 938 bp. The putative single (AATAAA) and multiple (AATAAAATAAA) polyadenylation signals were found at 8 nucleotide (single) and 437 and 432 nucleotide (multiple) from poly A tail. A predicted *N*-linked glycosylation site was observed (NCS, 43rd amino acid residues). Deduced amino acid sequence of the DMRT1 homologue of *H. asinina* -full length cDNA showed 57% sequence identity that of *Mus musculus* (E-value = 5 x 10^{-18}).

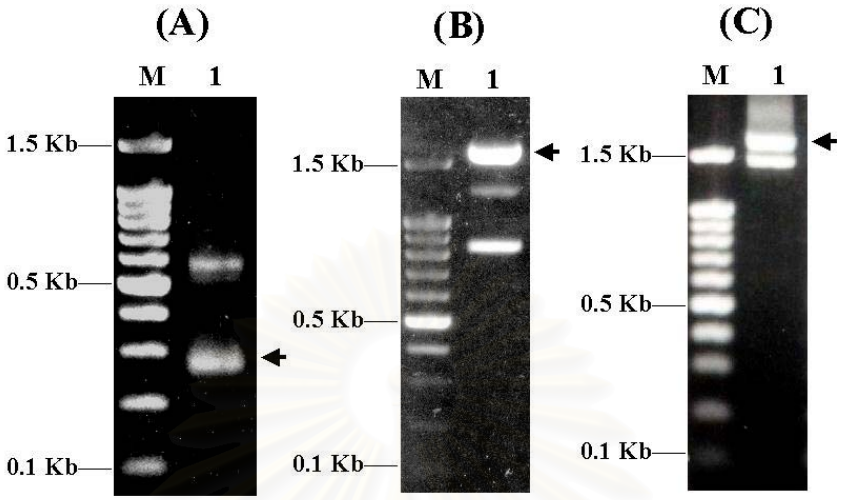


Figure 339 5′ RACE-PCR (A), 3′ RACE-PCR (B), and the full-length amplicons (C) of a *H. asinina* DMRT1 homologue using first strand cDNA of testes. Lanes M are a 100 bp DNA ladder. Arrows indicates the expected PCR product.

Multiple alignments of amino acid sequences of DMRT1 of *Homo sapiens* (AL162131), *Oryzias latipes* (AF319994), and *H. asinina* revealed that the DM domain of *H. asinina* DMRT1 was highly similar to that of *Homo sapiens* and *Oryzias latipes*. Conserved amino acid residues of the intertwined CCHC and HCCC zinc-binding sites were found in putative DM domain of *H. asinina* DMRT1 cDNA. In addition, the P/S domain (proline and serine rich region) in C-terminus of DMRT1 proteins was not conserved compared to that in other species (Figure 3.41). Apart from significant similarity due to the DM domain, the C-terminal region of *H. asinina* DMRT1 did not show significant homology to any previously deposited sequences in the GenBank of the other reported species. This transcript should be regarded as a new member of DM domain-containing genes.

Deduced amino acid sequence of the DM domain from *H. asinina* DMRT1 were compared to that of DMRT1 (NM_021951), DMRT2 (AF130729), DMRT3 (NM_021240), DMRT4 (AJ290954), DMRT5 (AJ301580), DMRT6 (AJ291671), and DMRT7 (AJ291669) of *Homo sapiens*, DSX (NP_524272) of *Drosophila melanogaster*, and AmDM1 (AF30064) of *Acropora millepora.*. The BLAST algorithm indicated identities and similarities of 52% and 67% to DMRT1, 42% and 62% to DMRT2, 40% and 58% to DMRT3, 48% and 64% to DMRT4, 48% and 64% to DMRT5, 27% and 44% to DMRT6, 47% and 62% to DMRT7 of *Homo sapiens*, 50% and 58% to DSX of *Drosophila melanogaster* and 54% and 68% to DM domain protein 1 of *Acropora millepora.*

After sequence analysis, the full-length cDNA (a 1727 bp) of *H. asinina* DMRT1 homologue was amplified by 5′UTRDMRT1-F and 3′UTRDMRT1-R for confirmation of the full transcript. The results showed two amplification bands of approximately 1700 bp and 1400 bp in length (Figure 3.39, C). The 1700 bp band presumably corresponded to the transcript represented by the 1727 bp full-length cDNA of *H. asinina* DMRT1. These fragments were cloned and sequenced. The nucleotide sequence of a 1700 bp PCR product was identical to RACE-PCR product. In contrast, a 1400 bp DNA fragment showed identical nucleotide sequence with that of a 1700 bp DNA fragment with the exception that it lacks 276 bases at the 5′ region. A 1400 bp fragment was regarded as a homologue of DMRT1 rather than a non-specific amplified product.

Nucleotide sequences of a 1400 bp DNA fragment did not revealed significant matching with any sequence previously deposited in the GenBank. The ORF of 1400 bp DNA fragment was analyzed. An ORF that corresponding with a ORF of the full-length DMRT1 of *H. asinina* at 88th amino acid residues was found. The putative initiating methionine of the non-DM containing transcripts (1400 bp) were found and located at 141st amino acid residues of ORF of the full-length DMRT1 *H. asinina* (Figure 3.40).

```
        10         20         30         40         50         60
GACTACAAGGTGCCTGCCTTTATGTTTTGATCGCAAGAGAGATTTCAGGTACAAAATGGC
                                                      M  A
        70         80         90        100        110        120
CGAAAGGTTTACCACTACCAAAGAAAAGAAGAAAAATACACAGCTACGTTCCTGTGCGCG
  E  R  F  T  T  T  K  E  K  K  K  N  T  Q  L  R  S  C  A  R
       130        140        150        160        170        180
ATGTCGTAACCATGCTATCATAAGCCCTCTGAAGGGCCACAAGCGCTTCTGTCTGTTCGC
  C  R  N  H  A  I  I  S  P  L  K  G  H  K  R  F  C  L  F  A
       190        200        210        220        230        240
CAACTGCAGCTGTAGCAGCTGTTGCCTGAGTGCACAAAGACAGAAGGTGTCCCGGGAACA
  N  C  S  C  S  S  C  C  L  S  A  Q  R  Q  K  V  S  R  E  Q
       250        260        270        280        290        300
GATTGCCCTCCGTCGTCGACAGGTACAGGATGAGGAGGCGGGACGGTCGATACCTGTACC
  I  A  L  R  R  R  Q  V  Q  D  E  E  A  G  R  S  I  P  V  P
       310        320        330        340        350        360
CGAGGAGGTGCTGACGGGTCGAGGAATCAGCTCGCTTGCCGGGGCCAACCTCCAGTCACT
  E  E  V  L  T  G  R  G  I  S  S  L  A  G  A  N  L  Q  S  L
       370        380        390        400        410        420
GAGACAACAGTTCCCTCACATCAGCGTCGGCAGACTCAACGCCGTCCTCGTGGAAACAGG
  R  Q  Q  F  P  H  I  S  V  G  R  L  N  A  V  L  V  E  T  G
       430        440        450        460        470        480
CAACCTACGCCAGACCATTATACGCCTTCAGAAAGAGGAGACAAGTCGAGAGGGAATGGA
  N  L  R  Q  T  I  I  R  L  Q  K  E  E  T  S  R  E  G  M  D
       490        500        510        520        530        540
TTTGCCAATAAACTCAGGCCCGACATTATCGACAGCAACGACGTCCACAGCTATGGGACT
  L  P  I  N  S  G  P  T  L  S  T  A  T  T  S  T  A  M  G  L
       550        560        570        580        590        600
TCATCAGGGACCTCAAAGAACGTTCAAAACCGAGCCATTCTTCCCCAGCGAGCTGTATCC
  H  Q  G  P  Q  R  T  F  K  T  E  P  F  F  P  S  E  L  Y  P
       610        620        630        640        650        660
ACTGGTTACTACCCGCAGGCTTCTCCCTACACCCATATGTACAGCACCCTACCCTCATAC
  L  V  T  T  R  R  L  L  P  T  P  I  C  T  A  P  Y  P  H  T
       670        680        690        700        710        720
ACCAGATCCACCCCATCAGCAGCAACCCCCTACAAGACCTATCACGCTGTCGAGCCCCAC
  P  D  P  P  H  Q  Q  Q  P  P  T  R  P  I  T  L  S  S  P  T
       730        740        750        760        770        780
TTCCAGCAGACCTCGCCCTTCCTTCCTCCACCCCCACCAACTACCAACCCCTACAGCTTT
  S  S  R  P  R  P  S  F  L  H  P  H  Q  L  P  T  P  T  A  F
       790        800        810        820        830        840
CAGCTAATGGGATCTGTGGCGCTTCCGGGGATCAGCACCGCCCCAGCAGATGAGCCTTCT
  S
       850        860        870        880        890        900
AGAGAGTTCAAGAAAAGCCTTGCCAAGTATTGTGGAGTTGGTGAACAGGGGGAGGCCAAC

       910        920        930        940        950        960
CAGCCCCCAGAGACCACCCCCGTGTCTGTAAGTTTCCCCAGCAGCAGCACAACAGGGATG

       970        980        990       1000       1010       1020
TCATCAGATGAGGAGACAGCACTACTTATTGACTGTTGTGACGACTAAGGGCTGAGCGTG

      1030       1040       1050       1060       1070       1080
CCCTCCCTGCTGTCATTAGCCTCGATTACCCACTGGATCCACTGGATAACCGAGAATGAA
      1090       1100       1110       1120       1130       1140
ACATTTCCTCGCTACTGTATGAGATATTTTCTTTTGATGCTTTGGTGTTTTCAGACTTGA
```

Figure 3.40

```
      1150      1160      1170      1180      1190      1200
GTCTGTGCGATAATTGACTTGAGTCTGTGCGATAATTGACTTGAGTCTGTGCGATAATTG

      1210      1220      1230      1240      1250      1260
ACTTGAGTCTGTGCGATGACTGTTAATGATTGGTTTCATGCTGATATCAGATAGTTTTGT

      1270      1280      1290      1300      1310      1320
AAAGGCCTTTTACATGTTTACTCAATAAAATAAAAATGAATCACGGACAGTTCTACCGGA

      1330      1340      1350      1360      1370      1380
CACAGTGAAGTCAAATATAGTTGTTCTTTATATCGGTTTTGCATTGTTTTTCTTAAGTTC

      1390      1400      1410      1420      1430      1440
CGTTAATTTGGTCGGTTCTTTAAATCTCAACTGAGTGATGTCACCAACTCACTGAACATG

      1450      1460      1470      1480      1490      1500
AAGTTACTAGGAGTAAATGTAATGCAGGAATGTCCATTTCCTTAGCGCCACTCGCGTCAA

      1510      1520      1530      1540      1550      1560
TGCAAATTCTCAAATCAAAGAAGTGACCAGCAGCAAGCTGGACATGAAATGTCATTGTGA

      1570      1580      1590      1600      1610      1620
CATCACGTCTGTCATGACAAAGGTAATGACTGTAAGTGAATATAATAACTAAACATGCGG

      1630      1640      1650      1660      1670      1680
AACGTGTGTGCGTATTGTTAACGTTGTTTGATTCTGTTCAAGGATTGTGACTTAACAAAA

      1690      1700      1710      1720      1730      1740
TGCTTTCAATGTTATGTTCAGCGTTGCTTTCAATAAATTAACTGCAAAAAAAAAAAAAAA
```

Figure 3.40 The full length of nucleotide and deduced amino acid sequences of DMRT1 homologue of *H. asinina*. The start codon, stop codon, putative polyadenylation signal (AATAAA) and *N*-linked glycosylation site (NXS/T) are bold-underlined. Nucleotide sequences of a 1400-bp DNA fragment are identical to full-length DMRT1 cDNA of *H. asinina* except a lack of a 276 bp (bold). Position of primers and nested primers for 5′ and 3′ RACE-PCR are italicized and underlined.

Figure. 3.41 Multiple alignments of deduced amino acid sequence of DMRT1 homologue of *H. asinina*, *Homo sapiens,* and *Oryzias latipes*. Asterisks indicated conserved amino acid across compared species.

3.4.4.2 Full length cDNA of tektinA1 homologue of *H. asinina*

For isolation and characterization of 5′ RACE and 3′ RACE products of a tektinA1 homologue, two successive rounds of 5′ RACE- and 3′ RACE-PCR amplification were carried out with following primers; first, with gene specific primers (TektinA1-R primer for 5′ RACE-PCR and TektinA1-F primer for 3′ RACE-PCR corresponding to nucleotide positions 468 - 491 and 210 - 229 relative to the translation initiation site of the full length cDNA of tektinA1, respectively) and Universal Primer A Mix (UPM); and second, with nested gene specific primer (TektinRACE5-R corresponding to nucleotide positions 134 - 158 for 5′ RACE-PCR and TektinRACE-3R corresponding to nucleotide positions 747 - 771 for 3′ RACE-PCR) and Nested Universal Primer A (NUP).

Results of 5′ RACE-PCR and 3′ RACE-PCR showed one major fragment of approximately 350 bp and 1300 bp, respectively (Figure 3.42). The PCR products were cloned and sequenced. After characterization, sequence of 5′ RACE-PCR and 3′ RACE-PCR of *H. asinina* tektinA1 were obtained and exhibited significant similarity to that of *Ciona intestinalis* tektinA1.

The full-length cDNA (2166 bp) of *H. asinina* tektinA1 (Figure 3.43) represented the complete ORF of 1350 nucleotides encoding for a polypeptide of 449 amino acid residues. The tektinA1 homologue of *H. asinina* contained a 5′ untranslated region of 150 bp and a long 3′ untranslated region of 643 bp, respectively. The putative polyadenylation signals were found at 15 nucleotides from the poly A tail. Five; 54^{th} (NPS), 62^{nd} (NYS), 88^{th} (NET), 95^{th} (NKT), and 361^{st} (NRT) predicted *N*-linked glycosylation sites (NXS/T) were found. The conserved RPNVELCRD sequence, likely to form a functionally important protein domain in tektin, was observed at position number 365 - 373 (RPGVDLCRD). The deduced tektinA1 of *H. asinina* showed 62% sequence identities to that tektinA1 of *Ciona intestinalis* (E-value = 1×10^{-149})

Since adapter ligation reactions may have a possibility to generate a chimera between cDNA fragments, the full-length cDNA was verified by RT-PCR using a set of gene-specific primer at 5′ and 3′ untranslated regions. Specific primers of 5′ UTRtektinA1 and 3′UTRtektinA1 were designed and used to amplify the full

nucleotide sequence (2064 bp) of tektinA1. The amplified product was cloned and sequenced (Figure 3.42). The result showed nucleotide sequence of this product was identical to full-length tektin A1 cDNA of *H. asinina*.

Multiple alignments of tektinA1 amino acid sequence of *Ciona intestinalis* (BAB86298) and *Strongylocentrolus purpuratus* (NP_999787), and *H. asinina* revealed that the *H. asinina* high similarity between this gene in these species (Figure 3.44).

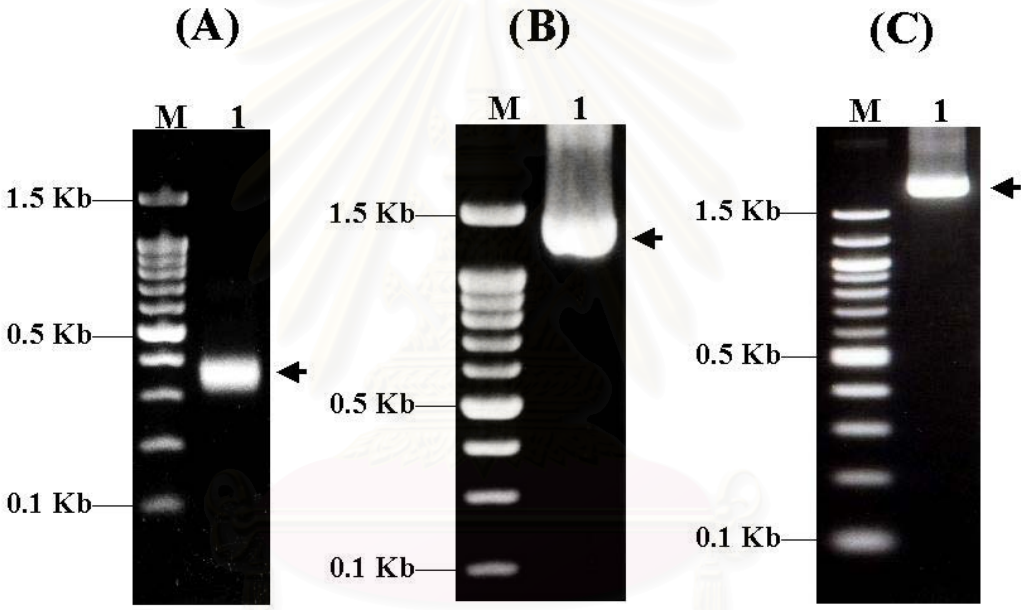


Figure 3.42 5/ RACE-PCR (A), 3/ RACE-PCR (B), and the full-length amplicons (C) of a tektinA1 homologue of *H. asinina* using the first strand cDNA of testes. Lanes M are a 100 bp DNA ladder. Arrows indicates the expected PCR products.

```
        10         20         30         40         50         60
GATGGCGACAAGAAACAAACTTCCGCTCATGGTTGCCGGGGATCTCTAAGCAACGTTTAT
        70         80         90        100        110        120
CCTGTTTCTCCTCTCGTCCAGCCCATGCAGATTACTAAAGGAGTAGAAAAGTAACTTTCA
       130        140        150        160        170        180
AAGTACACTCGAAGCACACAATAGGCAAGTATGGCGTCGGCAACACTTATGAGCCACGAG
                                  M  A  S  A  T  L  M  S  H  E
       190        200        210        220        230        240
ATTCCACCTGAGAGGCCTGTCATGAGGTCAGAATCAGGCACAGTAAATAATGAAATGGAG
 I  P  P  E  R  P  V  M  R  S  E  S  G  T  V  N  N  E  M  E
       250        260        270        280        290        300
ACGGTTTACACTGGGAACACTGGAAACAACATGGGCATATCCACAGAGGGATTTCGGGCG
 T  V  Y  T  G  N  T  G  N  N  M  G  I  S  T  E  G  F  R  A
       310        320        330        340        350        360
GCCAAGTACAATCCTTCAGAATGGCACGAAAGTAACTACTCGAAGTACTACCAGTCGTTT
 A  K  Y  N  P  S  E  W  H  E  S  N  Y  S  K  Y  Y  Q  S  F
       370        380        390        400        410        420
GTTGACAGGGATGGTGCGGAAAGGTTGCGCCATGAATCTAAGAGAGTTTCAAACGAAACT
 V  D  R  D  G  A  E  R  L  R  H  E  S  K  R  V  S  N  E  T
       430        440        450        460        470        480
GAAGCAACAACAAACAAAACACAGTCCGAGGTGACCAAGAAGCTGAACGAAAGAGTCCAA
 E  A  T  T  N  K  T  Q  S  E  V  T  K  K  L  N  E  R  V  Q
       490        500        510        520        530        540
GATATCAACTTCTGGAAATTCGAACTTGAAAGAGAAATTGGTGATGTTATTTCTGAAACA
 D  I  N  F  W  K  F  E  L  E  R  E  I  G  D  V  I  S  E  T
       550        560        570        580        590        600
GACTTGTTGCTTGCTCAGAAAAAGCGGCTGGAGAATGCCCTTCGTGCCACTGAAGTACCT
 D  L  L  L  A  Q  K  K  R  L  E  N  A  L  R  A  T  E  V  P
       610        620        630        640        650        660
CTCCATATTGCAACAGACAATCTCAACTGTAGGCAAAGGAGACAGGGTGTGGACCTGGTT
 L  H  I  A  T  D  N  L  N  C  R  Q  R  R  Q  G  V  D  L  V
       670        680        690        700        710        720
CAAGATGACCCAGAACTGTCATTATTAAAGGAGGTGGAGATAATAAACAATGTTCAAGAT
 Q  D  D  P  E  L  S  L  L  K  E  V  E  I  I  N  N  V  Q  D
       730        740        750        760        770        780
CTTCTCCAGAAAACAATCGGTCAAGCTGAGGCACAGATAAAGCTAAACCGAAATGCAAAG
 L  L  Q  K  T  I  G  Q  A  E  A  Q  I  K  L  N  R  N  A  K
       790        800        810        820        830        840
CAAGACCTGGAAATGGACTGGAGTGATAAGGAGGAAGCGCTGGAAATTGACACAAAATGT
 Q  D  L  E  M  D  W  S  D  K  E  E  A  L  E  I  D  T  K  C
       850        860        870        880        890        900
GCATCACTTAGGAATCACCACACTCACAAACAGTTCTACCCAGGAGCAGCTAAGTTCCAG
 A  S  L  R  N  H  H  T  H  K  Q  F  Y  P  G  A  A  K  F  Q
       910        920        930        940        950        960
AAATTCCTATCCAGCCCCGAGTCATGGGCTCAGTTTTCACATGACAACATTGTGAGAGCT
 K  F  L  S  S  P  E  S  W  A  Q  F  S  H  D  N  I  V  R  A
       970        980        990       1000       1010       1020
GAGCATGAGCGAATGGCCTCCATCCAGCTGCGAACTCTCATTGACAACATCTTGGAGGAC
 E  H  E  R  M  A  S  I  Q  L  R  T  L  I  D  N  I  L  E  D
      1030       1040       1050       1060       1070       1080
ACATCCCGTGACATGCGGGAACAGTGTGATGCCGTGGATGTTGCCTTCAACAAACGTGTT
 T  S  R  D  M  R  E  Q  C  D  A  V  D  V  A  F  N  K  R  V
      1090       1100       1110       1120       1130       1140
GAGGACATGGAGGATGCCAAGACCAAGCTTGAGGAGAATCTCCACAAGGTGTGTGTTGAG
 E  D  M  E  D  A  K  T  K  L  E  E  N  L  H  K  V  C  V  E
      1150       1160       1170       1180       1190       1200
ATCAGCAACCAGGAGAAGAATATTGAGAGCTTGAAGAGGGCCATAAGGGACAAGGAGGAT
 I  S  N  Q  E  K  N  I  E  S  L  K  R  A  I  R  D  K  E  D
```

```
      1210      1220      1230      1240      1250      1260
CCTCTTAAAGTTGCTCAAACTCGCCTTCAGAACAGAACCTTCCGTCCAGGCGTTGATCTG
 P  L  K  V  A  Q  T  R  L  Q  N  R  T  F  R  P  G  V  D  L
      1270      1280      1290      1300      1310      1320
TGCCGAGATCCTGTGCAATACCAGCTAGTTGGCGAAGTGAATGAGATCTGCCAGTCAATT
 C  R  D  P  V  Q  Y  Q  L  V  G  E  V  N  E  I  C  Q  S  I
      1330      1340      1350      1360      1370      1380
GATGCCCTCCAGGCCAAACTGAATGATGCTGTCAACTCCTTGAAGGATCTCCAGGATAAC
 D  A  L  Q  A  K  L  N  D  A  V  N  S  L  K  D  L  Q  D  N
      1390      1400      1410      1420      1430      1440
CCGTATGTCACTGAGAAGGAGATCTCTTGCAAGAAGAACAGTCTCTTCATTGACAGAGAC
 P  Y  V  T  E  K  E  I  S  C  K  K  N  S  L  F  I  D  R  D
      1450      1460      1470      1480      1490      1500
AAGTGCATGACCCATCGTACTCGCTGCCCAACCACCCTTAAACTCCAGGGCTATCAGTAA
 K  C  M  T  H  R  T  R  C  P  T  T  L  K  L  Q  G  Y  Q
      1510      1520      1530      1540      1550      1560
AGTCTGGGCATGCAAATCTGATTAAAGTATCAAAGAAACTAATACCGCTGAAGTTAATTA
      1570      1580      1590      1600      1610      1620
TTCAACGATTGTTATGTATTTAATGTTTGCCACCGATGGAGTTCACTCTGGCTCTGTCGG
      1630      1640      1650      1660      1670      1680
CCATCTATTTCACTGTTGCATGATTACTTGGAGATACATTTGTTGTTGTGATGTATATAT
      1690      1700      1710      1720      1730      1740
TATTTTTGTTTGTTTTTTGTTAAGAAATAACATACCAAATGATGCAAAACACACACTGTA
      1750      1760      1770      1780      1790      1800
ACATACCTTTTAGAAGATCCATAAATGGATCATATCTTTAAGTTTTAGTTCATGAATGTC
      1810      1820      1830      1840      1850      1860
ATTTTTGTTTAGGTATATGTGTGTAGTCCAAGGCTACCATGTTAAAATTATCAGTTGTAT
      1870      1880      1890      1900      1910      1920
TCCATTTCCATGAAGTAGTATTGCAGAAGTTTGCGGTTGATGTGTCATTACCCAGTTTGT
      1930      1940      1950      1960      1970      1980
TGATTTACTTCAGGTACTGTCTCCTCACTGGACATACATAACTATGACTTTCTGTGGCAT
      1990      2000      2010      2020      2030      2040
ATAAATCTTTCTAAGTGCAAACCTATGCTTGTGTTAGTTTTGTCACATGTATACATCAGT
      2050      2060      2070      2080      2090      2100
ATTATTGTTTTAACAAATTGGAGCCACACACAAAGATCTCTTAAAAGTGATTTATTATGC
      2110      2120      2130      2140      2150      2160
TTAATTGTATATAAAAGATGAATAAACCAACAAAAAACACGAAAAAAAAAAAAAAAAAA
      2170
AAAA
```

Figure 3.43 The full length of nucleotide and deduced amino acid sequences of tektinA1 homologue of *H. asinina*. The start codon, stop codon, putative polyadenylation signal (AATAAA), and *N*-linked glycosylation site (NXS/T) are bold-underlined. Position of primers and nested primers for 5′ and3′ RACE-PCR are italicized and underlined.

```
C. intestinalis MAAAEVLLKTEPAPQS---------------IPVDALGKKEHCIPINTGSYTSHGLATAG
S. purpuratus   MDAGATLLSRSYAPTIPVYPTQTTVGTKTDQALSQDLAKMSGLGETGVYGVPTGAPAAQG
H. asinina      M-ASATLMSHEIPPERPVMR------------SESGTVNNEMETVYTGNTGNNMGISTEG
                * *. .*:. . .*                    .   : .           . . :: *

C. intestinalis FRSSKYTPDEWHQNNYSKFYQSFSDRDNAERIQHESKKLANETLAITNRTQADVTKKLGE
S. purpuratus   FRSGKHTTQEWHESNYNKYFQSFTDRDNAERLCHESKQLSNETHALTMRTQADVTKKLGD
H. asinina      FRAAKYNPSEWHESNYSKYYQSFVDRDGAERLRHESKRVSNETEATTNKTQSEVTKKLNE
                **:.*:...***:.**.*::*** ***.***: ****:::*** * * :**::*****.:

C. intestinalis RIHDINFWKFELQREIDDLIAETELLFQQKVRLEKALDATEIPMQIATDNLECRFRRQGP
S. purpuratus   RMNDINFWKFELNREIEEMIEETDLLCAQKKRLENALDATEVPLKIARDNLTCRSRRQDI
H. asinina      RVQDINFWKFELEREIGDVISETDLLLAQKKRLENALRATEVPLHIATDNLNCRQRRQGV
                *::**********:*** ::* **:**  ** ***:** ***:*::** *** ** ***.

C. intestinalis DLVKDNVEIQLLREVDLLKSVQDLLKRTIEQTKNQIRMNRDAKQNLELDWSDKLDAHNID
S. purpuratus   DLVGDRVEMALNKEVDIITKVQDLLKRTLEQSDRQIKLNRGSKHKLTMDWSDKLSAFKID
H. asinina      DLVQDDPELSLLKEVEIINNVQDLLQKTIGQAEAQIKLNRNAKQDLEMDWSDKEEALEID
                *** *  *: * :**::::.*****::*: *:. **::**.:*:.* :***** .* :**

C. intestinalis DKCGRLNNQSTDIQYHPNSAKFEDNSSTPETWAQFTHDNIVRAERERMASINLRSLIDNV
S. purpuratus   EKCTGLNNNSTEIQYKEGSAKFEAVQTNPQSWAEFSHDNVVRAEHERLASQQLRNLIDQI
H. asinina      TKCASLRNHHTHKQFYPGAAKFQKFLSSPESWAQFSHDNIVRAEHERMASIQLRTLIDNI
                 **  *.*: *. *:  .:***:   :.*::**:*:***:****:**:** :**.***::

C. intestinalis LHDTSDDLREQCNSVNEAFAKRVEEMNDAKTKLENHLQKVLIEIGNQEHNIAMLKDAIND
S. purpuratus   LTDTSNDMREQCNTVNTEFARRIEEMNDAKTKMENHLLKTVEDIAGMEKNIKDLTQAVKD
H. asinina      LEDTSRDMREQCDAVDVAFNKRVEDMEDAKTKLEENLHKVCVEISNQEKNIESLKRAIRD
                * *** *:****::*:  * :*:*:*:*****:*::* *.  :*.. *:**  *. *:.*

C. intestinalis KEAPMQVSQTRLHHRTYRPSVELCRDPVQYRLVSEVGEITDSIEKLKFRLSEAERSLHEL
S. purpuratus   KEAPMKVAQTRLDHRTHRPNVELCRDPAQYRMVQEVGEIQDSIDKLQQKLAESKASLKDL
H. asinina      KEDPLKVAQTRLQNRTFRPGVDLCRDPVQYQLVGEVNEICQSIDALQAKLNDAVNSLKDL
                ** *::*:****.:**.**.*:*****.**::* **.** :**: *: :* ::  **::*

C. intestinalis EDTRMALEKEIAIKTNSLFIDREKCMTVRTRYPTVIKLSGYQ
S. purpuratus   MDTRMALEKEIALKKNSIFVDRDKCLKFRTRYPSTSKLVGYQ
H. asinina      QDNPYVTEKEISCKKNSLFIDRDKCMTHRTRCPTTLKLQGYQ
                 *.  . ****: *.**:*:**:**:. *** *:. ** ***
```


Figure 3.44 Multiple alignments of deduced amino acid sequence of tektinA1 homologue of *H. asinina*, *Ciona intestinalis* and *Strongylocentrolus purpuratus*. Asterisks indicated conserved amino acid across compared species.

3.4.4.3 Full length cDNA of axonemal p66.0 homologue of *H. asinina*

Likewise, 5/ RACE-PCR and 3/ RACE-PCR were performed to isolate the full-length cDNA of axonemal p66.0 mRNA homologues. PCR amplification of the RACE products was performed using gene specific primers (Axonemal-R for 5/ RACE and Axonemal-F for 3/ RACE corresponding to nucleotide position 1274 - 1292 and 967 - 988, respectively), and nested gene specific primers (AxonemalRACE3-F for 5/ RACE and axonemal RACE5-R for 3/ RACE corresponding to nucleotide position 1006 - 1030 and 1228 - 1252, respectively, Figure 3.46).

The expected amplicons of 1250 bp for 5/ RACE-PCR and 800 bp for 3/ RACE-PCR were obtained (Figure 3.45). These fragments were cloned and sequenced

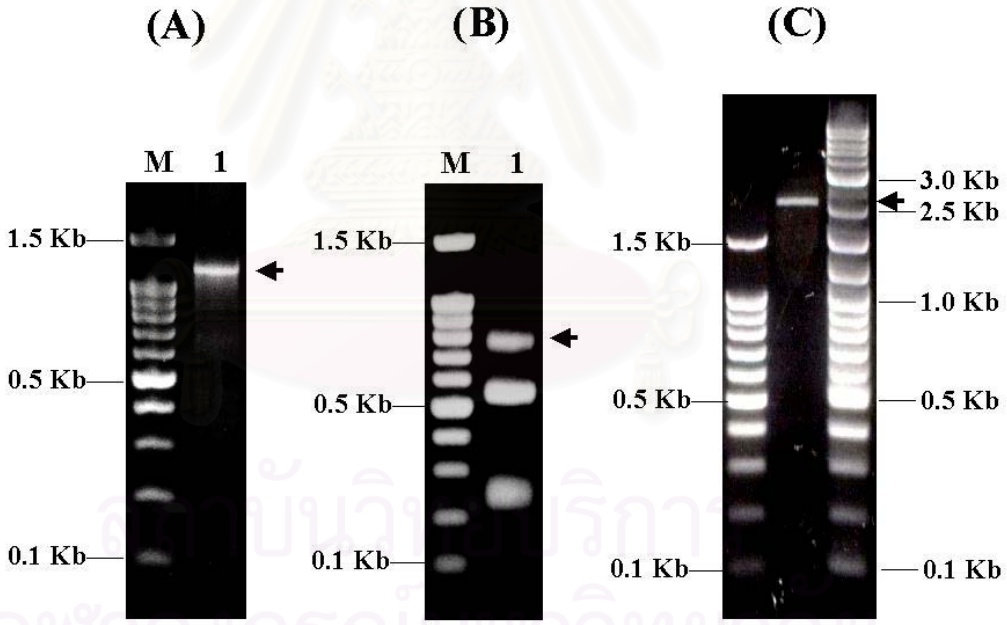


Figure 3.45 5 /RACE-PCR (A), 3/ RACE-PCR (B), and the full-length amplicons (C) of an axonemal p66.0 homologue of *H. asinina* using the first strand cDNA of testes. Lanes M are a 100 bp DNA ladder. Arrows (➜) indicates the expected PCR products.

The full length cDNA of an axonemal p66.0 homologue (2038 bp) of *H. asinina* contained an ORF of 1683 nucleotides encoding for 560 amino acid residues and a 5′ UTR of 81 bp and a 3′ UTR of 244 bp. Five (132nd, NGS; 417th, NRT; 461st, NLT; 532nd, NRS; and 542nd, NST) putative *N*-linked glycosylation sites were observed. The putative polyadenylation signal (AATAAA) was found at 20 nucleotides upstream from poly A tail (Figure 3.46). The deduced axonemal p66.0 of *H. asinina* exhibited 37% and 34% sequence identity with axonemal p66.0 of *Ciona intestinalis* (E-value = 4 x 10^{-97}) and hypothetical protein MGC6856 of *Xenopus laevis* (E-value 7 x 10^{-84}), respectively.

In order to amplify the full-length sequence of axonemal p66.0 homologues of *H. asinina*, 5′ UTRaxonemal-F and 3′ UTRaxonemal-R were designed from nucleotide sequence in 5′ and 3′ end of 5′ and 3′ RACE-PCR products, respectively. After amplification, the PCR product (1950 bp) was cloned and sequenced. Results showed that the nucleotide sequence of this product was identical to the inferred full-length axonemal p66.0 cDNA of *H. asinina*.

Deduced amino acid sequences of axonemal protein of *H. asinina* and *Ciona intestinalis* (BAB88833) and hypothetical protein MGC6856 of *Xenopus laevis* (AAH61941) were aligned using Clustal X. Results showed highly conserved among *H. asinina* axonemal protein and *Ciona intestinalis* axonemal p66.0 but distant relationships were observed between *H. asinina* axonemal protein and *X. laevis* hypothetical protein MGC6856 (Figure 3.47).

```
        10        20        30        40        50        60
GTGACGTCTTGTTACCATAACTTCCCAAGTAACAGCCGACCCTGAGTCCTCATCCGCACA
        70        80        90       100       110       120
CGGGCCACAACAGATTCAAAGATGCCGCGGGGAATAGCGGAGCGCCCCGACTCAGGAGAG
                      M  P  R  G  I  A  E  R  P  D  S  G  E
       130       140       150       160       170       180
GACCTCGGAGAGGAAGGGACACTGCGACAGGAACTGGCCCGGCTGCAACGACAACTACGA
 D  L  G  E  E  G  T  L  R  Q  E  L  A  R  L  Q  R  Q  L  R
       190       200       210       220       230       240
CTACTGGAAGGGGACAGGCGCGCATACACAGAGGAGACATGCAACGTCCTCATGAAACAG
 L  L  E  G  D  R  R  A  Y  T  E  E  T  C  N  V  L  M  K  Q
       250       260       270       280       290       300
AAGTCGGTCATCCGGGCTCTGCAGCAAGAGAATGGAGAGATGACCACGGTCCTGCGTCTG
 K  S  V  I  R  A  L  Q  Q  E  N  G  E  M  T  T  V  L  R  L
       310       320       330       340       350       360
TCTCAGAGCGACAGGAACGAGATGAAGGACGAGGGCAACACAGCTCGCCTGCAGGAACTA
 S  Q  S  D  R  N  E  M  K  D  E  G  N  T  A  R  L  Q  E  L
       370       380       390       400       410       420
ACAGAAGCTGCGGATACCTTGAACTCTCAGTGTGACACGGAGCGCAGCAAGATCGCCCAG
 T  E  A  A  D  T  L  N  S  Q  C  D  T  E  R  S  K  I  A  Q
       430       440       450       460       470       480
CTACAGTCAGAGCTAGTCCGAGTGAAGAAAGACATCTGTCACCAAAGGAGTCTCAACGGC
 L  Q  S  E  L  V  R  V  K  K  D  I  C  H  Q  R  S  L  N  G
       490       500       510       520       530       540
AGCAAATCGGGACCCATCATGTACAAGATGATGCTGAAGAAGACAGAGGTCATAGAGAAT
 S  K  S  G  P  I  M  Y  K  M  M  L  K  K  T  E  V  I  E  N
       550       560       570       580       590       600
AGACACGACAAGGCAATGGTCAAGTTCAACAACCAGCTGGCGGTGAACACACGTCTGCGC
 R  H  D  K  A  M  V  K  F  N  N  Q  L  A  V  N  T  R  L  R
       610       620       630       640       650       660
GAGGAGATCGACCACCTTCGGCAAGAGAAGTCTGTGTTCGACCTGCTGTTCCGGAAGCTC
 E  E  I  D  H  L  R  Q  E  K  S  V  F  D  L  L  F  R  K  L
       670       680       690       700       710       720
AACACTGACCTGGAGGACTATCGCAAAGACATGTCACTCATCATCGAGGAGGCGTCCCAG
 N  T  D  L  E  D  Y  R  K  D  M  S  L  I  I  E  E  A  S  Q
       730       740       750       760       770       780
GCGTATGAGGAAAGGGACGAGGCGTACAACAAGATGCTGGCGCTGAAGGAGAGGAGCGAG
 A  Y  E  E  R  D  E  A  Y  N  K  M  L  A  L  K  E  R  S  E
       790       800       810       820       830       840
AAGGACACAGCGCAGCACGAGATGGAGATGCGAGAGCTACAGCGTATCATCGAACACGAC
 K  D  T  A  Q  H  E  M  E  M  R  E  L  Q  R  I  I  E  H  D
       850       860       870       880       890       900
AACAAGCTGAAGGAGTTCATGATGATCAAGTCCAACGACCGCGGCGAGTACAAGGAGGAG
 N  K  L  K  E  F  M  M  I  K  S  N  D  R  G  E  Y  K  E  E
       910       920       930       940       950       960
GAGGACGCTAAGAAAAAGGGCATGAAGGGAGAGCGAGATGTGGACATCGAGAACAACCAG
 E  D  A  K  K  K  G  M  K  G  E  R  D  V  D  I  E  N  N  Q
       970       980       990      1000      1010      1020
ATCTTGTCGTACGAGGAGGCATTCGCTGAGATCAGGGAAGCCACGGGAGACGAGCACATC
 I  L  S  Y  E  E  A  F  A  E  I  R  E  A  T  G  D  E  H  I
      1030      1040      1050      1060      1070      1080
TTCAAAATCATGACCGAGTTCTGGGGCGCACGACGAGAACTTCGCCCTGTTCAATACGTC
 F  K  I  M  T  E  F  W  G  A  R  R  E  L  R  P  V  Q  Y  V
      1090      1100      1110      1120      1130      1140
AACGAGCTGAACGACCACCTCGACAAGCTCCAGGAAGAGATAGGGTTGATGAAGTCGGAG
 N  E  L  N  D  H  L  D  K  L  Q  E  E  I  G  L  M  K  S  E
```

Figure 3.46

```
      1150      1160      1170      1180      1190      1200
ATCGTCCGCTTTGAGGAGGAGGATGTCCAGATGGAGGCTGAGAGAAAGGTCTGGCTGAAG
 I  V  R  F  E  E  E  D  V  Q  M  E  A  E  R  K  V  W  L  K
      1210      1220      1230      1240      1250      1260
GAGTTGGAGGACAAGAGCACCCAGTCGGGGAAGGAGGCAGATGTAGCGGACAAGAAGATC
 E  L  E  D  K  S  T  Q  S  G  K  E  A  D  V  A  D  K  K  I
      1270      1280      1290      1300      1310      1320
GTGGAGATGTCACAGGTTCTAGACGAGCTGACGGATGGAGTTACCTCCCTGTTCCGGTGT
 V  E  M  S  Q  V  L  D  E  L  T  D  G  V  T  S  L  F  R  C
      1330      1340      1350      1360      1370      1380
CTGGGCTGCAATAGGACCCCGATCAACGAGATGCTGGGGAGCGAAGCGGGCGTCACGGAG
 L  G  C  N  R  T  P  I  N  E  M  L  G  S  E  A  G  V  T  E
      1390      1400      1410      1420      1430      1440
AAGAACATCCTTCTGTACATGGGCATCATCGAACAGCGGACGATGGAGCTTCTACACCTG
 K  N  I  L  L  Y  M  G  I  I  E  Q  R  T  M  E  L  L  H  L
      1450      1460      1470      1480      1490      1500
CAGCACTTCATAGAGATGAGGAATTTAACTCCCGAGAAGGAAAAGGAGAAAGGAAAAGAG
 Q  H  F  I  E  M  R  N  L  T  P  E  K  E  K  E  K  G  K  E
      1510      1520      1530      1540      1550      1560
GGAAAGACGAAGGAGCCAGATATGAAGAAGTTCCGCCGACCTCCCGCCGCAGTCACCGTC
 G  K  T  K  E  P  D  M  K  K  F  R  R  P  P  A  A  V  T  V
      1570      1580      1590      1600      1610      1620
ACTATAGCTGCTCCAGGTCTTGGGGAAGAAGAAGACTTGAATGAGATTAATGAGGAAATT
 T  I  A  A  P  G  L  G  E  E  E  D  L  N  E  I  N  E  E  I
      1630      1640      1650      1660      1670      1680
GAACTCCGTCCGCTGAGCCATGACGAACTGAAGCATATGGTGATGAGACGCATCAACCGC
 E  L  R  P  L  S  H  D  E  L  K  H  M  V  M  R  R  I  N  R
      1690      1700      1710      1720      1730      1740
TCAACCGGAAATACACAGTCGTTCAACTCCACTCCGGTTAAGTCCCCTCGAGTCCCAAAG
 S  T  G  N  T  Q  S  F  N  S  T  P  V  K  S  P  R  V  P  K
      1750      1760      1770      1780      1790      1800
GTCGAAGTCACTCAGCCAAAATAACCACTGATTGCTATTTAAAAAACACAAAAAACAAAA
 V  E  V  T  Q  P  K
      1810      1820      1830      1840      1850      1860
AAAAACCAGAGAGATTATGCATGTGGTATGAATGTAAAACTTTATATTCAGCTTATATAA
      1870      1880      1890      1900      1910      1920
TTTGATAATTGTAATTTAGACTGTGGATATATAATGTTTCTAACATTTACTGATTTATGG
      1930      1940      1950      1960      1970      1980
TTTCTTTGCTATTAGGTTACTTGTTGGTCTGAATATGTATATATATATATATATATCTAA
      1990      2000      2010      2020      2030      2040
TGTGCAATAAATTTATTGATATTGTTATGCCAAAAAAAAAAAAAAAAAAAAAAAAAA
```

Figure 3.46 The full length of nucleotide and deduced amino acid sequences of axonemal p66.0 homologue of *H. asinina*. The start codon, stop codon, putative polyadenylation signal (AATAAA), and *N*-linked glycosylation site (NXS/T) are bold-underlined. Position of primers and nested primers for 5′ and3′ RACE-PCR are italicized and underlined.

```
Ciona intestinalis      M---QRAASRRSESSEQEMDGLAEAELAKLQRQYRIMDGDRKAYCEESQNIIRRQRAQIA
Xenopus laevis          MPLRRSTSSARSDSSDVDVEGIADSELAKLQRMFRLKEVVHQEFTTKEKQKIRRQEKTIQ
Haliotis asinina        M--PRGIAERPDSGEDLGEEGTLRQELARLQRQLRLLEGDRRAYTEETCNVLMKQKSVIR
                        *   :  :.  ....:   :*    ***:***  *: :  :: :  :  : : :*.  *

Ciona intestinalis      SLQEEREEIQTNLLLAQSAQNTKKDSNNMEELCNLLSEQDSYERLIKEETEAVRALDTEI
Xenopus laevis          ALQKEQEELLKNLSVSESRRNQARDQGSRAKLHFLLEQKGELDGQLAEEKQAIANLEQEI
Haliotis asinina        ALQQENGEMTTVLRLSQSDRNEMKDEGNTARLQELTEAADTLNSQCDTERSKIAQLQSEL
                        :**:*. *: . * :::* :*  :*...  .*  * .  .  :     * . :  *: *:

Ciona intestinalis      RQIETKINTQHKSMGGVHSSHLRHVGTQKQIRVLENRLDKATVEFNKLLTSNSRLREEID
Xenopus laevis          KILEKKLSQQRK--HGRVSSGTTQKNAGTHVKVMENRLDRAVTRFNSQLSKNGNMREEIE
Haliotis asinina        VRVKKDICHQRS-LNGSKSGPIMYKMMLKKTEVIENRHDKAMVKFNNQLAVNTRLREEID
                          ::..:  *:.   *  *.        .: .*:*** *:* ..**. *: * .:****:

Ciona intestinalis      HLRSQRAVFDGLHKKLTKELGDQKRVMGEIIEQSTQAYDQRDDAQAKMMALKERNEKDLL
Xenopus laevis          ILRIEKSRFEQLYKRLEKELLQTRKEIGSVIDDSSAAYDARDEAQTKMLQLREKAEKDLN
Haliotis asinina        HLRQEKSVFDLLFRKLNTDLEDYRKDMSLIIEEASQAYEERDEAYNKMLALKERSEKDTA
                         ** ::: *: *.::* .:* : :: :. :*:::: **: **:*  **: *:*: ***

Ciona intestinalis      QYNMEYKELMRIIDHDAKLKLFMNMKSQERSELEEEEAAKRRAGE---EDKAEKTAEETM
Xenopus laevis          QHTAEIKELQRVIDHDRKLKEFMGAKTQERAISEEVLNARRKKEKDELERKKRDPTEETT
Haliotis asinina        QHEMEMRELQRIIEHDNKLKEFMMIKSNDRGEYKEEEDAKKKGMK---GERDVDIENNQI
                        *:  * :** *:*:** *** **  *:::*.  :*   *:::  :     :  .  ::

Ciona intestinalis      ETYQKAFEKIQEVTGEGDIYLLVSRFIETEDKNFALFNYVNELNNELELIQEQIDDVRMQ
Xenopus laevis          ETYERAFQQIQAVTGEDNLDILVTRFIEVEDRNFALFNFVNEQNNKIERLMEHIAEINKE
Haliotis asinina        LSYEEAFAEIREATGDEHIFKIMTEFWGAR-RELRPVQYVNELNDHLDKLQEEIGLMKSE
                         :*:.** :*: .**: .:  :::.*  .. :::  .::*** *:.:: : *.*  :. :

Ciona intestinalis      IEQFKNEGVQHQVERQNILKGLEEKLRKTTKEADMFDKQLKATEKILDQLKAGIESVFGK
Xenopus laevis          IEEFQAQGVRLDKEHEAILRSIEHKQEEAVQQADGYQQQLKGVMKILDQLKPGIDSLFKK
Haliotis asinina        IVRFEEEDVQMEAERKVWLKELEDKSTQSGKEADVADKKIVEMSQVLDELTDGVTSLFRC
                        * .*: :.*: : *::  *: :*.*  :: ::**  ::::    ::**:*. *: *:*

Ciona intestinalis      INCNRSTISDMLGDNDAVNENNMMQYLGIIEQKTNELLQIHTYLQLKELENKPEGTPGTS
Xenopus laevis          INCDRSVLDEMLGSSSSIRETNIMQYLGLIEQKTNELLATQSFLDSKNYDKPYD------
Haliotis asinina        LGCNRTPINEMLGSEAGVTEKNILLYMGIIEQRTMELLHLQHFIEMRNLTPEKEKEKGKE
                        :.*:*: :.:***.. .: *.*:: *:*:***:* ***  : ::: ::     :

Ciona intestinalis      PITTAALLGGPAAPPILAPIHIVPPSTEDDRDDDGGDSGEESVDFDRPLTQHELKARVMR
Xenopus laevis          PQETARLLLGPLTDMPTTVLQIKPPTTGEDHDSDIEFPSAE--EEERPLTHTELRERIMK
Haliotis asinina        GKTKEPDMKKFRRPPAAVTVTIAAPGLGEEED----LNEINEEIELRPLSHDELKHMVMR
                           .  :       . : * .*   ::.*       :     ***:: **:  :*:

Ciona intestinalis      NVAKKDLAGSNAVAGTKRPDKQMSTSEKTPRLSDAKKKPSK
Xenopus laevis          GVLKKEQKATRKSNSNETGTVKMAATTKK------------
Haliotis asinina        RINRSTGNTQSFNSTPVKSPRVPKVEVTQPK----------
                         : :.                  .  .
```

Figure 3.47 Multiple alignments of deduced amino acid sequence of *H. asinina* axonemal p66.0, *Ciona intestinalis* axonemal p66.0, and *Xenopus laevis* hypothetical protein MGC6856. Asterisks indicated conserved amino acid across compared species.

3.4.5 Expression of sex-related genes in different ages of juvenile abalone

Expression of axonemal p66.0, tektinA1, fertilization protein, sperm lysin, VCP1, VCP2, VCP3, VCP7, VCP49, VCP75 and VTG-1 in juvenile abalone (2, 3, 5-month-old *H. asinina*) was examined by RT-PCR. Results showed that tektinA1, FP, VCP1, VCP2, VCP3, VCP7, VCP49, and VCP75 were detected in all ages of juvenile *H. asinina* whereas the axonemal p66.0 and VTG-1 were not expressed in juvenile abalone (Figure 3.48). In addition, sperm lysin was expressed at the low level in 5-month-old juvenile *H. asinina*.

Expression of the newly isolated 1700 bp (contained DM domain) DMRT1 homologue and its variant (1400 bp, without DM domain) were examined in juvenile and testes and ovaries of adult *H. asinina* by 5′UTRDMRT1-F and DMRT1-1700-R (Figure 3.49, A), DMRT1Ha-F and DMRT1-1700-R (Figure 3.49,B) and DMRT1-1700-F and DMRT1-1700-R (Figure 3.49, C). The expected amplification products were 594 bp (1700 bp fragment) and 317 bp (1400 bp fragment), 423 bp (1700 bp fragment) and 221 bp (both fragments). Results indicated that these transcripts were only expressed in testes of adult. Alternative splicing of exons may play an important role for generating these gene products for a single locus gene.

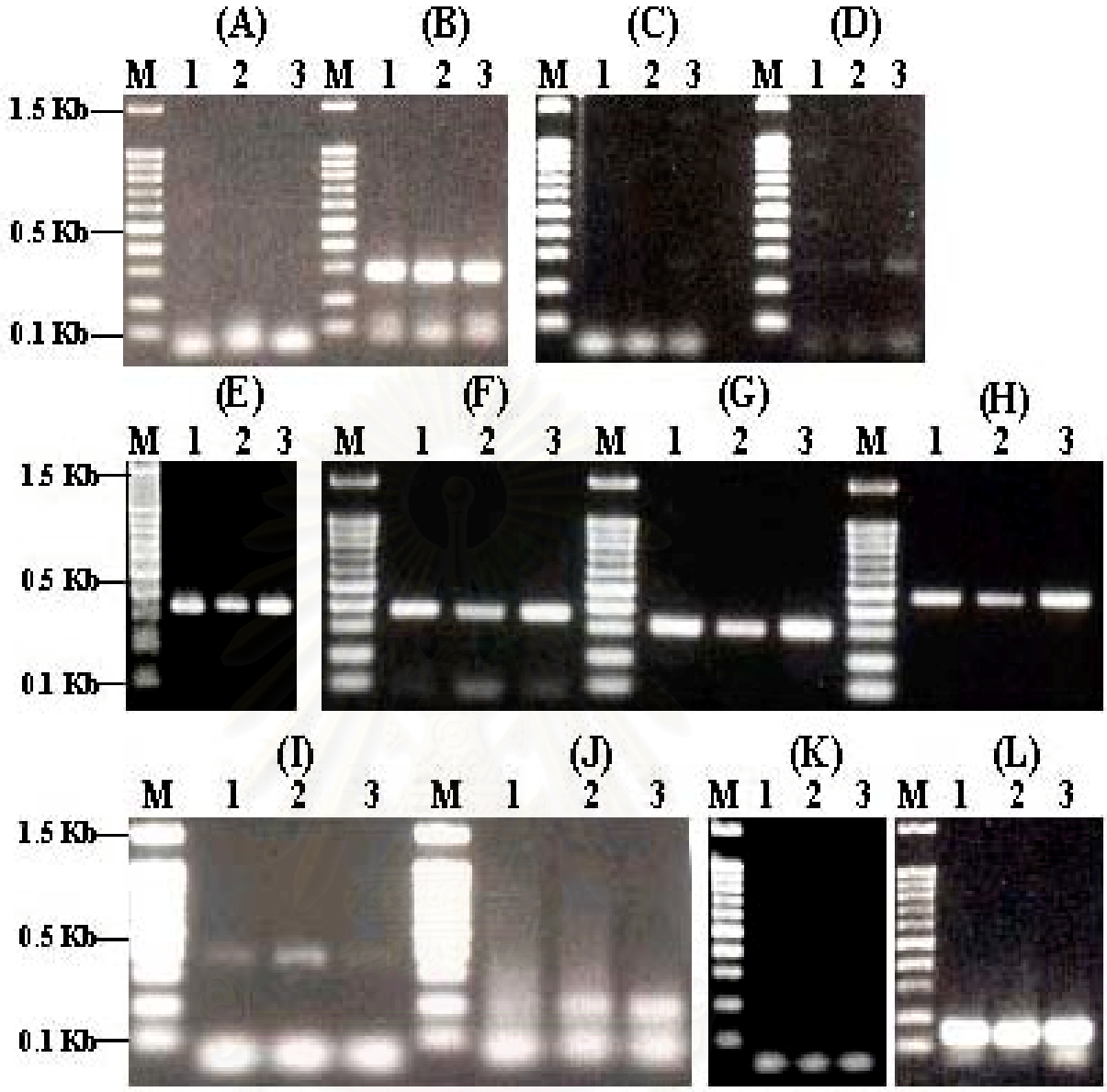


Figure 3.48 RT-PCR of axonemal p66.0 (A), tektinA1 (B), sperm lysin (C), FP (D), VCP1 (E), VCP2 (F), VCP3 (G), VCP7 (H), VCP49 (I), VCP75 (J); and VTG-1 (K). Expression of each gene and EF1α (149 bp; L) was examined in 2-, 3-, and 5-month-old juvenile *H. asinina* (lane 1, 2, and 3, respectively). Lanes M are a 100-bp DNA marker.

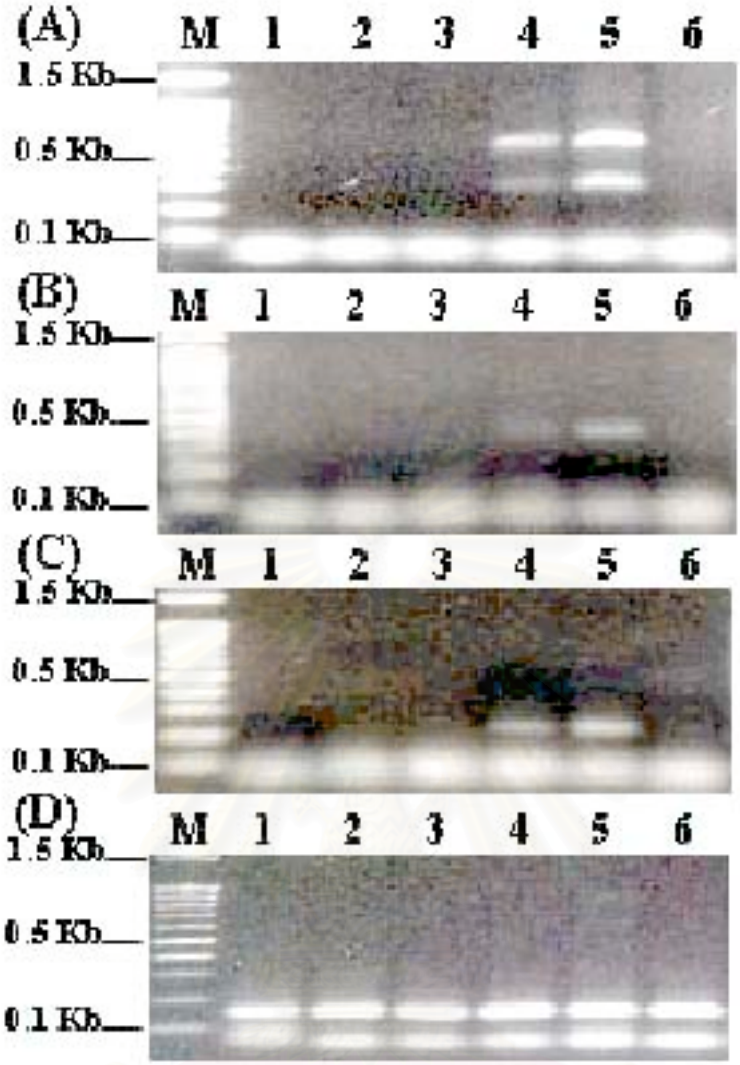


Figure 3.49 RT-PCR of a DMRT1 homologue using 5′UTRDMRT1-F + DMRT1700-F (A), DMRT1Ha-F + DMRT1-1700-R (B), and DMRT1700-F + DMRT1700-R (C) primers. Expression of each gene and EF1α (149 bp; D) was examined in 2-, 3-, and 5-month-old juveniles (lane1, 2, and 3, respectively), testes (lanes 4-5), and ovaries (lane 6) of adult *H. asinina*. Lanes M are a 100-bp DNA marker.

3.4.6 Determining polymorphism at the genomic level of sex-related genes of *H. asinina*

3.4.6.1 PCR amplification

Amplification of VCP1, VCP2, VCP3, VCP7, VCP49, VCP75, VTG-1, and sperm lysin using genomic DNA of *H. asinina* as the template was not successful, whereas sizes of GIOT1, SARIP, tektinA1, ADAMTS-9, and α-tubulin2 gene segments were identical to those expected from cDNA sequences. No length polymorphism was observed in both male and female of *H. asinina.* Non-specific amplification products were observed from amplification of FP (Table 3.17).

Only HSD and axonemal p66.0 gene segments contain a small intron (approximately 60 bp) and a large intron (approximately 575bp), respectively. Sizes of the amplification products were identical in both genders.

Restriction analysis of the amplified axonemal p66.0 fragment of male and female *H. asinina* provided similar patterns when digested with *Alu* I, *Hae* II, and *Ssp* I. Polymorphism was found when the amplified axonemal p66.0 was digested with *Hin*f I but this polymorphism was not fixed in each gender of *H. asinina* and therefore, cannot be used as a sex determination marker at the genomic DNA level.

3.4.6.2 SSCP analysis

Sex-related genes amplified from the genomic DNA of *H. asinina* were also characterized by SSCP analysis for identification whether SNP was existent in those genes. GIOT1, SARIP, tektinA1, ADAMTS-9, α-tubulin2, HSD and axonemal p66.0 gene segment did not exhibit sex-specificity in genomic DNA when analyzed by agarose gel electrophoresis and was further analyzed by SSCP.

Results indicated that all primers did not show sex-specificity. SARIP (Figure 3.50A), tektinA1, and α-tubulin2 were monomorphic across investigated individuals (Table 3.17) but several gene products including GIOT1, HSD, axonemal p66.0, ADAMTS-9 (Figure 3.50B) and DMRT1 were polymorphic but the pattern were not fixed for each gender.

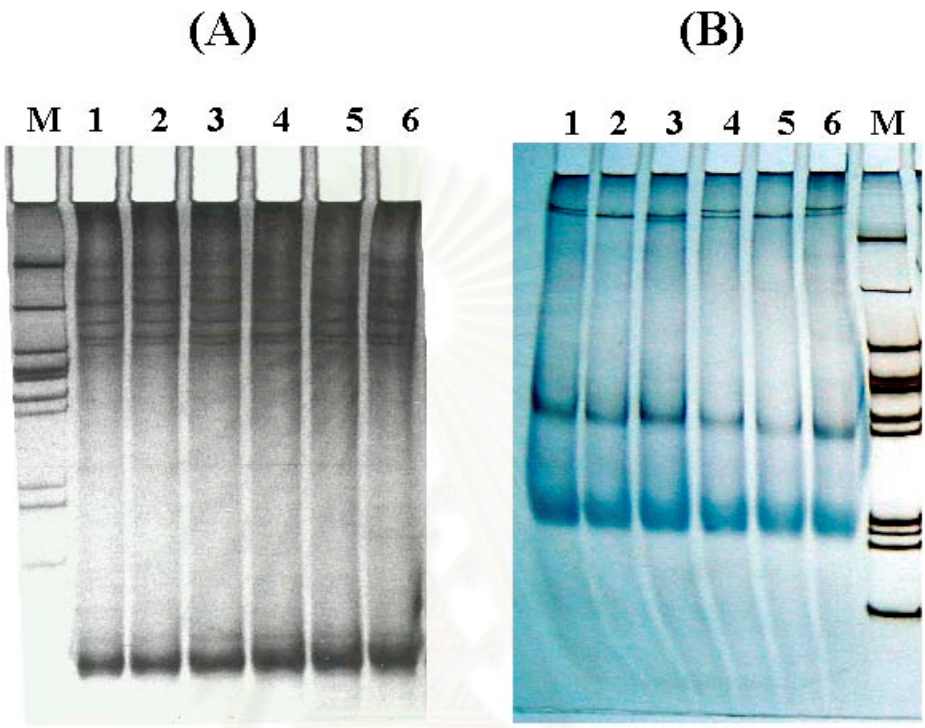


Figure 3.50 SSCP patterns of SARIP (A) and ADAMTS-9 (B) of male (lanes 1– 3) and female (lanes4 – 6) using genomic DNA of *H. asinina* as the template. The PCR product was denatured and electrophoresed through 15% non-denaturing PAGE (37.5%). Lane M is a 100 bp DNA marker. Lane D is non-denatured PCR product.

Table 3.17 Characterization of sex-related gene segments using genomic DNA of *H. asinina* as the template. The PCR products of each gene were analyzed by agarose gel electrophoresis and SSCP analysis.

Primer	Expected product (bp)	Agarose gel electrophoresis		SSCP	
		Male (*N*=3)	Female (*N*=3)	Male (*N*=3)	Female (*N*=3)
VCPg1	346 bp	-	-	ND	ND
VCPg2	363 bp	-	-	ND	ND
VCPg3	316 bp	-	-	ND	ND
VCPg7	467 bp	-	-	ND	ND
VCPHa49	425 bp	-	-	ND	ND
VCPHa75	189 bp	-	-	ND	ND
VTG-1	202 bp	-	-	ND	ND
ADAMTS-9	173 bp	+	+	Polymorphism	Polymorphism
α-Tubulin2	258 bp	+	+	Monomorphism	Monomorphism
Axonemal p66.0	327 bp	940 bp	940 bp	Polymorphism	Polymorphism
TektinA1	282 bp	+	+	Monomorphism	Monomorphism
GIOT-1	208 bp	+	+	Polymorphism	Polymorphism
HSD	231 bp	290 bp	290 bp	Polymorphism	Polymorphism
FP	248 bp	NS	NS	ND	ND
Sperm lysin	262 bp	-	-	ND	ND
SARIP	246 bp	+	+	Monomorphism	Monomorphism
DMRT1	200 bp	+	+	Polymorphism	Polymorphism

Abbreviations: +, successful amplification with the expected product; -, no amplification product; NS, non-specific fragments; ND = not determined.

3.4.7 Semi-quantitative RT-PCR analysis of sex-related genes in gonads of *H. asinina*

Expression levels of sex-related genes in testes (axonemal p66.0, tektinA1, DMRT1, sperm lysin, and fertilization protein) and ovaries (VCP1, VCP2, VCP3, VCP7, VCP49, VCP75, and VTG-1) of *H. asinina* were semi-quantitatively examined. This technique requires optimization of several parameters including concentration of primers, and $MgCl_2$, and the number of PCR cycles.

Twelve primer sets were designed. Primer for EF-1α was included as an internal control. The preliminary RT-PCR was carried out using the standard conditions and the annealing temperature of 55°C. The resulting product of each gene is shown by Figure 3.51 and 3.52. The single discrete band was observed. Therefore, optimization of the primer concentration was performed at the above annealing temperature.

3.4.7.1 Optimization of the primer concentration

RT-PCR of each gene was carried out with fixed components except primer concentrations (0.05, 1.0, 2.0, 3.0 and 0.4 μM). Lower concentrations may result in non-quantitative amplification whereas higher concentrations of primer may leave a large amount of unused primers which could give rise to non-specific amplification products. Results indicated that the suitable concentration of primers for the amplification of tektinA1, fertilization protein, sperm lysin, VCP1, VCP2, VCP3, VCP7. VCP49, VCP75, VTG-1, and EF1α was 0.1 μM and that for axonemal p66.0 and DMRT1 was 0.05 μM, respectively (Figure 3.51 and 3.52 and Table 3.18).

3.4.7.2 Optimization of the $MgCl_2$ concentration

The optimal concentration of $MgCl_2$ (between 1.0 - 5.0 mM) for each primer pair was carefully examined using the standard PCR reaction. The concentration of $MgCl_2$ that gave the highest yields and specificity for each PCR product was chosen (Figure 3.53 and 3.54 and Table 3.18). All primer pairs gave the most specific results when 1.0 mM $MgCl_2$ was used in the amplification reaction.

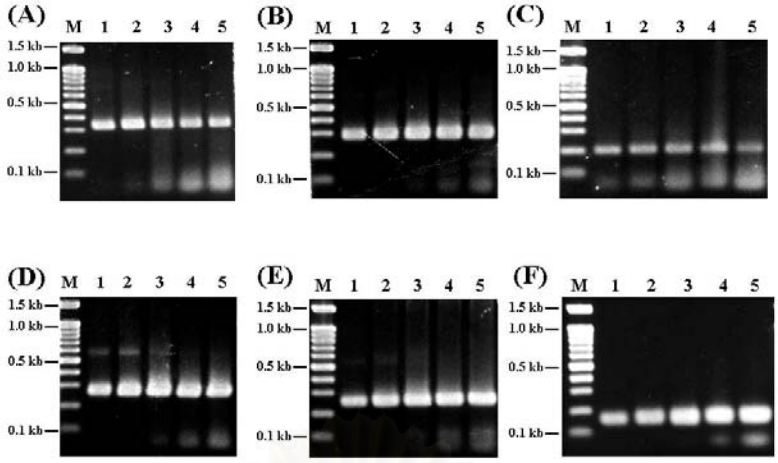


Figure 3.51 Optimization of primer concentrations for RT-PCR of axonemal p66.0 (A), tektinA1(B), DMRT1(C), sperm lysin(D), FP(E) and EF-1α (F) homologues. Primer concentrations were varied between 0.05, 1.0, 2.0, 3.0 and 0.4 μM corresponding to lanes 1 – 4 (A – G), respectively. Lames M are a 100 bp ladder.

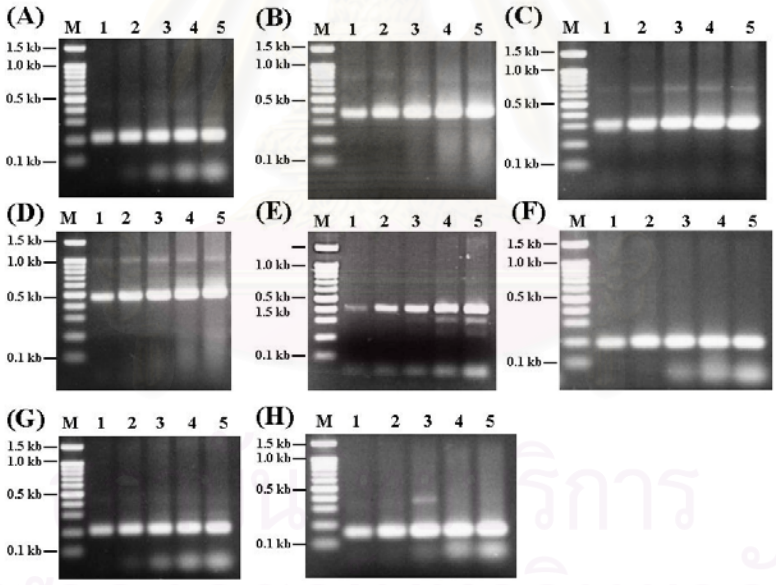


Figure 3.52 Optimization of primer concentrations for RT-PCR of VCP1 (A), VCP2 (B), VCP3 (C), VCP7 (D), VCP49 (E), VCP75 (F) and EF-1α (G) homologues. Primer concentrations were varied between 0.05, 1.0, 2.0, 3.0 and 0.4 μM corresponding to lanes 1 – 4 (A – G), respectively. Lames M are a 100 bp ladder.

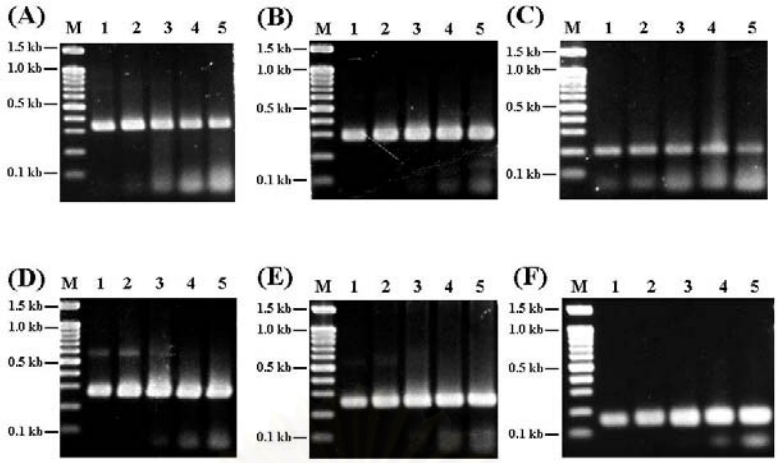


Figure 3.53 Optimization of $MgCl_2$ concentrations for RT-PCR of axonemal p66.0 (A), tektinA1(B), DMRT1(C), sperm lysin(D), FP (E) and EF-1α (F) homologues. The $MgCl_2$byconcentrations were varied between 1.0, 2.0, 3.0, 4.0 and 5.0 μM corresponding to lanes 1 – 5 (A – F), respectively. Lames M are a 100 bp ladder.

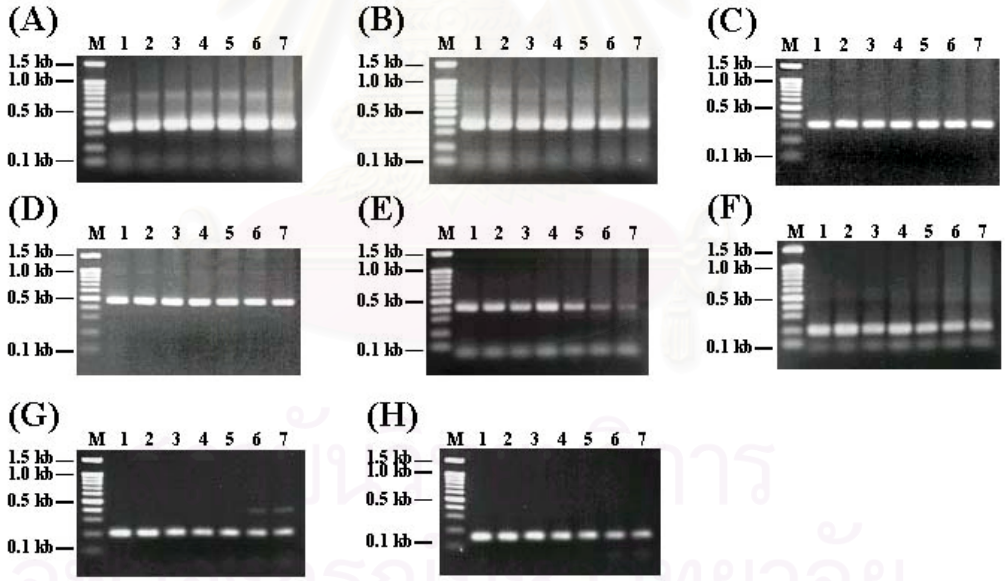


Figure 3.54 Optimization of $MgCl_2$ concentrations for RT-PCR of VCP1 (A), VCP2 (B), VCP3 (C), VCP7 (D), VCP49 (E), and EF-1α (F) homologues. The $MgCl_2$byconcentrations were varied between 1.0, 1.5, 2.0, 2.5, 3.0, 4.0 and 5.0 μM corresponding to lanes 1 – 7, respectively. Lames M are a 100 bp ladder.

3.4.7.3 Optimization of the cycle numbers

The number of amplification cycles was important because the product reflecting the expression level should be measured quantitatively before reaching a plateau amplification phase. At the plateau stage, transcripts initially present at different levels may give equal intensity of the amplification products.

In this experiment, RT-PCR of each gene was performed using the optimal conditions for 18 - 36 cycles with a 3 cycle increment. The number of cycles that gave the highest yield before the product reached a plateau phase of amplification was chosen. Homologues of axonemal p66.0, tektinA1, DMRT1, sperm lysin, and FP reached a plateau of amplification after 24, 24, 26, 17, and 20 cycles, respectively. In contrast, the most optimal cycle number for VCP1, VCP2, VCP3, VCP7, VCP49, VCP75, VTG-1, and EF-1α was 20, 21, 24, 18, 24, 25, 21, and 21 cycles (Figure 3.55 and 3.56), respectively.

3.4.7.4 A-single-tube semi-quantitative RT-PCR

Amplification of the target and the internal control can be carried out together or separately. The former refers to a-single-tube semi-quantitative PCR where competitive effects between target and control primers may be occurred and result in decreasing of the intensity of the target and/or the control products when compared with those separately amplified.

Primers for amplifications of axonemal p66.0, tektinA1 and DMRT1 did not compete with EF-1α primers when 1.0 mM $MgCl_2$ at was included in a single tube PCR and amplification for 24, 24, 26 cycles, respectively (Figure 3.57-3.58). For other primer, competitive effects of primers were observed.

To eliminate primer competition, $MgCl_2$ concentration was further adjusted (1.0 to 3.0 mM) (Figure 3.59-3.62). Although competition between VCP1, VCP2, VCP3, and VTG-1 and EF-1α primers was not observed, the single tube RT-PCR of these genes was not consistent. Therefore, only axonemal p66.0, tektinA1 and DMRT1 were quantitatively examined using the single tube RT-PCR. On the other hand, sperm lysin, FP, VCP1, VCP2, VCP3, VCP7, VCP49, VCP75 and VTG-1 were separately amplified and quantitatively estimated using EF-1α as the reference.

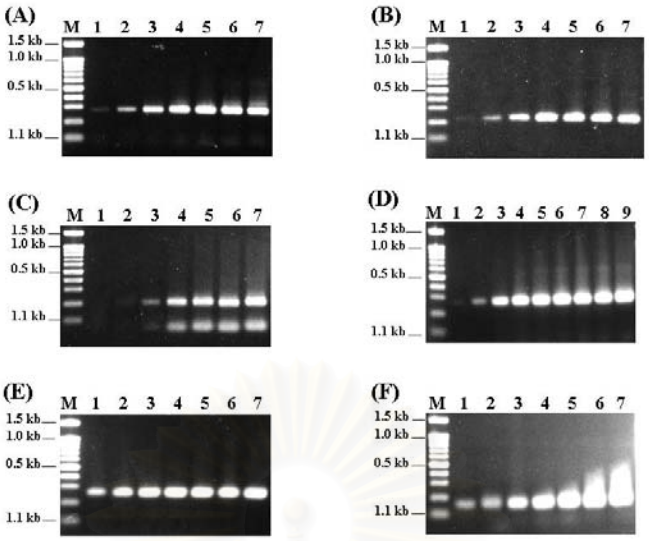


Figure 3.55 Optimization of the cycle numbers for RT-PCR of axonemal p66.0 (A), tektinA1 (B), DMRT1 (C), sperm lysine (D), FP (E) and EF-1α (F) homologues. RT-PCR of each gene was carried out for 18, 21, 24, 27, 30, 33 and 36 corresponding to lanes 1 – 7 (A – F), respectively. Lanes M are a 100 bbp ladder.

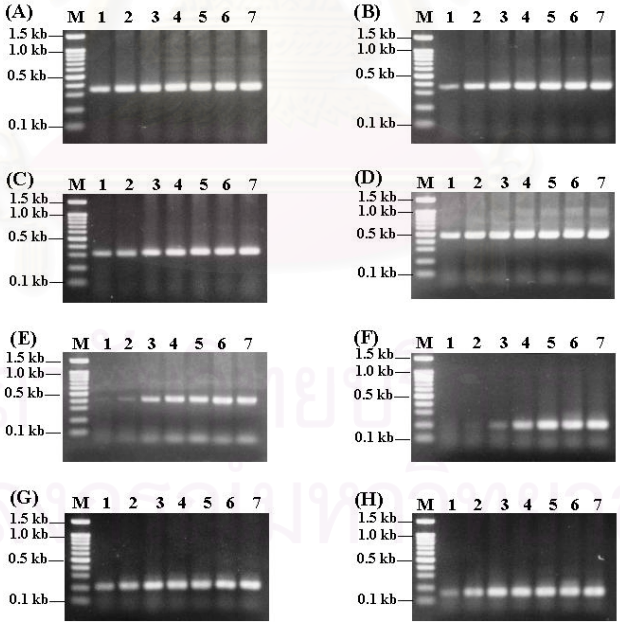


Figure 3.56 Optimization of the cycle numbers for RT-PCR of VCP1 (A), VCP2 (B), VCP3 (C), VCP7 (D), VCP49 (E), and EF-1α (F) homologues. RT-PCR of each gene was carried out for 18, 21, 24, 27, 30, 33 and 36 corresponding to lanes 1 – 7 (A – F), respectively. Lanes M are a 100 bp ladder.

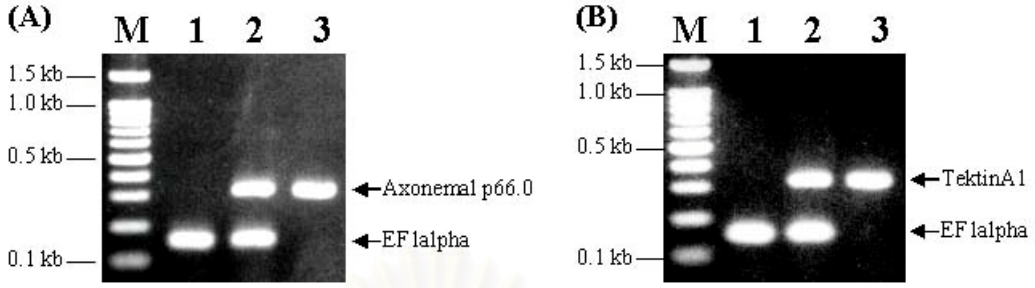


Figure 3.57 A single tube RT-PCR of the target; axonemal p66.0 (panel A, 1.0 mM $MgCl_2$) and tektinA1 (panel B, 1.0 mM $MgCl_2$) and internal control (EF-1α) was carried out. Lanes 1 and 3 = amplicons of the target and the internal control when amplified separately. Lanes 2 = amplicons of the target and the control when a single tube RT-PCR was performed. Arrows indicate the PCR product of axonemal p66.0 (327 bp), tektinA1 (282 bp) and EF1α (149 bp), respectively.

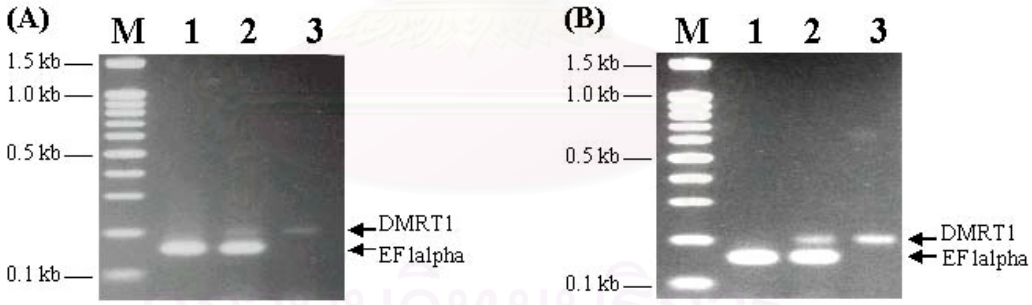


Figure 3.58 A single tube RT-PCR of DMRT-1 (panel A, 1.0 mM $MgCl_2$) and tektinA1 (panel B, 1.5 mM $MgCl_2$) and internal control (EF-1α) was carried out. Lanes 1 and 3 = amplicons of the target and the internal control when amplified separately. Lanes 2 = amplicons of the target and the control when a single tube RT-PCR was performed. Arrows indicate the PCR product of DMRT1 (200 bp) and EF-1α (149 bp), respectively. Competition between different set of primers were observed when 1.5 mM $MgCl_2$ was used.

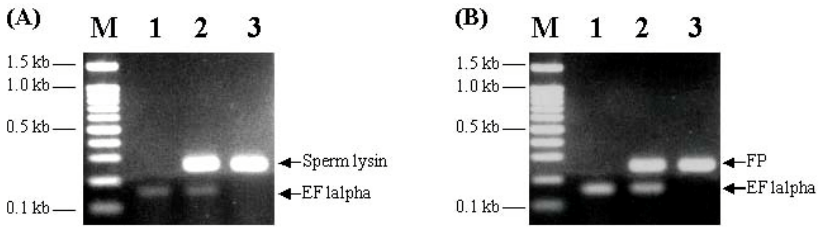


Figure 3.59 A single tube RT-PCR of sperm lysin (panel A, 1.0 mM $MgCl_2$) and FP (panel B, 1.0 mM $MgCl_2$) and internal control (EF-1α). Lanes 1 and 3 = an amplicon and the internal control primers when amplified separately. Lanes 2 = amplicons of the target and the control when a single tube RT-PCR was performed. Arrows indicate the PCR product of sperm lysin (262 bp), FP (248 bp) and EF- 1α (149 bp), respectively.

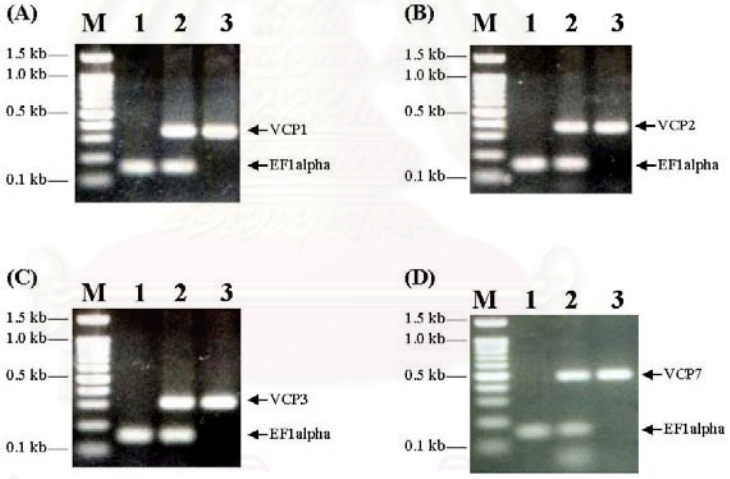


Figure 3.60 A single tube RT-PCR of VCP1 (panel A, 1.0 mM $MgCl_2$), VCP2 (panel B, 1.0 mM $MgCl_2$), VCP3 (panel C, 1.0 mM $MgCl_2$) and VCP7 (panel D, 1.0 mM $MgCl_2$) and the internal control (EF-1α). Lanes 1 and 3 = amplicons of the target and the internal control when amplified separately. Lanes 2 = amplicons of the target and the control when a single tube RT-PCR was performed. Arrows indicate the PCR product of VCP1(346 bp), VCP2 (363 bp), VCP3 (316 bp), and VCP7 (467 bp) and EF- 1α (149 bp), respectively.

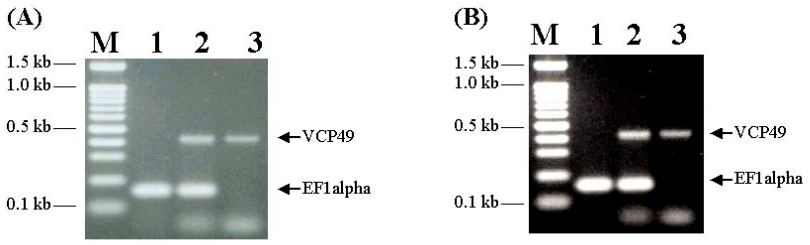


Figure 3.61 A single tube RT-PCR of VCP49 (1.0 mM $MgCl_2$ panel A, and 1.5 mM $MgCl_2$, panel B) and internal control (EF-1α). Lanes 1 and 3 = amplicons of the target and the internal control when amplified separately. Lanes 2 = amplicons of the target and the control when a single tube RT-PCR was performed. Arrows indicate the PCR product of VCP49 (425 bp) and EF1α (149 bp).

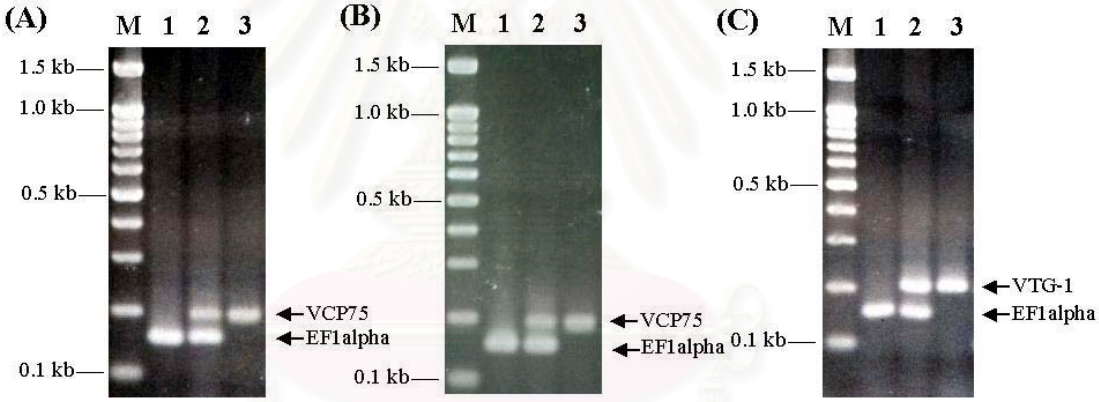


Figure 3.62 A single tube RT-PCR of VCP75 (1.0 mM $MgCl_2$, panel A, and, 1.5 mM $MgCl_2$, panel B) and VTG-1 (1.0 mM $MgCl_2$, panel C) and internal control (EF-1α). Lanes 1 and 3 = amplicons of the target and the internal control when amplified separately. Lanes 2 = amplicons of the target and the control when a single tube RT-PCR was performed. Arrows indicate the PCR product of VCP75 (189 bp), VTG-1 (202 bp) and EF-1α (149 bp), respectively.

The optimal primer and $MgCl_2$ concentrations, and the number of PCR cycles used for semi-quantitative RT-PCR analysis of sex-related genes in this study are summarized in Table 3.18.

Table 3.18 Optimal primer and $MgCl_2$ concentrations and the number of PCR cycles for semi-quantitative analysis of sex-related genes in *H. asinina*

Transcript	Expected amplicons (bp)	Primer concentration (μM)	$MgCl_2$ concentration (mM)	PCR cycles
Axonemal p66.0	327 bp	0.05 μM	1.0 mM	24
TektinA1	282 bp	0.1 μM	1.0 mM	24
DMRT1	200 bp	0.05 μM	1.0 mM	26
Sperm lysine	262 bp	0.1 μM	1.0 mM	17
FP	248 bp	0.1 μM	1.0 mM	20
VCP1	346 bp	0.1 μM	1.0 mM	20
VCP2	363 bp	0.1 μM	1.0 mM	21
VCP3	316 bp	0.1 μM	1.0 mM	24
VCP7	467 bp	0.1 μM	1.0 mM	18
VCP49	425 bp	0.1 μM	1.0 mM	24
VCP75	189bp	0.1 μM	1.0 mM	25
VTG-1	202 bp	0.1 μM	1.0 mM	21
EF1α	149 bp	0.1 μM	1.0 mM	21

3.4.7.5 Expression level of sex-related genes in different developmental stages of testes and ovaries of *H. asinina*

Semi-quantitative PCR was then carried out for determining the expression levels of axonemal p66.0, tektinA1 DMRT1, sperm lysin and fertilization protein of different developmental stages of testes and those of VCP1, VCP2, VCP3, VCP7, VCP49, VCP75, and VTG-1 of developmental stages of ovaries of *H. asinina*. Testes of male abalone exhibiting the 1st (N =4), the 2nd (N =4), the 3rd (N =4) and the 4th (N =4) testicular stages were colleted and used as the starting materials for total RNA extraction and subsequently, for the first strand cDNA synthesis. Likewise, the 1st (N =5) and 3rd (N =5) stages of ovaries were dissected out from female abalone and subjected to total RNA extraction and the first strand cDNA synthesis.

RT-PCR analysis revealed out against. The amplification products were size-fractionated through a 1.8% agarose gels (Figure. 3.63-3.65 and Table 3.19 and 3.20). The relative expression values were estimated from the ratio between the band intensity of the target and that of EF-1α.

Results indicated significant differences between the expression level of axonemal p66.0, tektinA1, DMRT1 of 1st developmental stage of testes with subsequent stages ($P < 0.05$) but not between stages 2, 3 and 4 of testes. Neither sperm lysin nor FP exhibited significant differences of expression at all stages of testicular development.

While VTG-1 did not show differential expression patterns between the early (1st) and the late (3rd) stages of ovarian development, VCP1, VCP2, VCP3, VCP7, VCP49, and VCP75 revealed greater expression levels in the early stage of ovarian development than did the late stage.

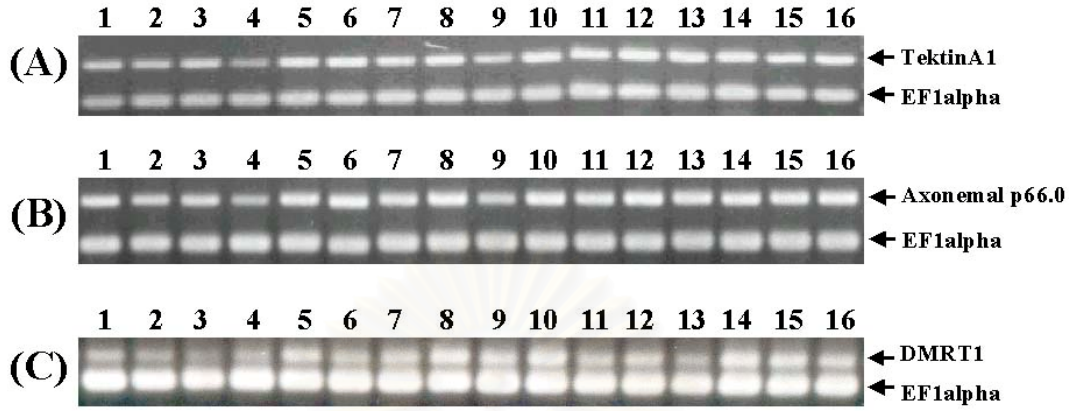


Figure 3.63 Semi-quantitative RT-PCR analysis of tektinA1 (A), axonemal p66.0 (B), and DMRT1 (C) of *H. asinina* individuals exhibiting the 1st (lanes 1 – 4), the 2nd (lanes 5 – 8), the 3rd (lanes 9 – 12), and the 4th (lanes 13 – 16) developmental stages of testes.

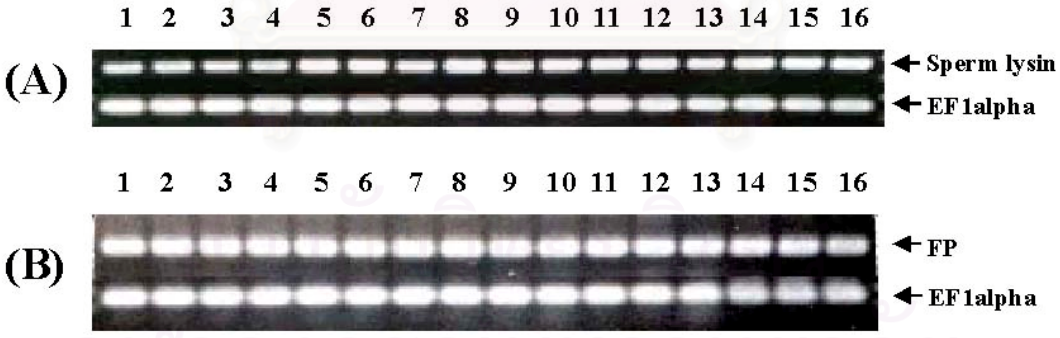


Figure 3.64 Semi-quantitative RT-PCR analysis of sperm lysin (A) and FP (B) of *H. asinina* individuals exhibiting the 1st(lanes 1 – 4), the 2nd (lanes 5 – 8), the 3rd (lanes 9 – 12), and the 4th (lanes 13 – 16) developmental stages of testes.

Figure 3.65 Semi-quantitative RT-PCR analysis of expression levels of VCP1 (A), VCP2 (B), VCP3 (C), VCP7 (D), VCP49 (E), VCP75 (F), and VTG-1 (G) of *H. asinina* individuals exhibiting 1st (lanes 1 – 5) and 3rd (lanes 6 – 10) developmental stages of ovaries.

Table 3.19 Relative expression values* of sex-related transcripts (axonemal p66.0, tektinA1, DMRT1, sperm lysin, and FP) in testes (stages 1 – 4) of *H. asinina*

Transcripts	State 1	State 2	State 3	State 4
Axonemal p66.0	0.917 ± 0.055^{a}	1.060 ± 0.019^{b}	1.069 ± 0.016^{b}	1.064 ± 0.009^{b}
TektinA1	0.685 ± 0.030^{a}	0.969 ± 0.155^{b}	1.019 ± 0.120^{b}	1.083 ± 0.004^{b}
DMRT1	0.200 ± 0.010^{a}	0.271 ± 0.016^{b}	0.273 ± 0.016^{b}	0.277 ± 0.022^{b}
Sperm lysin	1.146 ± 0.004^{a}	1.162 ± 0.012^{a}	1.159 ± 0.009^{a}	1.233 ± 0.123^{a}
FP	1.157 ± 0.008^{a}	1.166 ± 0.006^{a}	1.167 ± 0.009^{a}	1.166 ± 0.006^{a}

*Expression level of the target/expression level of EF-1α. Different superscripts indicated significant different expression levels ($P < 0.05$).

Table3.20 Relative expression values* of sex-related transcripts (VCP1, VCP2, VCP3, VCP7, VCP49, VCP75, and VTG-1) in ovaries (stages 1 and 3) of *H. asnina*

Transcripts	Stage 1	Stage 3
VCP 1	1.422 ± 0.153^{a}	1.385 ± 0.017^{b}
VCP 2	1.479 ± 0.008^{a}	1.331 ± 0.005^{b}
VCP 3	1.482 ± 0.006^{a}	1.346 ± 0.010^{b}
VCP 7	1.484 ± 0.006^{a}	1.358 ± 0.007^{b}
VCP 49	0.842 ± 0.138^{a}	0.725 ± 0.031^{b}
VCP 75	0.969 ± 0.094^{a}	0.705 ± 0.023^{b}
VTG-1	0.957 ± 0.009^{a}	0.933 ± 0.022^{b}

*Expression level of the target/expression level of EF-1α. Different superscripts indicated significant different expression levels ($P < 0.05$).

CHAPTER IV

DISCUSION

I: Development of species-specific markers of the tropical abalone (*H. asinina*, *H. ovina*, and *H. varia*) in Thailand using RAPD and SCAR markers.

RAPD-PCR has widely been used for population genetic studies and identification of molecular markers for various applications in several organisms (Heipel et al. 1998; Tassanakajon et al. 1998; Klinbunga et al. 2001a and 2001b). Analysis of genetic diversity and population differentiation of *H. asinina*, *H. ovina* and *H. varia* using PCR-RFLP (Jarayabhand et al. 2002; Klinbunga et al. 2002) and RAPD-PCR (Popongviwat 2001) approaches revealed a lack of genetic heterogeneity in *H. asinina* over vast geographic locations ($P > 0.05$) and significant genetic differences between *H. asinina*, *H. ovina* and *H. varia* ($P < 0.0001$). This suggested the possibility to develop a large number of species-specific RAPD markers in *H. asinina*.

Nevertheless, RAPD-PCR is sensitive to reaction conditions, including the requirement of a good quality DNA template for consistent results. False negatives may possibly occur. The candidate species-specific (and population-specific) RAPD fragments should be converted to sequence-characterized amplified region (SCAR) markers for consistent results (Weising *et al*., 1995).

Differentiation of a commercially exploited abalone, *H. midae* and a sympatrically congeneric *H. spadicea* was reported based on species-specific PCR and PCR-RFLP of the sperm lysin gene (Lee and Vacquier 1995). A 1300 bp fragment was specifically amplified from genomic DNA of *H. midae*. Discrimination of these abalone could also be carried out by digestion of the amplified fragment with *Cfo* I, *Dra* I, *Taq* I and *Hin*f I. Species-specific PCR provided a 146 bp fragment to verify canned *H. midae* where DNA quality was poor and only low molecular weight DNA template was available (Sweijd et al.1998).

In the present study, 10, 2, and 3 fixed RAPD markers were found in *H. asinina*, *H. ovina*, and *H. varia*. Three candidate population-specific RAPD markers were

observed in *H. asinina* that originated from Talibong Island (HATRAW) and the Philippines (HAPHIE), respectively. In addition, a 1,650 bp RAPD marker that was generated from UBC195 was specifically found in *H. ovina*, originating from the Gulf of Thailand (east) but not in the Andaman Sea sample (west).

Twenty-two RAPD fragments were cloned. Seventy-two clones were sequenced. Thirty-nine different sequences were found. These indicated that the RAPD fragments represented co-migrating fragments that had different nucleotide sequences but similar sizes. Almost all of the RAPD markers (33/39 accounting for 85%) were unknown sequences when compared with the data in the GenBank using Blast*N* and Blast*X* (E values >10-4).

Based on the preliminary screening, seven pairs of primers (CUHA1, CUHA2, CUHA4, CUHA11, CUHA12, CUHA13, and CUHA14) revealed species-specificity in *H. asinine*, while the CUHO3 and CUHV1 primers exhibited the expected product in *H. ovina* and *H. varia*, respectively. CUHA14 was not selected for the analysis against larger specimens of *H. asinina* because different sizes of the amplification products were observed (expected 473 bp in HACAMHE and HACAME and a larger 515 bp in HATRAW and HAPHIE, respectively).

Five sets of primers (CUHA2, CUHA12, CUHA13, CUHO3, and CUHV1, Fig. 3) were further examined against a large sample size of abalone (N = 216). Species-specificity was observed from CUHA2, CUHA12, CUHA13 and CUHV1 (100% without false positive/negative results). Although CUHO3 yielded a strong amplification product in the target species (100%), a very faint product was also observed in some individuals of *H. asinina* and *H. varia*, which suggests that problems may arise from the non-specific amplification of this primer pair. No heterozygotes that exhibited two different sizes (alleles) of the amplification products were observed across the overall specimens that were analyzed by these primers. This implies the retention of a dominant segregated fashion of the original RAPD markers.

The sensitivity of CUHA2, CUHA12, and CUHV1 was tested. The detection sensitivity was approximately 100 pg and 500 pg of the target DNA template for CUHA2 and CUHA12 and CUHA13 but a greater sensitivity (20 pg) was found with CUHV1. The sensitivity levels of species-specific PCR that was developed in this study were

sufficient for the identification of species-origins of abalone, beginning with the early development stages.

A limited sample size of *H. varia* was included in this study (*N* = 29 and 3 for HVPHAW and HVPHUW, respectively). This was due to a lack of *H. varia* that was found in our sampling sites. Accordingly, a *H. varia*-specific SCAR marker (229 bp from CUHV1) should be further tested against specimens that cover a larger geographic distribution of *H. varia* before practical implementation of this marker for the species-identification purpose.

Several candidate population (or region)-specific RAPD markers were found in this study. Nine SCAR primer sets were designed for development of specific SCAR markers in a particular sample of *H. asinina*. Neither population- nor region- specific SCAR markers were observed in *H. asinina*.

Recently, species-diagnostic markers of *H. asinina*, *H. ovina* and *H. varia* were developed using PCR-RFLP of 16S rDNA (Jarayabhand et al. 2002). Restriction of the amplified 16S rDNA with *Alu* I could differentiate these abalone unambiguously (patterns A and E in *H. asinina*, *N* = 100; B in *H. ovina*, *N* = 71 and C, D, F and G in *H. varia*, *N* = 23). Further digestion of 16S rDNA with *Bam* HI, *Eco* RI and *Hae* III yielded non-overlapping composite haplotypes in these abalone; AAAA and AAAE in *H. asinina*, ABBB, AAAB and AABB in *H. ovina* and BABG, BABC, BABD, BABF and AABG in *H. varia*, respectively (Klinbunga et al. 2002).

Moreover, sequences of 16S rDNA of representative individuals exhibiting those composite haplotypes were examined. Species-specific PCR based on 16S rDNA polymorphism in these abalone was developed. Like results from this study, species-diagnostic markers were successfully developed in *H. asinina* and *H. varia* (100% amplification success without any false positive) but not in *H. ovina* (68.75% amplification success with extensive false positives from other species) (Klinbunga et al. 2003).

Thaewnon-ngiw et al (2004) study genetic diversity and characterized species-specific markers of the introduced golden apple snail, *Pomacea canaliculata* (Lamarck, 1801) and four native apple snails; *Pila ampullacea* (Linneaus, 1758), *P. angelica* (Annandale, 1920), *P. pesmei* (Morelet, 1889) and *P. polita* (Deshayes, 1830) in Thailand

using RAPD analysis. The candidate species-specific RAPD markers found in *Pomacea canaliculata* (340 bp, OPB10), *P. ampullcea* (640 bp, OPA07), *P. angelica* (380 bp, UBC122) and *Pila* snails (430 bp, OPA07) were converted to SCAR markers by cloning and sequencing. A 259 bp SCAR marker was found in 95.0% of *Pila* apple snails (*N* = 163) but not in *Pomacea canaliculata* (*N* = 30). This indicated the successful development of a specific SCAR marker for detection of *Pila* apple snails in Thailand. Therefore, this SCAR marker could be used in coupling with a *Pomacea canaliculata*-specific RAPD marker to unambiguously differentiate the introduced and native apple snails in Thailand.

Interspecific hybridisation and gene introgression between *H. rubra* and *H. laevigata* was reported based on allozyme analysis (Brown 1995). Theoretically, hybridazation between male *H. asinina* with females of different species and backcrosses between subsequent generations of hybrids could have eliminated the accuracy of the developed markers. The existence of both 16S rDNA (mtDNA) and several nuclear DNA markers allow the possible to detect those phenomena, if occurred, when specimens from additional geographic samples are examined.

Based on immunological techniques, Lopata et al (2002) developed abalone-specific antibodies to identify species origin of different abalone from South Africa (*H. midae and H. spadicea*), Australia (*H. rubra*), and Japan (*H. diversicolor supertexa*, *H. diversicolor diversicolor*, *H. sieboldii*, *H. discus discus*, *H. discus hannai*, *H. gigantea*, and *H. laevigata*). Three developed IgG1 monoclonal antibodies (MoAb) against heat-stable proteins from *H. midae* including MoAb1.4 (45 kDa and 35 kDa), 2.11 (45 kDa and 35 kDa), and 2.12 (45 kDa) were identified. The combination of those enabled differentiation between South African abalone, *H. midae* (MoAb 1.4, 45 kDa, 2.11, 45 kDa, and 2.12 45 kDa) and *H. spadicea* (MoAb 1.4, 45 kDa and 35kDa, 2.11, 45 kDa and 35kDa) from other species. Moreover, *H. diversicolor supertexa* (MoAb 1.4, 45 kDa, and 2.11, 45 kDa) could be differentiate from *H. diversicolor diversicolor* (MoAb 1.4, 45 kDa and 35 kDa, 2.11, 45 kDa, and 2.12, 45 kDa) by these antibody detection.

Species-specific PCR in this study was performed at a high stringent annealing temperature (65 °C) to eliminate the possible problems from false positive results. A tedious and time-consuming phenol/chloroform DNA extraction method was simplified to a rapid 5% chelex-based method. The positive fragment was still consistently obtained

from various forms (frozen, ethanol-preserved, dried and boiled specimens) of *H. asinina* using CUHA2, CUHA12 and CUHA13.

DNA extraction was simplified for more convenient applications and implementation of the identification techniques. Results indicated that poor DNA quality extracted by a 5% chelex-based method also provided comparable detection ability with that from the phenol-chloroform exteaction method. Therefore, tedious and time consuming steps were eliminated.

A large number of the remaining clones possessing *H. asinina*-specific RAPD inserts would guarantee that additional SCAR markers could be further developed if diagnostic markers that are described here fail to provide species-specific results when used to examine the species-origins of new populations of *H. asinina*.

II: Isolation and characterization of sex-specific markers in *H. asinina* using AFLP analysis.

Sex determination is problematic in researches of many species. This can usually be solved by the application of DNA-based technology but this is only possible if a sex-specific markers located on the unique sex chromosomes is available. Several molecular genetic approaches can be used to identify sex-identification/differentiation markers. Basically, molecular genetic markers at the DNA level are ideal for sex identification purposes. When these markers are not available, sex-specific/differential expression markers have alternatively been studies in several species.

Identification of sex-specific markers at the DNA level in species that does not contain sex chromosomes is difficult and usually not possible. Delvin and Nagahama (2002) described that genetic determination of sex in fish involves monogenic and polygenic systems, with factors located on autosomes or on sex chromosomes. In the latter case, both male (XY) and female (ZW) heterogametic as well as several subtle variations have been described. Unfortunately, sex chromosomes are found in only 10% of fish species examined. Therefore, it is very difficult to identify sex-diagnostic markers (100% accuracy) even in fishes.

Li et al. (2002) collected the green spotted pufferfish, *Tetraodon nigroviridis* (*N* = 83) and used for identification of genomic sex markers using RAPD (600 primers and 1700 primers for the first and the second sets of pooled DNA), AFLP (64 primer

combinations) and representational difference analysis (RDA, 1 set of adaptors) methods. A total of 59, 126, 16 and 16 putative sex-specific markers were found after the primary screening. Nevertheless, secondary screening (re-testing of DNA from individuals for RAPD and AFLP and using the putative RDA markers as the probes for genomic Southern analysis of male and female DNA) did not demonstrate the presence of sex-specific marker in *T. nigroviridis.*

Sex-specific markers should be developed from fixed polymorphism in genomic DNA of males and females of the interesting species to avoid destruction of specimens. For species exhibiting well defined sex chromosomes, molecular approaches at the genomic levels for example, RAPD, AFLP and Genome-walk analysis, are useful for identification of sex-linked markers. The isolated marker can be converted to SCAR markers for direct application for the sex determination purpose.

For species that does not exhibit sex chromosomes like *H. asinina*, both molecular approaches at the genomic DNA (AFLP and RAPD) and cDNA level (RT-PCR, cDNA subtraction, typical differential display (DD)-PCR, RNA arbitrary-primed (RAP)-PCR and cDNA-AFLP may be used for isolation and characterization of sex-specific/differential expression markers.

AFLP is a powerful technique for identification of sex-specific markers where the species of interest possesses sex heterochromatic systems. Griffiths et al. (2000) isolated sex-specific markers of (*Gasterostus aculeatus*) by AFLP. Pre-sexed specimens from geographically different locations were collected from Edinburgh, Scotland (N = 16), Milngavie, Scotland (N = 5), Silverdale, England (N = 8) west coast of Canada (N = 6) and various population in Japan (N = 10). The nine-spined Pungitus pungitus (N = 8) and 15-spined sticklebacks Spinachia spinachia (N = 6) collected from Sweden were included as the control. Three males and female individuals were screened for sex-specific AFLP markers. Primers produced bands in one sex were re-screened with additional 6 individuals. Two candidate male-specific markers were found from primers E_{AAG}/M_{CAA} and E_{AAG}/M_{CAG}, respectively. SCAR markers were developed. Primers GalF/R yielded a 600 bp fragment in females and both 600 bp and 371 bp fragments in males. Therefore, it was concluded that this species possesses the XX/XY sex differentiation system rather than the ZW/ZZ system previously reported by Beatty (1964).

In the present study, 14 candidate sex-specific AFLP fragments (seven candidate fragments for males and females) identified from 214 primer combinations against 4 bulked DNA were cloned and characterized. Fragment-speecific primers were designed. The positive SCAR product with the identical size was amplified from genomic DNA of both males and females *H. asinina*. It was possible that the amplification product may contain fixed SNP between genders. SSCP analysis was then applied but polymorphism of the developed markers was not sex-linked.

Thumumrungtanakit (2004) isolated five (FE10M9520, FE10M10725, FE14M16340, FE15M14400 and FE16M8350) and one (ME10M8420) candidate female and male-specific AFLP fragments in *P. monodon* (species does not show sex heterochromatic chromosomes) using 256 primer combination against 6-10 bulked *P. monodon* genomic DNA. All candidate fragments were cloned and further characterized. SCAR markers derived from FE10M9520, FE10M10725.1, FE10M10725.2, FE14M16340 were successfully amplified but did not retain the original sex specificity. SSCP analysis was applied to identify whether fixed SNP between genders was existent in the amplified SCAR. Like results from this study, polymorphism of the SCAR fragments was found but it did not reveal sex-linked patterns in *P. monodon*.

The giant freshwater prawn (*M. rosenbergii*) is particularly interesting cultured species with respect to variation in growth and the role of size in the social structure of populations. Mature females grow more slowly than males of similar sizes. Therefore, the basic knowledge about the sex determination/differentiation system in this species, if discovered, allow further applications leading to significant increase the production through all male (or the majority of male) culture.

Sex chromosomes have not been cytologically identified in *M. rosenbergii* but the previous study on mating of neomales (females transplanted with the androgenic glands and the sex of female prawns are reversed to males) with normal females, the sex ratio of the offspring from each cross was biased towards the ZW/ZZ system. It was then expected that W-linked AFLP markers may be possibly isolated.

Preechaphol (2004) identified sex-specific AFLP markers of *M. rosenbergii* using 64 AFLP primer combinations against 2 bulked genomic DNA of small orange claw male (SOC1, $N = 10$ and SOC2, $N = 5$), 1 of orange claw male (OC1, $N = 15$), 2 of blue claw

male (BC1, *N* = 5 and BC2, *N* = 10), and 2 of female (F1, *N* = 10 and F2, *N* = 20). A total of 5 and 4 male-and female-specific AFLP fragments were consistently found. After characterization, they were regarded as the unknown DNA segments. SCAR markers were developed from all candidate sex-specific fragments. Among 13 primer pairs tested, 9 pairs of primers generate the expected amplification product in both male and female *M. rosenbergii*. Like studies in *H. asinina* and *P. monodon*, further analysis of the amplification product by SSCP analysis revealed polymorphism of all markers except $FE_{+3}8M_{+3}270.2$–F/R but polymorphic patterns did not show sex-specificity in *M. rosenbergii*.

Losses of sex-specificity of AFLP fragments in *H. asinina*, *P. monodon* and *M. rosenbergii* were resulted from shared sequences within the amplification regions of males and females in those respective species. Moreover, this opens the possibility that polymorphism between genders of *H. asinina* may have occurred at the restriction sites of *Eco* RI and *Mse* I during the AFLP approach. To verify this circumstance, a genome walk analysis (Siebert *et al*., 1995) should be applied. The 5' and 3' resulting fragments can be cloned and sequenced. Point mutations at the *Eco* RI and *Mse* I sites can then be determined whether they are fixed in males and females of *H. asinina*.

Reverse southern blot hybridization was applied for determining the presence/absence of candidate male- and female-specific AFLP markers in genomic DNA of *H. asinina*. Positive hybridization signals were found with the probe synthesized from *Hinc* II-digested genomic DNA of both male and female *H. asinina*. This indicated that the candidate sex-specific AFLP markers of male and female were existent in the genome of both genders. As a result, Southern blot analysis using the cloned AFLP fragment as the probe and genomic DNA of male and female *H. asinina* digested with various restriction enzymes as the target should be carried out to examine possible variability of the genomic regions covering candidate sex-specific AFLP fragments.

Sex determination markers are usually not conserved outside the genus (or in several cases for different species) in non-mammalian species. In fish, sex-specific markers successfully derived from AFLP fragments of the 3-spined stickleback (*Gasterosteus aculeatus*) failed to identify genders in 2 closely related stickleback species (Griffiths et al., 2000).

In addition, a growth hormone pseudogene (*GH*-Y) was studied and verified that it is located on Y-chromosome (*GH*-Y) and can be used as male-specific markers in the four salmon species; the chum salmon (*Oncorhynchus keta*), the pink salmon (*O. gorbuscha*), the coho salmon (*O. kisutch*) and the chinook salmon (*O. tshawytscha*). However, *GH*-Y has not been detected in some *Oncorhynchus* species (*O. nerka*, *O. mykiss* and *O. clarki*) suggesting deletion of this locus in some lineages during evolutionary processes (Devlin el al., 2001).

In the medaka *Oryzias latipes*, the *DMY* gene (derived from duplication of the *DMRT1* gene) is found in the Y chromosome and appears to be the sex-determining gene (Matsuda et al., 2002). This gene is also found in another species in the same genus, *O. curvinotus* (Matsuda et al., 2003) but is thought to have derived from DMRT1 immediately before speciation of *O. latipes* and *O. curvinotus*, and is absent in two other members of the same genus (Kondo et al., 2003).

Sex heterogametic chromosomes have not been identified in the Pacific oyster *Crassostrea gigas*. The genome project has been esbablished for identification of sex-determining and growth-regulating genes (and/or transcripts). A genomic BAC library containing 21,504 clones and the average insert size of 130 kb covering 4X of the oyster genome was constructed. In addition, several cDNA libraries at different developmental stages of fertilized eggs and sex gonads (containing approximately 500,000 clones) representing various stages of oyster maturation were also constructed. Over 4,200 independent cDNAs clones were determined. Homology search revealed a number of genes which revealed significant homology with known genes (cytoskeleton, transcription and translation, signal transudation, cell cycle regulation, membrane coponents, energy metabolism, mitochondrial components, and miscellaneous functions) from other well-characterized organisms. There are 45% unclassified cDNA sequences which were not homologous to any known genes. Only 2 cDNA clones showed significant homology with, FL(2)D and 1(1)10Bb in *Drosophila* genes, respectively. These genes are known to be involved in the insect sex-differentiation processes (Shimizu et al., 2000).

The other difficulty for development of sex-diagnostic marker is that sex-determiation markers may not be located at non-recombination regions resulting in the inability to apply the markers for general use. Ezaz et al (2004) identified sex-specific markers in the Nile tilapia (*Oreochromis niloticus*) using BSA and AFLP approaches.

Bulked genomic DNAs (XX and YY gynogenetic family pools) and individual DNA screening (XX and YY gynogenetics and XX and XY control individuals) provided 3 Y-linked (*Oni*Y425, *Oni*Y382, *Oni*Y227) and one X-linked (*Oni*X420) AFLP markers using 64 AFLP primer combinations (*Eco* RI and *Mse* I). Sequence analysis of *Oni*X420 and *Oni*Y425 were shown to be allelic. Locus-specific PCR was developed for these markers. Tight linkage was observed between AFLP markers and the sex locus within the source family. However, these markers failed to consistently identify sex in unrelated individuals, indicating recombination between the markers and the sex-determining loci. BAC clones of *O. niloticus*, containing the AFLP markers, hybridized to the long arm of chromosome 1. This confirmed previous evidence that chromosome pair 1 is the sex chromosomes in *O. niloticus*.

In the present study, development of sex-specific SCAR markers at the genomic DNA level was unsuccessful. SSCP also failed to detect sex-linked SNP in this species. Accordingly, identification of sex-linked markers was not carried out further.

III: Isolation and characterization of expressed sequence taq (EST) in ovaries and testes of *H. asinina.*

EST analysis is an effective method for isolation of interesting genes at different developmental stages, tissues and/or organisms. More specific transcripts can be identified by suppressive cDNA subtraction.

Using the above approaches, known cDNA (the number of clones sharing significant homology with known sequence in the GenBank at the E-value < 10^{-4}) found in the normal ovary and testis cDNA libraries are 59.5% (109 clones) and those in the subtractive ovary and testis cDNA libraries were 40.5% (73 clones) and 64.5% (71 clones) and 35.5% (56 clones), respectively.

Based on functions of their deduced proteins, 309 cDNA encoding known and 279 cDNA encoding unknown proteins were classified into seven broad categories: sex-related genes (114 clone); stress response, detoxification and cell defense proteins (27 clones); protein synthesis and ribosomal proteins (62 clones); cell division and DNA replication (35 clones); metabolism (41 clones); unidentified function (39 clones); and unknown (279 clones). A large number of sex-related genes (21.5%, 9.3%, 36.4% and

12.5%) and unknown genes (40.5%, 38.1%, 35.5%, and 65.0%) were found in normal ovary and testis and subtractive ovary and testis cDNA libraries.

Sex-related transcripts were predominant among known transcripts in all libraries but the most abundant transcript in ovaries and testes of *H. asinina* was VCP41 and sperm lysin ESTs, respectively. A total of five (VCP41, VCP42, VERL, VTG-1 and ADAMTS-9) and six (sperm lysin, GIOT1, axonemal p66.0, tektinA1, fertilization protein, and SARIP) different sex-related gene homologues were isolated from ovary and testis cDNAs, respectively.

Zeng and Gong (2002) isolated and characterizied gene expression in testis and ovary of the zebrafish using the EST approach. A total of 1025 ESTs (501 from the testis and 524 from the ovary libraries) were sequenced. A total of 641 ESTs matched with known sequences in the GenBank. Abundant cDNA clones in the two gonad libraries were identified and over 10% of which were homologues of egg membrane encoded proteins (ZP proteins). The testis library showed a more even distribution of cDNA clones with relatively fewer redundancy of EST in the library. Eleven potential cDNA clones (a male germ cell-specific protein; dynein IC3; tektin; a human protein involved in sexual delvelopment, PSD; and creatine kinase homologue for testis and chorion proteic component; ZPC-1; ZPC-2; ZPA; pregnancy zone protein, PZP; and a human protein expressed only in placental villi, PVP homologue for ovary) were used as molecular markers for the analysis of the gonad development in zebrafish by Northern blot hybridization. Most of them showed testis specific (dynein IC3 and tektin) and ovary specific (ZPC-1, ZPC-2, and ZPA) expression or testis predominant (PSD and creatine kinase) and ovary predominant (PZP, ZPC and PVP) expression as expected. Four (dynein IC3, and ZPC-1, ZPC-2, and ZPA) of these were further analyzed by *in situ* hybridization and found to be expressed specifically in the germ cells of testes and ovaries, respectively. Therefore, these promising molecular markers can be used for analyses of spermatogenesis and oogenesis in the zebrafish.

Several other functionally important transcripts including ATP synthase (synthesis of ATP), ATPase inhibitor (prevent the enzyme from switching to ATP hydrolysis during collapse of the electrochemical gradient, Gomez-Puyou et al 1983), heat shock protein 70 (play a vital role in mediating the folding of newly synthesized proteins and the refolding of denatured proteins), stress inducible chaperone GrpE-like 1 (play a role as a nucleotide

exchange factor for the stimulation of HSP70-ATPase activity, Tamura et al 2003), thiosulfate sulfertransferase (play a central role in cyanide detoxification, Hatzfeld and Saito, 2000), FKBP12 (cytosolic receptor for the immunosuppressive drug FK506 and rapamycin, Yazawa et al 2003), and TCTP (implicated in cell growth, acute allergic response, and apoptosis, Fiucci et al 2003) homologues were also isolated.

Nam *et al.* (2003) investigated gene expression in five cDNA libraries (gill, diverticula, hepatopancreas, skin/mucus and rectangular muscle) of the abalone (*Haliotis discus hannai*). Among 1235 analyzed EST, 738 exhibited significant similarity woth previously deposited sequences in the GanBank (E-value < 10^{-3}). The percent unique sequence (singleton) ranged from 55.6% - 73.5%. Analysis of the organisms represented by the best hit for each EST showed that 27% of EST were homologues of mammalian species followed by 18% of mollusks, 16% insects and 7% of fish. ESTs were classified into 10 functional categories and the expression patterns were judged by differences of the same category of each tissues. Most abundantly expressed ESTs in each library were chymotrypsin-like serine protease (3.1% in the diverticular library), actin (3.8%, 9.8%, and 4.2% in gill, diverticula and skin/mucus libraries, respectively) and ferritin (3.5% in the hepatopancreas library).

There have been no reports about EST analysis in gonads of abalone. In the medaga (*Oryzias latipes*), Kanamori (2000) used the subtractive hybridization technique to identify differences in gene expression during sex differentiation between males and females of the medaka. Fifty female-specific cDNA fragments were cloned. They were classified into three groups by virtue of whether their earliest expression is at 1, 5, or 30 days after hatching. All 15 near full-length cDNAs belonging to the first two groups were cloned. Many of these female-specific genes are coordinately expressed in oocytes at the earliest stages of oogenesis. Some gene homologues were identified by their sequences. These include egg envelope proteins, oocyte-specific RNA binding proteins, and a transcription factor containing a basic helix-loop-helix motif.

More recently, Leelatanawit et al (2004) used a suppressive cDNA subtraction to idenfy differential expressed genes in ovaries and testes of the giant tiger shrimp (*Penaeus monodon*). A total of 218 clones (157 clones from subtractive cDNAs of ovaries and 61 clones from those of testes) were unidirectional sequenced. Most of the expressed genes in ovaries were thrombospondin (TSP, 45 clones accounting for 28.7%),

peritrophin (17 clones, 10.8%) and unknown transcripts (78 clones, 49.7%). Conversely, almost all of the ESTs in *P. monodon* testes were unknown transcripts (58 clones, 95.1%). RACE-PCR was carried out for further characterization of TSP. Homologues of elongation factor-2 (EF-2), oxidoreductase, peritrophin, transketolase, hypothetical protein FLJ23251 and sex-linked ENSANGP00000010123 and XNP-1 were additionally isolated. Gender-specific expression of candidate sex-linked gene homologues was examined by RT-PCR. While XNP-1 and peritrophin were expressed in both ovaries and testes, TSP and ENSANGP00000010123 homologues revealed sex-specific expression in female *P. monodon*.

The full length cDNA of TCTP, sperm lysine and ATPase inhibotor wer obtained from EST anlysis of the present study. In addition, two isoforms of VCP (VCP1 and VCP7) were also obtained by clustering nucleotide sequences of different EST clones. Results from EST analysis indicated the potential for characterization of gene homologues in *H. asinina*, for which the numbers of well characterized genes is limited. RACE-PCR can be carried out against functionally important transcripts for isolation of the full length cDNA. Functional activities of interesting EST can be further studied by at the protein level.

IV: Isolation and characterization of sex-related genes in ovaries and testes of *H. asinina*

Sex-specific transcripts of ovaries (VCP1, VCP3, VCP7, VCP49, VCP75 and VTG-1) and testes (axonemalp66.0 tektinA1, sperm lysin, fertilization protein and DMRT1) of *H. asinina* were found from RT-PCR. In addition, ADAMTS-9 and α-tubulin 2 exhibited differential expression patterns between ovaries and testes. These expression markers are useful for further studies about egg and sperm development in *H. asinina*.

Gonads of *H. asinina* can be externally classified to 4 different stages (0 indicating no gonadal development, I indicating pre-mature gonads covering a little portion (25 %) of the hepatopancreas, II indicating partially mature gonads covering about 50% of the hepatopancreas and III indicating mature gonads covering about 75% of the hepatopancreas and IV indicating fully mature gonad covering the whole hapatopancreases).

Ovarian and testicular developmental processes of *H. asinina* were studied Based primarily on light and electron microscopy and indicated that they start and require different developmental periods of time. Gonial cells appear at 2 month of ages. At 4-month-old, early spermatocytes and spermatids were observed in testes whereas early oocytes arise at 6-7 months of ages. Mature spermatozoa could be found as early as 7-8 month old abalone. Longer time (11-12 months) is required for females to reach full sexual maturity (Sobhon, 1999 and 2000).

Transcripts exhibiting male- (axonemal p66.0, tektin A1, DMRT1, sperm lysine and FP) and female- (VCP1, 3, 7, 49 and 75 and VTG1) specific expression patterns in adults were re-analyzed against the first strand cDNA template of males and females adults of *H. asinina* having different ovarian (I and III) and testicular (I, II, III and IV) stages. Significantly different levels of all transcripts except sperm lysin and FP were observed between adult abalone possessing different stages of gonad development ($P < 0.05$).

Considering expression of sex-related transcripts in pooled cDNA of the whole body of *H. asinina* ($N = 6$) at 2, 3 and 5 month old, tektinA1, FP, VCP1, VCP2, VCP3, VCP7, VCP49, and VCP75 were detected in all ages of juvenile *H. asinina*. This

indicated non-selective expression of these transcripts in juvenile *H. asinina* where sexes could not be externally identified. Comparing to histological investigation of this species, only testes should be partially developed whereas ovaries should still be initial stage of development. The question arises about expression of VCP1, 3, 49 and 75 (restrictively expressed in ovaries of female adults) in juveniles with immature gonads, therefore it is interesting to carry out *in situ* hybridization of juvenile abalone with the probes from tektinA1, FP, VCP1, VCP2, VCP3, VCP7, VCP49, and VCP75.

In contrast, axonemal p66.0 and VTG-1 were not expressed in gonads of juvenile abalone. This indicated that both transcripts were expressed during the late developmental stages of testes and ovaries of *H. asinina*. The expression patterns of axonemal p66.0 and VTG-1 may be used to monitor maturation of *H. asinina* in the hatchery stock.

Like axonemal p66.0, expression of the newly isolated DM domain containing EST (approximately 1700 bp estimated from agarose gel elctrophoresis but the actual size is 1727 bp) and its variant (without DM domain, the size estimated from agarose gel electrophoresis is approximately 1400 bp but the actual size is 1351 bp) were restrictively found in testes of adult *H. asinina* indicating that these transcripts should play the important role for testicular development of this economically important abalone species.

Sperm lysin was expressed at the low level in 5-month-old juvenile *H. asinina*. Based on histological examination, spermatocytes and spermatids are existence in testes in male juveniles. Therefore, sperm lysin should have expressed in the early development of testes in male *H. asinina*.

SNP in genomic DNA of sex-related transcripts were also investigated because sex determination markers should be developed from fixed polymorphism in genomic DNA of male and female *H. asinina* to avoid destruction of specimens. Sweijd et al. (1998) illustrated successful development of species-specific markers for differentiation of a commercially exploited abalone in South Africa, *H. midae* and a sympatrically congeneric species, *H. spadicea* using primers designed from genomic DNA of sperm lysin. Therefore, a sex-specific marker based on genomic DNA polymorphism of *H. asinina* sperm lysin was not attempted. Amplification of VCP1, 2, 3, 7, 49, 75 and VTG1 using genomic DNA of *H. asinina* as the template did not yield the amplification

products. New primer pairs of cDNAs of these genes should be designed and further characterized to identify the existence of SNP between males and females of *H. asinina*.

The amplification products of axonemal p66.0 (containing a relatively large of approximately 575 bp) and tektinA1 gene segments were further characterized using SSCP analysis to determine whether fixed SNP between genders of *H. asinina* was existent. Tektin A1 exhibited a monomorphic SSCP pattern. Although axonemal p66.0 exhibited polymorphic SSCP patterns, they were not fixed between sexes.

RACE-PCR was then applied for identification of the full length cDNA of homologues of tektin A1, axonemal protein and DMRT1 where the full length sequences of these transcripts in abalone are not available at present.

Tektins were originally identified as a set of structural proteins of sea urchin flagellar microtubules (Hinchcliffe and Linck, 1998). The RPNVELCRD stretch is conserved and proposed to be important residues for tektin proteins. This consencus sequence was found between amino acid positions 365^{th} - 373^{rd} (RPGVDLCRD) of *H. asinina* tektinA1. In addition, four-cysteine residues conserved were found in the homologue in *H. asinina*. The position of the cysteines suggested their function to stabilize (via disulfide bonds) interactions within tektins and/or between tektins and tubulin (Xu et al, 2001).

Ota *et al*. (2002) isolated and characterized a testis-specific tektin in *Bombyx mori*. The complete open reading frame was obtained to encode 508 amino acid residues, whose sequence had 28, 28 and 30% identities with the *Strongylocentrotus purpuratus* tektins A1, B1 and C1, respectively. Expression analysis by RT-PCR with the cDNA and Western blotting with a polyclonal antibody indicated that the *BmTST* gene was expressed specifically in the testis during sperm maturation. The protein was immunologically detected exclusively in the fraction expected to contain the flagellar axonemes of sperms. This result infers that the *Bm*Tst protein is possibly involved in the spermatogenesis of *B. mori*.

Axonemal protein is a protein associated with axonemal structure of the sperm flagella. Eukaryotic cilia and flagella are elaborate molecular machines constructed of more than 250 proteins (Dutcher, 1995). DMRT1, which encodes protein with a conserved DNA-binding DM domain, has been recently isolated from a wide range of

vertebrates including fishes, reptiles, birds, and mammals. DMRT1 was expressed mainly in the adult gonads and/or in the developing gonads. Evidence has also been accumulated that the Dmrt1 genes play important roles in sexual (particularly testes) development in various organisms.

Results from the present study indicated that sex-specific markers in *H. asinina* were successfully developed at the cDNA level but not at the genomic DNA level. Gender-associated cDNA markers can be applied for studies of sex differentiation cascades and control of gonad development in this economically important species in the future. Long-term benefits of the present study include the possibility of improving reproductive performance of domesticated broodstock, inducing high-quality egg and sperm development in the hatchery stocks and producing monosex culture of *H. asinina*.

CHAPTER V

CONCLUSION

1. Species-specific SCAR markers of *H. asinina* (CUHA2, CUHA12, CUHA13) and *H. varia* (CUHV1) were successfully developed. Specificity of markers were tested against 216 individuals of abalone (N = 111, 73, and 32 for *H. asinina*, *H. ovina* and *H. varia*, respectively) and species-specific nature of all markers were found with the sensitivity of detection of 100 pg, 100 pg, 500 pg and 20 pg of the target DNA template, respectively.

2. Fourteen (seven fragments for each gender) candidate sex-specific AFLP markers of *H. asinina* were identified. These fragments were cloned and characterized. Eleven SCAR markers derived from candidate AFLP fragment generated the expected product in both male and female *H. asinina*. SSCP analysis indicated polymorphism in the amplified SCAR products for HaMale1, HaMale5, HaMale6, HaFemale2, HaFemale3, HaFemale4, and HaFemale5 but polymorphic SSCP patterns/fragments were not sex-linked.

3. Normal and subtractive cDNA libraries were established to investigate gene expression in ovaries and testes of *H. asinina* A total of 588 randomly selected clones (200 and 118 transcripts for normal and 110 and 160 transcripts for subtractive libraries from ovaries and testes, respectively) were unidirectional sequenced. Homologues of VCP (41 clones, 20.5%) were abundantly expressed in ovaries of *H. asinina*. In contrast, the most abundant transcripts in testes of *H. asinina* was sperm lysin (10 clones, 8.5%).

4. Three full-length cDNAs (sperm lysin, TCTP and ATPase inhibitor) were obtained. In addition, the full length cDNA of VCP (VCP1 and VCP7) were obtained from combining sequences of VCP-matched ESTs.

5. Homologues of VCP1, VCP3, VCP7, VCP49, VCP75 and VTG1 and axonemal p66.0, tektin A1, sperm lysin, FP, and DMRT1 exhibited sex-specific expression patterns in adults of female and male *H. asinina*, respectively.

6. Tissue distribution analysis of sex-related transcripts in adult *H. asinina* indicated abundant expression of sperm lysin and FP in testis followed hepatopancreas, and digestive gland of male adult. A tektinA1 homologue was mainly expressed in testes and a lesser level was observed in gill, digestive gland, left hypobranchial gland, and mantle of both males and females. Axonemal p66.0 and DMRT1 homologous of *H. asinina* were only detected in testis cDNA. Likewise, VCP1, VCP3, VCP49, VCP75, and VTG-1 were only expressed in ovaries. In addition, VCP2 transcript was detected in hemocytes of males and ovary of females. The VCP7 homologue was expressed in ovaries, eyes, and digestive gland of females.

7. Expressions of respective transcripts were examined in (whole) juveniles of *H. asinina*. TektinA1, sperm lysin, fertilization protein, VCP1, VCP2, VCP3, VCP7, VCP49, and VCP75 were all detected in 2, 3, 5-month-old juvenile abalone. However, expression of axonemal p66.0, DMRT1 and VTG-1 were not observed.

8. The full-length cDNAs of male-specific tektinA1 (2166 bp), axonemal p66.0 (2038 bp), and DMRT1 (1727 bp) were successfully identified by RACE-PCR. The ORF of tektinA1, axonemal p66.0, and DMRT1 was 1350, 1683, and 732 nucleotide encoded for a protein of 449, 560, and 243 amino acid residues, respectively.

REFERENCES

An, W., and Wensink, P.C. 1995a. Integrating sex- and tissue-specific regulation within a single *Drosophila* enhancer. *Genes. Dev.* 9: 256-266.

An, W., and Wensink, P.C. 1995b. Three protein binding sites from an enhancer that regulates sex- and fat body-specific transcription of *Drosophila yolk protein* genes. *EMBO J.* 14: 1221-1230.

Adams, M.D., Kelley, J.M., Gocayne, J.D., Dubnick, M., Polymeropoulos, M.H., Xiao, H., Merril, C.R., Wu, A. and Venter, J.C. 1991. Complementary DNA sequencing : expressed sequence tags and human genome project. *Science* 252: 1651-1656.

Adams, M.D., Soares, M.B., Kerlavage, A.R., Fields, C., Venter, J.C. 1993. Rapid cDNA sequencing (expressed tags) from a directionally cloned human infant brain cDNA library. *Nature Genetics* 4: 373-380.

Affara, N., Bently, P., Pelmear, A., and Jones, M.H. 1994. The identification of novel gene-sequences of the human adult testis, *Genomics* 22: 205-210.

Aoki, T., Nam, B.H., Hirono, I., Yamamoto, Y. 1999. Sequences of infected 596 cDNA clones (567,977 bp) of Japanese flounder (*Paralichthys olivaceus*) leukocytes infected with Hirame Rhabdovirus. *Mar. Biotechnol.* 1: 477-488.

Baker, B.S., and Ridge, K., 1980. Sex and the single cell: on the action of major loci affecting sex determination in *Drosophila melanogaster*. *Genetics* 94: 383–423.

Bangrak, P., Graidist, P., Chotigeat, W., and Phongdara, A. 2004. Molecular cloning and expression of a mammalian homologue of a translationally controlled tumor protein (TCTP) gene from *Penaeus monodon* shrimp. *J. Biotechnol*. 108: 219–226.

Beatty, R. 1964. Chromosome deviations and sex in vertebrates. *In Intersexuality in vertebrates including Man* (Armstrong, C.N. & Marshall, A.J. eds) 17-144.

Boag, P.R., Newton, S.E., Hansen, N., Christensen, C.M., Nansen, P., and Gasser, R.B. 2000. Isolation and characterization of sex-specific transcripts from *Oesophagostomum dentatum* by arbitrarily-primed PCR. *Mol. Biochem. Parasitol.*108: 217-224.

Brown, L.D. 1995. Genetic evidence for hybridisation between *Haliotis rubra* and *H. laevigata*. *Mar. Biol.* 123: 89-93.

Burtis, K.C., Coschigano, K.T., Baker, B.S., and Wensink, P.C. 1991. The doublesex proteins of *Drosophila melanogaster* bind directly to a sex-specific yolk protein gene enhancer. *EMBO J.* 10: 2577-2582.

Carding. S.R., Lu, D., Bottomly, K.A. 1992. A polymerase chain reaction assay for the detection and aquantification of cytokine gene expression in small number of cells. *J. Immunol. Method.* 151 : 277-287.

Coe, W.R. 1943. Sexual differentiation in molluscs. *Q. Rev. Biol.* 18: 154-164.

Connor, F., and Ashworth. 1992. Sequence of the cDNA for Xenopus XZFY-1. *Nucleic Acid Res.* 20: 5845.

Coschigano, K.T., and Wensink, P.C. 1993. Sex-specific transcriptional regulation by male and female doublesex proteins of *Drosophila*. *Gene and Dev.* 7: 42-54.

Cox, K.W. 1962. California abalones, family Haliotidae. California Division of Fish and Game. *Fish Bulletin* 118: 1-131.

Degnan, B.M., Degnan, S.M., Fentenany, G., and Morse, D.E. 1997. A Mox homeobox gene in the gastropod mollusc *Haliotis rufescens* is differentially expressed during larval morphogenesis and metamorphosis. *FEBS Letters* 411: 119-122.

Devlin, R.H., Biagi, C.A., and Smailus, D.E. 2001. Genetic mapping of Y-chromosomeal DNA markers in Pacific salmon. *Genetica* 111: 43-58.

Devlin, R., and Nagahama, Y. 2002. Sex determination and sex differentiation in fish: an overview of genetic, physiologycal, and environmental influences. *Aquaculture* 208: 191-364.

Dower, W.J., Miller, J.F., and Ragsdale, C.W. 1988. High efficiency transformation of *E. coli.* by high voltage electroporation. *Nucleic Acids Res.* 16: 612-617.

Dutcher, S. K., 1995. Flagellar assembly in two hundred and fifty easy-to-follow steps. *Trends Genet.* 11: 398–404.

Eckelbarger, K.J., and Davis, C.V., 1996. Ultrastructure of the gonad and gametogenesis in the eastern oyster, *Crassostrea viginica* I. Ovary and oogenesis. *Mar. Biol.* 127: 79-87.

Ezaz, M.T., Harvey, S.C., Boonphakdee, C., Teale, A.J., McAndrew, B.J., and Penman, D.J. 2004. Isolation and physical mapping of sex-linked AFLP markers in nile tilapia (Oreochromis niloticus L.). *Mar. Biotechnol.* 6 : 435-445.

Fallu, R. 1991. *Abalone Farming*. United Kingdom: Fishing News Books, Oxford, 195 pp.

FAO Fish and Fishery Products. 2000, Combined Yearbook of Fisheries 1970/2000, FISHSTAT PLUS.

Fiucci, G., Lespagnol, A., Cuvelette, P.S., Beaucourt, S., Duflaut, D., Susini, L., Amson, R., and Telerman, A., 2003. Genomic organization and expression of mouse Tpt1 gene. *Genomics* 81: 570–578.

Galindo, B.E., Moy, G.W, Swanson, W.J., and Vacquier. 2002. Full-length sequence of VERL, the egg vitelline envelope receptor for abalone sperm lysin. *Gene* 288: 111-117.

Garabedian, M.J., Shepherd, B.M., and Wensink, P.C. 1986. A tissue-specific enhancer from the *Drosophila* yolk protein 1 gene. *Cell* 45: 859-867.

Geiger, D.L. 2000. Distribution and biogeography of the *Haliotidae* (Gastropoda: Vetigastropoda) world-wide. *Bollettino Malacologico.* 35: 57-120.

Geiger, D.L. and Groves, L.T. 1999. Review of fossil abalone (Gastropoda: Vetigastropoda: *Haliotidae*) with comparison to Recent species. *J. Paleontol.* 73: 872-885.

Gibbons, I.R. 1981. Cilia and flagella of eukaryotes. *J. Cell. Biol.* 91:107s–124s.

Giusti, A.F., Hinman, V.F., Degnan, B.M., Morse, D.E. 2000. Expression of a Hox5/Scr gene in the larval central nervo is system of the gastropod *Haliotis*, a non-segmentted spiralian lophotrocozoan. *Evol. Dev.* 2: 294-302.

Goerke, C., Manfered, G. B., and Wolz, C. 2001. Quanification of bacterial transcripts during ifection using competitive reverse transcription-PCR (RT-PCR) and light cycler RT-PCR. *Clin. Diag. Lab. Immunol.* 4: 279-282.

Gordon, H.R. and Cook, P.A. 2001. *World abalone supply, markets and pricing. Fourth International Symposium on Abalone Biology*, Fisheries, and Couture, Cape Town, South Africa, February 6-11, 2000

Gordon, H.R. and Cook, P.A. 2003. *World abalone supply, markets and pricing: historical, current and future prospectives*, 5th International Abalone Symposium, Quingdao, China, October 2003.

Griffiths, R. and Tiwari, B. 1993 The isolation of molecular genetic markers for the identification of sex. *Proc. Natl. Acad. Sci. USA*. 90: 8324-8326.

Griffiths, R., Orr, K. 1999. The use of amplified fragment length polymorphism (AFLP) in the isolation of sex-specific markers. *Mol. Ecol.* 8: 671-674.

Griffiths, R., Orr, KJ., Adam, A. and Barber, I. 2000. DNA sex identification in the three-spined stickleback. *J. Fish Biol.* 57: 1331-1334.

Guan, G., Kobayashi, T., and Nagahama, Y. 2000. Sexually dimorphic expression of twotypes of DM (Doublesex/Mab-3)-domain genes in a teleost fish, the Tilapia (*Oreochromis niloticus*). *Biochem. Biophys. Res. Commun.* 272: 662-666.

Guo, Y., Li, Q., Gao, S., Zhou, Z., He, Y., Shang, X., Cheng, H., and Zhou, R. 2004. Molecular cloning, characterization, and expression in brain and gonad of Dmrt5 of zebrafish. *Biochem. and Biophy. Research Commun.* 324: 569–575.

Hahn, K.O. 1989. *Handbook of culture of abalone and other marine gastropods*. Boca Raton, Florida: CRC press, 348 pp.

Hatzfeld, Y., and Saito, K. 2000. Evidence for the existence of rhodanease (thiosulfate: cyanide sulfurtransferaser) in plants: preliminary characterization of two rhodanese cDNAs from Arabidopsis thaliana. *FEBS Letters* 470: 147-150.

Hayashi, K. 1992. PCR-SSCP: A method for detection of mutations. *GATA* 9: 73-79.

Heller, J. 1993. Hermaphroditism in molluscs. *Biol. J. Linn. Soc.* 48: 19-42.

Hinman, V.F., and Degnan, B.M. 2002. *Mox* homeobox expression in muscle lneage of the gastropod *Haliotis asinina*: evidence for a conserved role in bilaterian myogenesis. *Dev. Genes Evol.* 212: 141-144.

Iguchi, N., Tanaka, H., Fujii, T., Tamura, K., Kaneko, Y., Nojima, H., and Nishimune, Y. 1999. Molecular cloning of haploid germ cell-specific tektin cDNA and analysis of the protein in mouse testis. *FEBS Letters* 456: 315–321.

Inaba, K., Padma, P., Satouh, Y., Shin-I, T., Kohara, Y., Satoh, N., and Satou, Y. 2002. EST analysis of gene expression in testis of the ascidian Ciona intestinalis. *Mol. Reprod. Dev.* 56: 259-264.

Inoue, K., Hoshijima, K., Sakamoto H., and Shimura Y. 1990. Binding of the *Drosophila* sex-lethal gene product to the alternative splice site of transformer primary transcript. *Nature* 344: 461-463.

Jarayabhand, P., and Paphavasit, N. 1996. A review of the culture of tropical abalone with special reference to Thailand. *Aquaculture* 140: 159-168.

Jarayabhand, P., Pripue, P., Klinbunga, S., and Tassanakajon, A. (2002). Identification of species-diagnostic markers of abalone in Thailand using PCR-RFLP of 16S rDNA. *Fisheries Science* 68 (supplement II), 1091-94.

Jarayabhand, P., Yom-La, R., and Popongviwat, A. 1998. Karyotypes of marine molluscs in the family Haliotidae found in Thailand. *J. Shellfish Res.* 17(3): 761-764.

Jursnich, V.A., and Burtis, K.C. 1993. A positive role in differentiation for the male doublesex protein of *Drosophila*. *Dev. Biol.* 155: 235-249.

Kakhai, N., and Petjamrat, K. 1992. Survey on species and broodstock collection of abalone (*Haliotis* spp.) in Chon Buri, Rayong and Trad Provinces. *Technical Paper No. 6/1992*, Rayong Coastal Aquaculture Station, Department of Fisheries, Ministry of Agriculture and Cooperatives, Thailand, 31 pp. (in Thai, with English abstract).

Kanamori, A. 2000. Systematic identification of genes expressed during early oogenesis in medaka. *Mol. Reprod. Dev.* 55: 31-36.

Kang, H. S., Lee, M. J., Song, H., Han, S. H., Kim, Y. M., Im, J. Y., & Choi, I. 2001. Molecular identification of IgE-dependent histamine-releasing factor as a B cell growth factor. *J. Immunolo.* 166: 6545–6554.

Kim, S., Kettlewell, J.R., Anderson, R.C., Bardwell, V.J. Zarkower, D. 2003. Sexually dimorphic expression of multiple doublesexrelated genes in the embryonic mouse gonad. *Gene Expr. Patterns.* 3: 77–82.

Klinbunga, S., Ampayup, P., Tassanakajon, A., Jarayabhand, P., and Yoosukh, W (2000). Development of species-specific markers of the tropical oyster (*Crassostrea belcheri*) in Thailand. *Mar. Biotechnol.* 2: 476-484.

Klinbunga, S., Ampayup, P., Tassanakajon, A., Jarayabhand, P., and Yoosukh, W. (2001a). Genetic diversity and molecular markers of commercial oysters (Genera *Crassostrea*, *Saccostrea* and *Striostrea*) in Thailand determined by RAPD analysis. *Mar. Biotechnol.* 3: 133-44.

Klinbunga, S., Boonyapakdee, A., and Pratoomchat, B. (2000). Genetic diversity and species-diagnostic markers of mud crabs (Genus *Scylla*) in Eastern Thailand determined by RAPD analysis. *Mar. Biotechnol.* 2: 180-87.

Klinbunga, S., Khamnamtong, N., Tassanakajon, A., Puanglarp, N., Jarayabhand, P. and Yoosukh, W. 2003. Molecular genetic tools for three commercial cultured oysters (*Crassostrea belcheri*, *C. iredalei* and *Saccostrea cucullata*) in Thailand. *Mar. Biotechnol.* 5: 27-36.

Klinbunga, S., Pripue, P., Khamnamtong, N., Tassanakajon, A., Jarayabhand, P.and Menasveta, P. 2003 Genetic diversity and molecular markers of the tropical abalone (*Haliotis asinina*) in Thailand. *Mar. Biotechnol.* (in press).

Klinbunga, S., Siludjai, D., Wuthijinda, W., Tassanakajon, A., Jarayabhand, A., and Menasveta, P. (2001b). Genetic heterogeneity of the giant tiger shrimp (*Penaeus monodon*) in Thailand revealed by RAPD and mtDNA-RFLP analyses. *Mar. Biotechnol.* 3: 428-38.

Kovacs, B., Egedi, S., Bartfai, R., Orban, L. 2000 Male-specific DNA markers from African catfish (*Clarias gariepinus*). *Genetica* 110: 267-276.

Kondo, M., Nanda, I., Hornung, U., Asakawa, S., Shimizu, N., Mitani, H., Schmid, M., Shima, A., Schartl, M., 2003. Absence of the candidate male sex-determining gene dmrt1b(Y) of medaka from other fish species. *Curr. Biol.* 13: 416–420.

Koopman, P., Ashworth, A., Lovell-Badge, R. 1991. The *Zfy* gene family in humans and mice. *Trends. Genet.* 7: 132-136.

Kozak, M. 1986. Point mutations define a sequence flangking the AUG initiation codon that modulates translation by eukaryotic ribosomes. *Cell* 44: 283-292.

Kresge, N., Vacquier, V.D., and Stout, C.D. 2001. Abalone lysin: the dissolving and evolving sperm protein. *BioEssays* 23: 95–103.

Lee, Y.H., and Vacquire, V.D. 1995. Evolution and systematics in Haliotidae (Mollusca: Gastropoda): inferences from DNA sequences of sperm lysin. *Mar. Biol.* 124: 267-278.

Leelatanawit, R., Klinbunga, S., Tassanakajon, A., Jarayabhand, P. Hirono, I. Aoki, T and Menasveta, P. 2004. Isolation and Characterization of Differentiatially Expressed Genes in Ovaries and Testes of the Giant Tiger Shrimp (*Penaeus monodon). Mar. Biotechnol.* 6: S506-S510.

Lehnert, S.A., Wilson, K.J., Byrne, K., and Moore, K. 1999. Tissue-specific expressed sequence tags from the black tiger shrimp *Penaeus monodon*. *). Mar. Biotechnol.* 1: 465-476.

Liew, C.C., Hwang, D.M., Fung, Y.W., and et al. 1994. A catalog of genes in the cardiovascular-system as identified by expressed sequence tags. *Proc. Natl. Acad. Sci.* 91: 10645-10649.

Lindberg, D.R. 1992. Evolution, Distribution and systemetics of *Haliotidae.* In: Shepherd, S.A., Tegner, M.J. and Guzman Del Proo, S.A. (Editor). *Abalone of the World, Biology*, Fisheries and Culture. Fishing News Books, Oxford. p. 3-18.

Li, Y., Hill, J.A., Yue, G.H., Chen, F., and Orban, L. 2002. Extensive search does not identify genomic sex markers in *Tetradon nigroviridis*. *J. Fish Biol.* 61: 1314-1317.

Lopata, A.L., Luijx, T., Fenemore, B., Sweijd, N.A., and Cook, P.A. 2002. Development of a monoclonal antibody detection assay for species-specific identification of abalone. *Mar. Biotechnol.* 4: 454-462.

MacDonald, S. M., Bhisutthibhan, J., Shapiro, T. A., Rogerson, S.J., Taylor, T. E., Tembo, M., Langdon, J. M., and Meshnick, S. R. 2001. Immune mimicry in malaria: *Plasmodium falciparum* secretes a functional histamine-releasing factor homologue in vitro and in vivo. *Proc. Natl. Acad. Sci. USA*. 98: 10829–10832.

MacDonald, S.M., Rafnar, T., Langdon, J., and Lichtenstein, L.M. 1995. Molecular identification of an IgE-dependent histamine releasing factor. *Science* 269: 688–690.

Marchand, O., Govoroun, M., D'Cotta, H., McMeel, O., Lareyre, J., Bernot, A., Laudet, V., and Guiguen, Y. 2000. DMRT1 expression during gonadal differentiation and spermatogenesis in the rainbow trout, *Oncorhynchus mykiss*. *Biochimic. Biophys. Acta* 1493: 180-187.

Matsuda, M., Nagahama, Y., Shinomiya, A., Sato, T., Matsuda, C., Kobayashi, T., Morrey, C.E. Shibata, N., Asakawa, S., Shimizu,N., Hori, H., Hamaguchi, S., and Sakaizumi, M. 2002. DMY is a Y-speci.c Dmdomain gene required for male development in the medaka fish. *Narture* 417: 559-563.

Matsumoto, T., Nakamura, A.M., Mori, K., and Kayano, T. 2003. Molecular characterization of a cDNA encoding putative vitellogenin from the Pacific oyster *Crassostrea gigas*. *Zoological Science* 20: 3742.

McCombie, W.R., Adams, M.D., Keller, J.M., Fitzgerald, M.G., Utterback, T.R., Khan, M., Dubnick, M., Kerlavage, A.R., Venter, J.C., and Fields, C. 1992.

Caenorhabditis elegans expressed sequence tags identify gene families and potential disease homologues. *Nature Genet.* 1: 124-131.

Michelmore R.W., Paran I., and Kesseli R.V. 1991. Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *Proc. Natl. Acad Sci USA.* 88: 9828-9832.

Miller, S.W., Hayward, D.C., Bunch, T.A., Miller, D.J., Ball, E.E., Bardwell, V.J., Zarkower, D., Brower, D.L., 2002. A DM domain protein from a coral, Acropora millepora, homologous to proteins important for sex determination. *Evol. Dev.* 5: 251-258.

Mozingo, N.M., Vacquier, V.D., and Chandler, D.E. 1995. Structure features of the abalone egg extracellular matrix and its role in gamete interaction during fertilization. *Mol. Reprod. Dev.* 41: 493-502.

Muchmore, A.E., Moy, G.W., Swanson, W.J., and Vacquier, V.D. 1998. Direct sequencing of genomic DNA for characterization of a satellite DNA in five species of Eastern Pacific abalone. *Mol. Mar. Biol. and Biotechnol.* 7: 1-6.

Naganuma, T., Hisadome, K., Shiraishi, K., and Kojima, H. 1998. Molecular distinction of two resemblant abalones, *Haliotis discus discus* and *Haliotis discus hannai* by 18S rDNA sequences. *J. Mar. Biotechnol.* 6: 59-61.

Nanda, I., Zend-Ajusch, E., Shan, Z., Grutzner, F., Schartl, M., Burt, D.W., Koehler, M., Fowler, V.M., Goodwin, G., Schneider, W.J., Mizuno, S., Dechant, G., Haaf, T., and Schmind, M. 2000. Conserved synteny between the chicken Z sex chromosome and human chromosome 9 includes the male regulatory gene DMRT1: a comparative (re)view on avian sex determination. *Cytogenet. Cell Genet.* 89: 67-78.

Nateewathana, A. and Hylleberge, J.1986. A survey on thai abalone around Phuket Island and feasibility study of abalone culture in Thailand. *Thai. Fish. Gazette.* 39: 177-190.

Ngow, O. and Jarayabhand, P., 1993. Distribution and habitat selection of the abalone, *Haliotis ovina* (Gmelin, 1791), at the eastern coast of Thailand. *Presented at the 19th Congress on Science and Technology of Thail*and, 27-29 October 1993, Bangkok, pp. 472-473 (in Thai, with English abstract).

Norrander, J., Larsson, M., Stahl, S., Hoog, C., and Linck, R.W. 1998. Expression of ciliary tektins in brain and sensory development. *J. Neurosci.* 18: 8912–8918.

Norrander, J.M., Perrone, C.A., Amos, L.A., Linck, and R.W. 1996. Structural comparison of tektins and evidence for their determination of complex spacings in flagellar microtubules. *J. Mol. Biol.* 257: 385–397.

Obara, M.,Matsunaga, S., Nakao, S., and Kawano, S. 2002. A plant Y chromosome-STS marker encoding a degenerate retrotransposon. *Genes Genet. Syst.* 77: 393-398.

O'Brien, E.K., and Degnan, B.M. 2000. Expression of *POU*, *Sox*, and *Pax* genes in the brain ganglia of the tropical abalone *Haliotis asinina*. *Mar. Biotechnol.* 2: 545-557.

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. U S A.* 868: 2766-2770.

Orlando, C., Pinzani, P., and Pazzagli, M. 1998. Developments in quantitative PCR. *Clin.Chem. Lab. Med.* 36 : 255-269.

Ota, A., Kusakabe, T., Sugimoto, Y., Takahashi, M., Nakajima, Y., Kawaguchi, Y., and Koga, K. 2002. Cloning and characterization of testis-specific tektin in Bombyx mori. *Comp. Biochem. and Physiol. Part B.* 133: 371-382.

Pawlak, A., Toussaint, C., Levy, I., and et al. 1995. Characterization of a large population of mRNAs form human testis. *Genomics* 26: 151-158.

Pipe, R.K. 1987. Oogenesis in the marine mussel Mytilus edulis: an ultrastructureal study. *Mar. Biol.* 95: 405-414.

Pirner, M.A., and Linck, R.W. 1994. Tektins are heterodimeric polymers in flagellar microtubules with axial periodicities matching the tubulin lattice. *J. Chem. Biol.* 269: 31800–31806.

Popongviwat, A., 2001. *Genetic diversity of tropical abalone in Thailand using RAPD-PCR*. Master's Thesis, Programme of Biotechnology, Faculty of Science, Chulalongkorn University.

Preechaphol, R. 2004. *Sex determination markers and gene expression in androgenic gland of different male morphotypes of the giant freshwater prawn Macrobrachium rosenbergii.* M.sc. Thesis.Chulalongkorn university.

Raymond, C.S., Kettlewel, J.R., Hirsch, B., Bardwell, V.J., and Zarkower, D. 1999. Expression of *DMRT1* in the genital ridge of mouse and chicken embryos suggests a role in vertebrate sexual development. *Dev. Biol.* 215: 208-220.

Raymond, C.S., Shamu, C.E., Shen, M.M., Seifert, K.J., Hirsch, B., Hodgkin, J., and Zarkower, D. 1998. Evidence for evolutionary conservation of sex-determining genes. *Nature* 391: 391-695.

Roy, A., Yan, W.,Burns, K.H., and Matzuk, M.M. 2004. Tektin3 encodes an evolutionarily conserved putative testicular microtubules-related protein expressed preferentially in male germ cells. *Mol. Reprod. Dev.* 67: 295-302.

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

Sappington, T.W. and Raikhel, A.S. 1998. Molecular chharacteristics of insect vitellogenesis and vitellogenin receptors. *Insect. Biochem. Mol. Bio.* 28: 277-300.

Satou, Y., Kawashima., Kohara., and Satoh., N. 2003. Large scale EST analyses in Ciona intestinalis: Its application as Northern blot analyses. *Dec. Genes. Evol.* 213: 314-318.

Schneider-Gädicke, A., Beer-Romero, P., Brown, L.G., Nussbaum, R., and Page, D.C. 1989. ZFX has a gene structure similar to ZFY, the putative human sex determinant, and escapes X inactivation. *Cell* 57: 1247-1258.

Shan, Z., Nanda, I., Wang, Y., Schmid, M., Vortkamp, A., and Haff, T. 2000. Sex-specific expression of an evolutionarily conserved male regulatory gene, DMRT1, in birds. *Cytogenet. Cell Genet.* 89: 252-257.

Shen, M.M., and Hodgkin, J. 1988. *mab-3*, a gene required for sex-specific yolk protein expression and a male-specific lineage in *C. elegans*. *Cell* 54: 1019-1031.

Shepherd, S.A. and Hearn, W.S.(1983). Studies on southern Australian abalone (genus *Haliotis*) IV. Growth of *H. laevigata* and *H. ruber. Austral. J. Mar. and Fres. Res.* 34: 461-475.

Shimizu, N., Lee, J., Shimizu, Y., Ohtake, H., Sato, Y., and Asakawa, S. *2002 Aquatic Genomic*: Genomics of the Pacific oyster *Crassostrea gigas*. 128-132.

Singhagraiwan, T., and Doi, M. 1993. *Seed production and culture of a tropical abalone, Haliotis asinina Linne*'. The research project of fishery resource development in the Kingdom of Thailand. Department of Fisheries, Ministry of Agriculture and Cooperatives, Thailand, 32 pp.

Sobhon, P., Apisawetakan, S., Champoo, M., Wanichanon, C., Linthong, V., Thongkukiatkul, A., Jarayabhand, P., Kruatrachue, M., Upatham, S., and Poomthong, T. 1999. Classification of germ cells, reproductive cycle and maturation of gonads in *Haliotis asinina* Linnaeus. *ScienceAsia* 25: 3-21.

Sosnowski, B.A., Belote, J.M., and McKeown. 1989. Sex-specific alternative splicing of RNA from the *transformer* gene results from sequence-dependent splice site blockage. *Cell* 58: 449-459.

Sungthong, S., Ingsrisawang, V and Fujiwara, S., 1991. Study on the relative growth of abalone, *Haliotis asinina* Linne, off Samet Island. *Thai. Mar. Fish. Res. Bull.* 2: 15-20.

Swanson, W.J., and Vacquier, V.D., 1997. The abalone egg receptor for sperm lysin is a giant multivalent molecule. *Proc. Natl. Acad. Sci. USA.* 94: 6724–6729.

Swanson, W.J., and Vacquier, V.D., 1998. Concerted evolution in an egg receptor for a rapidly evolving abalone sperm protein. *Science* 281: 710–712.

Sweijd, N.A., Bowie, R.C.K., Lopata, A.L., Marinaki, A.M., Harley, E., and Cook, P.A. 1998. A PCR technique for forensic, species-level identification of abalone tissue. *J. Shellfish Res.* 17: 889-895.

Takeda, J., Yano, H., Eng, S., Zeng, Y., and Bell, G.I. 1993. A molecular inventory of human pancreatic islets: sequence analysis of 1000 cDNA clones, *Human molec. Genet.* 2: 1793-1798.

Tamura, S., Kinouchi, H., Izaki, K., Okubo, A., Sugawara, T., Kunizuka, H., and Mizoi, K. 2003. Induction of heat shock protein 40 and GrpE mRNAs following transient focal cerebral ischemia in the rat. *Brain Res.* 960: 277–281.

Tassanakajon, A., Pongsomboon, S., Jarayabhand, P., Klinbunga, S., and Boonsaeng, V. (1998). Genetic structure in wild populations of the black tiger shrimp (*Penaeus monodon*) using randomly amplified polymorphic DNA analysis. *J. Mar. Biotechnol.* 6: 249-254.

Thaewnon-ngiw, B. (2002): *Species diversity and molecular markers of alien apple snails (Pomacea sp.) and native apple snails (Pila sp.) in Thailand*. PhD Thesis, Interdisciplinary Graduate Program, Kasetsart University. 230 pp.

Thumrungtanakit, S. 2004. *Identification of molecular genetic markers involved in sex determination in the giant tifer shrimp (Penaeus monodon)*. M. Sc. Thesis, Chulalongkorn University

Tiersch, T.R., Simco, B.A., Davis, K.B., and Watchel, S.S. 1992. Molecular genetics of sex determination in the channel catfish–studies on *SRY*, *ZFY*, *BKM*, and human telomeric repeats. *Biol. Reprod*. 47: 185-192.

Tookwinas, S., Leknim, W., Donyadol, Y., Preedalampabuttra, Y. and Perngmak, P.,1986. Survey on species and distribution of abalone (*Haliotis* spp.) in Surattani, Nakornsrithammarat and Songkhla Provinces, Contribution No. 1/1986, National Institute of Coastal Aquaculature, Department of Fisheries, Thailand, 16 pp. (in Thai, with English abstract).

Vacquier, V.D., Carner, K.R., and Stout, C.D. (1990). Species-specific sequences of abalone lysine, the sperm protein that creates a hole in the egg envelope. *Proc. Natl. Acad. Sci. USA*. 87: 5792-5796.

Vacquier, V.D., Swanson, W.J., and Lee, Y. 1997. Positive darwinian selection on two homologous fertilization proteins; What is the selective pressure driving their divergenece ? *J. Mol. Evol.* 44 (Suppl 1): S15-S22.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. and Zabeau, M. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.* 23: 4407-4414.

Wahli, W., 1988. Evolution and expression of vitellogenin genes. *Trends. Genet.* 4: 227–232.

Waterston, R., Martin, C., Craxton, M., Huynh, C., Coulson, A., Hillier, L., Durbin, R., Green, P., Shownkeen, R., Halloran, N., Metzstein, M., Hawkins, T., Wilson, R., Berks, M., Du, Z., Thomas, K., Thierry-Mieg, J., and Sulston, J. 1992. A survey of expressed genes in *Caenorhabditis elegans*. *Nature Genet.* 1: 114-123.

Weising, K., Nybom, H., Wolf, K., and Meyer, W. (1995). *DNA Fingerprinting in Plant and Fungi.* Boca Raton, Fla: CRC Press.

Welsh, J. and McClelland, M. 1990. Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Res.* 18: 7213-7218.

Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A., and Tingey, S.V. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res.* 18: 6531-6535.

Wolkowicz, M.J., Naaby-Hansen, S., Gamble, A.R., Reddi, P.P., Flickinger, C.J., and Herr, J.C. 2002. Tektin B1 demonstrates flagellar localization in human sperm. *Biol. Reprod.* 66: 241–250.

Xu, M., Zhou, Z., Cheng, C., Zhao, W., Tang, R., Huang, Y., Wang, W., Jian, X., Zeng, L., Xie, Y., and Mao, Y. 2001.Cloning and characterization of a novel human *TEKTIN* 1 gene. *Internat. J. Biochem. and Cell Biol.* 33: 1172–1182

Yazawa , S., Obata, K., Iio, A., Koide, M., Yokota, M. Sasaki, S., Kagami, H., Ono, T. 2003. Gene expression of FK506-binding proteins 12.6 and 12 during chicken development. *Comp. Biochem. and Physi. Part A.* 136: 391–399.

Yi, W., Ross, J.M., and Zarkower, D. 2000. Mab-3 is a direct tra-1 target gene that regulates diverse aspects of C. elegans male sexual behavior. *Development* 127: 4469-4480.

Zeng, S., and Gong, Z. 2002. Expressed sequence tag analysis of expression profiles of zebrafish testis and ovary. *Gene* 294: 45-53.

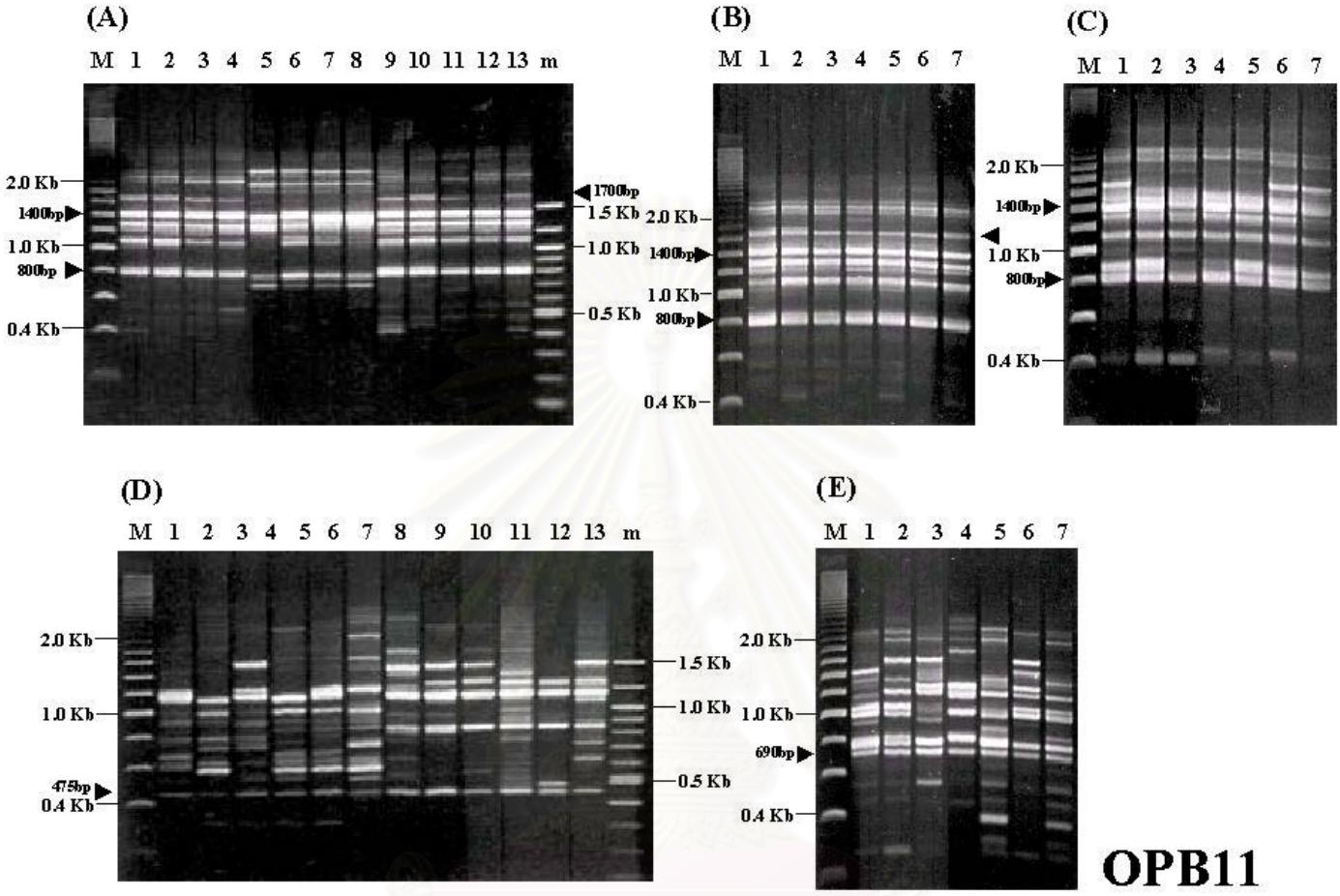
Zhiyong, M.A., Khatlani, T.S., Sasaki, K., Inokuma, H., and Onishi, T. 2000. Cloning of canine cDNA encoding tektin. *J. Vet. Med. Sci. Sep.* 62: 1013–1016.

Zhou, R., Zhang, Q., Tiersch, T.R., and Cooper, R.K. 2001. Four members of the *Sox* gene family in channel catfish. *J. Fish Biol.* 58: 891-894.

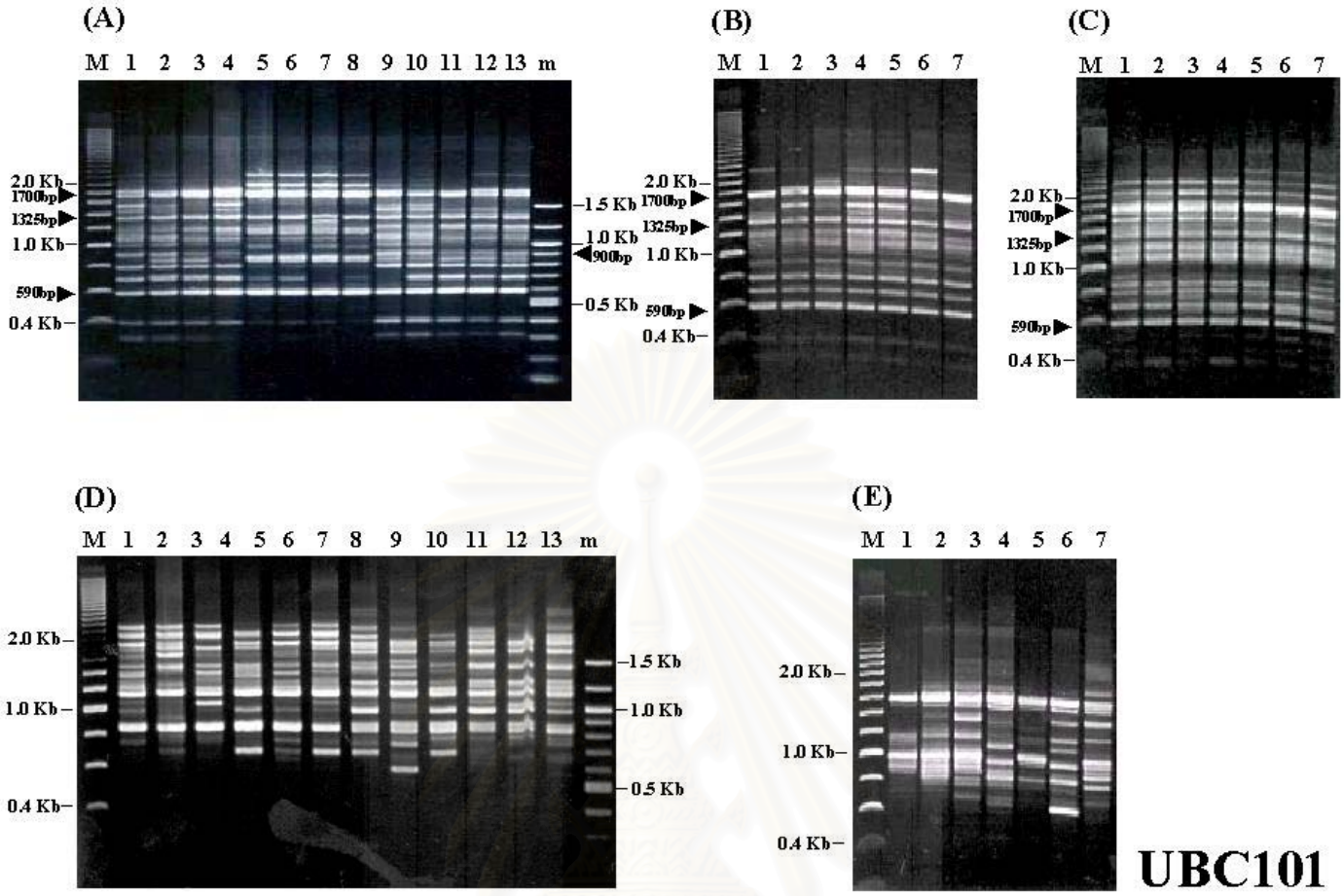
Zhu, L., Wilken, J., Phillips, N.B., Narendra, U., Chan, G., Stratton, S.M., Kent, S.B., and Weiss, M.A. 2000. Sexual dimorphism in diverse metazoans is regulated by a novel class of intertwined zinc fingers. *Gene Devel.* 14: 1750-1764.

APPENDICES

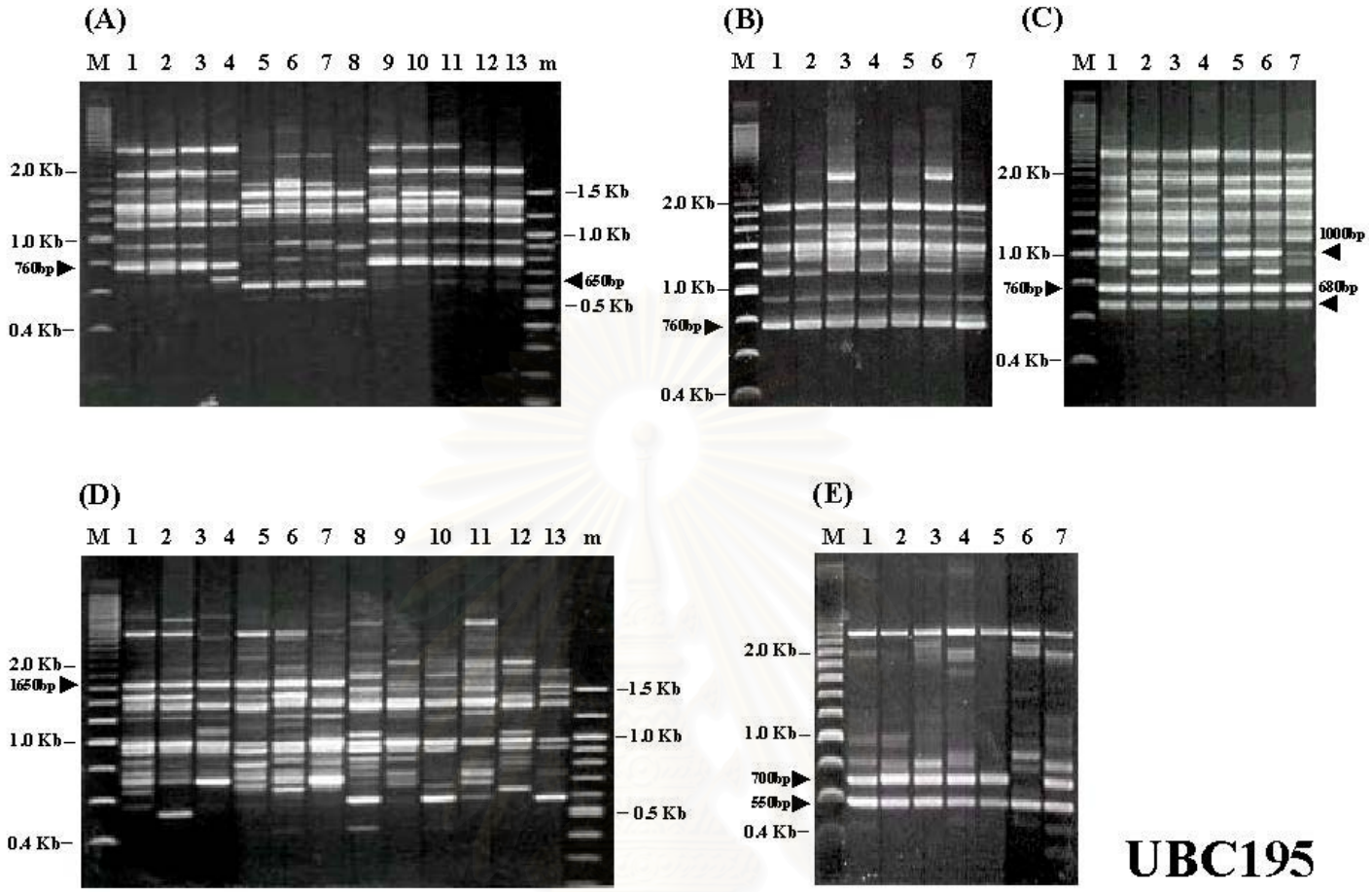
APPENDIX A



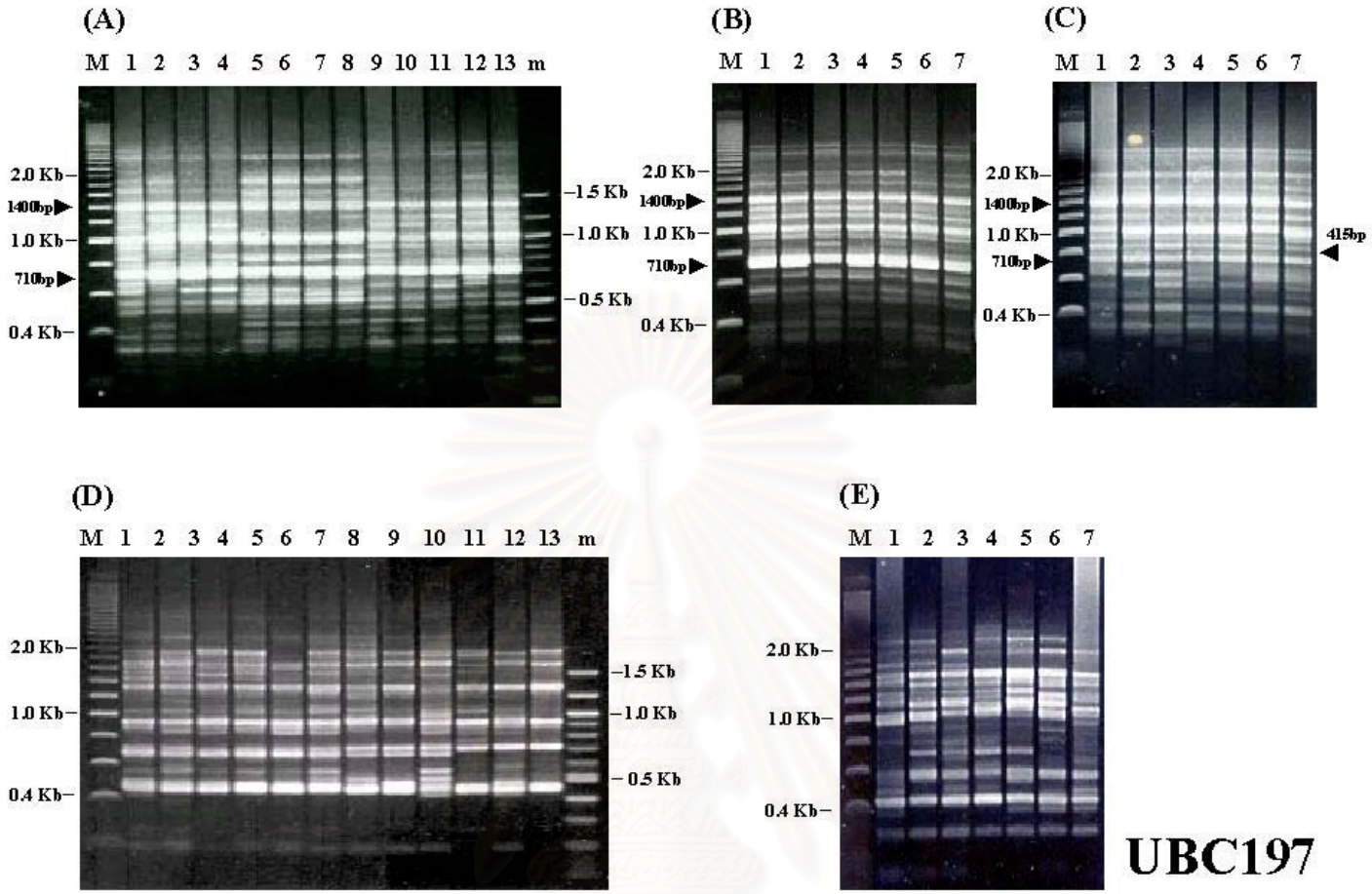
Appendix A1.1 RAPD patterns resulted from analysis of *H. asinina* originating from Samet Island, Rayong (lanes 1-4, A), Talibong Island, Trang (lanes 5-8, A), Cambodia (lanes 9-13,A), P_0, Cambodia (lanes 1-7, B) and F_1, Philippines (lanes 1-7, C), *H. ovina* originating from Sichang Island, Chon Buri (lanes 1-3, D), Samet Island, Rayong (lanes 4-6, D), Churk Island, Trang (lanes 7-9, D) and Similan Island, Phangnga (lanes 9-12, D) and *H. varia* originating from L-Island, Phuket (lanes 1-4, E) and Similan Island, Phangnga (lanes 5-7, E) with the primer OPB11. Lanes M and m are 200 bp and 100 bp DNA ladders, respectively. An arrowhead (◂) indicates a species – specific marker.



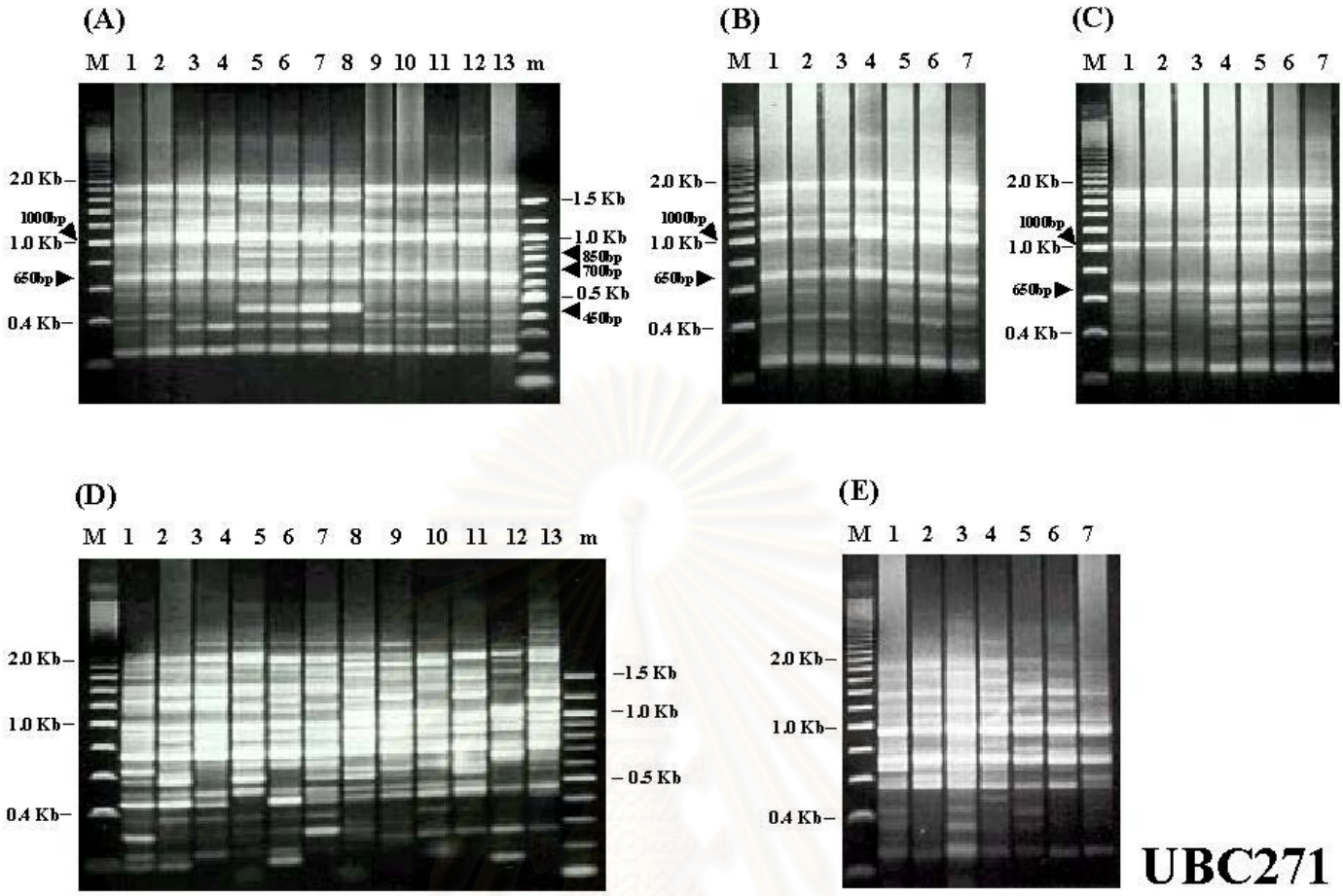
Appendix A1.2 RAPD patterns resulted from analysis of *H. asinina* originating from Samet Island, Rayong (lanes 1-4, A), Talibong Island, Trang (lanes 5-8, A), Cambodia (lanes 9-13,A), P_0, Cambodia (lanes 1-7, B) and F_1, Philippines (lanes 1-7, C), *H. ovina* originating from Sichang Island, Chon Buri (lanes 1-3, D), Samet Island, Rayong (lanes 4-6, D), Churk Island, Trang (lanes 7-9, D) and Similan Island, Phangnga (lanes 9-12, D) and *H. varia* originating from L-Island, Phuket (lanes 1-4, E) and Similan Island, Phangnga (lanes 5-7, E) with the primer UBC101. Lanes M and m are 200 bp and 100 bp DNA ladders, respectively. An arrowhead (◂) indicates a species – specific marker.



Appendix A1.3 RAPD patterns resulted from analysis of *H. asinina* originating from Samet Island, Rayong (lanes 1-4, A), Talibong Island, Trang (lanes 5-8, A), Cambodia (lanes 9-13,A), P_0, Cambodia (lanes 1-7, B) and F_1, Philippines (lanes 1-7, C), *H. ovina* originating from Sichang Island, Chon Buri (lanes 1-3, D), Samet Island, Rayong (lanes 4-6, D), Churk Island, Trang (lanes 7-9, D) and Similan Island, Phangnga (lanes 9-12, D) and *H. varia* originating from L-Island, Phuket (lanes 1-4, E) and Similan Island, Phangnga (lanes 5-7, E) with the primer UBC195. Lanes M and m are 200 bp and 100 bp DNA ladders, respectively. An arrowhead (◂) indicates a species –, population–, and region– specific marker.



Appendix A1.4 RAPD patterns resulted from analysis of *H. asinina* originating from Samet Island, Rayong (lanes 1-4, A), Talibong Island, Trang (lanes 5-8, A), Cambodia (lanes 9-13,A), P_0, Cambodia (lanes 1-7, B) and F_1, Philippines (lanes 1-7, C), *H. ovina* originating from Sichang Island, Chon Buri (lanes 1-3, D), Samet Island, Rayong (lanes 4-6, D), Churk Island, Trang (lanes 7-9, D) and Similan Island, Phangnga (lanes 9-12, D) and *H. varia* originating from L-Island, Phuket (lanes 1-4, E) and Similan Island, Phangnga (lanes 5-7, E) with the primer UBC197. Lanes M and m are 200 bp and 100 bp DNA ladders, respectively. An arrowhead (◂) indicates a species – and population– specific marker.



Appendix A1.5 RAPD patterns resulted from analysis of *H. asinina* originating from Samet Island, Rayong (lanes 1-4, A), Talibong Island, Trang (lanes 5-8, A), Cambodia (lanes 9-13,A), P_0, Cambodia (lanes 1-7, B) and F_1, Philippines (lanes 1-7, C), *H. ovina* originating from Sichang Island, Chon Buri (lanes 1-3, D), Samet Island, Rayong (lanes 4-6, D), Churk Island, Trang (lanes 7-9, D) and Similan Island, Phangnga (lanes 9-12, D) and *H. varia* originating from L-Island, Phuket (lanes 1-4, E) and Similan Island, Phangnga (lanes 5-7, E) with the primer UBC271. Lanes M and m are 200 bp and 100 bp DNA ladders, respectively. An arrowhead (◂) indicates a species – and population– specific marker.

TCCCCGTTCCCCATATCTGTAAAAATGCCTGAATTTATCCAACCTTTTACCTAAACCTGTAAACATGCTTTACTGCACCCTACTGC
ATGTGTGTGAGTGCATGTGTGCGTATGTGTGTTTTACTAGTAGTTTCTTGTGTATTTGTTGGGATGGTATACATGACTACCAGAGT
TGTTATTTCATTAATGTTGGGTAATATCATCAGTGAATATTTCCAGATGGTCACAAACACTCACTCAGTTTCAATATTTCACTATA
ATATTTTAACCGAACAAATTTTGTCGTAATTGTGAAAACATTATGAACAAAATGTATTGTAAAAACATGCCTTTTTCCAGTTTTAT
TTCAGATGAGACTGAAATTTGGTCCTGAAATTGTTTTCCTGTCATTCACATGCAATGAAACACATTTAA**GAATCCAACATGCGTCA**
AAGCTCGTTTGAAGAATGGGGCTGCAAGTTGACACCCCCCATGTTGCCGGTAGACTGTCTCCTACCCAATGGCATCATGGTCCCGA
TGTTTTGCTGCAGGGATGCCACTCTGGAGAAAATCAAAATTGATCTGTGGATAGAGTCCAAGAACTATCCCCTTCATCAACTCCTC
CAGAGTCGAGAGTCATACATCTTTGTCGGGGTTACCCAAGATTCTGAGCGAGAGGAGTTCTATGATGAAACTCGGAGACTGTTG**TG**
ACCTGCGATTGTTTCAAGCCACTCTCAAAACTCATGGAACGGGGGAAATCACTAGTGATTCGCGGCGCTGCAGGTCGACATATGGG
AGAGCTCCAACGCGTTGGATGCATAGCTGAGTATTCTATAGTGTCACCCTAAATAGCTTGGCG

Appendix A 2.1 Nucleotide sequence of pCUHA1. The locations and sequences of forward primers and those complementary to reverse primers are labeled in boldface and underlined.

GCCATCAAGATGTGACATGTTCATGGGCAACTAAAATGGATTGAAAAAACGTTACAAGTATTTCACAAACCTTTCCCTTTATATAT
TTATCCTGAATTAAATTGAACTCTGCTGAAAAGTGGCATATCATTAATCAAAATATAGACTGGTCAAACGTAATGACTTCTTCATG
ACTAAATATAACTGAAGATCAATGAACAAGGTAAATATTGCTGTAAACAATAAACAGGCACATGTTTCAAAACAATTCCACACAAA
GGTAACTTTTGCTATTTGGTTATGTAAACAAGCTGCAGAGCCCTCTTATGAAAGTTCACCTCCCAATGACTTCATTTATCAGTAAT
GTGTGATCATATCCAGGATGC**TTGTTCAGCATTCTGTGGCAGTTC**TTGACAACTTGATTTCAGTACTGAATGAACAAGAAAGACTG
ACAATTTTGTTCCATGTTGCCCCCAAGTTATTCTTGACAGATGTAGATTCTGCATTAGTGACAAAGATTTTTATATATC**TCCAAAG**
GGTCAGCAAAAAAGAAGACAGTATACCCTGGTGATATCTTCCACTGTTTGTCAGAGGGAACATATGGTACTGTATCCAGACACACG
TTATAACCAGGGTGCACTGTACATGCCCACATGGAAGCATGAATCATGACAGTCAATACCATAACTGTATTTAATTTACAAAAAAA
AATATATATGATCTTGATGGCG

Appendix A 2.2 Nucleotide sequence of pCUHA2. The locations and sequences of forward primers and those complementary to reverse primers are labeled in boldface and underlined.

GATTGCCATCAAGAATGCAGAGAATATAGAGTCATTTGTAATGAACGGATGGAACAACTTGGCTATTCACCAGTGGTCTAATACTA
GAGGGAGAT**CCTGGCAAAGGAAAGTGGTG**AAAATATTTGAAGAACTTAACCAAACTGATATACTGACGCTGTTAGCTAAAGTACAT
GGGACACACAACATATCTGTGAGCCAATATCTTTACCGATGTCTTTCGTGTAGCTGGTTTCATAATGGCATGTTTGTCTGGCCTTA
CTTTGATCTCACCTTCAATAATCTGATGTGAAGCCAGCGGCTAGTTGTAAGGTCTCTTTGCTTGCCGTCCGCGGGTATCATGTTGT
AAGGTAGGTGAGTTTAGCTTTACGCCGCTTTTTAGGAATATTCCAGCAATATCACGGCGGTCGACACTAGAAAATGGGCCTCACAC
ATTGTGAAGGAATCTAACCGGGTCTTCGGCGTG**ACGAGCGAACACTTTAACCA**CTAGGCTACCCCACAGACCCCATGTTGTAGGGT
CCATGTACGTCTTCTGTGGGTTTGACTAACGTCATGTCGTTGTTTGATATGTGGAAATGAGTGACGTTATACTGTTGTTGGGTGTA
TCGGTTGATATAGCCTTGGTGTGTTCCCTGTGTAAGGCATGATGATATAACATTGCTGAGTGTGTCCTCTGTGCAACGGCTGATGA
TGTGATGCTGCCTGGGTGTGTTCCCTCTTGATGGGCAATCACTAGTGAATTTCGCGGCCGCCTGCGGTCGACCCATATGGGAGAGC
TCCCAACGCGTTGGATGCATAGCTTGAGTATTCTAATAGTGTCACCTTAAAATAAGCTTGGGCGGAAATCCATGGG

Appendix A 2.3 Nucleotide sequence of pCUHA3. The locations and sequences of forward primers and those complementary to reverse primers are labeled in boldface and underlined.

GATC**TCAGCGAAACCAACCAACA**CTGCCGTATTTTTAAATTTAAAAACATAAATTTAAATATTTTGAAATAAAATTATATTTAACT
GAAGCCATGAAGAAGTACTCTCTCTCTCAGTGGTGACAAGATCCTTGTATAATGTATAATATATATTATGAGTATTATAAAATAGT
AAGTACCTAACTAAAAATATTTTCATAACTTTGAACGGTATTGCAAGCTGTAGAATATTTGCAGAAACGGTTGCCATTGATTTCCA
GTCTTTTATTCTATCAAT**GTGAATAGCTGCGTCCAA**CCCGTTGTTGAGAATTTGCTTCATAGGAGTATTAGCGTCGGAACCCAAAA
GAGCATTGATACAAGTATTAAATGAAATTCTAAAATCAATTCAAAGGAACAAATAGTGTCATTATGGTTTTTTTTCTGTATATGTG
CGGCCCAACTTAAAAGACATTCATAAAAGATTATCTGGGTTAATCGCCAGCCTATGTCACACGTGTTTTATTCGTAGACTATGATA
ATTGTATACTATCCCCCTACTACTTATGTATTGACACAGTGTGACATGTCCGATACCTATATATATGTCACACCTGGGCTGTTCGA
TGACAGGTAATAAAACACTGATCAGTCGCCTCACTGATGTTAACCCCCTACACGCTGAGATCA

Appendix A 2.4 Nucleotide sequence of pCUHA4. The locations and sequences of forward primers and those complementary to reverse primers are labeled in boldface and underlined.

GCCATCAAGACGAAATTTTAGAAAGGCTAACTGGGCTTTATATGAAACACTGTGTGCTGAAAAACTTAAACCTGAACGTTTTATTG
ACGTTCCTGATGCTATTAAATGCTTTTCTGATGAACTGAATTCCATAGCTGATGAGTGTATACCAAAGTCCTCTGTAGTTCCACAC
ATAAGAAAACCATGGTTCAACGATGACTGCAAACAAGCTAGGAAGGCAAGGGAAAAAGCTGAACAT**TATTTCCGTCGCCATCCTAT**
GGTGCATAATTTCAATAAATTTAAAGTTTTAAATGCTTAAAGCACGGCGTACTTTTAAACAGAACAAACGCCAATCTTGGCAAAAT
TATGTAGCCAAAATAAATTCTCGGACACCTATGTCCAAGGCATGGAACATGGTCCAGAAAATCAAAGGTAAAGGTACTAAATCTAG
TGTCCATCATCTTAAACATGGAGATCAATTACTTACTGATAAATCAGATATT**GCAAATAAACTGGGCGAATC**TTGCTAAACATTCT
TCCTCTTCTAATTATGTACCTAAATTTCAGCAATATCAAAAACAACAAGAAAAGAAAACTATTAATTTCAATTCTGATAATGGGGA
AGATTATAATGAAACATTTTCTATTCATGGACTCCATACTGCTCTTGATGGC

Appendix A 2.5 Nucleotide sequence of pCUHA5. The locations and sequences of forward primers and those complementary to reverse primers are labeled in boldface and underlined.

GCCATCAAGAGTGCATCATGAATAGTAACGTATGTCAAGGAACAGTCGTATATGGTAAACATCATGTCTTTAAAGACCAGTATCAT
ATCGTAAAAGACGTATCTGTTTCAGAAAGTGAGC**AAGTGAGTTTCGTTTTATGCC**GCACTCAGCAATATTCCTGCTTTATGGCGGC
AGTCTGTAAATAATCGAGTCTGGACCAGAC**AATCCAGTCATCAACAGCATG**AGCATCAATCGATACAATTGGGGACCGATGACGTG
TCAACCAAGTCATCGAGTCTGGCCACTTGATCCCGTTAGTCGCCTGTTATGACAAGCATGGGCTGCAGAAGACTAGTTGTAACCCG
TCTCTTGACGGGTGCATGTAACTTTTGACAGATATCATATTGTGACCCGTCTCTTTATGGGTGCATGTAACTTTTCACAGATATCA
TATTGTTGCCCGTCTCTTGACGGGTGCATGTAACTTTTGACAGATATCATATTTTGACCCGTCTCTTGACGGGTGCATGTAACTTT
TGACAGATATCATATTGTGACCCGTCTCTTGACGGTGCATGTAACTTTTGACAGATATCATATTGTTGCCCGTCTCTTGATGGGTG
CATGTAACTTTCGACAGATATCATATTGTGACCCGTTTCTTGATGGC

Appendix A 2.6 Nucleotide sequence of pCUHA6. The locations and sequences of forward primers and those complementary to reverse primers are labeled in boldface and underlined.

GCCATCAAGACCCAACGTCCAAAACAGATGAGTGAGTGAGTGAAGCGTTTTGACAATTAGTGTAGGTTCGATCCCCCAAAGCGTTA
GTTTTAAAGCTATTCTCAAGGATTCTCACGTTTAGTGAAGTCCTATATTATT**GTGTCATCATTTCGTGCAGTTG**ATCCAGTCTACC
GAATATGCCTTTCTCCGAAACAAGACACGAGAAGGGTCATTAAAACATCGGAATTTCACAAGAATATCTTTAAATTTTAAGGGAAT
AAATCGGAATTTGATACAATTATGTAAAACATGTATTAATTTATATGTTCTATTGTCTGCGACCAGATCTTAATACACGCATGGGC
AATGTACGATTATTCAAAACCAGATTTAATCAATAAATGTACAATCTACGATCATTTAGATGCAAATCTGAATAAACAGATGTACA

ATCTACGATCATTTAGATGCAAATCTGAATAAACAGATGTACAATCTACGATTCTATAAAACCAGATCTAAATCCACAAATGTTTA
ACCTACAATGTACGATGACTGGAGGATGCATCCACACCATGACTTGTATCTAGGTGACAGCTGATTCAGCAAAAGTACATGTAGGT
GAAAATGTAGCCAAGGGTATTCTAAACACTGGCTAGTGAATAATTATAACATCTTATCTTCT**GGAAAAATCGCTGACTAGGAGCAA**
GAATTAGCTCTGAAACGTAGTTTGAAGAATAAAAAGTGGTTCTGAAACGTGGTATAAAGAATAAAAAAGAAGTACTGTTAGTTTGC
ACAAGGGAAGGGTGAACAAACCCCCCCCTG

Appendix A 2.7 Nucleotide sequence of pCUHA7. The locations and sequences of forward primers and those complementary to reverse primers are labeled in boldface and underlined.

GCCATCAAGACTCATTGAGAGATATCTGATAATGCCGGGTATTACCTGGTATCAGCCCAACAAAATAACCCAGCAATACCAACACC
TTTCCCATGTTGAGAAACTATTTAACTGAATCACATTAAAGACGTTGAAATCCAAGAATATCCACATGTCATTACAGTAGCTTCTG
TGAGTTGGATAAATATTGTATATTCCATGGCAGACCTGCTAGACGAGATCAAACAGGACACATATTCTCTTAACTGTATGTTTCTG
GGAGTGTGATAGGAGGTGAATGGAGGCAATTGGCCATATCTGCAGTGTGAGATTACCAGATGAGTAGGGTTACCCAGTCTTCTGAG
GAGCAGCCCCAGGGGGTAGGGGGCACCTGTCAGTGTGGGGTGCTGACAGAAA**CATGGTTAGTTGTGAGGAGGAC**AGTATAAAGTTG
AGGGCACTGTTGTGCTTGCATGCAGGTATGGAACAGTACTAGTGTATGAAGA**CTGAGCTAGATAGGTTACATACTTCTCCTGAC**TT
AAACGGAAACAGTATCTGGTGTATTTCAAGGATTCAATAAGGACAGGTATTGCGTGTAAGACATTCTAATCAAATCATTTGTGACT
CAGTCTTAGACAACCTAAAGTCTATGTCAAAGAGGTTTAATTTACAAATTTTATTATTAAATTTTAGCTATATTTTCAGTAGATTT
TACTCTCAGATTATGGACATGTGTATGCTGATCTGAGAAGTTTTGTTGAAATTTAATTAGACTGAGATTTCATAGGACTACACCTT
GTGTAGGCTGTGAGTGAACACAACGTCCCATGACAATTAATTACGTCTGCTGATCTTGATGGC

Appendix A 2.8 Nucleotide sequence of pCUHA8. The locations and sequences of forward primers and those complementary to reverse primers are labeled in boldface and underlined.

GCCATCAAGAATGACATGCATCTATCTCCCTTAGTCTGGTACTACATATTTATGTCAACCAGAATCATACATATGACGCAAACGCT
GATACATTTTGCGAGATGCCTGCGTCATTTCACGTCATCTTTGGACTAATAACCATTAATTTTCTTCTGAAAATTCTGTCGTATAT
TTTCAGAAATTCTGTTGTAGTAGATGTGACCAGTTACTTCTTAACGATAAA**TCACCATTTAGCATAGTTCCTTTTAAA**CCTGTATC
CACAGAGTGCCCAATCTTGCTGATTGTTACAAAGTTTTAATGACATGAACA**ATCAGCAA**G**AAACAGCCAGAAAACGCAT**TGTGGCA
GCCATACCCACAATTATGTACTCATCTTCAGCTTCACCATCTCCATGCTGCAAAGCAGA**ACCTTGCTTCAGTCTTGAAA**ACAGTGC
TCAAAATACGTCTTGATGGCA

Appendix A 2.9 Nucleotide sequence of pCUHA9. The locations and sequences of forward primers and those complementary to reverse primers are labeled in boldface and underlined.

GATCTCAGCGTCCTTGACCTTGTAATAGTAATGGCGGGCCGTTAACTAAGTACCTGCTTGTCACGGTCCCCCACGCGACCATTTGT
ATGCTGAGCTGGATTCAAACT**GTCCGTCCCAGGTGCTTCATATC**TGGGTTAGTTATCACGATTCATCATTTCATTTATTATTGTTT
CATCAAATTATCGTTTTCACTTTGTAGACAGGACCACAAATCCTCATAGTTTTTATTTAAAGATTAAAAAATACCTTTGACACTTC
TGGTGTTTATTGATCTAAACTGTTAAGACGACCCCAAGAGTCTAAATTTGCGAAGCCGTCTTGGTGTTAAGATAGTCGTAAGGGTC
ATACGAAAGTGAACCTACGACTGTCTTAGCTCCAAGAGAGCTTCTAAACTCTACGCCTAGGCGCGTAGCATGCCATTTATTCATGT
TGCAAAGCTTCATATTTAAATAGCATTGATGGCTTTTCAGCCTTTTGTACTTACGACCGTCAGTGACCTATGATTAGGTCTGTTTG
TGGAACAGGCTTGACGAAACTTTCGTGCGTATCACACG**CAACCTATGACTGGCCGTGCGCT**GAGATCA

Appendix A 2.10 Nucleotide sequence of pCUHA10. The locations and sequences of forward primers and those complementary to reverse primers are labeled in boldface and underlined.

GATCTCAGCGTCACATCGTATTATCAATCTCAACT**CCCCGAGGAGTATACAACTCTTC**CTGCCACTAGGCGCACCGAGTTTATTGT
GTCTCTCTCATCCCAATTCACCTGCTGGGTGGACTTCGACAGATAGTCACAATACTTTGTCCAAGTTTACTGCACGTTGCTGTAAC
CGCCACTAGGTGTATGCACGTATTCATGTGGCCACCATCTGGTCATATACCAACGGATTCATGGGTTCTTTTAAGTGCACAAAGTT
ATGTAATGCACACTAGGGATTGTGACAACTCGGAAAGAGTCTGCGAACAATGTTGACTCCTAGGTTTTTACACCCGGTCACAGGTG
GGCTCGATCCAAACAACTCGCTGGATCACTAGCCTGGCGCTTTAGCCATCTCATCCACCGCGCACCCGAGGAATGAATGGCTTT**GG**
TGCATTGTGGAAAAGAACTCGAGTGAGAAAGGACATGGGTTGCTTCAATAGTCTCCCGCCAACTCCTGCAGTAAAAATAAACCTTC
ACTGTAAATAAATTCACTTTCGTTATGCAGTGATATATGACACTTAAGGGATTGAAATTCTTGATAATCATTGGCAAATATGAAAG
TGATCATTTCGCCCTTATAACTTGGAATCCTTCCTGAACA

Appendix A 2.11 Nucleotide sequence of pCUHA11. The locations and sequences of forward primers and those complementary to reverse primers are labeled in boldface and underlined.

GCCATCAAGAATAAGCGTTATGAAAACACTAACTCAAATCAGCCAAATAATCCCACTAGTGATATCCTCTTTGGACTGAAGTTGTT
TGTTTTAAGAGTAAGAGGATTATAAGGCATGCTCAACTGGGTTTGGTCGTTTAA**CTAATCCCACACAGCCATCACCAG**CAGAGGGA
ATGATTTACACATCTCATTTTTTTGTGCTTAATTATTGTGTTTGTTTTTTATTGAAAAAAAAATGAAATGGGAGTGTGAACAGCTTT
ACAGTTTATATCGACACTGGTACCATACAACATACAATACAAAATTTAAATGCAGCATGTGTTGTCACATATGTATATAAATTTGA
ATTACAAACATGGGTAAAGCTTGAAGACAATCTTATATGGAAATAGTGAATTATCTGATCTTGATAATCTGTCCATCTTAAGAT**CT**
GCCTACCTCTTCGTCACTTCTTCTTTTAGATTTGATACTAAACATGATTGAAGAATTAACCTCGTGGTTTGTCATATGACAATCTT
GTATTAGAACTAGGTTAGGTTTATTTACCTTTGTAAAGTGCATTACACACATACAAACTAACTGTATTAATATAGAGCATTTATGA
GAATGGTATTGGTAAGAAATTATTTTTGGTGTCTTGATGGC

Appendix A 2.12 Nucleotide sequence of pCUHA12. The locations and sequences of forward primers and those complementary to reverse primers are labeled in boldface and underlined.

GCGCCTGGAGACAAGGTTACGTGTAATTATATACTGACACATATCATTGTGAGATGATAAGAGTTAACACTGCATTCTGGATTATT
CAGTGTGTGTGAACAGTGTCAAACTTAATGTTAGCGCAGTACAGCACATCAC**TGACCTGTGTTGAGACTCTACGGA**AAAGTCTAAT
ATCGTAATGCACAACAATAAAAACCATGGTTACATTTCTATTACCAAAAAACTAGAAAAACAACATCTACCAGAATGAGTAAGTGT
GTGAAGTTTAGTTCTACACCACTTCTAGCAATAATCCAGCAATATGACGGGAGTGGGGGAGATAATAGAAACGCTAGAAATGGGCT
TTACACATTGCATCAATGAGATGGTTTGCACTTTGGTCTTGGACGTGACGTGCAAATGCTTTATCCAT**GCGGCTACTCCATCTCCC**
CTCATCTAACTGACACAAACACAGGAATTGTCAGCTGTCAAAGCTNCAGTACAGTGACTCTCTGCAACCATTGTTGTTTTGTGTTG
TCACAGGACCAGGAGCCCTATGGGTTAAAAAGTTGCCAAGTCCCACCCACCTTCAGCGC

Appendix 2.13 Nucleotide sequence of pCUHA13. The locations and sequences of forward primers and those complementary to reverse primers are labeled in boldface and underlined.

GTAGACCCGTGAGAAAGTTGGACAGTAAAATCGCGTGATTTTATATCTTTTCTAAATATTGAAACATGACAGGTATCGTTTTCGTC
TTTCAACG**CGTGAAGACAGTTACTGAAAGTGG**TATCAATATCCACGTCGAAAATTACTGACAGTCATATGAATACCACATGTCAAG
TATTCTTTATTATTTGTACGGTTCACTCACGTCTATATTTACATAAAATGTTAGCAGCGTGTTTACAGTGTTGGTTTCTTTTGTTT
GTTTGTGTTTTGTTGTTTTGGTTTTTTTTGTTTTTTGTAAGTTTGTTTGTTCGTTTATTTGTTTGTTTTTGTTTTTAAATAGCCTA

AATTTGAAGACATGTATTGAAATATGTATGGAATTTACTGGGCGAATGCTCTGGAAAGTTTGGGAGAATGCATGTGACTTCTTAAA
ATTGTTCAGCATGTTTAGCATATCTCTGTTTAACTAAGTCTGGCAAAACTGAACGTATTCAATTGCTAAAGCTGCTTTTTGGAATA
CTAACATTATCTAAGTAACTAAT**CAGAGGAGACATAACACAAACGAT**GCGTGATATAGATGCCAT

Appendix A 2.14 Nucleotide sequence of pCUHA14. The locations and sequences of forward primers and those complementary to reverse primers are labeled in boldface and underlined.

TCCCCGTTCCAGATGTAATCAGATCAAGTATGGTGAATGGTGATACGCCTACGTTGTATAGAACTTAACTCCGGGGAGTTAAGTTT
TAGGAAATTAATTATCTCTTGTGGAGATAATCTCTTGTGGAGATAATGTTTTAACTATTTGATATTTAGCATAGTTATCAAATCTG
TTCATTTTAGAAAGAAAAACAAATAATAAAATAATAGC**CGAAAGCAACAATAACTAAACC**AACCATGTCATCGGTGTATATATTTT
TTTCCTTCACAGTAAGCTAAGGCCAAGTTTCATATTGTAGATCTTTAGAAAAAGATTCATTTGGTAAGTTTGATAAAGAATACCAG
AAATCTATTTTGACCCG**CGACTTTAAGATCCGTGTTA**GAACTTCAGCAATCTATGCTTGTCGTAAGAGGCGACTATAACGGGAACG
GGGA

Appendix A 2.15 Nucleotide sequence of pCUHA15. The locations and sequences of forward primers and those complementary to reverse primers are labeled in boldface and underlined.

GTAGACCCGTCAACGAAGGCGTCTATGGATTTAGATTATACAGGTCGATAACCCTGAATTAAGTATCGATCAGTTATCGATGAATA
TCGCTGTCATAACTTTGAGCACAAGTGCAAAACACTGTCTGAAAGAAATGGATATACCATCGCATCGGTGACTGTGTGGTAACATG
TATTTGCGTGGGCTTTTAACAAAATAGCACCCTAAACATGACTTTTACGCAGACCATAAATAAAGTCGTAAATTTGTTTATTGTAT
GAGAAAGTTAAAAATTCTTTTTTCTGTCACAAAAAGTATGCACGATACTTTCGATAATGCTTAACATTTCTTTCCTGTGGGCGATA
TTTCCGATCGATCGATCATTTTAGTCATCCCTAGTAGGAATTCGATATTACACATAAGATATCCAGAAAATAATGTCACAACATGA
ATTCCCCTTCCATCTGTATATGTAAGCATGAGTGGATGAGTATTGTTTTATGCCGCATTTAGCATCATCACTGCTGGGGACACCAG
AAATGGGCTTCACAGATTGTACCCACGTGAGGAATCGAACCCGGGTCTTCATTGTGAGGACCCAACGCTTTAACCACTTGGCTACC
CTACCGCCCGCTTACTTAAAAGTGAAATTTACCCCGCTTTTAAAAATTTTCCAGCAATATGACGGCTGGGGACACAAAAATGGGCT
TCTCATATTGTACCCATATGAGGAATGGAACCCGGGTCTTCGTCGTGAGGAACCAATGCTTTAACCACTTGGCTACCCTACCGCCC
GCTTACTTAAAAGTGAGTGAGGAGTTTCAATCCGCTTTTAACGATATTCCAGCAATGGGGCTGGGGACACCAGAAATGGGCTTCTC
ACACTGTACCCATGTGAGGAATGGAACCCGGGTTTTCGCTGTGAGGAGCCAACGCTTTAACCACAACAGAATGCTACTCAACCCCC
CTTTTACGAAAGATGTCGACTGAGTTTCGTTACAGCAGACAATGCTTGCAACTTTCACAACAGCGCAAAACCATGTACTTACTCTG
CGATAGTCTCTGTTAGGATTCAAGCGTGCATGCCTTGACAATACGCAATAAATCTAAGACATTCGTAACTCTCAATCTGCAGCATA
TACCGAGAACTAAAAATGGTTATTTGGTAGTAACTATGTGCGATAAGCCTGAATAAGTATGAATTAATGCCTACCATCGTACTGTG
TGAATCGGAAATGAACGGTAATCATATACACGGGTCTAC

Appendix A 2.16 Nucleotide sequence of pCUHA16

GTAGACCCGTATGATGCTTTGACTCAAAACGACAACATGGCAGGATTTTATGTTGACAGTTGTTAATTTTTAAAAAAATCCAAACA
AAATAATCACTTTCATACAAAAACATGTCTGGCCCTGGAATATTTCATGTTTTATCTGTGAATGTTAAGTATGCATTATCATGTGC
TTCTCTTACACGATTCAAGATCAATAGGGTTAAAATCGGGGCACATTCAATTATGAGATTCTAAGTTCTTTGTATGTATTTCTAGA
TTACTAAGAACACAACATCACATGTAGTAGCCAGTCGGGAATATTACGCGCTATAACCAACAGGTTTTCGAACCACCTCCATTGTC
AAGCCGTAGCGTTCCTAAATAGTCAGGTACACCAGCTGTGCAATCTGAATTTAGTTGTCTAGCTGATTTGCCATAATCGTGTGTCC
GATGTTAGTTTGCATATCAATTGTATGCATGTCAAGTTCATTTATCATTTATACGACATCATGTTGAAGCCTGTGGTCGTTTAAAA
CGCCAACTGTGCAATTTAGAACTGTCATTACCACCTTCCTCGTTTGTGAGTGAGTTTAGTTTTACGCCGCACCCAGCAATATTCCA
GCTGTATGGTGGCGGTCTGTATATAATCGTGCTGGACAAGCTAATCCATTGATCAACAGCACGAACACCGTTCTCCGCAATTAGGA
ACCGATAACATATGCCAACCAAGACAGCGAGTTGTCCACCCGATCCCGTTAGTCGCCTCTTACGACCATCATTGTGTCAAGTGTGG
GTTGCTGAAGACCTATTCTACCCCTTATCTTCACGGGTCTAC

Appendix A 2.17 Nucleotide sequence of pCUHA17

GTAGACCCGTCACATACACATGCAAGGGAAAGGCATTGGTACACAACTAGGAACATACGTCGACTCGACTGCGTGTATCAAAGATT
TTCTCTTTATAACATTAATCTGTGTGATTAACAACTCGAGCACGAGACTGTCTACCGTCAGATCTAGCACACAAAACCTCCATAAA
TGTGTAGAATGAAGACTCTGAGCTTCTTGGCCACGGGCATATCCTCTTTGATATTATATAGTATCAATATGGACTGATTTGTAATG
CATCTCTTCCTATTAGTGTGAAATGCCATATGCCTTAATGTTCCAACAGTTCGGTTAATGGCATCGGAAATTAGACAGGAGCGAAC
CATGGCCGGTTTTGCCCGCCGGGTGTTGGTGATTTTCAGTCATGTGACAACAAACCAATCTAATCAAAAGACAGAACGTTACAGGA
AATTCTGATGTTGAGGTAACGTCTATAACTTTTAAAGCAAAGCAATTTTGTAAACATTTTGAGTCGGGAAGAAAGTATTAAAAAAA
CAAACACACCACCCGATAATGACAGCTGATTAAGGCATTCTGTTATGTTTAGAGGCAGTTCAAAGGCGCCTTCTGTTTAACAAACA
ACAAATAGTGTCAGATAGGGGGTGAGGCAGCTACTTGTTTAATGCATCACTCAGCAATGCCCCAGATATCTGTAAGTACTCGAATC
TGGTGTAGACAACCCAGTGCTCAAACGCATGAACATTGATCTACGCAATGGGGATACGGTGGCATATATAAACCATGTCAGCGAAC
CTGACCACTAGAAAAGATCATGTCAGACGAATTCCTTCCAATATCTTCACGGGTCTAC

Appendix 2.18 Nucleotide sequence of pCUHA18

GCGCCTGGAGGACATATCAGTCAGGACGAACCTCTGTCCATAAAATCCACATCGACACATACTTATTCTCTGAATGTATTGCCGAG
AGGAGTGAAATGCGCCATGGAGTTCACGTCACACTGTGGGTCTTGTGTTTCGTTGAACTTCAGTAAAGAACTGTTTCATTCGGCAG
TGTGATGTGGATTTACTGGACTCATATCTTACGATGGGAGTGATCTCACTGATGGCCTCTGTCAATTCAATTGTGTACATGTTATG
CGTGAAGTATGAACGCCAGGATTACAATTGTATTTGATGTTCTTTCATTTTTATAATTCTTTCCCATGTGTACCTATTGTATGGTT
GTTTTGCGAGAATAGTACTATTACATATTAATGTATTGCCATTTGAAGGTTTAAATATAATGTGATTTTGCAGATCGTTAATTAAA
CCCTGTTTGACGGAACATTTGAAGCGAAAATACTTGAGACAAAAGGACAAGAGGTTTTATGGATGAGCAACTTCTTTATTGTGAGG
TGTCACGACGTTTCGAAGAATATTCTTACTTCTTCATCAGGTGAATGAGATGAAGAAGTAATCATTCACCTGATGAAGAAATAAGA
ATATTCTTTGAAACGTCGTGACTCCTCACAATAAAGAAGTTGTCATCCATAAAAACTATTGTCTTTATGTGAAAATACTGGTATTC
ATTTGATTTCATTGTAACAGATGAATGCAGTATAATTTTGTTTTGATTCAACAGATCAAAGAGCGAGAGTTTCGCCGCCAGTCCCA
AGCATCCTTGCTGACTGTGATGAGTGACGGGGATTCCGAGCCCGACTAATAACACTCCACCCTCGGCTCCTTGGACAGCGGTAATG
CCTCATCGGACACCATTCACACCATCATGGACGCCAACCGAGGGAAGAGAATGTCCGTTCCAAATCTTCTCCATGTTCAACCCACT
CCGGATTTCTCCAAACGGAGGATATCAATGGCCAACATCACGGGAGATTTCGTCGGGGGCTCCCTAGACCGGAAACTGTTCCCAAC
CAGTTTGTCTGCATGTGACACTGACAGCAGACGACTTTCTGTGGCTGAGACTCTGGAACAGGACTTGAGTGAAAGTGTGTCTATGG
TGGGTACAGACTCACAGAGTCTGTATAACACAATCGGGGATGTGAAGTGTGCCGGTGGTGGGGACCCCGGGACCAGGGACGATGGG
AACATCAAGGAACCTCTCCCTGATCCAGAAGTTACATGTGAAGACGAGTCAGCAGAATCAGATAAGGAGGACGAGGATGATCTCTG
GTAAGACTTTGATCTAGTTGTAGGTTGTCTAGTTGGAGATGTTTTCATCAGCAATACCTGGTGTTTTGAATATCCTGTACGTCGGG
ATACGCACAGAGATACGCACTATGTGTGATGACTCCATGAAACGATCGAAGCGCTACGACAGTCCTTGGTCAGCATCAAACTATAA
GAGTTACGATCACTTTTTCTGTACGGTATCTTCGTTAAAAGGAAGGGCTGTCCGTGTGATCAAGTGCATATTGTTGACCAATATCT
TTACGATTTCCGTTGATCGGATTGCCTGTGGCCGTCGCCATGTGCTGGAAAGAAGCATCCAGTGTTCACTCCATCTTTTTTCAAAG

CTATGAAGCAATGGAAATAAGGAAAACTTATTCCTAGGCATAGAAAGATAGCATACGCACTTTTACTTTATTTATTGTTATCTCGT
GTTCCACTCCAGGCGC

Appendix A 2.19 Nucleotide sequence of pCUHA19

GCGCCTGGAGATATCAGGTGTCTGGTGCTGTGACAGCCTCAACTCGACGGACTCTGTAATTTACAGGAAATGGAAGAACGAAATTA
GAACATTGCACAAGGGATGGAGCCTTGAGGGCACACACTAGGACAGGACCTGGAATTGTTCTCTAATATGAAGCTGGTCATTACTG
CCCAGGCTATTTAGAAAGCGTCTGACATGCCGGGTGGCGCGAAGATATAGCTGGATGGTTCGAGGATATTGATGGGTGGTGCGGGA
ATATAGCTGGGTGATGAGAGGGTATAACCGGGTGGTCTCTCTGTCGAATAGAAGTTTAATAACAATGGAGGCGCACAAAGGCAGAT
TTCTTTGTGTTCCTCACATCATGAAATTTTCACCATCATATCGGTCCACCCGGGCGGACACGTACTCTTCCACCAACGTGATGAAGC
AGACATCGTGATTCGTCACAGAACCATAAGTTCGCCGCTAGTTCGAACGTTGTTGACCCTCTTGACCCGTCGATGTAGTCGTCGCA
CCACAGGGCCAAAATAGGGTTTTCTTGTACGATATTTTTTGCACCGATACTCTCCTCAAACCAGACATTGCACTGGCAGGTGCACC
CTAGTTTGCAGCTCAGTGGTCATCCGAGATCTTGCGTGGAATTGGTTTGTTGATAAACGATCGATGGTACTCTGATGAACATTGAG
GTGACGTGCAAGGCAGACGTTTAGACTTCCCAACCTCCAGACGTCGGATTGCAATGTTTAGATTTGCAGCAGTGAATCGCGGCATA
TTTCCTTCATCGAGCTAAAAAAAATGACGGGACAATGTCATAAATACTATATGCTTGACGCCCAGAAAGGGCTTCGAGATGGACACG
GCCATTAATACTCCATGCTTGACGCATAGCTACCATTATAAAAGAAACAAAAAATCAAAGCATGGGCTAACACATCATTGTTGTTG
GAAAGGCAGGGCAAGCCTTCGATACTTTTTAGATATATAGCTTTTCAACCTGTCGGGCTTACAAGAAACAACTGGCATGTTGGCTG
GCTTCTTTTCCAGAATGGTAGTGACGTTAGAAGTGACATTGAACAAAGGTATAACCTTCTGTCTTTCAAAGAATGAATTCAGCTTA
TTCTGACGGTTATTTTGTTATTTCTGTTGAAAGGCAAAGCGACAATACCATACATACCAACAATACGTGCTATACGTAATCTGACG
TAATGGTAATCTTTGTCAAAACTCGACCTTTTCCGTCTGTCACGGAATGACCCGAACCCTCATAGAAGGATCGCTTGCATTGTCGA
AAGGAAGCAAGGGTCCAAGTCTGGCTCCAGGCGC

Appendix A 2.20 Nucleotide sequence of pCUHA20

GATCTCAGCGTGATGGGCGAACGCTTAAACTATTTGGCTACTCCACCGATCCCGATTCAAGAAGGAACCGAGATACGAATATCAAA
TACATCTCTGTACGTACCACGGATATCGGCATTGCATTTCAAGCCACACGTCGGCATTTGAGATAAAAGGCAGACATCAAGCAACA
CCCAGAGACTGCAAGACCTTACATAACACCTGACTATATGCAATGAGATCCCTCTTACGGACTAGTCTCGTACAGTTCTAGAATCT
ATCTATCTAATTTGCCTCGGTATCTCAACGGACCATCCTCGTACATAACACTGTCTGATCGCTGTAGCTCAGCAAGGAATTATGGG
ATAAAATCGTACAATAATGTAATACTTGTATGGTATTAAGCAGCGCTATGATTCAGCACCTAATGGCTACCTTTATCATAGGATCA
GTATGCGAAGGATATTCAGTGATGCTTACTGTTGTGACGATGATATCGAGAAGGTAAATACGTCATATACGTCATGTAGCATAGCT
CAATAGCTACTATTGAAAATGATCTTGTCACCAATTCCTCAGATATATCAGGGCACATATGCAGCCTGAAGCCAGGCTCTAAAGTC
ACCAAAGGTGTGACCGTACCAGTGCACCAGTACCCGCTTCCACAACATTCTGGGATTATCAGCTCATTATCGCCCGTATGGGTTAC
AACAGAACTGGTAAACGTCAAACACAGGCATCACTTT

Appendix A 2.21 Nucleotide sequence of pCUHA21

GATCTCAGCGCGCGATTGTCTGTGGCGTTATACCACGCCTACAGAGATCCTCACTCACTTGGTAAGTTAACGATGTACATAGTAAG
CACCACATACGCAGGAACTGTCAGTGTACATCACCTACTCTATTCCAGTGTCCTTTCCCACTATCCATATTCATACTCCATACCCA
TAATCCAACCCTTTTCTACCTCCCACAGCTCTCATAACTCGATGCCTATCAATAACCATATTGCATCCCACCATGCTTGTGTTGTC
ATTTCCTGTGAAGTCGCTATTTCCTACGGCGGGCCTGTCATTAAACTACTTTTCGGGCCAATACAATTTTGGTTTATGAGGCGAAA
GTCAACCAAGCAGTACGGTCGGTATAACGTCTCATTTCAGTCTGATGGTAAAGGGGGGAATCAAAGTGGCGGATACAGGGTAGTAG
GTGTTGGGAACACTCGCGTCCATAAATATGCTGTATAGATACGACATTGTGCTAATATGGCCTTGACCAAAGGGAACTAAGCGAAA
GCACAGAGGACGGTATTCATATTTAAGTGCATTGTAACTTATAGTTACTGTTGACTCTAGACCAACGTAGAGATACTAAAGCCTGT
CATTTCAAGATTAAGCGTGCCCAGACGGGATTGAGTGAACACAGTCTCCGCTTCGCAGACAGTATTTAGGCTAGTTACCGCTGGTA
CCCAGTCTGTGTTTTGAAACACTAATGGTGTGTCGACACTAATGGCCCTACTTTGTCGTCCCCGTGTCCAGTGCGTTAGGAATTTA
TACATCCTAATACAGTCATATAACGCTGAGATC

Appendix A 2.22 Nucleotide sequence of pCUHA22

TCCCCGTTCCTCTTCAAAGTATTGTTCCTCGTCTGATGAAGACGTTGTTGTAGCAGAGGCATCCCCATGTTCAGCTTCTTCTTCTA
TGTCGTTGATGTTATTCGATCTCTTTGATCTTCTTGATCGTGATGACGAGCGTCGCCGTTGATGCTGGAATCCAGTACCCACATTA
CGCATGCCTGCTTCAACTTTATCCAGTTTCTCAATCACCTGCCTCAAAATTTCAGTATTGAGATCCTGGTGTCTTCCTCCACAAAT
CCGTCTAAAGAACCGAACAAATCCACTGCAGCACGCATTATCGTCTTCATGTCTGGTGAAAAAAGACATCCAACTGGCTTGATTGT
TCTTGGCGTTCGGCATCTCTGGAGGCTTTGGTCCCGGGACTCGAGTTGCCACTTCTCTAGTTCCCCAGGAGACGACGTGGAGGTTG
CACATGGCGTAGATCATCAACAAAACGTAGCCAGAGGGGATGCACAGAAAGTACAGAGCCCCCGGAATTAAACAAATGAACTCCTG
CGGATGGAAGCATGCCGATATTATGAAGATTGCCAGTAACATCAGCAAGAAGACTGCGTTGGGACTGACGATGGTATCTTCTGCAA
TCTGAACCAGTGTTCCGACCACTACAGCCATCATGAGCAAGGAGTAGACCGCACTCATGATGGCAGCTATGTTCAACTGCATACTG
CCTTTAGCTATGAAGCAGATGATGAGATACAAGACAGACGGAGCCACGGCTAGGATGTAGGACTGCCAGAGACTGGTCCGCATGAC
GGCGTTGAAGGCACCAGCCATCATAAGCAATACGGTTGCAGGTCCGAGAATGGTGGACAGCATCAATAGACCTTGGTATCCCATGT
AGAGGTAGGAGATGTTGTCGTTGATGTTGATAGTTGTTCGGTAACTCTGTAGAAGGTCCATGATGTTAGCGAGAGTGGACGGAATC
CATCGGCGTCGCTGGTTGAAGAACTCCTTGAATCGTTCAGGTGCGTGTGTCATAGCGTCCGAGGCTGCACAGTACTCCACTCGGTA
TCCTTGCTGCAGCATCAGTGTACATAGCCATCGGTCCTCACTGCACAAAGGATGACATTTGTGAATTAATAAAAGATGTTTACGTA
CAAGACATGAATGGAAAGAGTTTTCCTGTCACTGAGTGAGTTGGGTTGAACATCGCTTTGAATATCATTTAAATCATACCACCACT
TAAATACTCGTACAACCCAGTGGTCAACTCTCAAGATGAATCTGAGATTCGAGCCAATGCTCTGAGGTTTATGGCACCAGTAAGCA
TATACATCGTTGAGATAATCAAGGGGAATCAAACGGTGACACATTTGAGCATACCCTGATCGTACTGCAGGTAGTGTCTGGCCTCA
CTGGAACGGGGA

Appendix A 2.23 Nucleotide sequence of pCUHA23

GCCATCAAGACAGTGGACAACACAGATGTTTCTTGGGGTTTTGGCGTAACCACAGGTGCATGTTCTTGCATATGGCAAATAGGTTA
GCTAGTGTTGGAGCCTTCTTCTGAACCCAGCCACATTCAGACACCTGCAGGCAATGATACACGTTAAGTGTGTGACTGAATATATC
ATTCTGACGAACACAAGTGGTGTTCCATTAACTGAAATAATCAAACATTAGTCAAAGTGACCAAACAACCAAGCAATCAATGTTAT
AAACATTTTTTCATGATTGAACACAAATGATTTTGGAATTATTGTTAATAAGGTCTTGACAATGGAACATCTTAAAAGTTGTCAGG
GCCTGAAAAAATGTATAATTCTTGGAAAATCGATTCTAAAACTGAAATGTCATGACTGAATATTTGCTGACTCCAGGAAGTGTATG
TTCATGATATGTTAGTCATCACTTTAAAGGAGCAAGCTGTATCATGTTTTACATCTCTAAAACAACAACATGAATAGGAAAATCTG
ACAAATGACGCTAACCAACTTCTTTCAATGATTGTTTGTTGGTAATGAAGCAATCATCAGTTGTGAAATTTCAATGAATTCCCAAT
ACAAACAAACAAACACAGAATTTCCTTACATGCACACAGGAAATAGGCATTGAAAATGTTTAATGTGAATGTCACTAACTGATTTC
AATTTGATACTTTACTGAAAACATAATAGGTAAATAAATCTGAATATCCTCCTGTTTTGAAATGCTTAAAGCGCAATGTCAAATGA
AATTGATTATGTAAGTGTCAATGTATCTTTAAACAAATGTCCATTTCCACAAGTGCTGTTTGCCCTTATTGCTTGATCCTGTGGAT
GTATTCCATGACATCTGCACATCCTGTGGGCACAGCAATCTAGTTCACAGACTTACCCTATTTTCAAATTTAGCCACTCTATATGA

TCTCTGTGACAGGTTATAGACAGCATAGGAGTTTCGGTGCCGATTCTCCAGGAAGGCTCGTACATCATCTATATGGTTCTTGATGG
C

Appendix A 2.24 Nucleotide sequence of pCUHA24

GCCATCAAGACACCAAAAATAATTACCTTACCAATAACATTCTCATAAATGCTCTACATTGATACATGTCAATTAGTTTGTATGTG
TGGAATCTTTAGAAGGGGAAAGTCAGATGTAATACACACATATATATATGGTTAAAACACATTTAAACGTGCAGTCGAAAATGGAT
TGGTAAACAAGCTAGAAGCTTATGTTAATACCTCTCTAATGCACTTTACAAAGATAAATAAACCAAAATTCTGATACAAGATTATC
ATATGACAAACCACGAGCTTAATTCTTCAATCATATGTATTATCAAATCTAAAAGAAGCAGTGACGAAGAGTTAGGCAGGTCTTAA
GATGGACAGATTATTAAGATCAGATAATTCACTATTTCCGTATAAGATTGTCTTCAAGCATTACCCATGTTTGTAATTCAAATTTA
TATACATATGTGACAGCAAATGCTGCATTAAAATTTTGTATTGTATGTTGTATGGTACCGGTGTTGGTATAAACTGTAAAGCTGTT
CACAATCCCATTTCATTTTATCAATGAAAAACAAACACAATAAATAAGCACAAACAATGGAGATGTGTGAATCATTCCCTCTGCTG
GTGATAGCTGTGTGGGATTAGTTCAGTCCACTGGCAAAGAGGATACCACTAGTGGGATTATTTGGCTGATTTGGGTTAGTGTTTTG
ATAACGCTTACTCTTGATGGC

Appendix A 2.25 Nucleotide sequence of pCUHA25

GATCTCAGCGAAACCAACCAACACTGCCGTATTTTTAAATTTAAAAACATAAATTTAAATATTTTGAAATAAAATTATATTTAACT
GAAGCCATGAAGAAGTACTCTCTCTCTCAGTGGTGACAAGATCCTTGTATAATGTATAATATATATTATGAGTATTATAAAATAGT
AAGTACCTAACTAAAAATATTTTCATAACTTTGAACGGTATTGCAAGCTGTAGAATATTTGCAGAAACGGTTGCCATTGATTTCCA
GTCTTTTATTCTATCAATGTGAATAGCTGCGTCCAACCCGTTGTTGAGAATTTGCTTCATAGGAATATTAGCGTCGGAACCCAAAA
GAGCATTGATACAAGTATTAAATGAAATTCTAAAATCAATTCAAAGGAACAAATAGTGTCATTATGGTTTTTTTTCTGTATATGTG
CGGCCCAACTTAAAAGACATTCATAAAAGATTATCTGGGTTAATCGCCAGCCTATGTCACACGTGTTTTATTCGTAGACTATGATA
ATTGTATACTATCCCCCTACTACTTATGTATTGACACAGTGTGACATGTCCGATACCTATATATATGTCACACCTGGGCTGTTCGA
TGACAGGTAATAAAACACTGATCAGTCGCCTCACTGATGTTAACCCCCTACACGCTGAGATC

Appendix A 2.26 Nucleotide sequence of pCUHA26

GATCTCAGCGTTCAGAGAATGGCGAGTTTTAACATGAGAATCTTGGATTAAACCTTAGTCCTTCCAGAAGAGGGAATTTCGAATGT
CCACCGGATGTTCACTTTCCCTTCTAAAAGGAAGTAGCCTTTTTCGGGAACTTGGTACACTCGTAATCTTGTCTAATGGACACATG
GCGATGTGATATTACAAAATGGACAGAATCTTAAATCTCTTGGGATATTATTTCTTCTGCTTGTATTCTTTGTGATGTTTAATTCA
CCTAATCTTAAAAGTGATAAAAATCTTCTTATATCTATGTTATATAAATTTGACTAATATGATTCCATGTTAAACACACTTTACAT
GAATAGAACTCTAAATTTCAAACATCTGATATAGGACAGTGGGAGAAAGTTGTTTTTTTTAATTTATGAAGCTGTTTGATTGGAAG
GAGGGCTTTATGAGCTTTATGAAGCTGAAGCCTGTGCGGCTTTTGACCAAGATAGTCAAGATATAAATGCTTTGTTATACTATGAA
TTGCTCAATCAAACAGCCTGAGTTGTTTTGCATTGGCAAATGCTTTGACGTCATCTTAGTGAACATCACATCGGTGTAAACCTTGT
AAACCACTGCAAACCAAAGGCTCTCAAAGTCATGCTAGAAGCGTTAACACTTAACACTTTACCAAATGCTAAATATATCAGATCCA
TATTTAAAGACCTGAGATTACTATTGCAGATTGATATGTTTACCTGGCAGTGATTTGTTTATGTTCGGCGAGAACACGTGGCTACC
GATCGGTTTGTGAAAGAGCTGTGACACCAACGGGGCCATGTATTGTCCATTCTCGGGCACCACGTGGATACTCCTGAAAATGTTCA
CATGATCATGATGTATATGAAGATGAGGTCGTGAATAACCAGTTTGAAGACATTGGTGAAGGTAATGCTAGACCATCGAAGTGAAA
GAAATTTCAGCCTGTTGCGTAGTAATAAACACTGAAGTCAGATATCCCGCTGAGATC

Appendix A 2.27 Nucleotide sequence of pCUHA27

GTAGACCCGTGCAGATTCGGGTTTCGATTGGTATTTAGCTTGATCTTCCTGATCCTTCTATCACTAGATTATCTGGCCCAGACTCG
ATTATTTACAGACCGTCGCACTATAGCTAGAATATTGCTGAATGTGAAGTTTAACAATAAACTCAATTTCCAACGCAGTCGTTTGA
TTGAGAGGGAAACAGCATGGGTTCGATGCCCAGTCTTGTCTTATTGAAAGGCTTTAAAATACTTCTTGTAACGCTGCTTGGTGCAA
TGGAATACTGATAGGAGCAAGTATATTGCCTCCAATATCGATTCCTTCATGTCTGCTATTGTTGCCGATAATTGTGAGTCATTGTA
AGTTATCACATCGTGTCTCAGGAAACACGGGGTTTTATCAGTCCCTCGTAAGGCTGTATTGGTTGTATGGATACAACCTTACGAGG
GACTAATAAAATCCCGTATTTCCCTAGACACGATGTGATAACGCATTTATCTAGTTGAATGTTTTCATTAAATCAAAACAAAACAA
CACGCATCGAATCGACTTTCTGCCATGTTGACACCAGCTGAGCGTTTGACCTCAAGGCACCGCACGTTACTAAAGGCTTGTCGATG
CACGCTTTTCTCATTTCGCGTCGTCATCGAGGCGTAGGGGGGTGATATGACTTTCGTAGCGGGAGTACACGACTTTTATACTCCCG
CTCGCGGATGGTGATGGATGTACAtGGTGTttTATCAGAGTCACGTGGACCAATCAAAATTCGACATTCtTACAtGAGGCCAGATA
ATAGTTTTGGTCATTCTtAACGTGCATTTAGTACAAGTTCAATAACAAGTGACTATACGCAAGTACAGAGAAATAACGTGATTATC
TTTGATGGGAGGGTGCATATTGTTTAATTTCCCATTCGCAATTGTGCAGCGATTCGCGGGGAGAAATTTCGGTGTGAGTTGATGAT
ATAACTCGAGGTGTTAAGTTGTTGATATCACTCGAGGTGTAGGTTAGTGATATCACACCAGGTGTTAGATTGATGTATCACATGTC
TAGTCTCCCACATTGAAACGGTTAAACAGCTGAATTTCAGTTTGTCTTGTGCAAGGCAGTGTCCTTGCAATAATGCTGTTGGAAGC
ACCCATTTGTGACATTCCTGTGGCCCGTTCTAGTTGCGCCGACGTCAATCTTGACAGTTTCTCGTGAGATCTTTGTCAAGTGATTG
TTTTTCACGACCAACAGGGTTTATaCCCTTTGATCGAGAAGTACTCGTACGTGGGAAAATAATCACGATTTCGCAAACATTCGTGC
AAaCCACATAAAAATATGTTTGGTTGCATTTTGGCAAATGGAACATGATGATTGCGTTCTATGCAATCAGTGCTCTTCAATAATGT
TACTATTTAGGACACAAATTTGTCACTGCAGCAGTAATTACACCCATGTACTCTCTTTTGCTCATCAACATATATTAATTTAAATT
AGGTTCATGTAAATTGTACTATtCCAGACCTAACTGATAGTACTGTGAAAGAACGTTTCAGATTTGAATGTCTTAGCAATGGCAAG
ACAAGAGTGAAGGCTGCTATTGAAAGTTACTAGTCATTCACAGGTCTGAACAAATATTTATATGGTATGTATGTACATGGTATAAT
TGCTGTGTGTCTAAAATGTACGTACGGGTCTAC

Appendix A 2.28 Nucleotide sequence of pCUHA28

GTAGACCCGTTACAACATTCGAAACACTTTTGTGTTTAATGGTTATGGTGGGGGCTACTGCGCCGTAAAGTGAAAAGAATTGTGGA
GAGATTATCCTTTTTTTTATGTTACACCATTTCTAAATGGAACGCTCCATTTTCATTGGTTAGTGTCTGGTATCCCTGGCATAATC
TTCCACACTACAACTGTGTCCTACATTACCGCATCTGCCACACTACAACTTTCTCCTACATTACCGTATCTCCCATCTATACAGAC
AGCTGTCTCCTACATTACCGTATCTCCCACACTACAACTGTCTCCTACATTACCGTATCTCCCACACTACAACTGTCTCCTACATT
ACCGTATCTCCCACACTACAACTGTCTCCTACATTACTGTATGTTCCACACTACAACTGTCTCCCACATTACTCTATCTCCCACAC
TACAGACAGCTGTCTCCTACATTACCGTATCTCCCACACTACAACTGTCTCCTACATTACTGTATGTCCCACACTACAACTGTCTC
CCACATTACTCTATCTCCCACACTACAGAGAGCTGTCTCCTACATTACCGTATCTCCCACACTACAGACAGCTGTCTCCTACATTA
CCGTATCTCCCACACTACAGCCAGCTGTCTCCTACATTACCGTATCTCCCACACTACAGACAGCTGTCTCCTACATTACCGTATCT
CCAATTCTAACCTATCTACAACACTAGCCTATCTCCAACATTTACCTCTCTCTAACACTAACAGAAACGTTGACATTTCCTTTCAG
CTTGGAGACTAAAACTTCGAGTTTTCAAGATCTTAACCCTGTATACTGTGACTGTAAATTAGGTCTAAATCTTCATAGACAATGCA
ACGTCATAATCATAACAGAATAACACATAAAATTGTTTAATCTGTCGGTAACATAATCATAACGGATCCGGGTACTGACACGGGTA
TTCGGGTCCAGTTCATTACCCACACGTCCCGGGTATCCGTGCTACCCGTCCCAGCCCAAGACTATGTCGCTATATCAGCGTATTAT
TCACCTGAAGAAGAGACTAGAATCTGTCTCGAAACGTTGTGATCTTGTATAATAAAGAAGTTGAACCATGAAGCACTTTTGGTTTG
ATTCATCCAACTTCTTAATGCCTTTGAAACAAGCTATATCAGTGGGTAACAATTTTAACAAAGATATGACGTATAGTTCGTCCGAC
TACATAGCTCATGCCTGCCCGTGCACACAGGCCACAGCTGGCTGAACAGTCGAAGGGAGGTAACTGCAAATATTGCAGATTTCCAC

CATTCCGTGATGCCTTGTATGATCATGCCGTATGCGTTCACCTATTAATATATAGAGAGAACTTGCTTACTGTCACTGCGCGATAC
CAGTATGTCACGCCCAAAACGTATAATTGGAATATCTAATTTGCTGTGGTGGGAATGTGGCCGAATGAAAAACTGGCTCACACACT
AACTCATCTCTATCAGTTGATGCTGTATATATAGTTTGACGCCGCATTCAGCAGTGTTCGAACTATAATGCCGACAGTCTGTAAAT
AATCGACTCTGGACCAGACAATCCACTGATCAACATCATGAGCATCGATCTGCACAAATGGGAACCAATGATATGTGTCAACCAAG
TCAGTGAGCCTGACCACCAGATCCCGTTAGTCTTATGAAAAGCATGGGTTACTGAAGATACATTCTACCACGGGTCTAC

Appendix A 2.29 Nucleotide sequence of pCUHA29

GCGCCTGGAGTATATCTCTTTAATTTTACATGCTAGGTTCACCAAATTTCACACACATGTTAATCATCACCCATAGCTGTGCCTTT
TGCTGTTTTAATTTTTTTGTCAGAATTTGTTTTTGTTGTTGTTTCCATGGAAACATGACTTGGTGCAAAATGTGTTTTATGTTAGCT
GGCTACAGCTTATAACTGAAACTAAACATGCTAGATTCACCAAACTTAATACACAAGTTAATCTGATGATGACCCACAGATATGCC
ATATGCTGTTATTTAAGTGTGACGTTAACTTCTCACTTTTCTTGGTCTGATTCTTCATATCATATGTACCAATATATGGTTTATTC
TCTGTATCTAGATTTGTTATGTCCTTGATGTTATCAACATTGATTTATTTTCTGTATCTGCATGAGGTTGTTTTGTTCATGTCGTA
AGTACAAATATAGTTCTAATTGATTCTAAATAGTTGATCATTTTCATATCAATTTTCACAGTGCAAACAGTCCTTAGATTTTTGAG
GGTTCACCTATATTATCCCTCAACTTATAGGC

Appendix A 2.30 Nucleotide sequence of pCUHA30

GTAGACCCGTGGACGTTAAGGGGGCTGTATCGAGTTTACAGACCACCAACATTGTTCACTATACACTAATGACTGTCAATACTCTA
AGTCAAAA**GGTTCAACCGTTAAGTTCGCAGAT**GATGCAGCAACTGAACGACACTGATTCTGATGAACTGTGTAAAGGGGACACTAA
TCGTTTCACCACTTAGTGGGCAGAAAACCGCTTTCTCGTAAATGTCACAAAAAGAAAGGATCTGATATTTGGGTTTCGGGTTAAAC
CCACGGAAATGCTTCCCGTGATAATCAACAAGCAGTATATTGAGACAGTGTCGGAATATAATCACGGGGAACCTATATAATCCAGG
CTGGACTGGATTTCTTACACTGAAGGCCTATACTACAAGTGTCATACACGTCTTCATATCATGTGTAAACAGATCTTTTAATGTTA
ACCCAGCAGTGATGTTATGATTTTATGCAGCATTAATTCAGAGTATTCTAAAATTT**TCCTTCCATTGTTGGTACGGGT**CTAC

Appendix A 2.31 Nucleotide sequence of pCUHO1. The locations and sequences of forward primers and those complementary to reverse primers are labeled in boldface and underlined.

GTAGACCCGTGGCACAGACGTCTTCTTTGTCACATATTTCAGGGCAAAATTGATTGTTCCAAAATAAATCCTGAAAACAAATAAAA
AACATTATGAGCCACATACCAGTCATGAAC**ACCATGGGACTATTTCAATCTG**CAGATATCTCTAGTCATAAATAAACATACAAGAT
GAAGATATCTGATTACTCACCTTGCACATGTTAGGTGTATTGTATAGGTCTTCTAAATTCAAACAGGTG**TTATCGTCACCCAGACT**
TGGACAATCAACAACACACGTGAGGCAGTGTTGGAGAGCGACTCTCCTTCTCTTTGGTTTTCCTCCAAGACGTGGATGAAATCCTG
TCTGTTTTCTTGTATACTGCCCATTTTCTATCTTAAAGGTAACAAGCTACCAGGTTGTTAGCAGAGTGTGAGTAGAGTCAAGAACA
TAGAATGTATGGGAGGATACATTCTCTCTGGCGACTGAAGTTGGAAATGACCTGGAGGGTAGTTTGGGGCGCTTATATGTTACAGA
CGGGTCTAC

Appendix A 2.32 Nucleotide sequence of pCUHO2. The locations and sequences of forward primers and those complementary to reverse primers are labeled in boldface and underlined.

GTAGACCCGTCATT**GGGTATCTTCCCACAACAGC**TGTCTCGTGCTGTAGGTATCCTACAAATGAGAAAAACATCTGCTAGTTGTTA
TAATTATTTGGGTGATAAAGTTATGTATTACCCCCCCTGATGTTAGTTATTGGAATGTATACACACACATATACCATATACAAAGA
AAATACATACACACCAAACTGAACATGCTTAATCATAAAGTTTGTGGTGTCATATACTGAATGGCAAACAAGGATAAGTTGGACTG
GCACAGAGAACATGTTCATGGGAAGTAGGCAGGGCACAGCAGTTAGTAAAACCATTTGTTTTG**TGAAAGGATGTAGGCAAGTGC**AG
ATGCAGTGCATGGCGATGAAATTGATTCTGTGCATAGCTGCACATGCCTTACTCATAAGATATTCACCTGGTTGGTAACTACACTC
TATCATGGATGACTTGGGACCTCTTCACTCACAATAGCAAGAACCAGTTCATATGTGCTAATACGGGTCTAC

Appendix A 2.33 Nucleotide sequence of pCUHO3. The locations and sequences of forward primers and those complementary to reverse primers are labeled in boldface and underlined.

GATCTCAGCGGCACAAGTAGAAACAAACTTGAGATTGACTAATGGTTTCCTGGGCAAAGAGCAAAATAACGTATCCTTCTGAAATG
TGGCTGTGGCTATGAAACAGC**AGAACACGGAACACACGCACGCAC**GCACGCACCCACACACACTCATATGCCACGTGCAAGCACGT
GTGCTTGGCATCTATCAAGGATGCCTAAATGTTGGCCCCCTGCTAGCCAAGAGTTAAACACAGATCCGCAGACTAGCTAGTTATAG
TGTGTTAGCCCCATCGAACCACCACGTTTCCGCTCTCTAGGCAGATGCTCAATCCACTCCACAGCGAAGGCGGTATTTACATTTTG
AACTAATACTAATAGGTAATACCATTCATGCCAGTTGGGTGTAAATATCGTATGAATAATACCAGAAGTGTAACCTTGTACCACTT
GCATCTAAATGGCAGGAATGAGCTCCGCAAGGATATTAAAGACGCTAAGTCTTAAAGGTCATAGTTTGTAAGATTTTATACGAGGT
TGATAATGAATTCCATGACGTGCTGTATTTGCACAAATCGACATATCCAAGTCTCTGGATGAAATGTCACTAACTCTCACATATAC
GAAGTTATTATAGATTGAGAATTCAACATACAGGCTAAGTCAGCAGAAAGGTGAACACTGCAATCAGTTAGAAAGATGTAGTACAG
CAGCGTCTGTTAC**GTGAGTTGCGGAGGAGTGGTTATT**GCTCTCTCTGTGAACAAGAGGTAAAACCAACACGCGTATAGAGAACCGT
CACGACGGGACACGCACACAGGGGGACAACAAGACAAATTAATATTGTAATCGATTTT

Appendix A 2.34 Nucleotide sequence of pCUHO4. The locations and sequences of forward primers and those complementary to reverse primers are labeled in boldface and underlined.

GATCTCAGCGTATCGGCAAATGAAGCAGCCAATCGGAAGGCGTCGGTTACACCTGAGTGCACACTCACCCGCTTCGCTGGGCTTCG
GTCTCCCGGGTGC**TTCGTTTGGAGGGAAGTGAGCC**CAAATACAAAATATTGCACTTTTGATTCGCGATTGTACGCTTATAACTTTG
TTTATTTGTTATTTCACAAGCAGCAATGTATATTATATGTCATGAATAAGTGAAAATTGCGTTTTAATTATGTTTGATTTTTTGGT
TGCATGTGACCTTTAACATACTTACAGTGAAATAGCATCACAGTA**CTCAACACTATCACCATCCCCTGT**GCCACTCAAACTGTACA
AAAGAACTTTGTCAAACACCCCATGGCATACCCCCTTGTTAAATGAGCGCCGGACCCGTCAAAATCATATCTGGGACATGGTATAC
AATAATACAATACATAATTATTTCATACACAAGAGCCCTAGAAACACCAGTCTGTGTCACGTTCCCGTTTGCCTCTATAGTGGTAT
TATGTATGACAAAAATTCGCCTTCAGGGCCTCCTTTCACCACTAAGGTGATCGTAAGTCACTAAGGTAACCTTAGTGAGATGCCAA
TACATGCCATGAATGGGAAGTAAGGTTAACATTAGTAAAGGTTGCATCGTGAAAGGACCTCCACCTCATTTGGTTCAGTGCTGAAT
GTCTCACGCATCTGTTGACAATGGCTGAAGTTCACAAGACTGAAGTCCTTAAGTCATCACCACACATTTGGCTTTTGTCACCGGGA
AAATTCTGTGTCTTGTATCGGTCTGTTTCCAGGTCTGATGACATCACCATAGTAGCATTCGTTTCCATACTGCAATAAATAAAACG
TGCACTCTCAGTGTCTGATTCTTTGTGAAAATTAAATTAGTAATTCAATTTTATTTTGTTGCTGTACTCTCAAAGGGGAATAGAGT
ATCGTCAAAGTATTGTCCTACTGAGAGTGTACGTAAGTCCCAAAACTATTGATGTAAATTTAAGTCTACCAGGTTTATAATCGTAA
TATCTTTGTCAATTTAATGTTTGACTAGTATTTTTGCTGTCCTTCGTTGACAAAGTTTTACATATTATATGTATGTATTCTGAATT
TAGAACTAGATTTAGTCACGATGTGGCTGTGAATTTAAATCTTAATTCACCACTACCTTGTGAAAGTAACACTAAGACTGTGATAA
CCCGATTTTATATCTGAGTCTGTAAATCTGTTTTAGTGACTGAAGTGAGTGAGTTAAAATGTAACGTCACCTCAGCAATATTTCAG
CCATATAGTGACCTGAAAGGTTCTCATACATAATTAGTACAAGTTAAAGTCCAGTTATCAAAAGATCGTAGAAAAACAGAGAACCA
CGGAGGATGTTAAAAACTAACGTGTAAACTTAAAACTGTTTTACCATTGAAGTTAACATAATACCAAAACAGGCTATTCGTCGACA
ACAGTTGAAGTTACATCAGCCATCGGGCAGAATATGAGTGAGTGAGTTAAACTGAATGGTAACTTTAAGAATTTCTCGGCCATATT
ATGACTATAACAATACATACTTTATGATTAGTGCTACGAAGTGTTATACATTAACATTATCTTACCACTATCTTAGCGCTGAGATC

Appendix A 2.35 Nucleotide sequence of pCUHO5. The locations and sequences of forward primers and those complementary to reverse primers are labeled in boldface and underlined.

GATCTCAGCGCATACATAATCATATTCATTACAAAGGATATAACACCAGAAATGTCGCGCCATAATTCTTATGACATACAGAAACC
GGGGTAATTTTATCGGCATCGCTGCTATATAATTTCCCCTCCTGAAACTAATAGAATTTCCTAATTGGGGTTGTCATGGTAGGTTC
TTTGCCTAA**CCCCTTGTTTCTCCTTCTTG**TCGACATGAAAAGACCACTTTATCAAATATTTCATCCTCACTTCATTCACAATAATC
TCTTGATGAGCCCAACATTCGTCTGAAACATCCGCACGAGGCAGCCGGAACCACCAGGCTGTTGGCAACTTATTCAAAGTCTCTTC
TCTTCACCCACAGAGATTTTTTCACAATACCAACAAACAATATCCA**TCAAACCGCCTGCGTCATCG**TTTGTACTCAAGCCTCAATC
GAACTTCGCCAGTGCACTATGGACGTGCTTCGACGGTCGGCTGCAAACCTTTGCAACAAAACCACCGATTCTAGAAGTTTCGTTCA
TTGCATTCCCCCGACAACTGCTATTATAAAACATTTAGTTAGTGTAATTGTAACTGGTCTATCTCGAGTCTTCGTCTTCATTACGG
ATAATTAGCCCGTCGCTGAGGGGGGCACTTTCTACGCTCATCCATCCCGCTGAGATC

Appendix A 2.36 Nucleotide sequence of pCUHV1. The locations and sequences of forward primers and those complementary to reverse primers are labeled in boldface and underlined.

GATCTCAGCGGCGATGTAGGTATTCCATACCGTAAGATGTAGGCTGTGTTCCGGCACTCACTCACTCACTCACTCACTCAAACAAT
AGAAGACAACACAATATTGGCTATAGATCACCAACAACGGAAGGTAGATCACCATACTAGAGATCATGTTCACTTACAGTACCTAC
TTTCATAGACCCTAACTGGATTTACACCATCTCTTCAGGCACTGATAAGAAGTAGTTGATGCTCTTGCATGAAATCCAAACAGAGA
GCTACACATAGGTCTTCCAGTCTTTTTGACAAACTCTGGTGACATAAAACGCTTTTTTGTGACTGTGCCTACAGTGATTATCCCGT
TCGACAGTAACTCCGGTGACAGAGCCATATCGGTGATAACGTTGTCCGAAAGAAAGGTCTACACAGAGTTATGTTTTCCCAAATTC
ATCTGCCTTTCCTTGCCCTCCTCTCCAAGATAGGGAACAGATCTGAGCTGGTAAAATGATTCAGAGTCAGAGTTAAAACCGTATTT
AGCTGGCTTCATTGGCATATACTGGATAAAAGAGCATCTGTATCTATATGGCAGTAGGTGTTCGTCTACCATTAGATTGGCTCCAG
GAGAATAATAACGCATCAGTGATAGATGAAATACCTCCCACAGATCACGGATTGCTGCCAGTTTATGTTTTTCCTTGCGCTGAGAT
C

Appendix A 2.37 Nucleotide sequence of pCUHV2

GATCTCAGCGGGAAGGTAACATGGATGTTGTCCTGGAGAAACTGTGGGCATAGCTTAACACTATGGTAAGTATTAATAAATGTAAT
TTTACTGTCTCTCTGCATAATAAACTTGATTTTAATCCATTTGAAATAGATCTTATTTTTTCTTACGACACCTTTGCAATCCAAAT
CAGTCATCAAAGAAGTCAGTTTTGGGGGTTGCGTGCTGTAAATCCTCGTGTTTTGGTAACGTCAGAAGGAGGAGACTCATGTAATG
ATTTTTGGTTCACACATTGAGGAATAAACATCTACGTACTCTATCAGTATTGACAAAGGGACACCTGTAGGGCCATAAGAAGCCGA
TACCTTAAATAACACGTGTTGTAATCTTTGGAGAGAGAAATCGATAGTATGATATGGATTATTGCAAATCCCTTCAGCATTAGCAA
TGTTAGTTCAGGGATATCACATCATTCAAGGTTAAATGAATGTATCACAACACCGGATATTGTTATATACCACAAAACACCCTCGT
TCCAATGCGGAATTTCAAATATTTTGAATTTAATGCGCTGAGATC

Appendix A 2.38 Nucleotide sequence of pCUHV3

GTAGACCCGTCTCAGGAGTAACATTGAATCTTCCAACACCTCCGCCTAAACCCATCAATACGGAACTGACGTCACTCAAACTGATG
GCTCAAACGGCGAGCAGTCTCAGCGACGAAATGACATATACTATTGTAAGTATTACAGTTATAATATATTAACTTCAACTATTCGA
TTATATTCCTAATGATTGTTCGATTTACTTATATCGTACTTTTAAATTCTTAATTTTAGGTAACATTTTGTGCTCAATCAAATTTG
AAAAGTCAGCATAACGAATTACAAGTTAATATGTTTGAAATCTGATTGGCTTCGGAATCCTCTCAATCAAACAGGCAATCTGTGCT
GTAACGGGAATATGTTGGTTGCGGCGTTAAATTCCTTACTCTTTAGGTTATAATTTCACAATATTTAAAACTTGAATTATATTGTT
CAGCTACAAATTCAGGCATAGAATTCTTCCGGTGAACAGTTGATATTTCATTTCTGTTTTATTTCTGTCGTCGTTAGTTATTCATG
AAAGACTATACTAAGGGTATTTTTACGGGTCTAC

Appendix A 2.39 Nucleotide sequence of pCUHV4

GTAGACCCGTTAAGATACATACCTTTAATCCGGTAGTGAAGGAAGCGCCGGTATCGGGGAATCCTTCCAGTGGGACATGCGAATCA
AGCTCTGCGGCCACAGACGAGCCAATCCGGGGCATCTACAAAAATACAAAAATACCCTTAGTATAGTCTTTCATGAATAACTAACG
ACGACAGAAATAAAACAGAAATGAAATATCAACTGTTCACCGGAAGAATTCTATGCCTGAATGTGTAGCTGAACAATATATTTCAA
GTTTTAAATATTGTGAAATTATAACCTAAAGAGTAAGGAATTTAACGCCGCAACCAACATATTCCCGTTACAGCACAGATTGCCTG
TTTGATTGAGAGGTTTCCGAAGCCAATCAGATTTCAAACATATTAACTTGTAATTCGTTATGCTGACTTTTCAAATTTGATTGAGC
ACAAAATGTTACCTAAAATTAAGAATTTAAAAGTACGATATAAGTAAATCGAACAATCATTAGGGATATAATCGCAATAGTTGAAG
TTAATATATTATAACTGTAATACTTACAATAGTATATGTCATTTCGTCGCTGAGACTGCTCGCCGTTTGAGCCATCAGTTTGAGTG
ACGTCAGTTCCGTATTGATGGGTTTAGGCGGAGGTGTTGGAAGATTCAATGTTACTCCTGAGACGGGTCTAC

Appendix 2.40 Nucleotide sequence of pCUHV5

GTAGACCGTCCCATGTTTTATACTCTTGTTATAAACGTTTACCATCTGGTCGCAGGCACCTTTTACCAAGTGTAAAGTCCAATAGA
AATAGACACTCTTTTATACCGGTTAGCATTTCCTTTTTAAATAAGTCTTAATCATATTTTACTTATATACTCACTGCTTTGTATCC
TGTTTTAAATTCATCAAACTTTAGTCATCGTCATCTACTCATATCTTATTATTATTGTTAAACATGTAGACTATAAAACCATGTTG
CTACCATGTAAAACATGTGTACTTCCTATCTTAATACATGTACTTATGTCTTACTGTTGTCTTAACTCTGATGTACACAAAATTCC
CTCTTTGGGCAATAAAGCATTATCTTATCTATCCTATCTTATCTTATCTCTTCCTCTACCCACACAGGGCCGCAACCTACGGCAAA
TGGGTTGATAGATCAAATATTTCCCCGGGTCCACTTGAGGGGGAGGACGGACTCCTAGCATTGCCAACACCCACAATGAGGATGTG
GCACGCCACGGGTGCCACCGGGGCATGGAAGGAAGTTAAAATTTACATTGCACTTCTGAACCACCCTGGCTACACGAGGTCTGGTG
TTGTCCTGTTTGATGATCATTCTTGGACCATTCTGCTGGGAACAATTTCTGGCGAGATGTACGCGCTTGAACTCGGTACTCACACG
GGTCTAC

Appendix A 2.41 Nucleotide sequence of pCUHV6

GTAGACCCGTCCCCAGGTCCATCATGGTTAAGAAAGTAGGACTGTTGGCGAAAAACATTAATCACTCTGGATCAGTGCCCATGTAA
TGGAACATGTTTTTCAGAGATGCTCTTATCAGTGTCAATGAACCTAAGAGAGCATCCCCACATGGCAGCAGGTAGATTGCTTGGCA
GTTATTTTCCTCGAATTAGCAGTCTTCCACACATATACAATCAAAAGTATTCAGCTCACCATTAAATCGTGATGGTAAAAATACGC
TGAATGTGCTGAATTTGCTCATTAATGTCTATGTGTTGAAGAATGTATAAGCATTGTTAAAATCATGTGAGGACGTTGGTAACTCA
GCAGTGTATTTTCTGCGAAAGACTATTCTGGATAATAACCGAGAATAAATTTAAAATTATCTTCAAGATGAGGACCCACCTGCTTG
AGCTCTTGCTCCAAGGCAGCTCTCTCATCCACTTCCTTCTCCAGCTGCGTCACCCACTCCTTCAGTTCAGAGGCTGTCAGCTCCTT
GTCCTGAAGTACCAGACGTAATCTACGGGTCTAC

Appendix A 2.42 Nucleotide sequence of pCUHV7

APPENDIX B

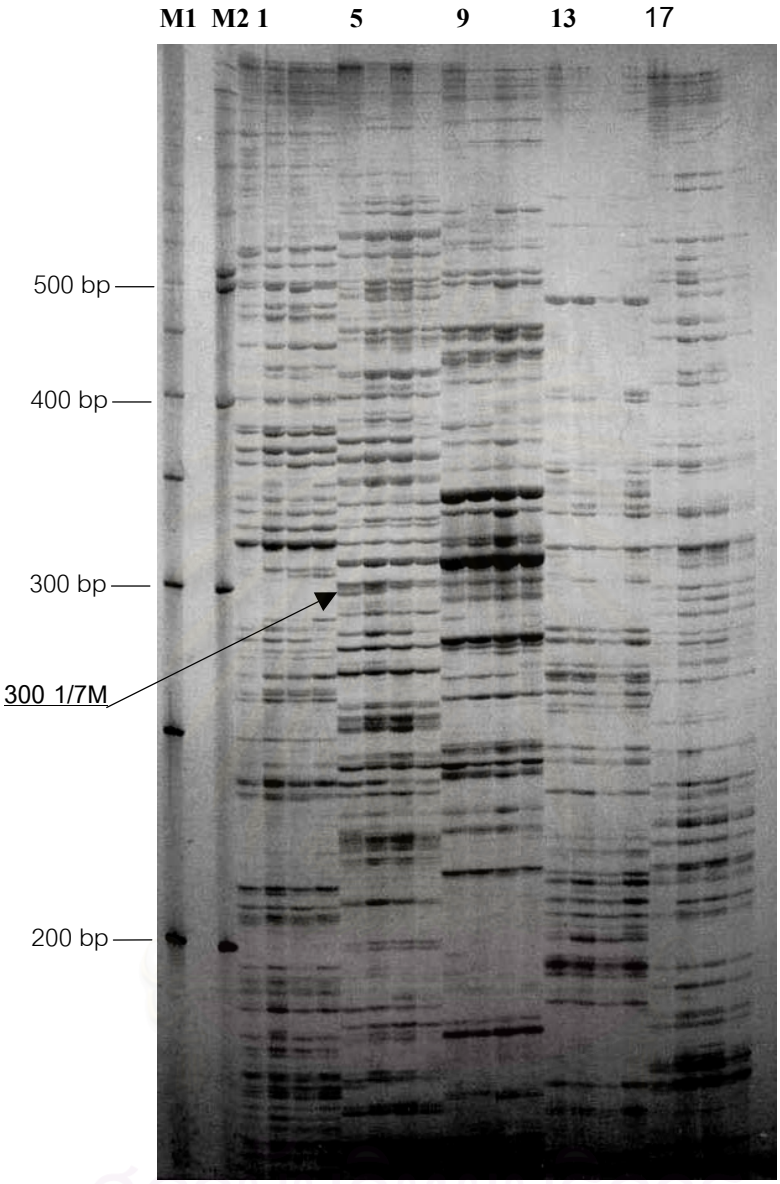


Figure B1.1 A 6% denaturing polyacrylamide gel electrophoresis showing AFLP products of 4 bulked ABM1 (lanes 1, 5, 9, 13 and 17), ABM2 (lanes 2, 6, 10, 14 and 18), ABF1 (lanes 3, 7, 11, 15 and 19) and ABF2 (lanes 4, 8, 12, 16 and 20) using primers E_{+3}-1/M_{+3}-1 (lanes 1 – 4), E_{+3}-1/M_{+3}-7 (lanes 5 – 8): E_{+3}-2/M_{+3}-5 (lanes 9 – 12) E_{+3}-2/M_{+3}-6 (lanes 13 – 16), E_{+3}-2/M_{+3}-8 (lanes 17 – 20). An arrowhead indicates a candidate sex-specific AFLP marker from E_{+3}-1/M_{+3}-7.Lanes M1 and M2 were 100 bp and 50 bp DNA markers, respectively.

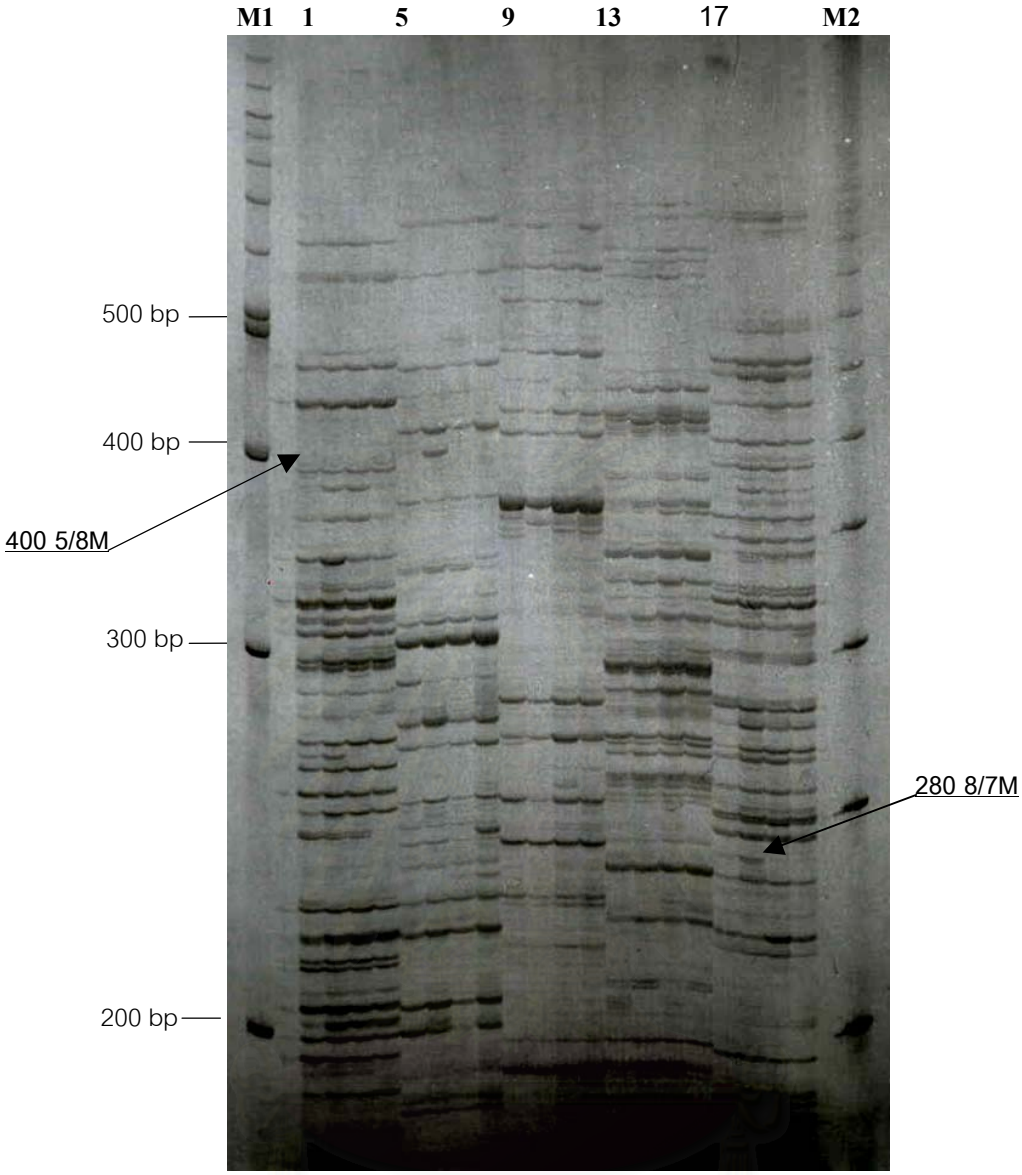


Figure B1.2 A 6% denaturing polyacrylamide gel electrophoresis showing AFLP products of 4 bulked ABM1 (lanes 1, 5, 9, 13 and 17), ABM2 (lanes 2, 6, 10, 14 and 18), ABF1 (lanes 3, 7, 11, 15 and 19) and ABF2 (lanes 4, 8, 12, 16 and 20) using primers E_{+3}-5/M_{+3}-8 (lanes 1 – 4), E_{+3}-6/M_{+3}-5 (lanes 5 – 8): E_{+3}-6/M_{+3}-8 (lanes 9 – 12) E_{+3}-7/M_{+3}-7 (lanes 13 – 16), E_{+3}-8/M_{+3}-7 (lanes 17 – 20). An arrowhead indicates a candidate sex-specific AFLP marker from E_{+3}-5/M_{+3}-8 and E_{+3}-8/M_{+3}-7.Lanes M1 and M2 were 100 bp and 50 bp DNA markers, respectively.

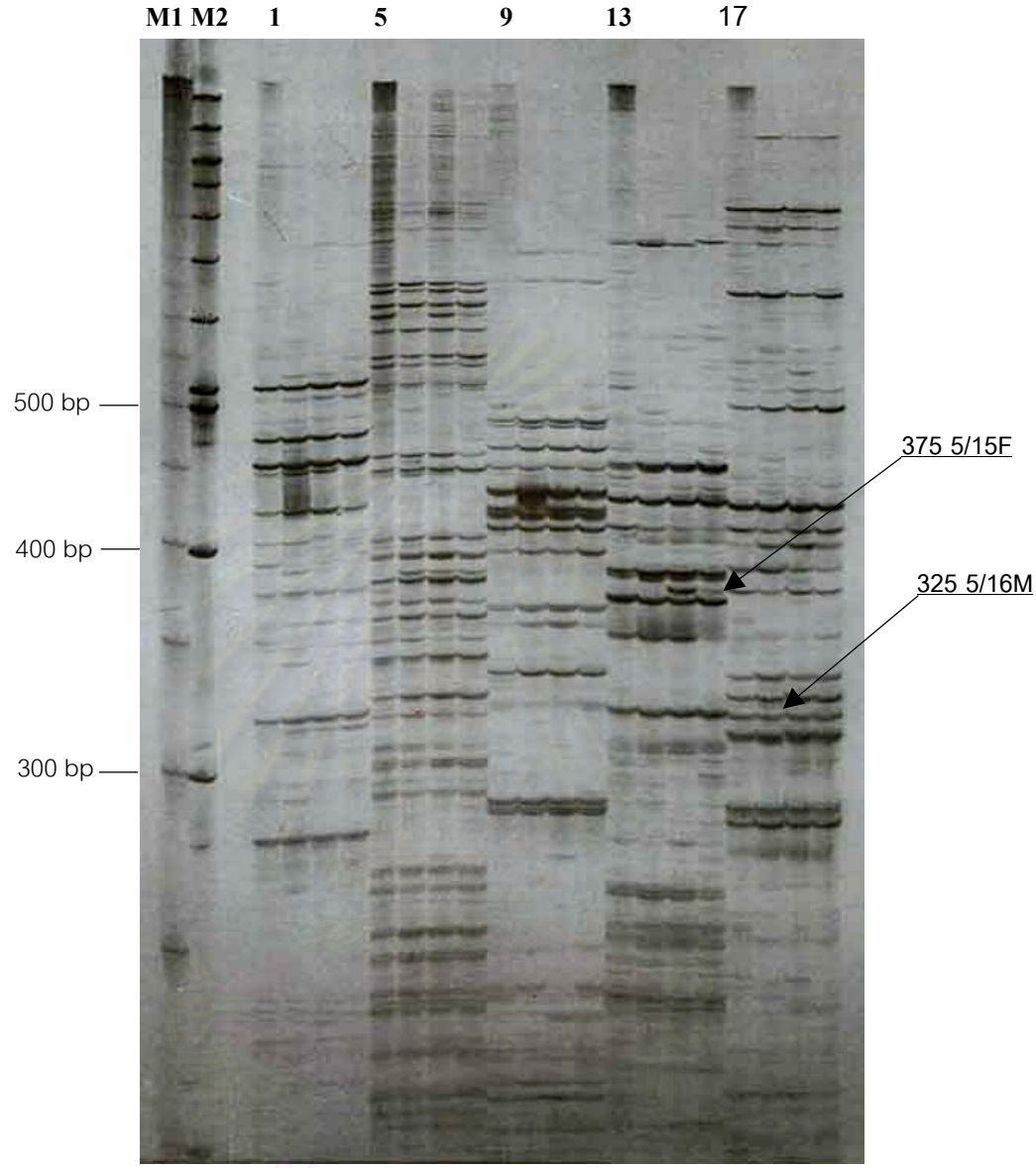


Figure B1.3 A 6% denaturing polyacrylamide gel electrophoresis showing AFLP products of 4 bulked ABM1 (lanes 1, 5, 9, 13 and 17), ABM2 (lanes 2, 6, 10, 14 and 18), ABF1 (lanes 3, 7, 11, 15 and 19) and ABF2 (lanes 4, 8, 12, 16 and 20) using primers E_{+3}-5/M_{+3}-12 (lanes 1 – 4), E_{+3}-5/M_{+3}-13 (lanes 5 – 8): E_{+3}-5/M_{+3}-14 (lanes 9 – 12) E_{+3}-5/M_{+3}-15 (lanes 13 – 16), E_{+3}-5/M_{+3}-16 (lanes 17 – 20). An arrowhead indicates a candidate sex-specific AFLP marker from E_{+3}-5/M_{+3}-15 and E_{+3}-5/M_{+3}-16. Lanes M1 and M2 were 100 bp and 50 bp DNA markers, respectively.

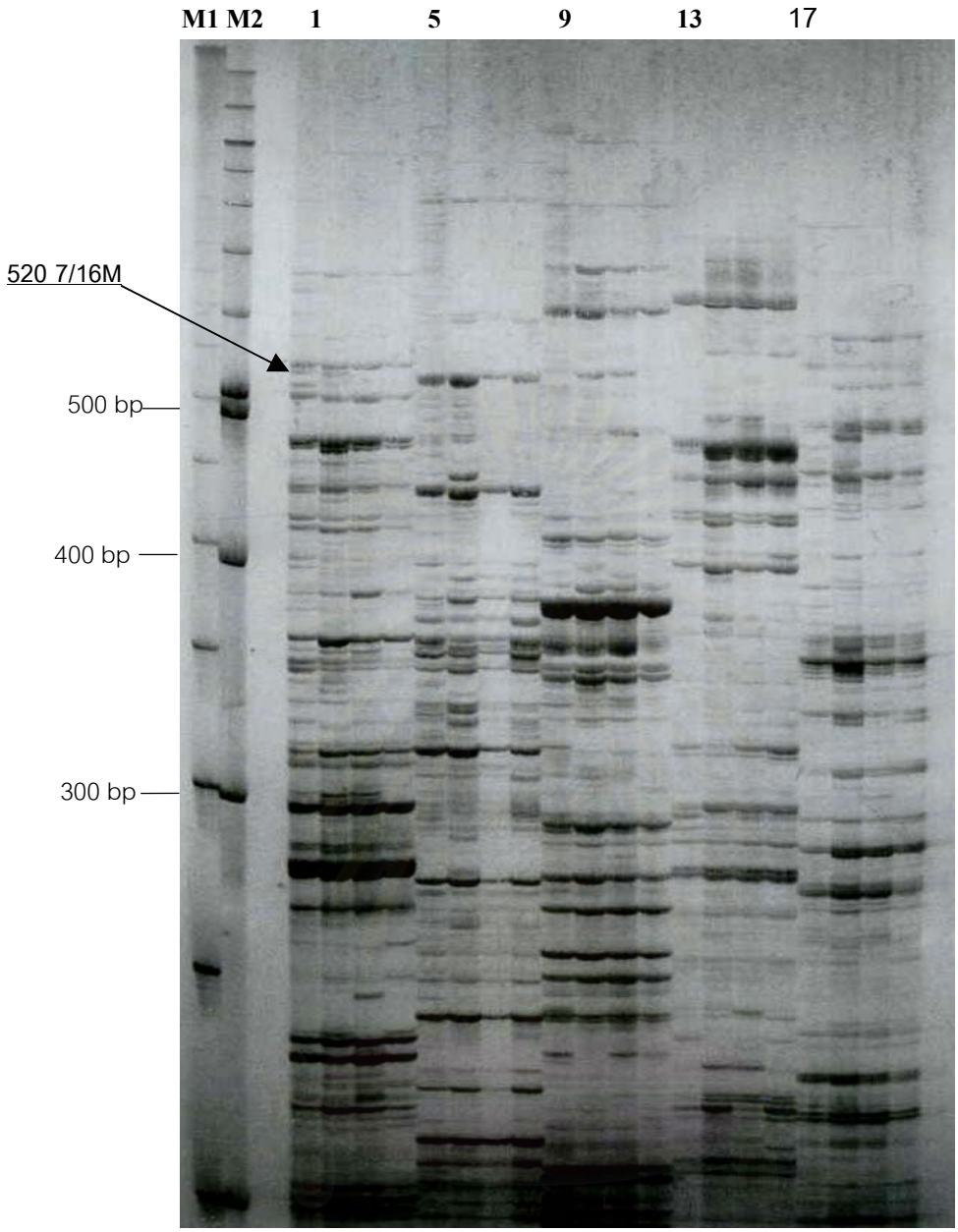


Figure B1.4 A 6% denaturing polyacrylamide gel electrophoresis showing AFLP products of 4 bulked ABM1 (lanes 1, 5, 9, 13 and 17), ABM2 (lanes 2, 6, 10, 14 and 18), ABF1 (lanes 3, 7, 11, 15 and 19) and ABF2 (lanes 4, 8, 12, 16 and 20) using primers E_{+3}-7/M_{+3}-16 (lanes 1 – 4), E_{+3}-8/M_{+3}-9 (lanes 5 – 8): E_{+3}-8/M_{+3}-10 (lanes 9 – 12) E_{+3}-8/M_{+3}-11 (lanes 13 – 16), E_{+3}-8/M_{+3}-12 (lanes 17 – 20). An arrowhead indicates a candidate sex-specific AFLP marker from E_{+3}-7/M_{+3}-16. Lanes M1 and M2 were 100 bp and 50 bp DNA markers, respectively.

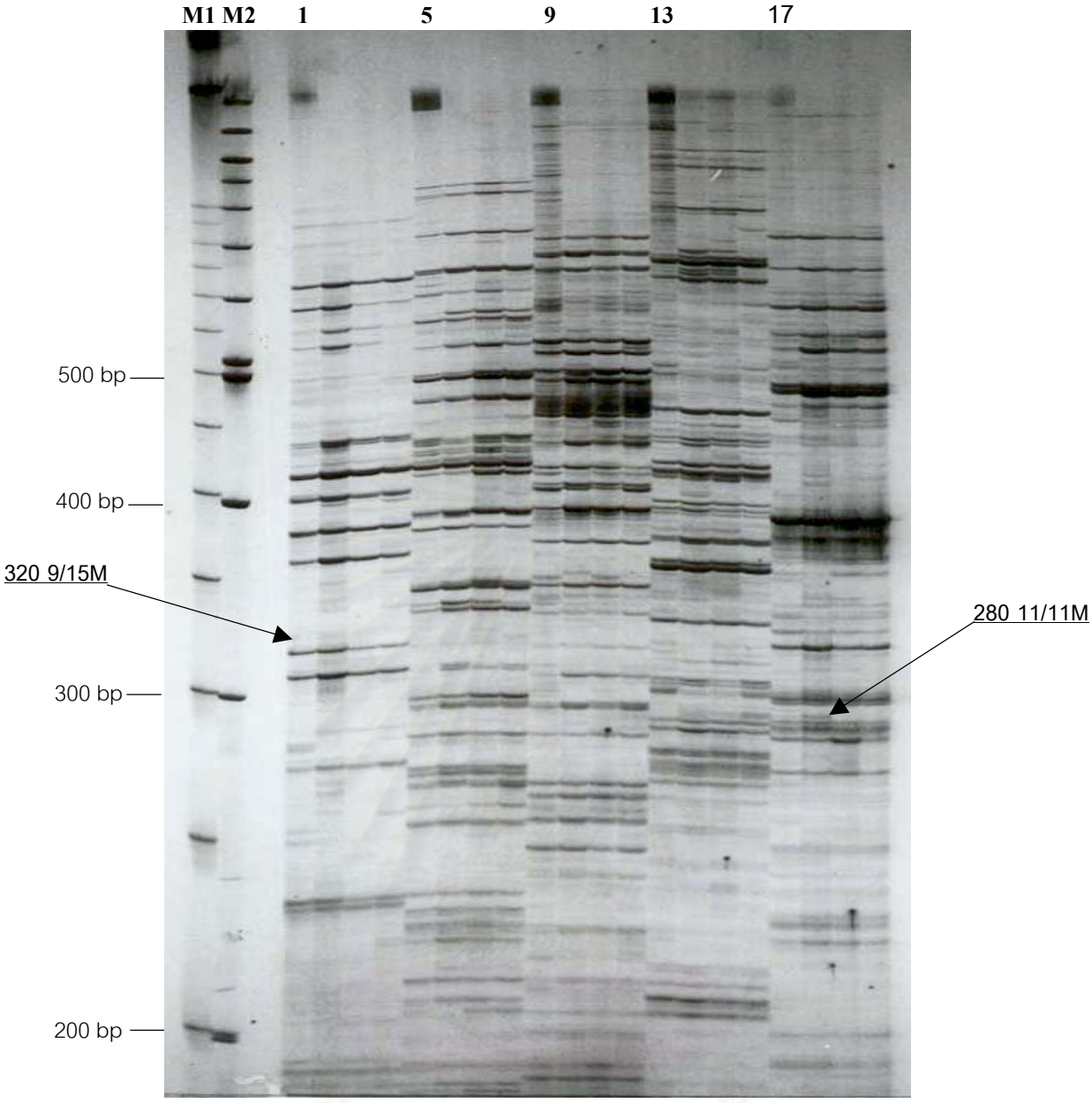


Figure B1.5 A 6% denaturing polyacrylamide gel electrophoresis showing AFLP products of 4 bulked ABM1 (lanes 1, 5, 9, 13 and 17), ABM2 (lanes 2, 6, 10, 14 and 18), ABF1 (lanes 3, 7, 11, 15 and 19) and ABF2 (lanes 4, 8, 12, 16 and 20) using primers E_{+3}-9/M_{+3}-15 (lanes 1 – 4), E_{+3}-9/M_{+3}16 (lanes 5 – 8): E_{+3}-11/M_{+3}-9 (lanes 9 – 12) E_{+3}-11/M_{+3}-10 (lanes 13 – 16), E_{+3}-11/M_{+3}-11 (lanes 17 – 20). An arrowhead indicates a candidate sex-specific AFLP marker from E_{+3}-9/M_{+3}-15 and E_{+3}-11/M_{+3}-11. Lanes M1 and M2 were 100 bp and 50 bp DNA markers, respectively.

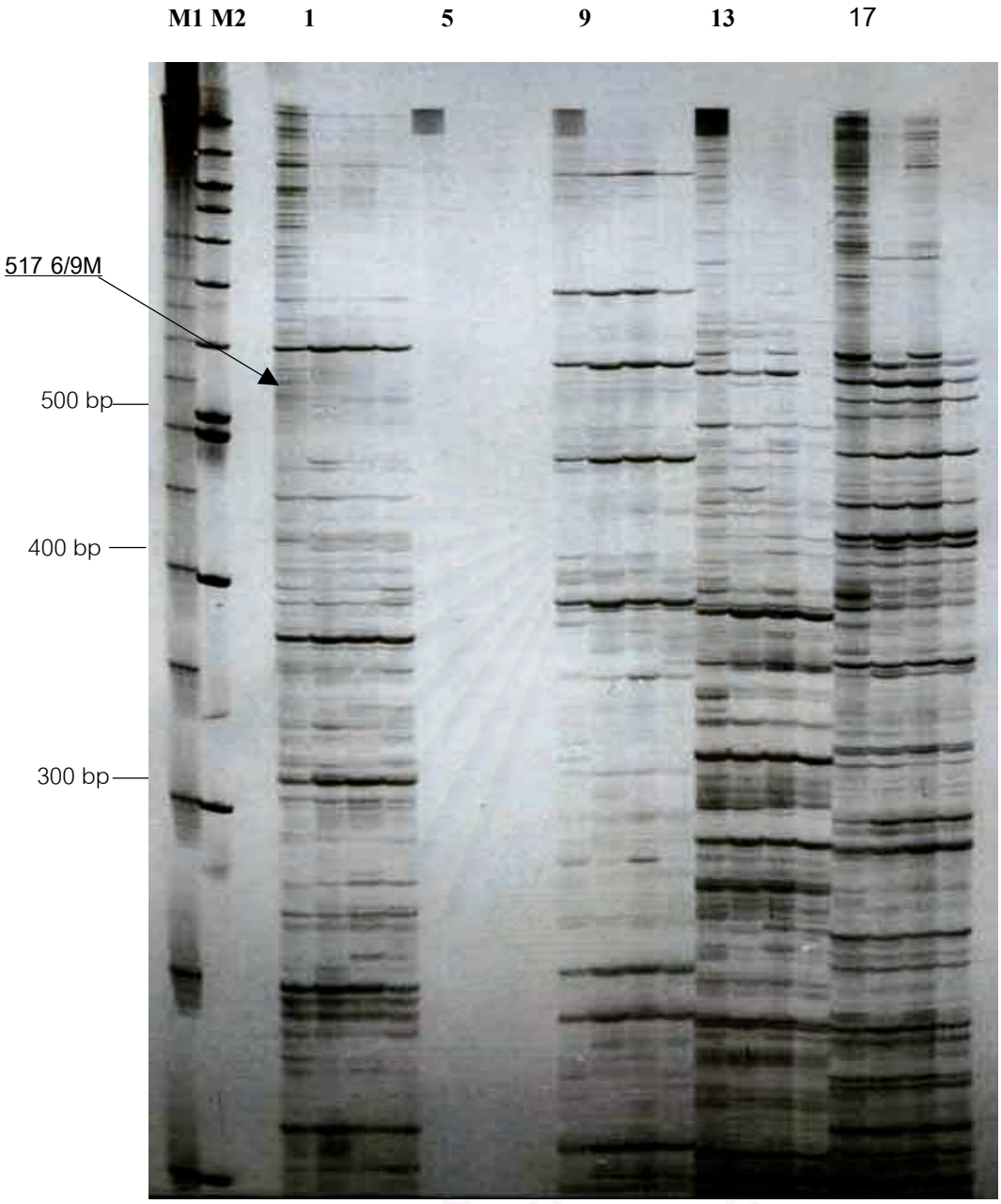


Figure B1.6 A 6% denaturing polyacrylamide gel electrophoresis showing AFLP products of 4 bulked ABM1 (lanes 1, 5, 9, 13 and 17), ABM2 (lanes 2, 6, 10, 14 and 18), ABF1 (lanes 3, 7, 11, 15 and 19) and ABF2 (lanes 4, 8, 12, 16 and 20) using primers E_{+3}-6/M_{+3}-9 (lanes 1 – 4), E_{+3}-6/M_{+3}10 (lanes 5 – 8): E_{+3}-6/M_{+3}-11 (lanes 9 – 12) E_{+3}-6/M_{+3}-12 (lanes 13 – 16), E_{+3}-6/M_{+3}-13 (lanes 17 – 20). An arrowhead indicates a candidate sex-specific AFLP marker from E_{+3}-6/M_{+3}-9. Lanes M1 and M2 were 100 bp and 50 bp DNA markers, respectively.

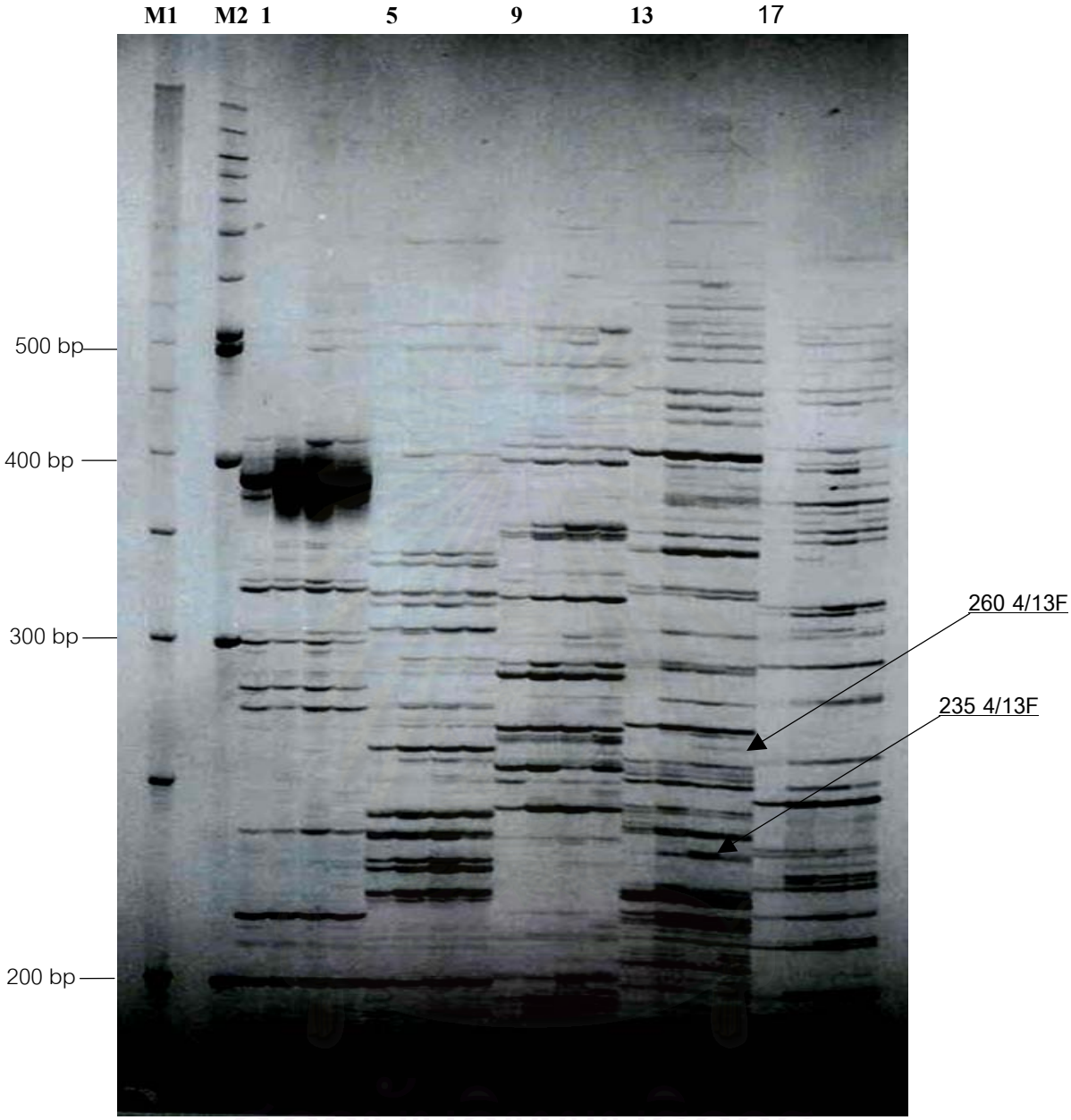


Figure B1.7 A 6% denaturing polyacrylamide gel electrophoresis showing AFLP products of 4 bulked ABM1 (lanes 1, 5, 9, 13 and 17), ABM2 (lanes 2, 6, 10, 14 and 18), ABF1 (lanes 3, 7, 11, 15 and 19) and ABF2 (lanes 4, 8, 12, 16 and 20) using primers E_{+3}-4/M_{+3}-10 (lanes 1 – 4), E_{+3}-4/M_{+3}11 (lanes 5 – 8): E_{+3}-4/M_{+3}-12 (lanes 9 – 12) E_{+3}-4/M_{+3}-13 (lanes 13 – 16), E_{+3}-4/M_{+3}-14 (lanes 17 – 20). An arrowhead indicates a candidate sex-specific AFLP marker from E_{+3}-4/M_{+3}-13.Lanes M1 and M2 were 100 bp and 50 bp DNA markers, respectively.

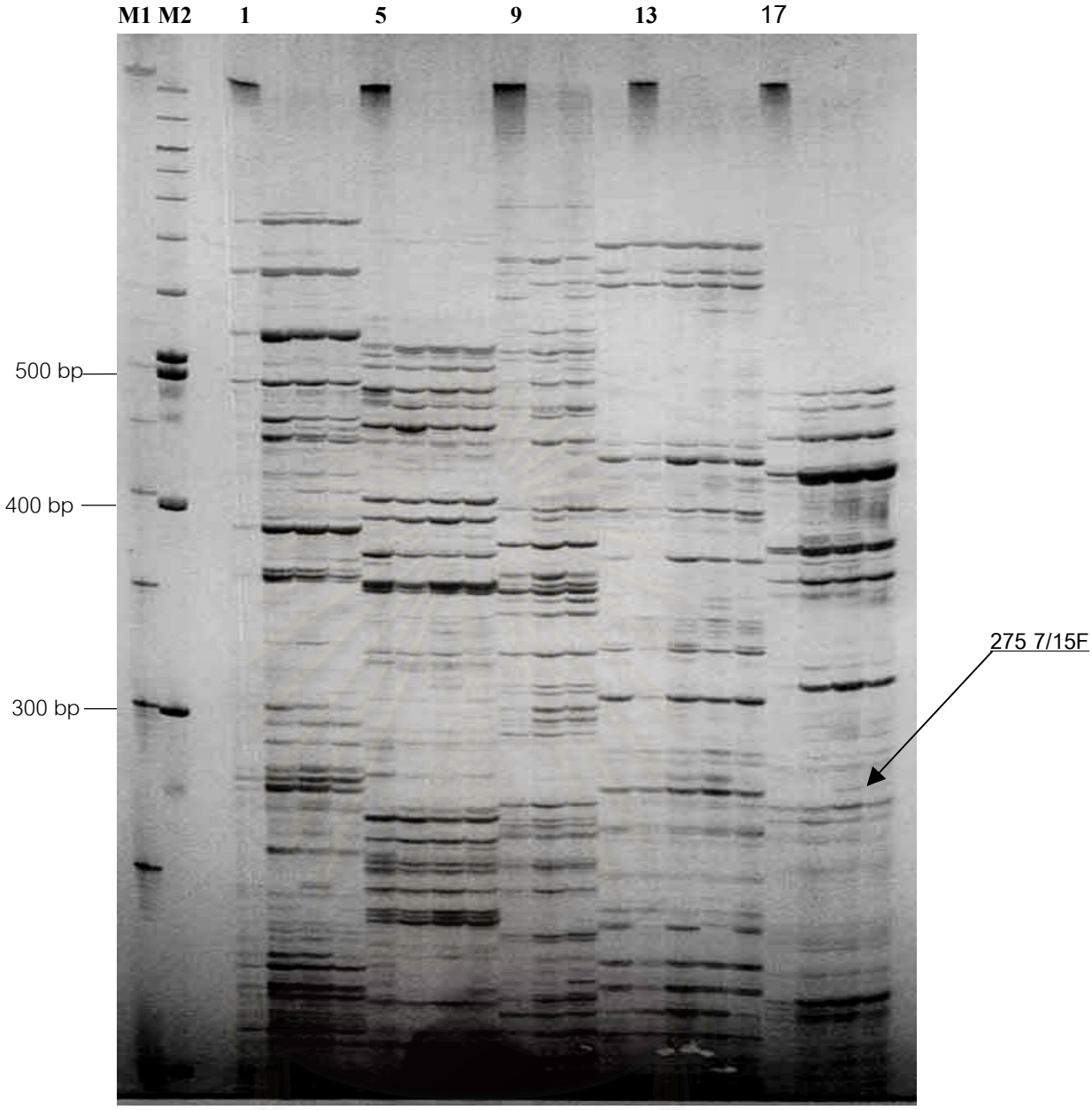


Figure B1.8 A 6% denaturing polyacrylamide gel electrophoresis showing AFLP products of 4 bulked ABM1 (lanes 1, 5, 9, 13 and 17), ABM2 (lanes 2, 6, 10, 14 and 18), ABF1 (lanes 3, 7, 11, 15 and 19) and ABF2 (lanes 4, 8, 12, 16 and 20) using primers E_{+3}-7/M_{+3}-11 (lanes 1 – 4), E_{+3}-7/M_{+3}12 (lanes 5 – 8): E_{+3}-7/M_{+3}-13 (lanes 9 – 12) E_{+3}-7/M_{+3}-14 (lanes 13 – 16), E_{+3}-7/M_{+3}-15 (lanes 17 – 20). An arrowhead indicates a candidate sex-specific AFLP marker from E_{+3}-7/M_{+3}-15. Lanes M1 and M2 were 100 bp and 50 bp DNA markers, respectively.

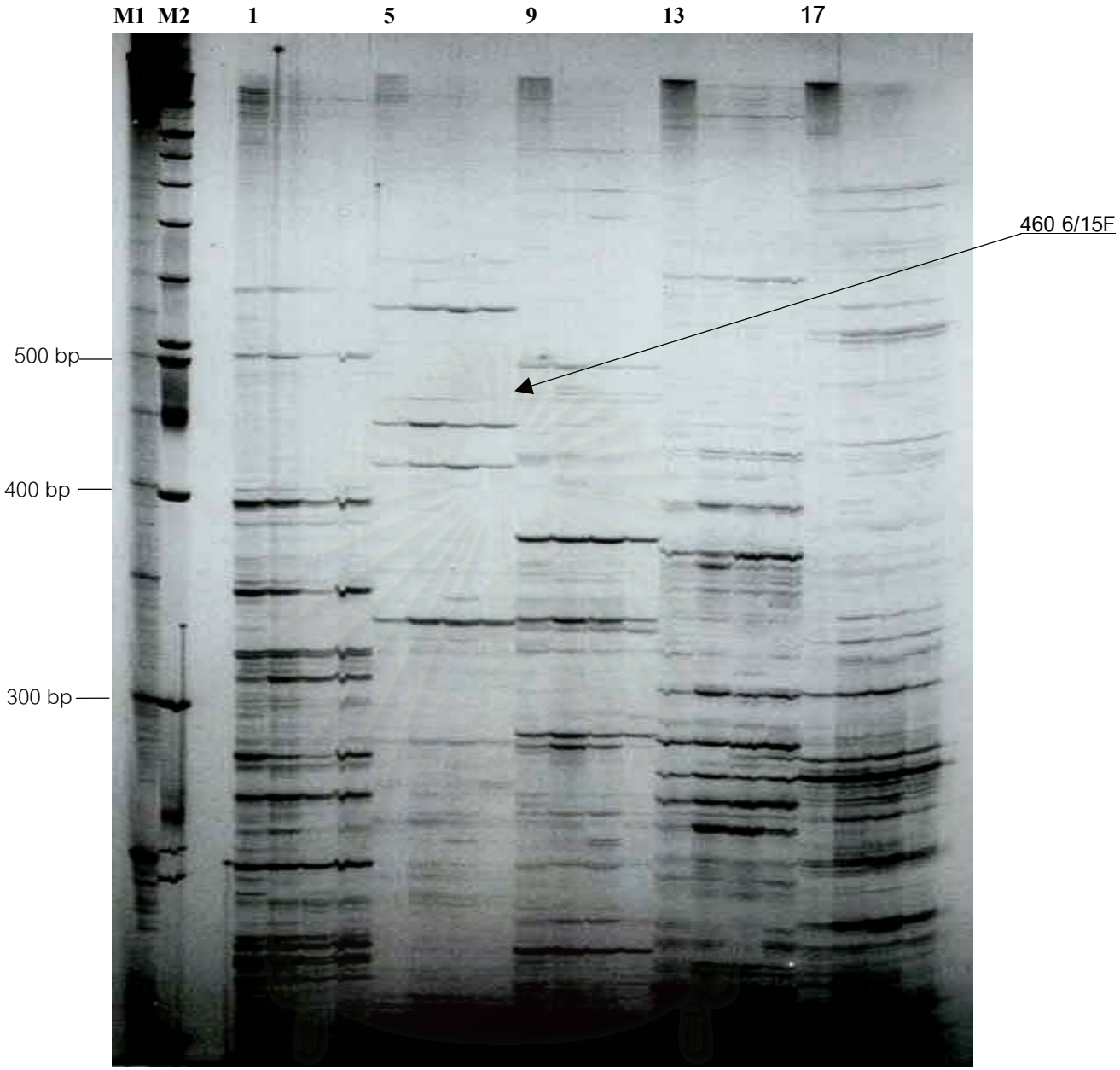


Figure B1.9 A 6% denaturing polyacrylamide gel electrophoresis showing AFLP products of 4 bulked ABM1 (lanes 1, 5, 9, 13 and 17), ABM2 (lanes 2, 6, 10, 14 and 18), ABF1 (lanes 3, 7, 11, 15 and 19) and ABF2 (lanes 4, 8, 12, 16 and 20) using primers E_{+3}-6/M_{+3}-14 (lanes 1 – 4), E_{+3}-6/M_{+3}15 (lanes 5 – 8): E_{+3}-6/M_{+3}-16 (lanes 9 – 12) E_{+3}-7/M_{+3}-9 (lanes 13 – 16), E_{+3}-7/M_{+3}-10 (lanes 17 – 20). An arrowhead indicates a candidate sex-specific AFLP marker from E_{+3}-6/M_{+3}-15.Lanes M1 and M2 were 100 bp and 50 bp DNA markers, respectively.

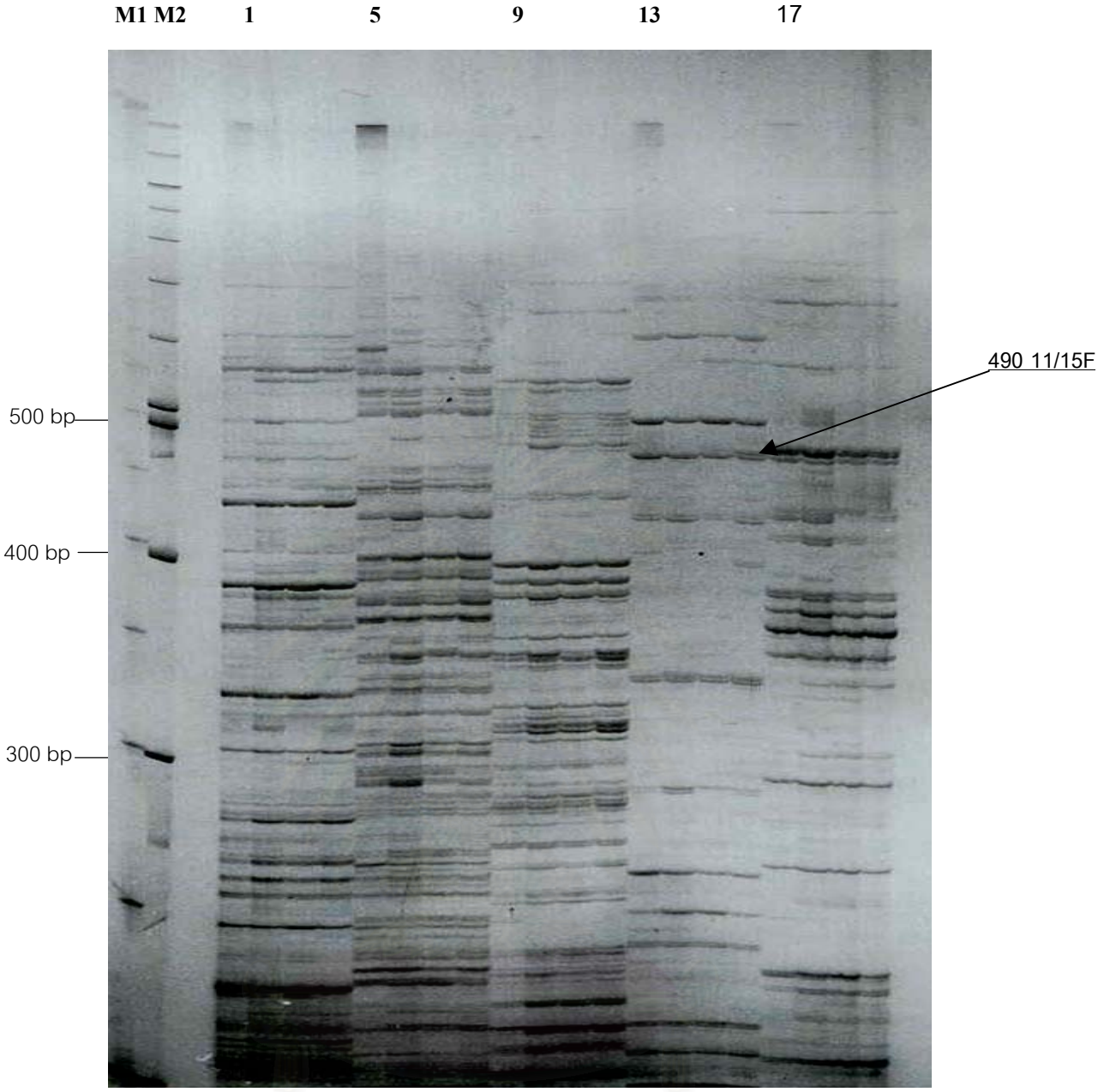


Figure B1.10 A 6% denaturing polyacrylamide gel electrophoresis showing AFLP products of 4 bulked ABM1 (lanes 1, 5, 9, 13 and 17), ABM2 (lanes 2, 6, 10, 14 and 18), ABF1 (lanes 3, 7, 11, 15 and 19) and ABF2 (lanes 4, 8, 12, 16 and 20) using primers E_{+3}-11/M_{+3}-12 (lanes 1 – 4), E_{+3}-11/M_{+3}13 (lanes 5 – 8): E_{+3}-11/M_{+3}-14 (lanes 9 – 12) E_{+3}-11/M_{+3}-15 (lanes 13 – 16), E_{+3}-11/M_{+3}-16 (lanes 17 – 20). An arrowhead indicates a candidate sex-specific AFLP marker from E_{+3}-11/M_{+3}-15.Lanes M1 and M2 were 100 bp and 50 bp DNA markers, respectively.

3001/7M1:

GACTGCTTACCAATTCAACATCACCAAGGGGTCGAGCTTCTCCAAAAGCCGCTGA**ATTGGGTTCTCCTGTTGTCA**AACTTGAAAAA
TTGGAGATGGATAGTGGAAAGAGAGAGTCTGGTCGGAGATCAGATGAGTGGAAATCCCCTAAGTGTCCTTCCTTCCCAGAAGATAA
ACCACCAACCTGCAAGAAAATATCTATAATTTTGAGGCCAGTGAGCAAGATCAAGAGGGGAAAACTGAACAATATCGAACACTCAG
AA**GCAGAGATAAAAAATCGGCAGA**TGCAGAACCAGTTACTCAGGACTCATC

3001/7M2:

GACTGCGTACCAATTCAACACTGGTGTTATCAGCTATGTGGGATTACACTGTTCTTTCAG**GAAGCAACTGGCACCATCAAT**ATTTC
ATAGAAATTCCTAAACACATGATATAGTTTGTTGTCTGGACAATGTCTCATGTGCATCATAATGTCCAATATCAGCCTAAAAAACT
GCCAAAATCACTTTGCTGGTAGTCTAGACACTG**CAGATAGTGATGTGGGGACAC**AAAACTGTTGGTGGTCTAGACACAGCAATTAG
TGCAATGTGAAACTGTTCAACAGTTACTCAGGACTCATC

3001/7M3:

GATGAGTCCTGAGTAACTGAAAAAAAGATCTCTTTGTACCTTT**CATTGCATACTTATGAAGGGAA**TTGAGTTCAAAATCATGAGAA
GAAAAGTCTATTTCCTAAATGAGTACTCAATGCGTATTCAGGGGTTGTGAAATTTTCATTACGATAAATTGATGCTAGGCAGGTGC
GCGTGTTGCTCACAATAAGATATGTAGTAAAACCGTGTAATTGTTGATATATTCAG**GACCCTATTTCAGTGATAAAACCA**TTCATT
TTTCATTTTGGCTAAAAAGTGTTGAATTGGTACGCATC

Appendix B2.1 Nucleotide sequences of 3001/7M1, 3001/7M2 and 3001/7M3. Sequences and positions of the forward primers and those complementary to the reverse primers are underlined and boldfaced

280 8/7M1:

GACTGCTTACCAATTCAGGTCAAGGGTCATCACGCCCCGGATGTTATCAAGAGTAGTGGAAAATTATCCGTTAAGAACCATTTTAT
GAAACTGCTATTCATTTCTGTAGCTCAACATTACGAGAGAATTGAAACACTGGCGAAATATTGCGGGGACACGCGCGAACAGTTTC
CAATACCGGTTTGAGCAGCACCGTAGACTCAACGGTTAGACTTTGACAGTTACTCAGGACTCATC

280 8/7M2:

GACTGCTTACCAATTCAGGATTTACAATGAGGCCACAATCATGCATATGAGCAGACTGAAATGTTTCCAAATATTCTGCACTTAGC
AGAATCTTCATGTTTTTTGCCATACTCTAGAGACTTTGCATATACTGTTTGTTTCCCCTTAGGAAATATGGTTTGCAGGAAGGCAG
TAGTTGGATTTTTGTTTTTCTGTTGATAATCTTTCCAAAATTTGAAGCAGTTACTCAGGACTCATC

280 8/7M3:

GACTGCTTACCAATTCAGGGGAAATTCAGCCTCAAACCTATAAATATCAATCCTTATGTTTCTGATGTGATTACCATCATATGTTG
GTCTACATCTAAATGATTATTTCTCATGCCAATCCAGTCTCTTCAATATCCAATAAACTTCATCTTCCAGTAAGGAGACCCCAAGA
AGCTTTGAAGTAGTTTTCAAAAATAAACATAAAATCACTCAAATGAAACTGAACAGTTATCAGGATCATC

Appendix B2.2 Nucleotide sequences of 280 8/7M1, 280 8/7M2 and 280 8/7M3.

325 5/16M1:

GACTGCGTACCAATTCACCGAGACAGGATAAAGAGGTGATAATGAAGGCCATTTTCAGTCACAGAATCACTTGGAATCTACTGGAT
CTGTCTCTACAGCTACCTTCTACGATTTTGCGAACAACTAAGTGCTCTGGCGAGCACAAAATTTACTTCGGGAAGTCTGGAACTAA
AGACAAACTTGCTGTGATATCAATGGAGTGGCGATATAAAAGTACAAGTAATCAAAGCCTTGCAATCTTCGGTACTGTATTCCAGA
GCTGAAACATGTATTCCACAATAACGACGCGACAAAAACAATTTCGGGTTACTCAGGACTCATC

325 5/16M2:

GACTGCGTACCAATTCACCGAGACAGGATAAAGAGGTGATAATGAAGGCCATTTTCAGTCACAGAATCACTTGGAATCTACTGGAA
CTGTCTCTACAGCTACCTTCTACGATTTTGCGAACAACTAAGTGCTCTGACGAGCACAAAATTTACTTCGGGAAGTCTGGAACTAA
AGACAAACTTGCTGTGATATCAATGGAGTGGCGATATAAAAGTACAAGTAATCAAAGCCTTGCAATCTTCGGTACTGTATTCCAGA
GCTGAAACATGTATTCCACAAAAACGACGCGACAAAAACAATTTCGGGTTACTCAGGACTCATC

325 5/16M3:

GACTGCGTACCAATTCACCACAGCAAACGTAAATGTTATGTTGTTAGTCTGACGCTTCTTGGAATCAAGATTTACCTGTGTGTCCA
TTAGGAATGGTTCTCTGGTTGCCATAGTTACCATCATCCTGATGGGTCTGATACCTCATGTTTATGACAGGCATGCCATCAGACGC
CATTGACAGACCCTGGAATATACAAAGGAAAGACTGTGATAGAGCATTTACGAAGATGGTGCTGTCAGAGTGAACTCTGCGTGATG
ATGTTTCAGTCATTTGGTTACCAAGTCACCTCAGAGGGTTACTCAGGACTCATC

325 5/16M4:

GACTGCGTACCAATTCACCTTGTCAAAGTCATATTGCCAGTCTCAAACCCTGCTTGTTACTTTGACGAACGACAGTCTCCAGCCAC
ACGGTTGTTCACCATGTCAGTATTTGCTGACTGGATCGAGCGCAAGTACAGCAACCATGTACTTACAAAACGTAACAAGTCTATAT
ACTTAGTGAAAGCAGGGATACTCTACAACATAAGCTCCCTGGAGAAAGTGAGTTGTGATTGGTACTCTCAAATTATGGGGCCGTAG
ATGCATCCAGGGCAGGCAGACTATATAGAGTAGATTATCCCTGGTGCAGGGTTACTCAGGACTCATC

Appendix B2.3 Nucleotide sequences of 325 5/16M1, 325 5/16M2, 325 5/16M3 and 325 5/16M4.

520 7/16M:

GATGAGTCCTGAGTAACCCTTTGCTGTCAAACCTAAACACGGGGTTCGAACCCGAAAGAAATTTATGAAATAATGCATATGGTAAT
GATTATGACGATCGGAGCTGACAATAAGATCTGAATCAGTTACTGAATATTAGCAT**GCCTCATTGAAGCGATTCCCTAT**CGCTAAG
TTTTACTCATTTGAGACAAACACACACATCATGTGAAAGTGAAGTAGTAAGAGTTGTTTACAAACCAATATACGTTGCTGCCAAAG
TGACAAGAAAGTTTCCCTCAATAGAGGTCGTTACTCTGAAAATATA**GACCAATCTAAACGAGACGAAAGGG**ACCTATAAACATTGA
CGGTGATGATTTATACGAAGGTCCCACAGACTTCCACGTCGGATAAACAATGGCAATGAGTACTTCTCAAGAGAAGCACCGAACAG
GGTTTGAAATGCTTATGAGAAATTCACGCTGAGCACGCACAAATGAAGAGCGGTGATTTTGATATCTTGCTTCGCTGAATTGGTAC
GCAGTC

Appendix B2.4 Nucleotide sequences of 520 7/16M. Sequences and positions of the forward primer and those complementary to the reverse primer are underlined and boldfaced

275 5/15 F1:

GACTGCATACCAATTC**AGCTGGGGCTTTTATGATGAGAC**TTTCATGATAAACTAGGATGTAAGAAAATCTGTATGCAATATTTGTG
ACATTAACAAGCTGATTTTCTCATAAGGATGCTAGATATGTTTTATGACGAGTAGTATACGGAATGTTAGAAAGAGACCAGGTTCC
TCTTTGTTCAGGCCCTTA**TCCCGTTCCACCGATCCCCGTACT**ACTAGTGTGTTACATAACAGTTACATGCAAGCGGACATAGGGTC
GGTTACTCAGGACTCATC

275 5/15 F2:

GATGAGTCCTGAGTAACCGACCGGAA**CACAGGCCCCCAAACGTGGAGCAT**TGTTTCTCTTTAGTCATCTGTTTGACAGTTTAGGAG
TTGATGATAAAAACAGAAACCTTACAGACGACAGATGTTTCTTTATTGTGATGCAATGTCCACGTATACCAGATATTGCATCTTTA
CTGAAATACCTTTTCAGGGTAGAATTGATCTTCAGTAACCCATGCTTGTCATATACTACC**AGAATGGGGTGGTCAAGCTCGCTG**AA
TTGGTACGCAGTC

Appendix B2.5 Nucleotide sequences of 275 5/15F1 and 275 5/15F2. Sequences and positions of the forward primers and those complementary to the reverse primers are underlined and boldfaced

280 11/11M1:
GACTGCATACCAATTCATGGTAGTGTCTTTCCGATCGCAAATGTCAGGGTGGTATTTTGAACTCCAGGTTCTGCCATGGAATATCT
TGTTGTAGGAGAAGGTCTCACTTGGTTTCGGCTGGCATTCTCAGACTTTTGAAAGGTGACAAAATGTCATGAAATGGTTGGTGATA
TTTTGAAGGCCTATATATATCATGAACTATACATATAGCGGCACAGTTGTCCAGAGCGGGAGAGCTGCTTGCACTAATAATGAAGG
CTTGGATTAGGATTTGCTTGCGTTACTCAGGACTCATC
280 11/11M2:
GACTGCATACCAATTCATGAATGCCTCGAGTGAGAGAGTTGGATGGAGAGCCATGATCGTCAATGCCTCTTACA**GGCAGGACGCTA**
CGTGATAATGAACCCATGTTTGTTGTAACAGTCGACGGGACGGGTGGTCAGGCTCGCTAACTCAGTTGACACATGTCATCGTATTC
CAAGTGCCTAGATCGATGCGCATGCTGTTG**GTCACTGGATTGTCTGCTCC**AGTCTCGATTATTTACGTCATATAGCTGGAATATTG
CTGGGCGAACGTTACTCAGGACTCATC
280 11/11M3:
GACTGCGTACCAATTCATGATGGGTAGATAGTACTAGATACTGGCAGATACTAGTATTCTGATCCACACATGTAAACCCATCGTCT
TCCAGGTACGTGCACCAAGCCCTCAGGCAAGATGACAGCGATGGCCGGACTCAGAGGGAGATCAAGACATCGATACGGTTAGATAG
ATGGGGTCAGTGGTTGGACTCCTCATCCCCATCCCTCCTGGTCACTCCGCGCACCCATGCCATGACATAATAGATGTTGATAGAGA
CCCATACGTGTGAGCGTTACTCAGGACTCATC
Appendix B2.6 Nucleotide sequences of 280 11/11M1, 280 11/11M2 and 280 11/11M3. Sequences and positions of the forward primer and those complementary to the reverse primer are underlined and boldfaced

320 9/15M:
GACTGCGTACCAATTCAGTGCGACCCTGATTTCTGACGCCGTAGATATAACTTAGTTTCACGCGTGACACAAGTTATGAACGATAA
GTATGGCCTTTCATACGATACATATGGTTTTCGTACTCTGTTCAATATCCGTAG**TAGGTTCTGGTGACACTTTGCCCGA**TGGAATT
GTCTGGATGCAATGTATCTTGAGGAAAATACCCAGTAATAATATGGTGTTCTAGAACTGTGTATGATTTTCTGGTTCTCTCAGCTG
CAAA**CAGCAACATTCCAGTTCTCACACAT**CTGGGAATAGCACGGTTACTCAGGACTCATC
Appendix B2.7 Nucleotide sequences of 320 9/15M. Sequences and positions of the forward primer and those complementary to the reverse primer are underlined and boldfaced

517 6/9M1:
GACTGCATACCAATTCACGGCATTATGTCATATACAGAATGAGGTACATAGACAATTCTGCATCAAATGTTATACTGCATGTAAAA
TGGAGCACCAAATCTTTCTAACAGCCTGTGTTATCAGGCTTGAATCAGTCAAAACTTCAGTAGTGTATGCCTGGATTGTGTAGGAT
GTCCAAGCTTTTCCAAGATGTAAGTATAGGTGCCCCAGGTGTTTGAGGTGCTCCACTGTGCAGTGTTGCATCACTACGACATGATT
GAAGCCTATGGTTGTTGGCTTGAGCATCAGGTGTCCACAGTAATTTGTCTGGACCAGCCTTCACACTGAACCCACAAGTTCAGATC
TTGTGGTAGCTGAAATTCAAACATGACTTTGTTAGGCTCACTTTCTTCACTCTTGTATTGTGTAGCATATCCTACTACTGCTTGTA
GTCTTATACATTACTTTAGTGTAGGATCATTGACAGCATAGAGTGTATTTAGAGGTGCATTTTCATGTCTTCGTTACTCAGGACTC
ATC
517 6/9M2:
GACTGCATACCAATTCACGTCATATTTTAGTTGTAGTTGTTTGTCGTAGTCATAGGACCAGATGGAATAGTCATTTCTATGGTAGA
AGCATGAAAATACAGTCAAATGCATTGTCACAAATCCACTTCGAACATGTTGTAAGCTGTTGTCTTCACATTCTTGTTTGTGGTTG
CTGAGGGTTTATTGAAGCAGTAGAAATTGTTATAAAAGCATAAGAGTCATCTTTGGGGCAGGACTGCAACCTGAAAAGGATGTGAG
GGACATATATATAAGGCCATTTATGACAGTGAAAGTCGTTCACAGAGCATGTACTGTAAACTGGAAATATTCTGCATCAGTTTATA
TTTTTGAATTTTGCATTTCTCTTCCCCACGCAAAAAAATACTTGGAGCAGTAACAGCAACTCTTTCGAAGCCAGTGAACCAGATGCA
TTGCTTATTATAAGATGATTTCTCAAGTGATATGTATAGTTGACAATTTCCCAATATCAGTGGCTTTCCGCTCGTTACTCAGGACT
CATC
Appendix B2.8 Nucleotide sequences of 517 6/9M1 and 517 6/9M2.

375 5/15F:
GATGAGTCCTGAGTAACCGGAAGCGTGCGTACCTAAATAGCATTCGGAAACGTTCGGGTACAACCTTTTCGAGATAAAAAGTTCAA
AAGGTATGTGCGACTAACCACTTTTGCTATT**TGGTAAGCTGTGTTCTGATCGGTC**AATCTCAAAGGTTACCTGACACGACCTCCCA
TAAGGTTTTGTTAGACTCAGTAGTGGAGTGTAAAGTAAAGTACTGAGGATAAGTTTTCTTTTCTATTTTATTCTAGGATACAGAAC
TCAGTTCA**AGAATGACTGGGAATGTTTGGAG***GGCC*TCAGCTCCAGAACTTGACCTGGAATCCTGAGGTGTGTTTGTTTCTGAAGTA
AAGAGAGGGTGAATTGGTACGCATC
Appendix B2.9 Nucleotide sequences of 375 5/15F. Sequences and positions of the forward primer and those complementary to the reverse primer are underlined and boldfaced

235 4/13F:
GACTGCTTACCAATTCACTTCTGTGCCAGTAAACGGAGCACTCAAGAATGACTGTACAGACACTGGGTCGGCAACATCGGATAACT
CTTTGTCGTACAGCTGCTGAATGGATGTGGTGTCCTCCTGTAATGCTGCAGTCTGCTAGGGAGTCGAATCATGCCCTCAGATTTCA
CTGGAACATGGGGCGGTGGGGTAGCCTAGTGGTTACTCAGGACTCATC
Appendix B2.10 Nucleotide sequences of 235 4/13F.

260 4/13F1:
GATGAGTCCTGAGTAACCATGCTTGTTGTAAGA**GGTGACTAACGGGATCAAATGGT**GAGGCTCACAGTGGTTGACACATGTCCAAA
GCTTGTTTAGAAAGACTAATGTCAGGTGTGGAATATCCAGATTAGCTACCAAGGTTGGGAGGCAGAATCATGCAAGTTCAAAGTGT
AGCTCAAAGCGATTACATGACACTGCTTGACAGGCAGCCAAAAGAAAGATCACT**TTACCCGTCGGCAAGATTCC**AAGCAGTGAATT
GGTACGCAGTC
260 4/13F2:
GACTGCGTACCAATTCACTCGTGCTGGCACGATACGCTTGAACGACTCGATTAGAACGTACGTGTGGGCGCTGTTGGGTGCGCCAC
ACGTCCAACCACGTGAGAAGGGTATAGTGGCGCCACCCACGGCTCATAAGGTTCAAGTGGAGGTACCCAAGGTGGGTGTGGGGGAA
ACCACCACCTATCAGGAACATATTGATAATAAAACGGCCCTGGTTGTTGGTGGGGTAGCTTTGGGGCTTATTTGGTTACTCAGGAC
TCATC
Appendix B2.11 Nucleotide sequences of 260 4/13F1 and 260 4/13F2. Sequences and positions of the forward primer and those complementary to the reverse primer are underlined and boldfaced

400 5/8F1:

GACTGCGTACCAATTCACCCTCTCTTTACTTCATAAACAAACACGCCTCATAATTCCACGTCAAATTCTCGAGCTGAGGCCCTCCA
AACATTCCCAGTCATTCTTAAACTGAATTCTGTATCCTAGAATAAAATAAAAAACAAAACTTATCCTCAGTACTTTTCATTACACT
CCACTACTGAGTCTAACAAAACCTTATGGGAGGTCGTGTCAGGTAACCTTTGATATTGACCGATCAGAACACAGCTTACCAAATCG
CGAAAGTTGTTAGTCGCACATACCTTTTGAACTTTTTATCTCGAAAAGTTTCGTACCCGAACGTTTCCGAATGCTATTTAACGTAC
GCACGCTTCCGGTTACTCAGGACTCATC

400 5/8F2:

GACTGCGTACCAATTCACCATGTGT**TCTTGTGCCGAAAAGTGGAG**GTGGTTGGCGCTGCTGTGCTGACATGAAACTGATTCGTATC
CTATCCCGAGAGTGGACGATTGCATTGATCGCATTGGTCAAACGAAGTATATATGTAAGTTTGACTTACTGAAGCGTTTCTAGCAG
ATTCCTCTCACAGATCGCGCTTAGAAAATATCGGCATTCGCTACACCAGGTGCACTCAATCAGTTCAAGGTGATGCCATTTGGTCT
GAAGAACGCTCCAGCGACGTTCCAGAGAC**TCGTCAACAATGTCACCTCAGG**TCTTGATCATGTCGAGGCTTACTTGGGTGATGTCA
TTGATTATGACATGGAGTGGGATGACCATCTGAGAAGGACTCGTGCTTTCATCGAAAGACTGTCTGAGGCAAAGTTACTCAGGACT
CATC

400 5/8F3:

GACTGCGTACCAATTCACCAAAGGAATACAAGGCATGACGTAGCGGGAGGCCGTATTCTGCTCTAGATTCAACACAGGGAGATGAC
AAGCATGACGTACTGGAGACCTGATTCTGCTCTAAATTCTCAGAACGCTTATTTTGACATCGTTATATTATGGAAGCTTCATCTTT
CACGCGATTACGTTCTCCTCATTCATTAGTGATTGAGAAAGAACGATAAGCTCGGCCAATATTAGCTTCAGAAACTGATTATGTCT
TTTTCGCAACAATGACATGTATATCACATTATCATGGTTTGTCCTAGGTATTTGACATTTAGACGTGGCTGTTTGCCTCTGGATAT
ATGCAAGTCTCCGAACAAGATGAGGTTTACAGATCTTATTCAAACCAAAGTTACTCAGGACTCATC

Appendix B2.12 Nucleotide sequences of 400 5/8F1, 400 5/8F2 and 400 5/8F3. Sequences and positions of the forward primer and those complementary to the reverse primer are underlined and boldfaced

T**460 6/15F:**

GACTGCGTACCAATTCACGGGGAGTACCTACAACATTGTATCGTCCGGCTGGGTGTTATGCAGCAATGTTAGAGCCATAAGTGATC
GGGTTCAAAATAAGAAATCACCCACATCCTTCCAACAACTAAGTGTGGGGAACCAAATCTTAGAAAGTATTGACGTCTCATTTGCC
ATCTTTTCTTTTCTATTATCTCTTGTCATTGTTGTGTGGACGTGTTGTAAGTGTTGTAGAGGAACAAACGAGCGCCCTGTTGTGCA
GATCCAACATTTGTCAAAAGAACTTATGCGATGATCTGTTCCTTTTAGCAGATTTGTCTCTACCTGCCCGCCAGTCTAAGTAAACG
TAGTCTCAGTGTGTTTCGTTACACTCCACCACTGAGTCTAACAAAACCTAGTAGAAGGTCGCATCAGGTAACCTTTGAGCGACCAA
TGACCGTCCAATCAAACGCGAGACGGTTACTCAGGAC CATC

Appendix B2.13 Nucleotide sequences of 460 6/15F.

490 11/15F1:

GACTGCGTACCAATTCATGACAACATCCCTGTCGATAATCCAGTAGGAAAATAACGTAAGAGTAATAACTAATTACATCCAACGAA
TTATTCCGCATCTCTGAAGGTCTAAAGGTAAATATTACGCGTTATCGGTGACTCCATTTCCTTCATGTTTCTTGTTAGGCAAAATC
TTCTTGTAGTTAGGTGCGCGCACGAGTTTTCCTGTTATGATAACAACTGAGATGTGAGCATCTACATCATAGTCGATCGATCCATA
GGGTATGCCCCTATACTAATATGAGAGCAAGGTAATTCACCCGCCATATTGGTGACTTCGATAAACTTCGATTCTCGCTAATTTCT
CGCTTGACATGGATAAGCAGGGATATATGCTGAGGACGGAAGTTACCCCTATGAGAAGAGATTTTCCAGTGCCAGCCTGTCCTAAG
ACGACTACGTTTTGACCTTTTCTTATCAATTCATATGGTAGTGTATGTTCAGCGGTTACTCAGGACTCATC

490 11/15F2:

GACTGCGTACCAATTCATGACAACATCCCTGTCGATAATCCAGTAGGAAAATAACGTAAGAATAATAACTAATTACATCCAACGAA
TTATTCCGCATCTCTGAAGGTCTAAAGGTAAATATTACACGTTATCGGTGACTCCATTTCCTTCATGTTTCTTGTTAGGCAAAATC
TTCTTGTAGTTAGGTGCGCGCACGAGTTTTCCTGTTATGATAACAACTGAGATGTGAGCATCTACATCATAGTCGATCGATCCATA
GGGTATGCCCCTATACTAATATAAGAACAAGGTAATTCACCCGCCATATTGGTGACTTCGATAAACTTCGATTCTCGCTAATTTCT
CGCTTGACATGGATAAACAGGGGTATATACTGAAGACGGAAGTTACCCCTATGAGAAGAGATTTTCCAGTGCCAGCCTGTCCTAAG
ACAACTACATTTTGACCTTTTCTTATCAATTCATATGGTAGTGTATGTTCAGCGGTTACTCAGGACTCATC

Appendix B2.14 Nucleotide sequences of 490 11/15F1 and 490 11/15F2.

APPENDIX C

Publication from This Thesis

1. Klinbunga, S., Amparyup, P., Leelatanawit, L. Aoki, T., Hirono, I, Tassanakajon, A., Jarayabhand, P. and Menasveta, P. (2004). Species identification of the tropical abalone (*Haliotis asinina*, *Haliotis ovina* and *Haliotis varia*) in Thailand using RAPD and SCAR markers. *J. Biochem. Mol. Biol.* 37: 213-222.

2. Amparyup, P., Klinbunga, S., Preechaphol, R., Tassanakajon, A., Hirono, I., Aoki, T. and Jarayabhand, P. (2004). Expressed Sequence Tag (EST) Analysis of Ovaries and Testes from the Tropical Abalone (*Haliotis asinina*). *Mar. Biotechnol.* 6: S365-S370.

3. Klinbunga, S., Amparyup, P., Thamrungtanakit, S., Tassanakajon, A., Hirono, I., Aoki, T., Jarayabhand, P. and Menasveta, P. (2004). Population Genetics and Species-Specific Markers of the Tropical Abalone (*Haliotis asinina*) in Thailand. *Mar. Biotechnol.* 6: S360-S364.

Conferences

1. Amparyup, P., Klinbunga, S., Tassanakajon, A. and Jarayabhand, P. (2002). Species-specific markers of the tropical abalone (*Haliotis asinina)* 28th Congress on Science and Technology of Thailand. 24-26 October 2002, Bangkok, Thailand, p.515.

2. Amparyup, P., Khamnamtong, B., Klinbunga, S., Puanglarp, N., Tassakajon, A., Jarayabhand, P. and Menasveta, P. (2003) Development of species-diagnostic markers of the tropical abalone (*Haliotis asinina*) in Thailand. 29th Congress on Science and Technology of Thailand 20-22 October 2003, Khon Kean, Thailand (Oral presentation).

3. Amparyup, P., Klinbunga, S., Preechaphol, R., Tassanakajon, A., Hirono, I., Aoki, T., Jarayabhand, P. and Menasveta, P. (2004). Isolation and characterization of sex-specific expression of cDNAs from ovaries and testes of the tropical abalone (*Haliotis asinina*). The 15th Annual Meeting of the Thai Society for Biotechnology. 3-6 February 2004, Chiang Mai, Thailand .

4. Amparyup, P., Klinbunga, S., Tassanakajon, A., Aoki, T., Jarayabhand, P. and Menasveta, P. (2004). Isolation and characterization of gene expressed in ovaries and testes of the tropical abalone *Haliotis asinina* by cDNA subtraction. 30th Congress on Science and Technology of Thailand, 19-21 October 2004, Bangkok, Thailand (Oral presentation).

Biography

Acting 2LT. Piti Amparyup was born on August 3, 1976 in Bangkok, Thailand. He graduated with the degree of Master of Science (Biotechnology) at the Program of Biotechnology, Chulalongkorn University since 1999. In 2000, he has studied in Doctor of Philosophy (Ph.D.) degree of Science at the Program of Biotechnology, Chulalongkorn University.