BSPE9921C-10706-3

# 볼락류(*Sebastes spp.*)의 체색관련 기능유전체 연구

Functional genomic analysis of pigment patterns in rockfishes (*Sebastes* spp.)

2015. 3

# 한 국 해 양 과 학 기 술 원

## 제 출 문

한국해양과학기술원장 귀하

본 보고서를 "볼락류(Sebastes spp.)의 체색관련 기능유전체 연구"과제의 최종보고서 로 제출합니다.

#### 2015. 3.

- 총괄연구책임자 : 장 요 순
- 참여연구원 : 이은경, 황선완,
  - 유재명,명정구
  - 오 승 용, 김 종 관
  - 이 정 아, 장 영 석
  - 권 순 열, 윤 진 영
  - 오지나, 최해영
  - 조혜경, 박서윤

### 보고서 초록

과제고유	DE0091C	해당단계	2014.0	9.01.~	단계			
번호	FE9921C	연구기간 2014.12.31.		구분				
여그시어머	중사업명		미래 해양생물자원 개발					
친구가협령	세부사업명	볼락류(Sebasi	볼락류( <i>Sebastes spp.</i> )의 체색관련 기능유전체 연구					
ਨੀ — ਹੀ ਸੀ ਸੀ	대과제명							
친가과제정	세부과제명							
		체다다게	총 :	15 명	체다다게	정부: 229,659,000원		
		에 8 번게 차여여구위스	내부:	7 명	에 8 년 개 여구비	기업: 원		
여그채이기	자이스	섬에신티현기	외부:	8 명	친구별	계 : 229,659,000원		
친구적 감사	0 11 1	초여그기가	총 :	15 명	ネ	정부: 229,659,000원		
		중연구기신 차성성그이스	내부:	7 명	중 하구비	기업: 원		
		삼여연구권구	외부:	8 명	연구미	계 : 229,659,000원		
연구기관명								
및	동해여	연구소	참여	기업명				
소속부서명								
국제공동연구								
위탁연구								
요약(약	친구결과를 중심.	으로 개조식 500	자 이내)		보고서 면수	83페이지		
- 동해서식 볼	⊦락류 4종(우력	볼락, 흰꼬리	볼락, 황	볼락, 노혁	랑볼락)의	유전자 정보를 분석		
하였고, 미토	트게놈 염기서 9	결 정보를 Gen	Bank에	등록하였	음			
- RNA-seq a	<i>le novo</i> assen	nbly 방법으로	개볼락	근육조직	전사체 :	정보를 체색타입별로		
조사하여 1	01,112개의 ui	nigene을 확보	하였고,	염기서의	열 및 아디	기노산 서열 정보를		
뱅크화 하였	<u>.</u>	_						
- 개복락 체색	타입벽 근육격	~직 저사체의	박혀량읰	- 분석하여	여 164개 :	의 차등박혀유저자륵		
화이하여고	개보라 color	type 근유조지	니에서마	트이전이	- 반혀되	느 51개의 유저자와		
wild type	, 개골의 COIOI 그은즈지에서미	나 뜻이 전 이 군 니	바처디느	- 68기 으	- 고 고 고 이 - 저 기르 히	비치여으		
ישות type י	_ 픽그 픽 테시니	. ㅋㅋㅋㅋ프로 이제 피르 키기	르 킨 ᅴ ᆫ 드 바 쳐 ㅇ	ןן ווי-100. - (מנכובר	·관/기관 ㅋ 기느거머리	포이从 <sub>ロ</sub> 고기최 거고 이게느		
- 개골닥 근표	소식의 세색티	-입에 따른 자-	허탈 연규	·선사의 /	기등상모들	소사안 결과, 2개근		
molecular f	molecular function, 46개는 biological process, 6개는 cellular component 기능그룹에							
해당하였음								
색인어	한 글	볼락류, 미토기	╢놈, 기능	등유전체,	전사유전체	∥, 차등발현유전자		
(각 5개 이상)	역 어	<i>Sebastes</i> , Mi	togenom	ne, Trans	scriptome,	RNA-seq analysis,		
	-9 A	Differentially expressed gene						

### 요 약 문

#### I.제 목

볼락류(Sebastes spp.)의 체색관련 기능유전체 연구

### Ⅱ. 연구개발의 필요성 및 목적

다양한 해양생명자원의 유전체 정보 해독은 신품종 개발과 식의약 소재 개발에 도움이 되고 식량안보 및 기후변화 등의 위기 대응에 도움이 되는 것으로 인식되고 있다. 해양생명자원의 유전체 정보를 분석하여 확보한 유용한 물질은 신약, 신소재, 바이오 에너지 생산을 위한 자원으로 개발할 가치가 매우 높다.

방대한 양의 염기서열을 신속하게 분석할 수 있는 NGS(next generation technology) 기술 개발로 기존 유전체 연구의 패러다임이 기술이나 시간, 비용 측면에서 급격하게 변하였다. NGS 분석 방법으로 생산된 대량의 염기서열 데이터는 유전체 재분석 연구, 유전자 발현 연구, allele variation 및 association mapping 등에 적용할 수 있는 새로운 연구 플랫폼을 제공하고 있다. 유전체 분석의 가속화로 다수의 생물 유전체가 해독됨에 따라 많은 양의 새로운 유전체 정보, 이른바 빅데이터 처리를 위한 기술 및 유전체 정보 실용화 기술 개발이 필요하다.

본 연구를 통하여 해양어류 자원의 기능유전체 정보뱅크를 구축하고 미이용 어류 자원의 유용자원화 및 고부가가치 산업화 연구에 활용하고자 한다. 또한, 양식이 가능한 수산생물의 생산성 향상 관련 분자유전학적 도구를 개발하고 생명자원의 유전체 정보 활용연구 기반을 마련하고자 한다.

#### Ⅲ. 연구개발의 내용 및 범위

1. 동해서식 볼락류의 유전자 정보 탐색

- 미토게놈 정보 분석

- 2. 동해서식 볼락류의 체색관련 기능유전체 탐색
  - 개볼락 체색타입별 근육조직 전사체 정보 분석
  - 개볼락 체색타입별 차등발현유전자 분석

### Ⅳ. 연구개발결과

1. 동해서식 볼락류의 미토게놈 정보 분석 및 Genbank 등록

- 동해서식 볼락류 4종(우럭볼락, 흰꼬리볼락, 황볼락, 노랑볼락)의 미토게놈을 분석하고 염기서열 정보를 NCBI (National Center for Biotechnology Information)의 GenBank에 등록하였다. 우럭볼락 미토게놈의 전체 크기는 16,453 bp (GenBank Accession No. KJ 525745), 흰꼬리볼락 미토게놈의 전체 크기는 16,445 bp (GenBank Accession No. KJ 834061), 황볼락 미토게놈의 전체 크기는 16,465 bp (GenBank Accession No. KJ 834063), 노랑볼락 미토게놈의 전체 크기는 16,450 bp (GenBank Accession No. KJ 834060)이었다.

2. 동해서식 볼락류의 체색관련 기능유전체 탐색 및 차등발현유전자 기능 분석

- 개볼락의 체색 타입별 근육조직 전사체를 RNA-seq 방법으로 분석하여 101,112개의 unigene을 확보하였다. BLASTx 검색과 InterProScan(v5) 검색으로 unigene의 기능을 예측하였다. 기능이 밝혀진 35,482개 유전자와 기능이 밝혀지지 않은 유전자 65,630개의 염기서열 정보 및 아미노산 서열을 확보하였다. 개볼락의 체색타입별 근육조직 전사 유전체의 발현량을 측정하여 두 개의 근육조직 시료에서 발현량 차이를 나타내는 164개의 차등발현 유전자(DEG, differentially expressed gene)를 확보하였다. 개볼락 근육조직의 체색 타입에 따른 차등발현 유전자의 기능 정보를 파악하기 위하여 Gene Ontology (GO) enrichment 분석한 결과, 164개의 DEG가 1,613개의 GO ID에 해당함을 확인 하였으며, DEG 2개는 molecular function, 46개는 biological process, 6개는 cellular component 기능그룹에 해당하였다.

### V. 연구개발결과의 활용계획

본 연구를 통하여 확보한 볼락류 미토게놈 정보 및 개볼락의 전사유전자 정보는 유전자 기능분석을 통하여 어류의 체색관련 생명현상 규명에 활용할 계획이다. 또한, 어류의 기능유전체 정보를 이용하여 분자유전학적 도구를 개발하고 더 나아가 생물의 유전체 정보 자원을 산업화 할 수 있는 기술 개발을 위한 참고자료로 사용하고자 한다.

## S U M M A R Y & KEYWORDS

A development of genome sequence Big-data application technology is needed for bio-industrial utilization of marine fish functional genome sequence. Rockfishes are members of the genus *Sebastes*. Currently, we know of about 110 rockfishes species worlwide. We analysed the complete mitochondrial genome of 4 rockfish species were determined using next-generation sequencing, including Sanger sequencing.

Hybrids of *Sebastes hubbsi*, considered a valid species, have been found in the ocean. The complete mitogenome of *S. hubbsis* was 16,453 bp (GenBank Accession No. KJ 525745) in length and comprised 13 protein-coding genes (PCGs), 2 rRNAs, 22 tRNAs and 1 control region. The complete mitochondrial DNA had 27.9% A, 26.7% T, 17.2% G and 28.3% C, showing a slight AT bias (54.5%). The start codon of all PCGs was ATG, except COX1 (GTG). The stop codons in the PCGs were mainly TAA, except ND1 and ND3 (TAG), ND4 (AGA) and Cytb (incomplete termination codon, T). The tRNAs had a typical cloverleaf structure, except tRNASer (AGY) which lacked the DHU arm.

Sebastes longispinis is considered to be a valid species despite its incomplete mitochondrial lineage. The complete sequence *S. longispinis* is 16,445 bp (GenBank Accession No. KJ 834061) in length and comprised of 13 protein-coding genes (PCGs), 2 rRNAs, 22 tRNAs and a control region. The complete mitogenome contains 28.0% A, 26.7% T, 17.1% G and 28.3% C nucleotides, with a slight AT bias (54.6%). The start codon of each PCG is ATG, except for COX1 (GTG). The stop codons found in the PCGs are mainly TAA, except for ND1 (TAG), ND3 (TAG), ND4 (AGA), and Cytb (incomplete termination codon, T). The tRNAs have a typical cloverleaf form,

except for tRNASer (AGY), which lacks the DHU arm.

Sebastes owstoni is popular fishery species in the northern western Pacific Ocean. The assembled mtDNA was 16,465 bp (GenBank Accession No. KJ 834063) in length, comprising 13 protein-coding genes (PCGs), 2 rRNAs, 22 tRNAs and 1 control region. The complete mitogenome comprised 27.7% A, 26.6% T, 17.3% G and 28.4% C, indicating a slight AT bias (54.3%). The start codon of all PCGs was ATG, with the exception of COX1 (GTG). The stop codons in the PCGs were mainly TAA, with the exception of ND1 and ND3 (TAG), ND4 (AGA) and Cytb (incomplete termination codon, T). All tRNAs had a typical cloverleaf shape, with the exception of tRNASer (AGY), which lacked the DHU arm.

The mitogenome of *S. steindachneri* was 16,450 bp (GenBank Accession No. KJ 834060) in length, and comprised of 13 protein-coding genes (PCGs), 2 rRNAs, 22 tRNAs, and a control region. The mitogenome is 27.4% A, 21.1% T, 17.5% G, and 28.0% C, showing a slight AT bias (54.5%). ATG was the start codon in all of the PCGs, except COX1 (GTG). The termination codon in the PCGs was TAA, except in ND1 and ND3 (TAG), ND4 (AGA), and Cytb (incomplete termination codon, T). All of the tRNAs had a typical cloverleaf shape, except tRNASer(AGY), which lacked the DHU arm. The sequenced mitogenome of *Sebastes* can be used in genetics-based management initiatives or for constructing the mitochondrial lineages within Scorpaenidae.

We constructed cDNAs from total RNA isolated from the skeletal muscles of two blass bloched rockfish (*S. pachycephalus*) genotypes : wild type and color type. We use next-generation sequencing (Illumina) RNAseq analyses to compare skeletal muscle transcriptome-wide expression levels in two distinct color pattern in blass bloched rockfish. Sequencing generated 156,145,718 raw reads and this number was reduced to 146,847,238 sequences after the cleaning pipeline was implemented. The latter sequences were used to build the de novo assembly. Trinityt produced a high number of transcripts that were clustered by TGICL into 101,112 unigenes. Using a combination of differential expression analyses (DE) we identified 164 candidate genes that showed DE between the color morphs. We assigned Gene Ontology (GO) terms to the 1,613 annotated loci. Fourty-six 'Biological Process' categories were assigned among 1,191 blass bloched rockfish loci; Six 'Cellular Component' categories were assigned among 192 blass bloched rockfish loci; and Two 'Molecular Function' categories were assigned among 230 blass bloched rockfish loci. Our study provides valuable transcriptome sequence data for *S. pachycephalus.* 

Our transcriptomic sequences will enhance the quality of gene annotation and functional analysis of the blass bloched rockfish genome and serve as a material basis for future genomic research on marine fishes. Using transcriptomic analyses we successfully identified key expression differences between different color morphs of blass bloched rockfish. These differentially expressed genes have important implications for our understanding of the molecular mechanisms underlying pigmentaion. The sequencing data from this study can be used to develop molecular markers and to identify the extreme differences among the phenotypes of different species in the genus *Sebastes*.

KEYWORDS : 볼락류, 미토게놈, 기능유전체, 전사유전체, 차등발현유전자 Rockfish, Mitogenome, Functional genome, Transcriptome, Differentially expressed gene

표목차 ••••	• • • • • • • • • • • • • • • • • • • •	•••••		• • • • • • • • • • • • • • • • • • • •	
그림목차 ••	•••••		•••••	•••••	

.2	ㅓ론	· 서	17	제
.3	설 연구 필요성 ·····	1절	제	
.4	설 연구 목표 ·····	2절	제	
.5	<u></u> 연구내용 및 범위 ·····	3절	제	

۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰۰	
황	
ब्र <del>]</del> -	

		수행내용 및 결과 '	제 3장 연구개발	제
	정보 탐색	배양어류 유전자원	제 1절 유용	
51	체 정보 이용기술 개발 탐색	류자원의 기능유전	제 2절 해양여	

제 4장 연구개발 목표 달성도	및 대외기여도75
제 1절 연구개발 목표 달성	도 ····································
제 2절 대외기여도	

제	5장	연구개발결과의	활용계획	 79

제	6장	참고문헌	 80

## List of Tables

Table 1 Summary of the mitochondrial genome of S. hubbsi

Table 2. Summary of the mitochondrial genome of S. longispinis

Table 3 Summary of the mitochondrial genome of S. owstoni

Table 4. Summary of the mitochondrial genome of S. steindachneri

Table 6. Unigenes transcripts detected in this study

Table 7. List of color pattern type-specific differentially expressed genes

Table 8. Known genes of color-type skeletal muscle specific expressed genes

Table 9. Known genes of wild color-type skeletal muscle specific expressed genes

Table 10. Molecular function group of the differentially expressed genes

## List of Figures

Figure 1. Sequences of the mitogenome of S. hubbsi (GenBank Accession No. KJ525745)

Figure 2. A gene arrangement in the complete mitogenome Sebastes hubbsi.
Figure 3. Sequences of the mitogenome of S. longispins (GenBank Accession No. KJ834061)
Figure 4. A gene arrangement in the complete mitogenome Sebastes longispinis.
Figure 5. Sequences of the mitogenome of S. owstoni (GenBank Accession No. KJ834063)
Figure 6. A gene arrangement in the complete mitogenome Sebastes owstoni.
Figure 7. Sequences of the mitogenome of S. steindachneri (GenBank Accession No. KJ834060)
Figure 8. A gene arrangement in the complete mitogenome Sebastes steindachneri.

Figure 9. Workflow of amplicon library construction and Hiseq2500 sequencing.

Figure 10. The skeletal muscles of S. pachycephalus for RNA-seq analysis.

Figure 11. The result of quality assessment of total RNA.

Figure 11. The result of quality assessment of total RNA.

Figure 12. Results of the blass bloched rockfish (S. pachycephalus) de novo assemly

Figure 13. Frequancy of the blass bloched rockfish unigenes BLAST annotation

- Figure 14. Predict of CDS of unigenes by InterProScan
- Figure 15. Annotation consensus betwwen Blastx and InterProScan of unigenes
- Figure 16. DNA sequences of unigene transcripts
- Figure 17. Amino acid sequences of unigene transcripts

#### Figure 18. Expressiom levels of unigene transcripts

#### Figure 19. Differentially expressed genes (DEG) analysis of unigene transcripts

Figure 20. Histogram of the Gene Ontology classification

# 제 1장 서 론

### 제1절 연구 필요성

생명공학 관련 기술의 발달로 생물의 유전체 정보를 신속하게 대량으로 생산할 수 있게 됨에 따라 유전자의 기능과 생명현상 규명 등 유전체 정보 활용 연구가 활발히 진행되고 있다. 유전체 데이터 분석 및 이용기술 개발 연구의 필요성 증가에 따라 다양한 생물의 유전체 정보를 활용하기 위한 맞춤형 데이터베이스가 구축되고 있으며 생명체가 가진 유전체 정보는 매우 유용한 연구자산으로 개별 유전자의 기능 및 유전자의 상호 네트워크를 이해하기 위한 유전체 정보 활용기술 개발이 요구되고 있다. 방대한 양의 염기서열을 신속하게 분석할 수 있는 NGS(next generation technology) 기술 개발로 기존 유전체 연구의 패러다임이 기술이나 시간, 비용 측면에서 급격하게 변하였고, NGS 분석 방법으로 생산된 대량의 염기서열 데이터는 유전체 재분석 연구, 유전자 발현 연구, allele variation 및 association mapping 등에 적용할 수 있는 새로운 연구 플랫폼을 제공하고 있다. 유전체 분석의 가속화로 다수의 생물 유전체가 해독 되었으며, 방대한 양의 새로운 유전체 정보, 이른바 빅데이터를 처리하기 위한 기술 및 유전체 정보 실용화 기술 개발이 시급하다.

해양생명자원의 유전체 정보 분석 및 이용기술 개발을 통하여 유용한 물질 및 기능을 활용한 신약, 신소재, 바이오에너지 등을 생산하는 해양바이오산업 육성을 위한 기반이 마련되어야 한다. 맥킨지(2013년 보고서)는 유전체학을 인류의 삶을 근본적으로 바꿀 미래기술로 분류하였으며, 2025년까지 글로벌 경제 파급효과를 1.6조 달러로 추정하였다. 해외 선진국들은 유전체 정보 활용기술의 막대한 경제효과를 인식하고 유전체 산업 자체가 미래 성장동력이 될 것으로 보고 치열한 경쟁을 벌이고 있다. 다양한 해양 생명자원의 유전체 정보 해독은 신품종 개발과 식의약 소재 개발에 도움이 되고 식량안보 및 기후변화 등의 위기 대응에 도움이 되는 것으로 인식되고 있다.

NGS 유전체 해독 기술 등 생명공학기술의 발전으로 다양한 생물의 유전체 정보 및 오믹스 정보가 축적되었고, 이들 정보를 활용하기 위한 데이터베이스 및 프로그램이 구축되고 있다. 『포스트 게놈 신산업 육성을 위한 다부처 유전체 연구사업』(보건 복지부, 농림축산식품부, 해양수산부, 미래창조과학부, 산업통상자원부, 농촌진흥청 / 2014년~2021년)이 '유전체 연구를 통한 국민건강 증진과 산업화 연계시스템 구현'이라는 비젼으로 동·식물, 해양생물등 유전정보를 활용한 고부가가치 생명자원을 개발하기 위한 프로젝트가 시작되었다. NGS 기술 개발로 생물의 유전체 정보가 매우 빠른 속도로 축적되고 있으나 유용한 연구자산인 유전체 정보를 효율적으로 활용할 수 있는 준비가 부족하다. 매우 빠른 속도로 발전하고 있는 생명공학은 다양한 분야와 융복합하여 많은 새로운 첨단기술이 개발되고 있으나 생명공학 연구의 활용은 매우 미흡한 실정이다.

#### 제2절 연구 목표

해양어류 자원을 효율적으로 활용하기 위하여 볼락류(Sebastes spp.)를 중심으로 기능유전체 정보뱅크를 구축한 후, 유전체 빅데이터 정보 데이터베이스를 융복합한 기능유전체 이용기술을 개발하고자 한다. 볼락류는 전 세계에 걸쳐 약 110종이 분포 하고 있으며, 우리나라에는 19종이 출현하고 있다. 체색이나 생태 및 행동이 매우 다양하게 나타나는 특성을 가진 볼락류는 특정 해역에서만 출현하는 지리학적인 분포 특성을 가진 종(species)도 있다. 본 연구사업에서는 우리나라에 서식하는 볼락류를 중심으로 유전특성을 조사하고 기능유전체를 분석하여 어류자원을 효율적으로 활용하기 위한 기능유전체 정보뱅크를 구축할 계획이다. 또한, 볼락류 기능유전체 연구 결과는 미이용 어류자원의 유용자원화 및 고부가가치 산업화 연구에 활용하고자 하며, 양식이 가능한 수산생물의 생산성 향상 관련 분자유전학적 도구를 개발하고 생명자원의 유전체 정보 활용연구 기반을 마련하고자 한다.

## 제3절 연구내용 및 범위

구 분	목 표	내 용 및 범 위
	○유용 해양어류 유전자원 정보 탐색	
	•유용 어류자원 탐색 및 볼락류	• 동해서식 유용 어류자원 조사
	유전자 정보 탐색	- 동해서식 볼락류의 자원특성 조사
		• 동해서식 볼락류의 유전특성 조사
1차년도		- 동해서식 볼락류의 유전정보 확보
(2014)		- 개볼락 시료의 유전특성 조사
	○해양어류자원 기능유전체 이용기술	탐색
	•볼락류 기능유전체 이용기술 탐색	• 동해서식 볼락류의 기능유전체 정보 탐색
		• 개볼락의 체색관련 유전특성 분석
수 문	<u> </u>	내 용 및 멈 위
	· 유용 해양어류 유전자원 확보	
	•유용 어류자원 탐색 및 볼락류	• 농해서식 볼락류 시료 확보
	유전자 정보 탐색	- 동해 고유서식 볼락류 확보
		- 체색 및 spot 패턴이 다양한 개
		볼락 확보
		• 동해 고유서식 볼락류의 유전특성 조사
		- 개체 또는 집단수준에서 유전자 변
		이 군식
		- 중해 고뉴적적 플덕류 미도세금 군적
2차년도		• 세색 및 spot 패턴이 나장한 개골
(2015)		- 기보라이 미드개노 부서
	○ 해양어류자원 기누운저체 이용기숙	· 개코릭ㅋ 티노게늄 판국 · 타새
	• 보라르 기느은저체 이용기수 타새	• 도체서시 보라르 기느은저체 저버배
		- 이에지ㅋ 필ㅋ 11 기 이 1 전세 18 포 18 
		•개복락의 체색 과려 기능유저자 확
		보 및 정보 분석
		•개볼락의 다양한 체색과 기능유정자
		발현과의 상관성 조사
		•개볼락 기능유전체 활용방안 탐색

구 분	목 표	내 용 및 범 위				
	○유용 해양어류 유전자원의 활용 연	구				
3차년도 (2016)	• 볼락류 유전자 정보 활용 연구	<ul> <li>볼락류의 유전특성 비교분석</li> <li>양식이 가능한 수산생물의 생산성 향상 관련 유전자마커 개발</li> </ul>				
	o 해양어류자원 기능유전체 이용기술 개발					
	<ul> <li>●볼락류 기능유전체 이용기술 개발</li> </ul>	<ul> <li>개볼락 기능유전체를 활용한 유전자 마커 개발</li> <li>볼락류 기능유전체 이용기술 개발</li> <li>분자유전학적 도구 개발 / 적용기술 개발</li> <li>바이오산업소재화 가능성 탐색</li> </ul>				

# 제 2 장 국내외 기술개발 현황

### 제1절 국내 기술개발 현황

국내 해양생물 유전체 연구는 해양생명공학사업으로「해양극한생물분자유전체 연구단」에서 유전체 연구를 추진(2004년~2013년)하였으며, 해양극한생물을 대상으로 유전자 및 단백질 수준에서 생명기능과 현상을 해석하고 유용유전자 및 단백질을 발굴하여 활용하기 위한 기술 개발 연구가 진행된 바 있다. 이 연구사업을 통하여 밍크고래 전장유전체가 해독되었고, 해양 동식물 12종과 해양미생물 43건의 유전체 정보가 확보되었다. 국립수산과학원에서는「양식생물 및 수산생물의 육종(분자육종 포함) 연구」가 넙치, 전복, 돌돔, 멍게를 대상으로 2004년터 2013년까지 10년간 진행 되었다. 이 연구사업을 통하여 넙치 유전체를 완전해독하고 유전자 지도를 작성하였으나, 유전자 정보 활용 연구 결과는 없는 상태이다.

국내에서 진행되고 있는 어류의 기능유전체 연구는 제주대학교에서 까막전복과 돌돔을 대상으로 수행하였으며, 한국해양과학기술원에서는 강도다리와 볼락을 대상으로 근육성장과 관련된 기능유전체를 분석하고 기능유전자를 이용한 어류의 생산성 향상 관련 유전자마커 개발 연구를 수행하였다.

#### 제2절 국외 기술개발 현황

해양생명자원의 유전체 연구는 인간유전체를 비롯하여 육상생명자원을 대상으로 다른 바이오 분야에 비하여 늦게 시작되었으나, 최근 미국 등 선진국을 중심으로 대 규모 연구가 진행되고 있다.

미국은 「21세기 해양청사진」과 같은 해양에 대한 지속적인 정책을 선언하고 세계 해양바이오에 대한 연구를 주도하며 다양한 해양시료 수집 활동 및 해양생물종 유전체 연구를 활발하게 진행하고 있다. 일본에서는 산학연 협력 연구활성화를 통하여 해양유전체 연구를 추진하고 있으며, 문부과학성, 국토 교통성 등 부처 공동으로 해양유전체 연구 분야 지원을 확대하였고, 해양과학기술연구센터(JAMSTEC)와 해양생명공학연구소(MBI), 수산종합연구센터 (FRA)를 중심으로 유전체 연구에 많은 예산을 투자하고 있다.

중국에서는 해양환경변화와 해양생물자원의 지속가능한 이용 및 대양 심해 연구에 대한 국가 지원을 강화하고 중국과학기술부를 중심으로 서해, 동중국해, 남중국해에 서식하는 어류, 원핵생물, 식물, 부유동물, 무척추동물 등 주요 해양생물 2천종의 유전자 정보를 담은 DNA 바코드를 데이터베이스화하는 사업인 '중국 근해 해양생물 DNA 바코드 DB구축 프로젝트'를 착수하였다.

프랑스에서는 로스콥해양연구소와 프랑스 국립시컨싱센터(Genoscope)가 중심이 되어 해조류 유전체 분석연구를 진행하였고, 캐나다에서는 수산해양부(DFP, Department of Fisheries and Oceans)를 중심으로 해양생명공학 분야에서 유전체학 및 생명공학 발전계획을 수립하였으며, 대서양연어 유전자지도 작성 등 연어 유전체 연구를 수행 하였다.

# 제 3 장 연구개발 수행내용 및 결과

#### 제1절 유용 해양어류 유전자원 정보 탐색

#### 1. 연구 재료 및 방법

#### 가. 어류시료 확보

동해에 서식하는 볼락류 시료를 확보하였으며, 임원항과 죽변항 일대에서 개볼락, 우럭볼락, 흰꼬리볼락, 황볼락, 노랑볼락을 구입하였다. 유전자 분석시료는 근육조직 및 지느러미 조직을 이용하였으며 절단하여 냉동 보관하였다.

나. 볼락류 어류의 genomic DNA 분리 및 정제

우럭볼락, 흰꼬리볼락, 황볼락, 노랑볼락의 꼬리지느러미를 0.5cm X 0.5cm(가로X세로) 크기로 잘라서 lysis buffer[10mM Tris-HCl pH7.5, 125mM NaCl, 10mM EDTA, 0.5% SDS, 5M Urea, 0.1mg/mℓ proteinase K]를 첨가하여 충분히 용해한 후, binding buffer와 isopropanol을 사용하여 genomic DNA를 분리하였고, *Accuprep*<sup>®</sup> Genomic DNA Extraction Kit(Bioneer Co., Korea)의 column을 이용하여 정제하였다. 분리·정제한 genomic DNA는 전기영동으로 확인한 후, NanoDrop<sup>®</sup> ND-1000 Spectrophotometer(NanoDrop Technologies, USA)를 사용하여 농도를 측정하고 -80℃에 보관하였다.

다. 미토게놈 정보 분석 및 GenBank 등록

분리·정제한 genomic DNA의 순도를 확인한 후 HiSeq2000(Illumina Inc.)을 이용하여 염기서열을 분석하였다. 볼락류의 미토게놈 정보를 SOAPdenovo2 software로 어셈블리 한 후, 빠진 영역의 염기서열은 PCR 증폭산물을 Sanger sequencing하여 확보하였다. 어셈블리가 끝난 미토게놈 정보는 DOGMA와 tRNA-SE를 활용하여 annotation 하여 GenBank에 등록하였다. 2. 연구결과

가. 동해서식 볼락류의 자원특성 조사 및 시료 확보

볼락류(Sebastes spp.)는 전 세계에 걸쳐 약 110종이 분포하고 있으며, 우리나라에는 19종이 출현하고 있다. 체색이나 생태 및 행동이 매우 다양하게 나타나는 특성을 가진 볼락류는 특정 해역에서만 출현하는 지리학적인 분포특성을 가진 종(species)도 있다. 볼락류는 대부분 난태생으로 연안의 암초가 많은 수심 20~170m 이내 지역에 서식하며, 돌삼뱅이(Sebastes baramenuke), 황볼락(Sebastes owstoni), 노랑볼락(Sebastes steindachneri) 등 일부 종은 100~400m 이내의 다소 깊은 바다에 서식한다. 동해안에는 17종의 볼락류가 출현하는 것으로 알려져 있으며, 돌삼뱅이, 황볼락, 노랑볼락, 탁자볼락 (S. taczanowskii), 말락볼락(S. wakiyai), 좀볼락(S. minor)은 동해에서만 서식하는 것으로 보고되었다. 체색과 관련된 기능유전체를 분석하기 위하여 다양한 체색 패턴을 나타내는 개볼락(S. pachycephalus), 우럭볼락(S. hubbsi) 및 흰꼬리볼락(S. longispinis)을 확보 하였다. 또한, 동해 중부 이북에 서식하는 황볼락(S. trivittatus), 돔블락을 확보하여 유전정보 분석을 위한 시료로 활용하였다.

#### 나. 동해서식 볼락류의 미토게놈 정보 분석 및 GenBank 등록

우럭볼락 미토콘드리아 DNA의 전체 염기서열(미토게놈) 정보를 분석하여 세계유전자 은행인 NCBI의 GenBank (http://www.ncbi.nlm.nih.gov/genbank/)에 등록(Accession number KJ525745)하였다(Figure 1). 우럭볼락의 미토게놈은 13개의 단백질 암호화 영역, 2개의 리보좀, 22개의 tRNA 및 1개의 조절영역으로 이루어져 있었고, 전체 크기는 16,453 bp 이었다. 우럭볼락의 미토콘드리아 DNA의 염기 조성비는 27.9% (A), 26.7% (T), 17.2% (G) 및 28.3% (C)이었고, AT 비율이 약간 높았으며(54.6%), COX1 (GTG)을 제외한 모든 단백질 암호영역의 개시코돈은 ATG이었고, 종결코돈은 TAA가 주를 이루었으나 ND1과 ND3는 TAG, ND4는 AGA였으며 Cytb는 종결코돈이 불완전 (T)하였다(Table 1). tRNA형태는 tRNA<sup>Ser(AGY)</sup>(lacked DHU-arm 구조)을 제외하고 모두 전형적인 cloverleaf 구조로 되어 있었다(Figure 2).

1	GCTAGCGTAG	CTTAATTAAA	GCATAACACT	' GAAGATGTTA	AGATGGACCC	' TAGAAAGTCC
61	CGCCCGCACA	AAGGCTTGGT	CCTGACTITA	CTATCAACTT	TAGCCAAATT	TACACATGCA
121	AGTATCCGCC	CCCCTGTGTGAG	AATGCCCTAC	AGCTCCCTGC	CCEGGAECAA	GGAGCTGGTA
181	TCAGGCACAC	ATCTGTAAGC	CCATGACACC	TTECTTAGCC		GGGAACTCAG
2/1	САСТСАТААА	САТТААВСССА	TAAGTGAAAA	CTTGACTTAG	ТТААЗССТАА	GAGGGCCGGT
201	AAACCTCITC	CARECALCE	CRECTATACE	ACACACCCAA	GTTGATACCA	TTCCCCCTAA
261	ACACTCOID	TUCADOCACCO	ACACTANACC	CCCACACCER	CAAACCTCTT.	ATACCONTCC
301 401	CARCICULA	13GAAAAIAA AGATCAACCA	CONCIARAGE	TTTNONACCOL		CAAACCHICC
441	GAAGGUIAGA	AGAILAALLA daleelaati	CUARJUIAU	ADTACAACUU.	A A COMMONDATION	ATATATCICIJ
481	BUAUAAA.IG	GOOTOOD NO	TACAGAGAC	LUIAGULLIA ACCETANANC	AAUUIIGBIA GONNOONOT	TRAIAICACA T
041 CO1	TAULUIGUUU	GULIBBBAAU	I ALGAGUAUU	AGUITAAAAU		
601	TIAGACUCUU		CIGITCIAGA	ACCGATAACC	CCCGIICAAC	CICACCUITC -
- 661 - 861	CITEITTAIC	CCGCCTATAT	ACCECCETCE	TCAGCITACC	CIGIGAAGGA	CTAAAAGTAA
721	GCACAACIGG	CACAACCCAA	AACGICAGGT	CGAGGTGTAG	CGCATGGAAG	GGGAAGAAAT
781	GGGCTACATT	CCCTACACTA	GGGAACACGA	ACGGCGCACT	GAAACACGCG	CCTGAAGGAG
841	GATTTAGTAG	TAAGCGGAAA	ATAGCGTGTT	CCGCTGAAAT	CGGCCCTGAA	GCGCGCACAC
901	ACCGCCCGTC	ACTCTCCCCA	AGCCTACCAC	TTTAAGTAAT	TAAAAACCCA	AAAATCGCAG
961	AGGGGAGGCA	AGTCGTAACA	TGGTAAG3GT	ACCGGAAGGT	GCACTTGGTA	ATATCAGAGT
1021	GTAGTTAAAA	TAGAATAACA	CTTCCCCTTAC	ACTGAAGAGA	CATCCGTGCA	AATCGGATCA
1081	CCCTGATGCC	CAACAGCTAG	CCCACAAACA	CAACAACAAC	CAACCATTAT	TTATAACCCC
1141	AAATGCACGA	ATGTTTCAAT	TAAACAAACC!	ATTTTTCCCCC	TTTAGTATGG	GCGACAGAAA
1201	AAGGACTTAG	GAGCAATAGA	GAAAGTACCG	CAGGGGATCG	CTGAAAGAGA	AATGAAACAA
1261	CCCAGTGAAG	CTAAGTAAAG	CAGAGATTTA	TTCTCGTACC	TTTTGCATCA	TGATTTAGCC
1321	AGCGTGACCC	AAGCAAAGAG	TGCTTTAGTT	TGACACCCCG	AAACTAGGGG	AGC'TAC'TC'C'A
1381	AGACAGCCTA	TTTATAGGGC	GAACCCGTCT	CTGTGGCAAA	AGAGTGGAAT	GAGCTTTGAG
1441	TAGAGGTGAT	AAACCTACCG	AACCTAGTTA	TAGCTGGTTG	CCCGAGAAAT	GGATAGAAGT
1501	TCAGCCTCTC	AGATTCTTTA	TTCACCTCAG	TATTACCCCA	CCTGATACCA	CAAGAAACTG
1561	TGAGAGTTAT	TCAAAGGGGG	TACAGCCCCT	TTGAAACAAG	ATACAACTTT	TOOGGGAGGA
1621	AAAAGATTAT	ΔΑΤΤΔΔΑΤΔΑ	AGGTAAGTAT	TTEGETEGEC	(TAAAAGCAG	CCACCCAGE
1691	AGAAAGATGAT		ТАСАТСАСТА		ТССТЕАТСАТ	TAATTOTTAO
1741	TCCCCCCTTC	COTACCERC	CATCCCATGC	ACACATESEA	GAGACCCTGC	TAATATGAGT
1801	AATAAGAGAG	CCAACCCTCT	CTCCTCCCAC	ACATCTAATT	CERANCEANC	CCCCACCAC
1861	CATTAACCC	CCCAAACCAA	GAGGGACCTG	AACAACAACC	CAAACAACCA	CAAAAAATT
1001	САТАТААА	COGTTAACCC	TACACAGGTG	TGCATTTAAG	GAAAGACTAA	GAGAAAGAGA
1001	ACCARCAIAAA	CONTRACCO	ARCHEROSIO	TETTACTAA	AAACATCCCC	
2041			COTROCTET	CACTATTACT.		
2041		COTACCICLU	TOACTTOTOT	TTTAAATTAA	CACCTOTATO	ANTERENCIA A
2161	CONCOLOURAN	ACTOTOTOTO	OTTTO A ACTOR	ATCAANTCC	A TOTOOOCOT	CONCINCICAL ARTICULAR
2101	COATATALA	ACIGICIUU	MANAGAGE IT	CONCERTING.	AILIUUUGI	BLAGAAGUGG Managara
2221	JORIAIAAAU	ATAAGAUGAG	AAGAGUUTAT	GGAGUIIIAG	ACALLAAAGA aamaamamama	RJAIUUIDIU TTaaTTaaaa
2281	AAGIAAUUUI	CIAIAAABBU.	UIGAAUIAAI Ngggggangga	GGAAUUUIIU	CURRENCE	ב בובותו בנתודו בובודו המתריק מריים
2341	ACCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC		AULUULAUGI GAATATGIIGA		CAULULUIUU .	IACAACIAAS Galega
2401	AGUUGUAGUI	CIAATIAACA	GAATAIUIGA	CUCAATAASAT	LUGGLAAIGU	CGAILAAUGI
Z401	AUUSABIIAU	CUTABBBBATA	ACAGUGUAAT	CUCULIIIIAG	ASCUCATATE	GAUAABBBBB
	TITAUGACUT	CJAIGIIGGA	TUAGGAUATU	CLAATEETEE	AGUUGUTATT	
Z581	TIGITCAACG	GITAAAGICC	TACGIGAICI	GAGIICAJAC	CEGAGIAAIC	
2641	TCTATCTATG	GIGIGCICIT	TTTTAGTACG	AAAGGACCGT	AAAIGAAIGAGIG	CCCCTGCTTT
2701	AAGCAAGCCT	CACCCCCACC	TAGTGAAAAC	AACTAAAGTA	GGCAAGAGGG	CATACCCCCA
2761	ATGCCTGAGA	GAACGGCATG	TTGGGGTGGC	AGAGCCCGGT	GAATGCAAAA	GACCTAAGCC
2821	CTTTTTACAG	AGGTTCAAGT	CCTCTCCTTA	ACTATGATTT	CAGTCCTTAT	TACCCATATT
2881	CTCAATCCCT	TGGCCTTCAT	TGTCCCCATC	CTCTTAGCCG	TEGECTTEET	CACACTTTTA
2941	GAACGTAAGG	TACTAGGGTA	TATACAACTA	CGAAAAGGCC	CAAATATTGT	AGGGCCTTAC
3001	GGACTATTAC	AGCCTATCGC	TGAC3GT3TG	AAGCTCTTTA	TTAAAGAGCC	CITTCICCCC
3061	TCCACTTCTT	CCCCAGTACT	ATTCCTCCTC	GCCCCACTAC	TCGCACTCAC	ACTTGCCTTA
3121	ACCCTTTJAG	CCCCTATACC	TCTCCCATAC	CCAGTAATTG	ACTTAAACCT	TGGAATCCTA
3181	TTTATTTTGG	CCCTATCAAG	CCTCGCTGTC	TACTCCATTC	TAGGCTCAGG	TTGAGCATCA
3241	AATTCAAAAT	ACGCTCTCAT	CGGGGGCCCTT	CGGGCTGTAG	CCCAAACCAT	TTCATATGAA
3301	GTTAGTCTAG	GCTTAATCCT	ATTAAGTACC	ATTATCTTTA	CAGGAGGTTT	CACACTACAA
3361	ACCTTCAACA	TTGCCCAAGA	AAGCGTCTGA	ATACTACTCC	CAGCCTGACC	ACTAGCCGCA
3421	ATGTGATATA	TTTCAACCCT	TGCAGAGACA	AACCGTGCAC	CETTTGACCT	TACTGAAGGC
3481	GAATCCGAAC	TAGTCTCTGG	${\rm CTTCAATGTC}$	GAGTATGCAG	GTGGCCCATT	CGCCCTATTT

Figure 1. Sequences of the mitogenome of S. hubbsi (GenBank Accession No. KJ525745).

OE 44	TTATAAAAA	3 3 T 3 T C C T 3 3	armanarm		TTACAT	der der en
3541	TICCIESCOS	AATAIGUTAA	CALICIECTI	ATAAACACGC	TITCUBUTAC	COLUMNA
3601	GGGGCCTCCC	ATTTTCCAAT	GCTACCTGAA	CTCACCGCAG	TGAACCTAAT	AACCAAAGCA
3661	GCCCTTCTGT	CCGTCTTATT	TCTATGAGTC	CGAGCCTCTT	ACCCACGATT	CCGGTACGAC
2721	CAACTTATAC	Α ΤΩΤΑ Α ΤΤΤΩ	AAAAAACTTC	CTCCCACTTA	CACTACCCT	COTTATCTCA
0701	CARCITATAC	TOCANTTOC	ATTTOCTO	CICCCACIIA CTROCALCIA	ACTINGUES	AACAACCOCT.
5/01	CAUCTAGOOL	LOCCAATIGE .	Alliaciado	CIGULAUUUU	AGUTATABAT COMPANYA	
3841	GCCTGAAGTA	AAGGGCCACT	TTGATAGAGT	GACTTAT3GG	GUITCAAATC	CCCCCCCGCTT
3901	CTTAAAAAAG	GGGGACTCGA	ACCCCGCCTA	AGGAGAGCAA	AACTCCTGGT	GCTCCCACTA
3961	CACTATTTCC	TAGTAAAGTC	AGCTAATTCT	AAGCTCTTGG	TCCCATACCC	CAAACACGAA
4021	GGTTAAAATC	COTCOTTAC	TAATGAACCC	ΤΤΑΓΙΔΤΓΙΤΤΑ	ACCRECCTRC	TATTTCCTAT
1001	TOOTTANATO	ACTACTACCA	COTTOCCARC	CTCCALCTIA	TTACTACCCT	CAATACCCCT
4001	CONTRACT.	ACTACTACCA	TOTICSUARS	ATTACTOR A	- IIAUIABUUI - araing	GARIAGGULI
4141	GGAAATAAAT	ACIULIGUCA	ICATICUSCI	AATAGUIJAA	CAUCATUACU	CCCGAGCAGI
4201	TGAAGCAGCC	ACAAAATATT	TCTTAATTCA	AGCTGCA3GA	GCAGCCATAC	TACTCTTTGC
4261	CAGCACCACC	AACGCTTGAT	TAAC'IGGAC'A	ATGGGACCTT	TTACAAATTG	CCCACCCTTT
4321	CCCAACTGCC	CTTGTCACTT	TGGCCCTCGC	ACTAAAAGTG	GGACTTGCAC	CTGTGCACTC
4381	ATRACTRCCT	GAAGTACTTC	AGGGCCTAGA	CICITAACICIACIA	GGACTTATTT	TGTCTACCTG
1001	GCALAACTT	GUCCCATTE	CCTTATTAGT	CCAASCCCCC	TETECCAACA	CCACCTTCT
4541		dayarayaar	CONTRINCT	Addyddaday	1010CCARCA	ACCARCECTICT
4501	AATCALLCIA	GGAUIUAUUI	CAACCAILGI	AGGAGGUIGA	GGAGGUUTTA	AUCAAACUCA
4551	ACTICUCAAA	AITCITECIT	ACTURTURAT	CUCACACUTA	GECTEAATAG	TAATTGTGCT
4621	ACAATTCTCC	CCCTCCTTGA	CTATTTAAC	ATTATTCACA	TACTTCATTA	TAACATTCTC
4681	AGCATTTCTT	ATGTTTAAAC	TTAATAAAGC'	AACCAACATT	AATACTCTAG	CAACCTCATG
4741	AGCAAAAACC	CCCGCCCTAA	CGGCCCTTGC	ACCCTTTTA	TTATTATCCT	TAGGGGGACT
4801	CCCACCACTT	ACAGGCTTTA	TGCCAAATG	ACTTATCOTT	CAAGAACTTA	CITAAGCIAAGA
4001	CCCRCCRCII	COTOCITIA	TACCACCANT	Addadate	CITCIACCOUTAT	ATTTTTACC
4001	NGRNGENTER	TACIDUA AURO		THELEUCIC	UICAGUUTAT	ATTITACCI
49Z1	ACGACIAICA	TACGUAATGG	CATTAACTAT	TICACCCAAC	AACCICACCE	CAAIIICCUU
4981	ATGACGCCTC	CCCTCCTTAC	AACTAACACT	GCCACTTGCT	ACCTCAGCCA	TAGCTACGCT
5041	ACTGCTTCTA	CCCCTAACAC	CCGCCGCAAT	AGCACTAATA	ACCCTTTAAG	GGACTTAGGT
5101	TAAAACAAGA	CCAAGGGCCT	TCAAAGCCCT	AAGTGAGGGT	GGAAGTCCCC	CAGTCCCTGA
5161	TAAGGCTTGC	GGBACACTAC	CCCACATOTO	CTGTATGCAA	AACAGGTACT	TTAATTAAGC
5221	TAAACCCTTC	CTACAACCC	ACCONTRAT	CCTCCAACAT	CTTACTTAC	AGCTAACCCC
5001	TRANCETTE	CIACATCONC.		CORPORADAL		
0201	ILAAALLAGU NNAMMOQQQQ	GABUAIUUAI	UTALLITIUL NOOMOONNOOM	TOLUGULIAAA		
5341	AAAGTCCCGG	CAGACGACTA	ACCIGCATCT	TCAGATTIGC	AATUTGATAT	GIATAACACC
5401	TCAAGACTTC	TGGTAAGAAG	AGGATTCAAA	CCTCTGTTTG	TGGGGCTACA	ATCCATCGCT
5461	TAGAAACTCA	GCCATCCTAC	CTGTGGCCAT	CACACGTTGA	TTTTTTCTCCA	CTAATCACAA
5521	AGACATC3GC	ACCCTTTATC	TAGTATTTGG	TGCCTGAGCC	GGTATAGTAG	GCACAGCCCT
5581	CARCETACTC	ATTCGAGCAG	AACTAAGCCA	ACCEGGECCE	CTECTTGGAG	ACGACICIAAAT
5641	TTATAATTA	ATCGTTACCC	CACATGCCTT	CGTAATGATT	TTOTTTATAG	TAATGCCAAT
5701	TATATOTA	COTTTTCCAA	ACTCATTAAT	TOCOTTANTO	ATTOCACCCC	CACATATACC
	ATTAALUJUA	ADA ADA ADA	ACIGATIAAI MAAGEMEAND	ACCULTARIG		
5701	ATTICUTUGI	ATAAATAATA	TAAGITICIG	ACTICITCC	CONTENTIO	IJUTACIACI
5821	TGCCTCCTCT	GGAGTAGAGG	CAGGCGCCGG	GACCGGCTGA	ACAGTGTACC	CGCCCCTGGC
5881	TGGCAACTTA	GCCCACGCAG	GAGCATCAGT	CGACCTAACA	ATCTTTTCAC	TCCACTTAGC
5941	AGGTATCTCC	TCAATTCTTG	GGGCAATCAA	TTTTATTACC	ACAATTATTA	ATATGAAGCC
6001	TCCGGCCATC	TCCCAATACC	AGACGCCCCT	ATTTGTGTGA	GCCGTCTTAA	TTACCGCTGT
6061	TOTTOTOT	CTCTCTCTAC	CAGTTOTOGO	TGCAGGTATT	ATAATGCTCC	TTACEGACCE
6121	AACCTTAAC	ACCACCTTCT	Traverse		GATCCAATCC	TTTACCAACA
G120 G101				ATATATTTT		COTTROCASCA
DIOI O	CUTATICIGG	111111GGGC	AUUUSGASGI	ATATATIST	ALIUIGUUIG	GUIIIGGIAI
6241	GATTICACAT	ATTGICGCCT	ATTACTUIGG	CAAAAAA.JAA	CCCTTTGGAT	ATATAGGCAT
6301	GGTCTGAGCA	ATAATGGCTA	TTGGCCTCCT	AGGATTTATT	GTATGGGCTC	ATCACATATT
6361	CACAGTTGGT	ATGGACGTAG	ACACACGTGC	TTACTTTACG	TCTGCCACAA	TAATCATCGC
6421	AATTCCCACC	GGTGTTAAAG	TATTTAGCTG	ACTTGCAACT	CTACATGGGG	GCTCTATTAA
6481	ATGAGAGACA	CCCCTTTTAT	GGGCCCTTGG	CTTTATCTTC	CTATTACAG	TAGGCGGGCT
6541	TACACCTATT	CTTCTCCCCA	ATTCATCTCT	ACATATTTA	CTACACCATA	CATACTATCT
0041	JACAGGIAII	TTCLUCTIO	TATICATURE	AGAIAIIJIA GGGGGGGGTA	TTTOGOATA	TOOOCOTT
0001	AGIAGUUUAU	TICCACIACS:	TAUTAIUTAI	totototototititi A	A GOLGATIG	
6661	CGIGCACIGA	TICCCGCTAT	ICICAGGATA	TACACTCCAC	ASCAUTTGGA	CAAAAATCCA
6721	TTTCGGCATT	ATGTTCTTGG	GGGTAAACTT	AACCTTCTTC	CCACAACACT	TCCTCGGATT
6781	AGCCGGAATG	CCCCGACGAT	ACTCCGACTA	CCCTGACGCC	TATACCCTAT	GAAATACAGT
6841	CTCCTCAATC	GGATCACTTA	TCTCGTTAGT	GGCTGTTATT	ATGTTCTTAT	TTATTATTTG
6901	AGAGGCATTC	GCCGCCAAAC	GTGAAGTTCT	AGCAACAGAT	ТТААСААСАА	CCAATGTAGA
6061		COCCUTECCIERCE COCCUTEC	CCCCCTACCA	CACATTCIAC	GARCETRICE	TTGTACAAGT
70901	ACANCOCAL	TAATCACAL	COULCELACIA	CACALLOAD		CANCERCANCE
7071	ACAAGLAGAC		DODADDADIC	GAAULUAII	Rota Lutra III	CAAGUUGAUG
7081	ACATAACCGC	TCTGCCACTT	ICTITATAAG	ACACTAGTAA	AAGAGAACAT	TACAUCGCCT

(Figure 1. Continued)

7141	TGTCAAGGCG	GAAGTGTGGG	TTAGACCCCC	GCGTGTCTTG	CTTTTAATGG	CCCATCCATC
7201	ACAGCTT3GA	TTTCAAGATG	CAGCTTCACC	TGTTATAAAA	GAACTTCTTC	ATTTTCACGA
7261	CCATGCTTTA	ATAATCGTCT	TCCTGATTAG	CACACTTGTG	CTTTACATTA	TTCTTGCTAT
7321	AGTTACCACT	AAATTAACGA	ΑΓΆΑΑΤΑΤΑΤ	TTTAGATTCA	CAAGAAATTG	AAATTATCTG
7391	AACAATTOTO	CTACCTATTA	TTTTAATTCT	AATTICCACTC		COMPCCTCTA
7301	COTTATION	CANATTANCA	ACCCCCTATT	ANCIATTAN	COCOUNTRACTO	AGGAICCICIA
7441		GAAAIIAALA GAAMAMA GMG	ACCULUTATI	AAUAAIIAAA		A.LAAIGAIA Maagaala
7501	CIGAAGUIAI	GAATATACIG	ACTACGAAGA	CULIGGUIII	GATICATACA	TAATUUUUAJ
7561	CCAAGACCTA	ACCCCTGGAC	AATTCCGCCT	ATTAGAAGCT	GACCATCGCA	TGGTTATCCC
7621	AGTTGAATCC	CCGATCCGAG	TTTTAGTATC	TGCAGATGAT	GTACTCCACT	CATGGGCAGT
7681	CCCAGCCCTA	GGGGTAAAAA	TGGACGCAGT	ACCAGGACGC	CTTAATCAAA	CAGCCTTCAT
7741	CGCATCCCGA	CCAGGCGTAT	TCTACGGACA	ATGCTCTGAA	ATCTGCGGAG	CAAATCACAG
7801	CTTTATGCCT	ATTGTAGTGG	AAGCAGTTCC	CCTAGAACAC	TTTGAAAACT	GATCATCTCG
7861	AATACTTGAA	GACGCCTCGC	TAGGAAGCTA	AATAGGGTAT	AGCGTTAGCC	TTTTAAGCTA
7921	AAGATTGGTG	GETETETE ACC	ACCENTAACE	ACATGCCCCCA	ACTEVACECCE	GCACCTTGAT
7981	TTGCTATTTT	AGTOTTOTOG	TGAATGGTCT	TCCTGGCCGT	TATTCTCGCT	AAAGTTACAG
20/11 	CCCACACCTT	СССУУУСТССС	CCTACTCTCC		λλλασστλλλ	AGAGACCCCT
0041	CANCETERACE	ATEACACTAA	COTTTTTCA	CCACTTTATA		ATCTTCCAAT
01.01	CHARLINGAUL	AIGALALIAA aqaammadaa	TTACATION		TABULULAULI TAGGGGAGAGAG	ATCIIGGAAL ATVIIGGAAL
0101	UCCATTANTA	BULLIIGUUL	TIALULIAUL	UIGACIUCII NGOTTONTTT	I AUUUUAUAU	UTACAACIUS
ŏZZ1	AIGAIIAAAI	AACCGATICC	TCTCGCTTCA	AGGIIGAIII	ATTAACCGII	TTACTCAACA
8281	GCTTCTCCTC	CCCCTAAATA	TGGGAGGICA	CAAGTGAGCC	GCCCTCCTAA	CITCATTAAT
8341	GATCTTTTTA	ATTTCCCTAA	ATATATTAGG	ACTTCTTCCC	TATACTTTTA	CCCACACCAC
8401	CCAATTATCA	CTTAATTTAG	GGCTCGCGGT	ACCTCTCTGA	TTAGCAACTG	TTATTATTGG
8461	CATECGAAAC	CAACCAACCC	ACGCCCTAGG	ACACCTCCTA	CCAGAAGGCA	CACCCGGCCC
8521	CCTTATTCCA	GTGCTTATTA	TTATCGAAAC	AATTAGTCTC	TITATICGCC	CCCTTGCCCT
8581	AGGGGTACGG	CTAACAGCCA	ATTTAACAGC	CGGTCACCTA	TTAATTCAAT	TAATTGCTAC
8641	AGGTGCGTTC	GTACTTCTTC	CCCTAATACC	AACCGTCGCA	TTCATCACAA	CAACAGTATT
8701	GGTTCTCCTT		AAGTTGCCGT	AGCIAATAATT	CARCCTACE	TOTTOGTTOT
0761	COTICICOTI	CTETACCTAC	AAGAAAACCT	CTANTECCCC		COOTTACCAC
0701	ATACTACIANCE			CLARIGUEL		COULTACCAC
0021	AIABIIGAUU Gaggmagaaa	COROLUCIIS	RUCULIANCA TTTTTCCCTCA		TRACETAC	DAIAAUAIUA AAGAGGTGTA
0001	GUULIUUUA	CUIGATIICA	ITTILGUICA	ACAAULIIAA	TAAUUIIAGG	AACAGUIUIA
8941	CIGCITCITA	CAATATATCA	ATGATGACGA	GATATCGTAC	GAGAAGGCAC	ATTCCAAGGA
9001	CATCACACGC	CCCCCGTACA	AAAAGGTCTT	CGATACGGAA	TAATTCTTTT	CATTACCTCC
9061	GAAGTATTCT	TTTTCCTAGG	GTTCTTCTGA	GCCTTTTACC	ACGCAAGCCT	CGCTCCCACT
9121	CCTGAGTTAG	GGGGCTGCTG	GCCCCCCACG	GGCATTACAA	CCCTTGACCC	ATTTGAAGTC
9181	CCCCTCCTTA	ATACAGCTGT	CCTGCTTGCC	TCCGGGGTAA	CAGTTACCTG	GGCCCACCAC
9241	AGCATTATGG	AAGGTGAACG	AAAACAGACC	ATTCAATCAC	TAGCCTTAAC	TATTCTTCTA
9301	GGCTTTTACT	TTACATTTCT	TCAAGCCTTG	GAATACTATG	AAGCCCCCTT	TACAATTGCA
9361	GACGGCGTAT	ACEGCTCTAC	ATTTTTCGTA	GCCACCGGAT	TECACGGACT	ACACGTTATT
9421	ATTEGCTCCA	CATTTTTAGC	CGTTTGCCTC	CTACGGCAAA	TCCAATACCA	TTTTACATCC
Q191	GARCACCATT	TUBBERTUGA	AGCAGCTGCC	TGATACTGAC	ATTTCGTAGA	CETTETCTEN
0541	TTATTOCTAT	ATATOTOTAT	CTACTCATCA	CCCTCTTAAT	CTTTCTACTA	TTANACTAC
0601	TINICCINI	TTCCATCAC	CINCIGNION	TTANATOON		ATCANCING
9001		ANTTACOATCAC	ACTATTTCC	TIAAAAIA	AJUARAUAIA COTOCONTT	ATGAAUGIAS OTATOOTTOT
9001		AATTALUATU.	ACTATITICU.		UCTOBLUATT.	GIATUUTUT.
9721	GAUTUUUUUA	AAIGAUUUUU	GACCACGAAA	AGUIUIUUUU	ATAIGAAIGI	GETTTUGAUG
9781	CITTAGGATC	AGCCCGCCTA	CCATTITCCC	TCCGCTTCTT	CETAGTEGEC	ATTCTCTTCC
9841	TACTTTTCGA	TTTAGAAATT	GCCCTTCTCC	TCCCCCTCCC	ATGAGGAGAC	CAATTAACCT
9901	CCCCCTTTATT	GACACTCTTC	TGAGCCGTAG	CCGTGCTTAT	TETTETCACC	CTTGGCTTAG
9961	TCTACGAGTG	AATTCAAGGA	GGGTTAGAAT	GAGCCGAATA	GCCAATTAGT	ТТАААААААА
10021	TATTTGATTT	CGGCTCAAAA	GCTTATGGTT	AAAGTCCATA	ATTGTCTAAT	GACTCCCGCT
10081	CACTTCGCTT	TETCATEGGE	CTTTACTCTA	GGATTGACAG	GCCTAGCATT	CCATCGAACC
10141	CACCTCCTCT	CCGCTCTTTT	ATGCTTAGAA	GGGATGATGC	TCTCTTTATT	TATTGGACTT
10201	TCAATTTGAA	CCCTTCAACT	GGGCTCCACA	AGTTTCTCTG	CGGCTCCTAT	GETTETETA
10261	GCTTTTTTAG	CTTGTGAAGC	AAGCGCGCGGGG	CTTECCUTEC	TGGTAGCCAC	AGCTCGCACA
10221	CATEGOTORE	AUGGOULTON	GACCTTAAAC	СТСТТАСААТ	GOTAAAATO	CTAATTOOCA
10201		TOTOCOLICA	COOTCOLLARD	COCOTECCAN	ATCATTCTCA	COTACTACTO
10744	TATAAAAAAA	COTACTOR		CCCCTGCCAA adamaxaama	AUTALIUIUA	ACTRUIAULU ACATCCAAAA
10501	- ICIUULAJAG	TROUCTABLE	CONTRACTOR	CONCINCICALLIS	ACTARARARI COTOTONIO	ACAILUSAAA degerrerre
LUDUL	CAGGCIGGIC		CCULTUATAG	TTTTTOOON	CULTUICAACC	LULUIIUIIJ
10561	TCUTCAUCTG	CIGACTACIC	CCCCTTATGA	TITGGCAAG	CCAAAACCAC	ACAGUACTUG
10621	AACCTATTAA	CCGCCAACGA	ACCTACATTA	GCCTATTGAC	ATCACTGCAA	GIGITCCTTA

(Figure 1. Continued)

10001	mm3 m3 (0 (13 mm	maamaaaaaaa	0110T10TD1	<b>mamma</b> and	alminment.	
10981	TTATAGCATT	LEGIECCACI	GAACIACIIA	TATTTAIGI	CATATITIGAA	GCTAUTCITA
10741	TCCCCACGCT	AATTATTATT	ACTCGCTGAG	GTAACCAAGC	AGAACGTCTT	AATGCAGGAG
10201	TATATTTTT	CTTTTATACC	CTCCCACCCT	OTOTOCOUTT.	ACTACTOCCO	CTATTACTTC
10001	TTCLCLARA	TALAGGERGE		TANCOUNT	ATATOICOCC	TOTACIOCIC
LUSDI	ТТСАБААВБА	TACABBUTCE	ULLICUL	TAAULATULA	АТАТЬССАВС	TOTACULU
10921	TTTCATCTTA	TGCTGATAAG	ATTTGATGGG	CAGGCTGCCT	GATTGCATTT	TTAGTAAAAA
10981	TACCCCTATA	CGGAGCACAT	CTATGGCTGC	CAAAAGCACA	TGTAGAAGC'C'	CCAGTTGCAG.
11041	COTCAATACT	COTCOTCOL	CTTCTTCT AA	AACTACCCC	CTACCCATA	λΤΟΟΓΆΛΤΑΛ
4.1.1.01	TOLICHAINOI	da Macrata	AGEARAGANE	TACINOUUU	ATTANTA	areadaarar
TTIUT	IGACCATATI	GGAACUICIC	ACTARAGART	TAAGITAICC	ALLIALIAIC	CICECCICI
11161	GAGGTGTAAT	TATAACTGGC	TCCACCTGCC	TTCGCCAAAC	AGATCTTAAA	TCCCTCATCG
11221	CTTACTCATC	CGTAAGCCAC	ATGGGCCTAG	TCGTTGGAGG	TATTCTTATC	CAAACACCAT
11281	CACCOUNTER	CERCECCETA	ΔΤΤΩΤΩΛΤΑΛ	TTCCACACCC	COTANCATOO	TARCOTTT
	Taraar	COOCOCCOIN	TICICAIAA	TICCACACOO		TCAUCCUIII
11.341	TCTGCTTGGC	CAACACAAAT	TACHAACHCU	TCCATAGECG	GACAATACTA	TTAGUTUGAG
11401	GGTTACAAAT	AGTACTTCCA	CTCATAGCAA	CATGATGATT	TATTGCCAGC	CTCGCAAACT
11461	TAGCCCTTCC	CCCCCTGCCC	AAC'C'TC'ATAG	GGGAACTTTT	AATTATTACC	TCATTATTTG.
11571	CTTCATCTC	A TRA ACCOTC	GTACTCACAC	CCCCACCAAC	COTCATTACC	CCAACCTATT
44504	- OTTOATCOTO - OTTOATCOTO	ATTACCOLL	Jagalagaa	- addalamadad	AGGICATIACC	ATTACIOTA A
11281	CACICIAIAI	ALICUICAIG	ACCUAGUSGG	GUUUAUTUUU	ASCACATATI	ATTAGUUTAA
11641	ACCCTTCCTA	CACGCGGGGAG	CACCTAGTCA	TAGCCCTTCA	CETCETCCC	CTGCTTCTAC
11701	TTGTCTTAAA	GCCCGAACTA	GTATGAGGCT	GAACCACCTG	TAGATATAGT	ТТААСААААА
11761	TATTACATTC	ΤΊΛΤΤΩΤΑΛΑ	CACARACTT	AAATCCCC	татесассаа	CARACCTCC
11001	TULINOVIIO	JANGTAAN	TATAAAAAAAA	Ammaammaal	addaladaa	CACADOCICU.
TTQT	UUABUAAUBA	ASAUTSUTAA	ICICCELEAC	CIIGGIIGA	CCCCAGGGCI	CACIUGGUUI
11881	GCTCCTAAAG	GATAACAGCT	CATCCATTGG	TCTTAGGAAC	CAAAAACTCT	TGGTGCAAAT
11941	CCAAGTAGCA	GCTATGCACT.	CCTCATCACT	TATTATATCA	TCCAGCTTAG.	TCATTATCTT
12001	TTTACTTA	CONTATCOTA	TTTTTACIONC	TOTACACOOT	COCCETCICAA	ACCCCACTO
100C1		- CORTAICCIA			TTATATA	TABLECCORCID TAATTOOOOD
TZOOT	GUUGIUIUA	CAIGIIAAGA	CAGCAGICGC	CUTAGUITTU	TICATCAGUE	TRAFICCULT
12121	ATTCCTGTTT	CTTAACGAAG	GAGCAGAAAC	AATCATCACC	TCATGAAATT	GAATAAACAC
12181	ACTGACCTTC	GACGTGAACA	TTAGTTTCAA	ATTTGACCAC	TACTCAGTTA	TETTTGTACE
12241	TATTGCCCTC	TATGTCIACCT	GATECATTET	AGAATTTGCA	THATGATATA	TACACIGCIAGA
10001	COUNTACATA	ATCOATTT	TTANATATOT	COTATTTC	CTTATTOCCA	TATTATTOT
17.301	CUCATACATA	AATUGATITT	TIAAATAICI	CUTAATTIC.	CLIATIGULA	
12361	TGITACAGCA	AATAATCTAT	TCCAGCTITT	CATTGGTTGG	GAAGGAGTAG	GCATTATGTC
12421	ATTTCTTCTC	ATTGGCTGAT	GGTACGGACG	AGCAGATGCC	AACACAGCGG	CCCTCCAGGC
12481	CGTTGTGTAT	AACCGGGTCG	GAGACATTGG	ACTECTATTC	ACAATAGCAT	GAATAGCAAC
10541	CANCECTANC	TOOTONONOO	TACAACAAAT	TTTTCTCCA	ACANANCIACC	TACATOTOAC
		AILA GOOGEON	TACAACAAAI	TITIGIGGCA	MORAGOO M	TAGAICIGAS
12601	CCTACCGCTA	CTAGGCCTGA	TTATEGEEGGE	TACAGGTAAA	TCTGCCCAAT	TIGGICICCA
12661	CCCTTGACTC	CCCTCTGCTA	TAGAGGGTCC	TACACCGGTC	TETGECETAC	TGCATTCAAG
12721	CACAATAGTC	GTCGCCGGTA	TTTTTCTTTT	AGTGCGAACA	AGCCACTCC	TAGAAAATAA
12721	TCAAACT2CC	CTCACCACCT	COCTATECCT	ACCACCCTA	ACIAACIACITAT	TCACACCCAC
10041		ACCALCT	ATTATOLOGI	AUDROCCCIA		ANATANAT.
12841	CIGIGCUCIG	ACCUAAAAIG	ATALCAASAA	AATIGTASCA	TICICCACAI	CAAGICAACI
12901	TGGCCTAATA	ATAGTCACTA	TTGGACTAAA	TCAGCCCCAA	TTAGCCTTCC	TACACATITIG
12961	CACCCACCCC	TTCTTCAAGG	CAATATTATT	CCTCTGTTCT	GGCTCAATTA	TTCACAGTCT
13021	сарсізар	с а адататес.	GAAAAATAGG	AGGCIATACIAT	CACCTTACCC	COTTTACATO
10001	CTCCTCCCTC	ACTATTCCTA	CTTTACCCCT	CACACCCACC	COOTTOOTOO	CACCETTCTT
TOADT		ACIALIGUIA	GITTAGULUT	CALAGGUAUU	CULTICUIGG	CAGGGIICII
13141	CTCCAAAGAT	GCCATCATTG	AAGCACTAAA	CACATCCCAC	CTAAACGCCT	GAGCCCTAGT
13201	CCTAACCETT	CTAGCCACCT	CATTCACGGC	CATCTACAGT	CTCCGCGTAG	TATTCTTTGT
13261	CTCTATEGGC	TACCCACGAT	TTAACCCTAT	TTCTCCCATC	AATGAGAAC'A	ACCCARCART
12201	TATTAACCCC	TTANAACCAC	TTECATEACE	ANCONTRATO	CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	TAATCATTTC
1.3.341	ATTARUACU Argannaga		IIGUAIDAGU	ARGUALISIU		IAAICAIIIC.
13381	AAGCATTACC	CCCCTCAAGA	CCCCIGIJAI	ATCUATACCC	CCUTTAUTCA	AATTAGCIGC
13441	CCTCGGAGTT	ACAATTACAG	GATTATTCAT	TGCCCTCGAG	CTGGCAACAT	TGACCAATAA
13501	ACAGTACAAA	ATCACCCCTA	ATCTAGTTAC	TCACCACTTC	TCCAACATAC	TAGGCTTTTT
13561	CCCCTCCATC	ATTCACCGAT	TTATACCOAA	ΔΟΤΔΔΔΤΟΤΔ	GTTTTAGGAC	ABCACTTON
10/01	- CCCCICOAIC	ATTCACCOAL	ammalamlal	ALLCINARIO A	dodryyddyy	TOUCHCIICC
13021	CAGUCAACIA	ATTGACCAAA	CITGACTAGA	GAAAGIIJGU	UJUAAAGUAA	ICIUIICAIC
13681	AAATATTCCC	TIGATTACAA	CAACAAGTAA	TACCCAACAA	GGAATAATTA	AAACATATCT
13741	CACCCTATTC	CTTCTCACCT	TAACCCTCGC	TGCCTTATTA	TTTACCCGTT	AAACTGCCCG
13801	AAGAGEEEEE	CGACTTAGTC	CACGAGTTAA	CTECAACIACIA	ACAAATAAGG	таассассаа
12041	A A CCC A CCCCC	CTANCTACCA	CTATACCICC	COOTANTON	TACIATTA ACC	CAACCCCCCC
10001	AAUUUAUSUA	CIAAGIAUCA	GIAIAUUIUU MAMGEMENEM	CUCTAAISAA	TACATTAACG	CAACUUUIUU
13921	AATATCGCCT	CUCAATACAG	AGAGCTCACT	AAGUTCATCA	GCCGGCACCC	AIGAGGACTC
13981	ATATCACCCC	CCTCAAAATA	CACTAGAAGC	CACCCCCACC	CCTACTAGGT	ACATCAACAT
14041	ATCACCTACA	ACAGGACCAC	TTACCCAACT	CTCCGGGTAC	GGCTCAGCGG	CGAGCGCCGC
1/101	CIGALATATICA	<u>886'86'886'T8</u>	ATATACCACC	CAGATAAATC	00000C100C10	CTAATGATAG
1 / 1 / 2 /	- CUARTALUCA - XXXX007000		CONTRACTOR			OTACTA CONTROL
14101	RAAAGGIJJCC	CUAIGACUIA	CCAATAUTCU	GUAUUUUAIG	CUUBUUAUAA	CIACIAACUU

(Figure 1. Continued)

14221	CAAGGCAGCA	AAGTAGAGAG	AAGGGTTGCA	GGCCTTG3CA	CCCCCCCCA	GAACTAATCC
14281	AATTAAAAAT	AAAGACATAA	TGTAAATCAT	AATTCCTGCC	GGGACTTTAA	CCAAAACCAA
14341	TGGCGTGAAA	AACCCCCGTA	GTTATTCAAC	GACAACAATC	CTGTAATGGC	AAGTCTACGA
14401	AAGACACACC	CTCTCCTCAA	AATCGCAAAC	AATGCCCTAG	TTGACCTACC	CGCCCCCTCA
14461	GATATTTCAG	TGTGATGAAA	CTTCGGATCT	CTCTTAGGAC	TCTGCTTAAT	TATTCAAATC
14521	CTCACAGGAC	TATTTC'TAGC'	CATACACTAC	ACCTCTGATA	TTGCGACAGC	TTTTTCTCC
14581	GTCGCCCATA	TTTGCCGAGA	CGTAAATTAC	GGCTGACTCA	TCCGAAACCT	TCACGCCAAC
14641	GGCGCATCTT	TCTTCTTCGT	ATGTATTTAT	GCCCACATTG	GCCGCGGGACT	TTACTACGGC
14701	TCATACCTCT	ATAAAGAGAC	ATGAAACATC	GGAGTAGTCC	TGCTACTTCT	AGTTATAATA
14761	ACTGCTTTCG	TCGGCTATGT	TCTACCCTGA	GGCCAAATGT	CCTTTTGAGG	TGCAACCGTT
14821	ATTACCAACC	TACTCTCTGC	AGTGCCCTAC	GTAGGTG3CT	CTTTAGTTCA	ATGAATTTGA
14881	GGCGGATTCT	CAGTAGACAA	TGCAACCCTT	ACCCGATTCT	TTGC/TTTCC/A	CTTCCTATTC
14941	CCCTTTGTAA	TTGCAGGCGC	AACCATAGTC	CACCTTCTCT	TECTTEATEA	AAC'AGGGTC'A
15001	AATAATCCCC	TCGGCCTAAA	TTCAGACGCA	GATAAAATAA	GCTTTCACCC	CTACTTCTCA
15061	TACAAAGACT	TATTAGGGTT	TGCAGTACTT	GTTATCGCCC	TCACATCTTT	AGCTTTATTC
15121	TCACCCAACC	TGCTAGGAGA	CCCAGACAAC	TTCACCCCAG	CCAATCCGCT	AGTTACTCCT
15181	CCCCACATCA	AACCAGAGTG	ATACTTCCTG	TTCGCATATG	CAATTCTGCG	CTCCATCCCC
15241	AACAAACTAG	GAGGAGTATT	GGCCCTACTG	GCCTCAATCC	TCATTCTGAT	GCTCGTACCA
15301	TTTCTACACA	CGTCTAAACA	ACGAAGTCTT	ACCTTCC3GC	CGCTTACACA	ATTCTTGTTT
15361	TGAACCCTAA	TTGCAGATGT	TATTATTCTC	ACCTGAATTG	GAGGTATGCC	TGTATCACAC
15421	CCGTTTGTTA	TTATTGGACA	AATTGCGTCC	TTTTTTATACT	TTTTCCTCTT	CCTAGTCCTT
15481	ACACCATTAG	CAGGTTATGC	AGAGGATAAA	GCACTTGAAT	GAACTTGCAT	TAGTAGCTCA
15541	GCGTCAGAGC	CCTGGTCTTG	TAAACCAGAT	GTCGGGGGTT	AAAATCCTCC	CTACTGCTCA
15601	AAGAAAGGAG	ATTTTAACTC	CCACCCCTGG	CTCCCAAAGC	CAGGATTCTT	ААТТАААСТА
15661	TTCTTTGTAG	TGTATGTACA	ATAATTTTAA	ATACATATAT	GTATTATCAA	CATTAATTTA
15721	TATTAACCAT	ATCATAGGGC	ATTCAAGTAC	ATATATGTTT	TATCACCATA	TCTAGGGTTA
15781	CACCATTCAG	GAATTACATG	ATACGAAAAT	TTTACATAAA	GCATGATAAT	AATAATAAC
15841	AAGTACTTAT	AAACACCAGG	CGAAATTTAA	GACCTAACAC	AAAAATCCAT	AGGTTAAGTT
15901	ATACCTTTAC	TCAAAATCCC	GCCAAACTCA	AATATTTAAT	GTAGTAAGAG	CCGACCAACA
15961	AGTCCATTTC	TTAATGCCAA	AGTTTATTGA	AGGTGAG3GA	CAAAAATTGT	GGGGGTTTCA
16021	CAGAGTGATT	TATTCCTGGC	ATTTGGTTCC	TATTTCAGGG	CCACAAATTG	TAAACATCCC
16081	CATAAATTAT	TTCACCGGGC	ATAGGTTAAT	GGTGGAATAC	AATAGGGGGA	GGGGCCCCCA
16141	TGCCGAGCCT	TCTTTTTATA	GGGCAATGGG	GTTTTTTTTT	TTCTTTTTCT	TTTCAATAGA
16201	CATTTCACAG	TGCACGAAAT	CTAGTTAACA	AGGTGGGAAT	CATCCTAGGA	AGCATGGAAA
16261	TAGTATGAGT	GATGAAAAGT	CTTTACTAAA	GAATTACATA	TAAAAATCTC	AAGGACATAA
16321	TATAGTGAAA	TTTAGTCGGA	AGATATCTAT	ATGACCCCCT	TTTGGCTTTT	TCGCGTTAAA
16381	CCCCCCTACC	CCCCTAAACT	CCTGAGATAG	CTAACGCTCC	TGTAAACCCC	CCGGAAACAG
16441	GAAAACCTCG	AGT				

(Figure 1. Continued)

Table 1. Summary of the mitochondrial genome of S. hubbsi

\*H and L denote heavy and light strands, respectively.



Figure 2. A gene arrangement in the complete mitogenome Sebastes hubbsi.

횐꼬리볼락의 미토게놈은 13개의 단백질 암호화 영역, 2개의 리보좀, 22개의 tRNA 및 1개의 조절영역으로 이루어져 있었고, 전체 크기는 16,445 bp 이었다(GenBank Accession number KJ834061)(Figure 3). 흰꼬리볼락의 미토콘드리아 DNA의 염기 조성비는 28.0% (A), 26.7% (T), 17.1% (G) 및 28.3% (C)이었고, AT 비율이 약간 높았으며(54.6%), COX1 (GTG)을 제외한 모든 단백질 암호영역의 개시코돈은 ATG 이었고, 종결코돈은 TAA가 주를 이루었으나 ND1과 ND3는 TAG, ND4는 AGA였으며 Cytb는 종결코돈이 불완전(T)하였다(Table 2). tRNA형태는 tRNA<sup>Ser(AGY)</sup>(lacked DHU-arm 구조)을 제외하고 모두 전형적인 cloverleaf 구조로 되어 있었다(Figure 4).

1	GCTAGCGTAG	CTTAATTAAA	GC'ATAAC'AC'T	GAAGATGTTA	AGATGGGCCC	TAGAAAGCCC
61	CGCCCGCACA	AAGGCTTGGT	CCTGACTITA	CTATCAACTT	TAGCCAAATT	TACACATGCA
121	AGTATCCGCC	CCCCTGTGAG	AATGCCCTAC	AGCTCCCTGC	CCGGGAGCAA	GGAGCTGGTA
181	TCAGGCACAC	ATCTGTAAGC	CCATGACACC	TTGCTTAGCC	ACACCCTCAA	GEGAACTCAE
241	CAGTGATAAA	CATTAAGCCA	TAAGTGAAAA	CTTGACTTAG	TTAAAGCTAA	GAGGGCCGGT
301	AAACCTCGTG		CGGCTATACG	AGAGACICITAA	GTTGATACCA	TTCGGCGTAA
361	AGAGTGGTTA	TICADATAA	AGACTAAAGC	CREACACOTT	CAAAGCTGTT	ATACIGCATCO
/21	GAAGGCTAGA	AGATCAACCA	CGAAGGTAGC	TTTACAACCC	CTGACCCCAC	GAAAGCTCTC
181	GUACIANATTE	GIATTAGATA	CCCCACTATE		AACCTTEGTA	
5/1	TACCORDANCIO	CONTRACAIN	TACCACINIC	ACCTTANAAC	CONANCONCE	TARACACA
- 5 <u>41</u> - 601	TTACACCOCC	CTACACCACC	CTATICTARA	ACCENTANCE	CORRECTOR	CTCACCCTTC
661	CTTCTTTATC	CTROROOROC	ACCEPTEINCE	TCACCTTACC	СТАТАЛАСАЛ	СТАЛЛИТАЛ
701	CITCITATC	CACAACCCAA	ACCOLCUICG	CONCEPTION	CIGIGARGGA	CIARAAGIAA
701	CCCCCTACATT		CCCAACACCA	ACCOCCICIAN	CALAIGGAAG	COTONNOCACI
0/1	CATTACALL	TARGERALIA	ATACCOTOTT		CARACAUGUG	CULUARUUAU
041	JOGGOCOCOTO	I AAGUGGAAA	ALAGUGIGII	TTTANT	TAAAACCCA	ANATCOCAC
901	ACCOULT	ACTUIUUUA	TAGULIAUSAU	ACCCC ACCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC		ARAAIUGUAS ATATCACACT
1001	AJUUUUUAUUUA	ASILGIAACA		ACCEBAAJEI	GLACIIGBIA	ATAICAGAGI
1021	GIAGIIAAAA	I ALFAATAATA GI AGUAGETAG	CITUCUTIAU	AUTISAAGAGA	CATCUSTECA	AAIUSSAIUA TTATA AGOOG
1081	CUUTGATSCU	CAACAGUIAG	CUCACAAACA		CAACCAITAI	TIATAACCCC
1141	AAAIGCACGA	AIGITICAAL		ATTTTCCC	TITAGIAIGG	GCGACAGAAA
1201	AAGGACTTAG	GAGCAATAGA	GAAAGTACCG	CAAGGGATCG	CIGAAAGAGA	AATGAAACAA
1261	CCCAGTGAAG	CTAAGTAAAG	CAGAGATITA	TICICGIACC	TITTGCATCA	TGATTIAGCO
1321	AGCGTGACCC	AAGCAAAGAG	TGCTTTAGTT	TGACACCCCG	AAACTAGGGG	AGCTACTCCA
1381	AGACAGCCTA	TITATAGGGC	GAACCCGTCT	CTGTGGGCAAA	AGAGTGGAAT	GAGCTTTGAG
1441	TAGAGGTGAT	AAACCTACCG	AACCTAGTTA	TAGCTGGTTG	CCCGAGAAAT	GGATAGAAGT
1501	TCAGCCTCTC	AGATTCTTTA	TTCACCTCAG	TATTACCCCA	CCTGATACCA	CAAGAAACTG
1561	TGAGAGTTAT	TCAAAGGGGGG	TACAGCCCCT	TTGAAACAAG	ATACAACTTT	TCCGGGAGGA
1621	AAAAGATCAT	AATTAAATAA	AGGTAAGTAT	TTGGGTGGGC	CTAAAAGCAG	CCACCCCAGT
1681	AGAAAGCGTT	ATAGCTCAAA	TACATCACTA	CCCCTCTCTA	TCCTGATCAT	TAATTCTTAC
1741	TCCCCCCTTC	CETACEGGE	CATCCCATGC	ACACATGGGA	GAGACCCTGC	TAATATGAGT
1801	AATAAGAGAG	CCAAGCCTCT	CTCCTCGCAC	ACATGTAATT	CIGAACGAAC	CCGCACCGAG
1861	CATTAACGGC	CCCAAACGAA	GAGGGACCTG	AACAACAACC	CAAACAACCA	GAAAAAATT
1921	CAAACATAAA	CCGTTAACCC	TACACAGGTG	TGCATTTAAG	GAAAGACTAA	GAGAAAGAGA
1981	AGGAACTCGG	CAAACAAATC	AAGCCTCGCC	TGTTTACCAA	AAACATCGCC	TCTTGCAAAG
2041	CTAAAGAATA	AGAGGTCCCG	CCTGCCCTGT	GACTATTAGT	TTAACGGCCG	CGGTATTTTG
2101	ACCGTGCAAA	GGTAGCGCAA	TCACTTGTCT	TTTAAATGAA	GACCTGTATG	AATGGCACAA
2161	CGAGGGCTTA	ACTGTCTCCT	CTTTCAAGTC	AATGAAATTG	ATCTCCCCGT	GCAGAAGCGG
2221	GGATATAAAC	ATAAGACGAG	AAGACCCTAT	GGAGCTTTAG	ACACCAAAGA	AGATCCTGTC
2281	AAGTAACCCT	CTATAAAGGC	CTGAACTAAT	GGAACCCTTC	CCTAATGTCT	TIGGTIGGGG
2341	CGACCGC3GG	GAAACAAAAA	ACCCCCACGT	GGAAAGGGAG	CACCCCCTCC	TACAACTAAG
2401	AGCCGCAGCT	CTAATTAACA	GAATATCTGA	CCAATAAGAT	CCGGCAATGC	CGATCAACGG
2461	ACCGAGTTAC	CCTAGGGATA	ACAGCGCAAT	CCCCTTTTAG	AGCCCATATC	GACAAGGGGG
2521	TTTACGACCT	CGATGTTGGA	TC'AGGAC'ATC'	CTAATGGTGC	AGCCGCTATT	AAGGGTCCGT
2581	TTGTTCAACG	GTTAAAGTCC	TACGTGATCT	GAGTTCAGAC	CIGAGTAATC	CAGGTCAGTT
2641	TCTATCTATG	GTGTGCTCTT	TTCCAGTACG	AAAGGACCGA	AAAGAAGAGG	CCCCTGCTCT
2701	AAGCAAGCCT	CACCCCCACC	TAGTGAAAAC'	AACTAAAGTA	GGCAAGAGGG	CATACCCCCA
2761	GTGCCTGAGA	GAACGGCATG	TTGGGGTGGC	AGAGCCC3GT	GAATGCAAAA	GACCTAAGCC
2821	CTTTTTACAG	AGGTTCAAGT	CCTCTCCTTA	ACTATGATTT	CAGTCCTTAT	TACCCATATT
2881	CTCAATCCCT	TEGCCTTCAT	TGTCCCCATC	CTCTTAGCCG	TEGECTTEET	CACACTTTTA
2941	GAACGTAAGG	TACTAGGGTA	TATACAACTA	CGAAAAGGCC	CAAATATTGT	AGGGCCTTAC
3001	GGACTATTAC	AGCCTATCGC	TGACGGTGTG	AAGCTCTTTA	TTAAAGAGCC	CGTTCGCCCC
3061	TCCACTTCTT	CCCCAGTACT	ATTCCTCCTC	GCCCCCACTAC	TCGCACTCAC	ACTTGCCTTA
3121	ACCCTTTJAG	CCCCTATACC	TCTCCCCTAC	CCAGTAATTG	ACTTAAACCT	TGGGATCCTA
3181	TTTATTTTGG	CCCTATCAAG	CCTCGCTGTC	TACTCCATTC	TAGGCTCAGG	TTGAGCATCA
3241	AATTCAAAAT	ACGCTCTCAT	CGGGGCCCTT	CGGGCTGTAG	CCCAAACCAT	TTCATATGAA
3301	GTTAGTCTAG	GCTTAATCCT	ATTAAGTACC	ATTATCTTTA	CAGGAGGTTT	CACACTACAA
3361	ACCTTCAACA	TTGCCCAAGA	AAGCGTCTGA	ATACTACTCC	CAGCCTGACC	ACTAGCCGCA
3421	ATGTGATATA	TTTCAACCCT	TGCAGAGACA	AACCGTGCAC	CCTTTGACCT	TACTGAAGGC
3481	GAATCCGAAC	TAGTCTCTGG	CTTCAATGTC	GAGTATGCGG	GTGGCCCATT	CJCCCTATTT
3541	TTCCTGGCCG	AATATGCTAA	CATTCTGCTT	ATAAACACGC	TTTCCGCTAC	CCTCTTTTTA

Figure 3. Sequences of the mitogenome of S. longispins (GenBank Accession No. KJ834061).

3601	GGGGCCTCCC	ATTTTCCAAT	GCTACCTGAA	CTCACCGCAG	TTAACCTAAT	AACCAAAGCA
3661	GCCCTTCTGT	CCGTCTTATT	CCTATGAGTC	CGAGCCTCTT	ACCCACGATT	CCGCTACGAC
3721	CAACTTATAC	ATCTAATTTG	AAAAAACTTC	CTCCCACTTA	CACTAGCCCT	GGTTATCTGA
3781	CACCTAGCCC	TCCCCATTGC	ATTTGCTGGC	CTGCCACCCC	AGCTATAGAT	AAGAAGCCGT
3841	GCCTGAAGTA	AAGGGCCACT	TTGATAGAGT	GACTTATGGG	GGTTCAAATC	CCCCCCGCTT
3901	CTTAGAAAAG	GGGGACTCGA	ACCERCICTA	AGRAGARTAA	AACTCCTGGT	GCTCCCACTA
3961	CACTATTTCC	TAGTAAAGTC	AGCTAATTCT	AAGCTCTTGG	TOCCATACCO	C'AAAC'ACGAA
4021	GGTTAAAATC	COTCOTTAC	TAATGAACCC		ACCECCTEC	TATTTGGTAT
4021	TEETTTAGEC		COTTOGONG	CTCCCACTGA	CTACTACCCT.	GAATAGGCCT
4001	GGAAATAAAT	ACTRCTRCCA ACTCTTCCCA	TCATTOCCOT		CIACINOCCI	CTERARCART
4201	Таллаглаг	ACALANTATT	TOTTAATTOA	ACCTCCCC	CACCATCACC	TACTCTTTCC
4201	CARCACCACC	ACCONTRACT	TAACTCCACA	ATCCCSOR	TTACAAATTC	COCACCTT
4201	CABUALUALU	ARCOUTION ATTOTOLOGY	TAAUIGBAUA	AUGUDALLII	CONCEPTION	CUCACCULTI CTCTCCUCTT
40/1	ATENTER	CINCICACIT		COTAGORIA	COACT ICCAC	TATOTACIA
4381	AIGALIGULI	GRAGIACIIC	AGGGUUTAGA	ULIAAUUAUA.	COACTIAITI Tataaa	ISICIACUIS
4441	GUAAAAAUII AAmeaammeeta		CULTAITAGE	ACCARACULUU	ISIGUCAACA	ACCALCULIUI
4501	AATCATICIA	GGACICACCI	CAACCATIGE	AGGAGGUIGA	GGAGGCUTTA	ACCAAACCCA
4561	ACTICECAAE	ATTCTTGCTT	ACTUCTUCAT	CGCACACCIA	GSCIGAATAG	TAATTGTGCT
4621	ACAATTCTCC	CCCTCCTTGA	CTATITTAAC	ATTATTCACA	TACTICATIA	TAACATTCTC
4681	AGCATTTCTT	AIGTTTAAAC	TTAATAAAGC	AACCAACATT	AATACTCTAG	CAACCTCATG
4741	AGCAAAAACC	CCCGCCCTAA	CAGCCCTTGC	ACCCCTTTTA	TTATTATCCT	TAGGGGGACT
4801	CCCACCACTT	ACAGGCTTTA	TGCCAAAATG	ACTTATCCTT	CAAGAACTTA	CTAAACAAGA
4861	CCTCGCCCCA	GCTGCAACAC	TAGCAGCAAT	AACCGCCCTC	CTCAGCCTAT	ATTTTTACCT
4921	ACGACTATCA	TACGCAATGG	CATTAACTAT	TTCACCCAAC	AACCTCACCG	CAATTTCCCC
4981	ATGACGCCTC	CCCTCCTTAC	AACTAACACT	GCCACTTGCT	ACCTCAGCCA	TAGCTACGCT
5041	ACTGCTTCTA	CCCCTAACAC	CCGCCGCAAT	AGCACTAATA	ACCCTTTAAG	GGACTTAGGT
5101	TAAAACAAGA	CCAAGGGCCT	TCAAAGCCCT	AAGTGAGGGT	GGAAGTCCCC	CAGTCCCTGA
5161	TAAGGCTTGC	GGGACACTAC	CCCACATCTC	CTGTATGCAA	AACAGGTACT	TTAATTAAGC
5221	TAAAGCCTTC	CTAGAAGGGC	AGGCCTCGAT	CCTGCAAGAT	CTTAGTTAAC	AGCTAAGCGC
5281	TC'AAAC'C'AGC'	GAGCATCCAT	CTACCTTTCC'	CCCGCCTAAA	AAGCGGGCTG	AAGGCGGGGG
5341	AAAGTCCCGG	CAAACAACTA	ACCTGCATCT	TCAAATTTGC	AATCTGATAT	GTATAACACC
5401	TCAAAACTTC	TGGTAAAAAA	AGGATTCAAA	CCTCTGTTTG	TEGGECTACA	ATCCATCGCT
5461	AAAAAACTC'A	CCCATCTTAC	CTGTGGCCAT	CACACGTTGA	TTTTTCTCCA	CTAATCACAA
5521	AAACATCGGC	ACCCTTTATC	TATTATTTGG	TGCCTGAGCC	GGTATAGTAG	GCACAGCCCT
5581	CARCETACTE	ATTCGAGCAA	AACTAAGCCA	CCCGGGGCGCT	CTCCTTGAAG	
5641	TTATAATGTA	ATCGTTACGG	CACATGCCTT	CGTAATGATT	TTCTTTATAG	TAATGCCAAT
5701	TATAATCIGA	GGTTTTTGGAA	ACTGATTAAT	TCCCCTANG	ATTEGACCCC	CAGATATAGC
5761	ATTTOOTOGT	ΔΤΔΔΔΤΔΔΤΔ	TAAGTTTOTG	ACTTCTTCCC	COTTOTTTTC	TICTACTACT
5701	TRECETCET	CONCEACE	CARCTRCTCC		ACACTCITIC	CIRCINCIACI CIRCIPCITACI
5921	Таасиластта	CORCINCECAC	CACCATCACT	CRACCERANCA	ATCITTTCAC	TOCLETION
50/1	ACCTATOTO		CCCCANTCAN	TTTTATTACC	ACAATTATTA	ATATCAACC
6001	TOOTAICICC	TOCCANTACC	ACACICCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	ATTTALIACC	COCCTCTTA	TRACCCC
6061	TOTTOTOTT	CTCTCTCTAC	CACTTOTOR	TCCACCTATT	ACAATCOTCO	TIACCOULDI
6121	- 10 110 10.011 - AAAC/2777AAC	ACCACCTERC	TCAGIICICGC	CCCAGGIAII	CATCOANTCC.	TIACGGACCS
0171 6101	- AAAUUTTAAU - aamaamaa	TTTTTTTTCC			ATTUCATULE	COTTRUCTAL
0101 6041	CATTRONGAT		AUCUSUASUI	CINNNAN CAN	COUTTTOCK	ATATACCCAT
6241 6201	GATIICALAI	ALIGIUGUUI	TTACIUIGE		CULTIIGBAI GENEGOGOEG	ATATADJI, AT
63UL	GUILIGAJLA GRANGTEROT	ATAAIGGUIA	A CARGE CITE	AUGATITATI	GIAIGGGUIU	AICAUATATT
6361 5401	CACAGIIGGI	AIGGACGIAG	ACACACGIGU	TIACITIACE	TCIGUCACAA	TAATCATCGC
64Z1	AATICCCACC	GGIGIIAAAG	TATTIAGUIG	ACTIGUAACT	CIACAIG6666	GUIUIAIIAA Taggggggg
6481	ATGAGAGACA	CCCCTTTTAT	GGGCCCTTGG	CITTATUTIC	CTATTTACAG	TAGGCGGGCT
6541	TACAGGTATT	GITCTGGCCA	ATTCATCICT	AGATATTJTA	CTACACGATA	CATACTATGT
6601	AGTAGCCCAC	TTCCACTACG	TACTATCTAT	GGGGGCCGTA	TTTGCCATTG	TUGUUGUCTT
6661	CGTGCACTGA	TTCCCGCTAT	TCTCAGGATA	TACACTCCAC	AGCACTTGGA	CAAAAATCCA
6721	TTTCGGCATT	ATGTTCTTGG	GGGTAAACTT	AACCTTCTTC	CCACAACACT	TCCTCGGATT
6781	AGCCGGAATG	CCCCGACGAT	ACTCCGACTA	CCCTGAC3CC	TATACCCTAT	GAAATACAGT
6841	CTCCTCAATC	GGATCACTTA	TCTCGTTAGT	GGCTGTTATT	ATGTTCTTAT	TTATTATTTG
6901	AGAGGCATTC	GCCGCCAAAC	${\rm GTGAAGTTCT}$	AGCAACAGAT	TTAACAACAA	CCAATGTAGA
6961	ATGACTGCAT	GGCTGCCCTC	CCCCTTACCA	CACATTC3AG	GAGCCTGCCT	TTGTACAAGT
7021	ACAAGCAGAC	TAACGAGAAA	GGGAGGAGTC	GAACCCCCAT	AGGTCGGTTT	CAAGCCGACC
7081	ACATAACCGC	TCTGCCACTT	TCTTTATAAG	ACACTAGTAA	AAAAGTACAT	TACACCGCCT

(Figure 3 Continued)

7141	TGTCGAGACG	GAAGTGTGGG	TTAAACCCCC	GCGTGTCTTG	CTTTTAATGG	CCCATCCGTC
7201	ACAGCTT3GA	TTTCAAGATG	CAGCTTCACC	TGTTATAJAA	GAACTTCTTC	ATTTTCACGA
7261	CCATGCTTTA	ATAATCGTCT	TCCTAATTAG	CACACTACTACT	CTTTACATTA	TTCTTGCTAT
7201	ACTTACCACT	ADATTANCOL	ACANATATAT	TTTACATION	CINCATIN	AAATTATOTT
- 7.374 - 120.04	AGTIACCACT AACAATCOTC		TTTTA		CARDAGATIG	CANTERICIS.
/381	AACAAICSIC	CCAGUTATCA	TITTAATICT	AATIGUASIA	CULTUUTIU	GAALICIGIA
7441	CCITATAGAT	GAAATTAACA	ATCCCCT3TT	AACAATTAAA	GCCGTTGGCC	ACCAATGATA
7501	CTGAAGCTAT	GAATACACTG	ACTACGAAGA	TCTTGGATTT	GATTCATACA	TAATCCCCAC
7561	CCAGGACCTA	ACCCCTGGAC	AATTCCGCCT	ATTAGAAGCC	GACCATCGCA	TGGTTATTCC
7621	AGTTGAATCC	CCCATCCGAG	TTTTAGTATC	TGCAGACGAT	GTACTCCACT	CATGEGCAGT
7681		GGGGTAAAAA	TGGACGCAGT	CCCAGGACGC	ΓΤΤΔΔΤΓΔΔΔ	CARCICITTAT
77/1	CCCATCCCCA	CCACCCTAT	TOTACCCACA	ATCOTOTOAC	ATCTCCCCAC	CAAATCACAC
7001	CUCHICCOR	ATTOTACTO	ANDENETTEE	COTACANCAC	TTTCANANTT	CARAICACAS
- 76UL - 2001	UTITATALUT AREACTERIAA	ATTOTACTOC ALGORATICS	TAGUAGIICU.	ULIADAAAL	1115AAAA11 Magammaaga	BAILAILILE TTTN MOOTA
7801	AATACTIJAA	GAUGUUTUGU	TAGGAAGCTA	AATAGGGTAT	AGUELLAGUU	TITIAAGUTA
7921	AAGATTGGTG	GCCCCCAACC	ACCCCTAACG	ACATGCCCCA	ACTCAACCCC	GCACCTTGAT
7981	TTGCTATTTT	AGTETTETEG	TGAATGGTCT	TCCTGGCCGT	TATTCCCGCT	AAAGTTACAG
8041	CCCACACCTT	CCCAAACACT	CCTACTCTGC	AAAGCGCAGA	AAAACCTAAA	ACAGACCCCT
8101	GAACTTGACC	ATGACACTAA	GCTTTTTTGA	CCAGTTTATA	AGCCCCACCT	ATCTTGGAAT
8161	CCCATTAATA	GCCCTTGCCC	TTACCCTACC	CTGACTCCTT	TACCCCACAC	CTACAACTC3
8221	ATGATTAAAT	AACCGATTCC	TCTCGCTTCA	AGGTTGATTT	ATTAACCGTT	ТТАСТСААСА
8281	GOTTOTOTO	CCCCTAAATA	TTGGGGGTCA	CAAGTGAGCC	GCCCTCCTAA	CTTCATTAAT
82/1	CATCTTTTTA	ATTTCCCTAA	ATATATTACC	ACTICTICC	TATACTTTA	COCCACCAC
0.041	CONTRACTOR	ATTICCCIAR ATTICCCIAR	adamaadaaa	ACTICITECE	TATACITITA	TEACCACCAC
8401 8401	CUAACIAICA	CIAAAIIIAG	10011001001	ACCICICIGA	TIAGUAAUIG	TIATIATIGS
8461	CATECEAAAC	CAACCAACCC	ACGUUCTAGG	ACACCICCIA	CUAISAAIGISCA	CACCUBGCCC
8521	CCTTATTCCA	GIGCTTATTA	TTATCGAAAC	AATTAGTCTC	TITATTCGCC	CCCTTGCCCT
8581	AGGGGTACGG	CTAACAGCCA	ATTTAACAGC	CGGTCACCTA	TTAATTCAAT	TAATTGCTAC
8641	AGGTGCCTTC	GTACTTCTTC	CCCTAATACC	AACCGTC3CA	ATTATCACAA	CAACAGTATT
8701	GGTTCTCCTT	ACCCTGTTAG	AAGTTGCCGT	AGCAATAATT	C'AAGC'C'TAC'G	TETTEGTTET
8761	CCTGCTAACA	CTGTACCTAC	AAGAAAACGT	CTAATGGCCC	ATCAAGCACA	CCCTTACCAC
8821	ATAGTTGACC	CTARCCCTTR.		GGGGCAATTG	CTRECCCTCCT	GATAACATCA
2221		COTCATTCA	TTTTCCCTCA	ACAACCTTAA	TAACCTTACC	A ACACCTCTA
0001		CONTINCA	ATCATCACCA	CATATOTIA	CACAACCTINOU	ATTOCIÓN
0941			AIGAIGAGGA	GATAICGIAC	TAGAAGGIAC	ATTUUAAUUA
900I	CAICACAUGU	USUUUBIACA	AAAAJUIUII	CGATACGSAA	LAATICITII	CATTACCICC
9061	GAAGTATICT	TITICCIAGG	GTTCTTCTGA	GCCTTTTTACC	ACGCAAGCCT	CGCTCCCACT
9121	CCTGAGTTAG	GBEGCTECTE	GCCCCCCCACG	GGCATTACAA	CCCTTGACCC	ATTTGAAGTC
9181	CCCCTCCTTA	ATACAGCTGT	CCTGCTTGCC	TCCGGGGTAA	CAGTTACCTG	GGCCCACCAC
9241	AGCATTATGG	AAGGTGAACG	AAAACAGACC	ATTCAATCAC	TAGCCTTAAC	TATTCTTCTA
9301	GGCTTTTACT	TTACATTTCT	TCAAGCCTTG	GAATACTATG	AAGCCCCCTT	TACAATTGCA
9361	GACGGCGTAT	ACEGCTCTAC	ATTTTTCGTA	GCCACCGGAT	TECACGGACT	ACACGTTATT
0/21	ATTGGCTCCA	CATTTTTAGC	CGTTTGCCTC	CTACGGCAAA	TOCATACCA	TTTTACATCC
Q191	GAACACCATT	TUBBERTUGA	AGCAGCTGCC	TGATACTGAC	ATTTCCTACA	CETTETCTEN
0541	TTATTOCTAT	ATATOTOTAT	CTACTCATCA	COCTOTTANT	CTTTCOILOL	TTAAAACTAC
9041	TIALICULAL	TTALLILIAL	CIACIGAIGA	TTANATOON	VILLEN	TIARAACIAJ
9001		IIUUAAIUAU		TIAAAAI.LA		AIGAAUGIAG OTATOOT
9661	CAATAGUIGI	AATTACCATC	ACTATITIGU	TITCCGTAGT	CONGECCATI	GIATUUTUT
9721	GACTECCECA	AATGACCCCC	GACCACGAAA	AGCTCTCCCC	ATATGAATGT	GGTTTCGACC
9781	CTTTAGGATC	AGCCCGCCTA	CCATTTTCCC	TCCGCTTCTT	CCTAGTCGCC	ATTCTCTCC
9841	TACTTTTCGA	TTTAGAAATT	GCCCTTCTCC	TCCCCCTCCC	ATGAGGAGAC	CAATTAACCT
9901	CCCCTTTATT	GACACTCTTC	TGAGCCGTAG	CCGTGCTTAT	TETTETCACC	CTTGGCTTAG
9961	TCTACGAGTG	AATTCAAGGA	GGGTTAGAAT	GAGCCGAATA	GCCAATTAGT	ΤΤΑΑΑΑΑΑΑΑ
10021	TATTTGATTT	CGGCTCAAAA	GCTTATEGTT	AAAGTCCATA	ATTGTCTAAT	GACTECCGET
10081	CACTTCCCTT	TOTOATORCO	CTTTACTCTA	CCATTCACAC	GOCTAGCATT	CONTRELACO
10141	CACITOGUI	COCONTRAT	ATCCTTACIAN	CORVERTENCE	TOTOTAT	TATTOCACTT
1.0004	CAUCIULUE TANA ATTITA A	COURCEANCE	ALGUITADAA aaaamaaaaa	ACTEMPTOTOL	ISIGIIIAII Adaamaamaa	IAIIGBAUII aarrana
TUZUI 10061	ICAALLISAA OOTTTTTTTCLC	ULUTULAAUT	ADDREEDED A	AUTICICIÓ	THEFT	ACTICICITA
10261	GCTTTTTCAG	CIIGIGAAGC	AAGUGUGGGG	CINGCCUIGC	TGGTAGCUAC	AGCICGCACA
10321	CATEGCTCGG	ACCECCTTCA	GACCITAAAC	CTCTTACAAT	GCTAAAAATC	CTAATTCCCA
10381	CTCTAATGCT	TCTCCCCACA	GCCTGGCTTG	CCCCTGCCAA	ATGATTGTGA	CCTACTACTC
10441	TCTCTCACAG	CCTAGTCATT	GCATTAGCCA	GCCTCACCTG	ACTAAAAAAT	ACATCCGAAA
10501	CAGGCTGGTC	TTGCCTCACA	CCCTTCATAG	CCACAGACCC	CCTCTCAACC	CCCCTTCTTG
10561	TCCTCACCTG	CTGACTACTC	CCCCTTATGA	TTTTGGCAAG	CCAAAACCAC	ACAGCACTCG
10621	AACCTATTAA	CCGCCAACGA	ACCTACATTA	GCCTATTGAC	ATCACTGCAA	GIGITCCITA

(Figure 3. Continued)

10001	mm3 m3 (0 (13 mm	maamaaaaaaa	0110T10TD3	<b>D</b> 1 <b>D D D D D D D D D D</b>	<pre>dim.mmmail.</pre>	
10981	TTATAGCATT	LEGIECCACI	GAACIACIIA	TATTTAIGI	CATATITIGAA	GCTAUTCITA
10741	TCCCCACGCT	AATTATTATT	ACTCGCTGAG	GTAACCAAGC	AGAACGTCTT	AATGCAGGAG
10201	TATATTTTT	CTTTTATACC	CTCCCACCCT	OTOTOCOUTT.	ACTACTOCCO	CTATTACTTC
10001	TTCLCLARA	TALAGGERGE	CINCLOSEI	TANCOUNT	ATATOICOCC	TOTACIOCIC
LOSOT	LICAGAAAGA	TACABBUTCE	CILICULIUI	TAAULATULA	ATATISULAGU.	TSTAUUUUU
10921	TTTCATCTTA	TGCTGATAAG	ATTTGATGGG	CAGGCTGCCT	GATTGCATTT	TTAGTAAAAA
10981	TACCCCTATA	CGGAGCACAT	CTATGGCTGC	CAAAAGCACA	TGTAGAAGC'C'	CCAGTTGCAG.
11041	COTCAATACT	COTCOTCOL	CTTCTTCT AA	AACTACCCC	CTACCCATA	λΤΟΟΓΆΛΤΑΛ
4.1.1.01	TOLICHAINOI	da Macrata	AGEARAGANE	TAACTTATCC	ATTANTA	areadaarar
TTIUT	IGACCATATI	GGAACUICIC	ACTAAAGAAT	TAAGITAICC	ALLIALIAIC	CICECCICI
11161	GAGGTGTAAT	TATAACTGGC	TCCACCTGCC	TTCGCCAAAC	AGATCTTAAA	TCCCTCATCG
11221	CTTACTCATC	CGTAAGCCAC	ATGGGCCTAG	TCGTTGGAGG	TATTCTTATC	CAAACACCAT
11281	CACCOUNTER	CERCECCETA	ΑΤΤΟΤΟΛΤΑΛ	TTCCACACCC	COTANCATOO	TARCOTTT
	Taraar	COOCOCCOIN	TICICAIAA	TIGCACACGG		TCAUCCUIII
11.341	TCTGCTTGGC	CAACACAAAT	TACHAAUGCU	TCCATAGECG	GACAATACTA	TTAGUTUGAG
11401	GGTTACAAAT	AGTACTTCCA	CTCATAGCAA	CATGATGATT	TATTGCCAGC	CTCGCAAACT
11461	TAGCCCTTCC	CCCCCTGCCC	AACCTCATAG	GGGAACTTTT	AATTATTACC	TCATTATTTG.
11571	CTTCATCTC	A TRA ACCOTC	CTACTCACAC	CCCCACCAAC	COTCATTACC	CCAACCTATT
44504	- OTTOATCOTO - ALAMAMATATA	ATTACCOLL	Jacobadada	doddalamadad	AGGICATIACC	ATTACIOTA A
11281	CACICIAIAI	ALICUICAIG	ACCUAGUSGG	GUUUAUTUUU	ASCACATATI	ATTAGUUTAA
11641	ACCCTTCCTA	CACGCGGGGAG	CACCTAGTCA	TAGCCCTTCA	CETCETCCC	CTGCTTCTAA
11701	TTGTCTTAAA	GCCCGAACTA	GTATGAGGCT	GAACCACCTG	TAGATATAGT	ТТААСААААА
11761	TATTACATTC	ΤΊΛΤΤΩΤΑΛΑ	CACARACTT	AAAATCCCCT	татесассаа	CARACCTCC
11001	TULINOVIIO	JANGTAAN	TATAAAAA	Ammaammaal	addaladaa	CACADOCICU.
TTQT	UUABUAAUBA	ASAUTSUTAA	ICICCGISAU	CIIGGIIGGA	CCCCAGGGCI	CACIUGGUUI
11881	GCTCCTAAAG	GATAACAGCT	CATCCATTGG	TCTTAGGAAC	CAAAAACTCT	TGGTGCAAAT
11941	CCAAGTAGCA	GCTATGCACT.	CCTCATCACT	TATTATATCA	TCCAGCTTAG.	TCATTATCTT
12001	TTTACTTA	CONTATCOTA	TTTTTACZAC	TOTACACOOT	COCCETCICAA	ACCCCACTO
100C1		- CORTAICCIA	- ANACNACIAC	ACTRONOUCT COTACACTEC	TTATATA	TABLECCORCID TAATTOOOOD
TZODI	GUUGIUIUA	CAIGIIAAGA	CAGCAGICGC	CUTAGUUTIU	TICATCAGUE	TAATICCCCT
12121	ATTCCTGTTT	CTTAACGAAG	GAGCAGAAAC	AATTATCACC	TCATGAAATT	GAATGAACAC
12181	ACTGACCTTC	GACGTGAACA	TTAGTTTCAA	ATTTGACCAC	TACTCAGTTA	TETTTGTACE
12241	CATTGCCCTC	TATGTCIACCT	GATEMATTET	AGAATTTGCA	THATGATATA	TACACIGCIAGA
10001	COUNTACATA	ATCOATTT	TTANATATOT	COTATTTC	CTTATTOCCA	TATTATTOT
17.301	CUCATACATA	AATUGATITT	TIAAATATUT	CUTAATITIC	CLIATIGULA	
12361	TGITACAGCA	AATAATCTAT	TCCAGCTITT	CATTGGTTGG	GAAGGAGTAG	GCATTATGTC
12421	ATTTCTTCTC	ATTGGCTGAT	GGTACGGACG	AGCAGATGCC	AACACAGCGG	CCCTCCAGGC
12481	CGTTGTGTAT	AACCGGGTCG	GAGACATTGG	ACTECTATTC	ACAATAGCAT	GAATAGCAAC
10541	CANCECTANC	TOOTONONOO	TACAACAAAT	TTTTOTOOOA	ACANANCIAC	TACATOTOAC
12041		ISCIGAGAGC	TACAACAAAI	TITIGIGGCA	RCAARAGACU	TAGAICIGAS
12601	CCTACCGCTA	CTAGGCCTGA	TTATEGEEGGE	TACAGGTAAA	TCTGCCCAAT	TIGGICICCA
12661	CCCTTGACTC	CCCTCTGCTA	TAGAGGGTCC	TACACCGGTC	TETGECETAC	TGCATTCAAG
12721	CACTATAGTC	GTCGCCGGTA	TTTTTCTTTT	AGTACGAACA	AGCCACTCC	TAGAAAATAA
12721	TCAAACT2CC	CTCACCACCT	COUTATCOOT	ACCACCCTA	ACIAACIACITAT	TCACACCCAC
10041		ACCALCT	ATTACANCA	AUDAOCCEIA		ANATANAT.
12841	CIGIGCUCIG	ACCUAAAAIG	ATALCAASAA	AATIGTASCA	TICICCACAI	CAAGICAACI
12901	TGGCCTAATA	ATAGTCACTA	TTGGACTAAA	TCAGCCCCAA	TTAGCCTTCC	TACACATITIG
12961	CACCCACCCC	TTCTTCAAGG	CAATATTATT	CCTCTGTTCT	GGCTCAATTA	TTCACAGTCT
13021	сарсізар	с а адататес.	GAAAAATAGG	AGGCIATACIAT	CACCTTACCC	COTTTACATO
10001	CTCCTCCCTC	ACTATTCCTA	CTTTACCCC	CACACCCACC	COOTTOOTOO	CACCETTCTT
TOOL		ACIALIGUIA	GITTAGULU	CACAGGCACC	CULTICUIGG	CAGGGIICII
13141	CTCCAAAGAT	GCCATCATTG	AGGCACTAAA	CACATCCCAC	CTAAACGCCT	GAGCCCTAGT
13201	CCTAACCETT	CTAGCCACCT	CATTCACGGC	CATCTACAGT	CTCCGCGTAG	TATTCTTTGT
13261	CTCTATEGGC	TACCCACGAT	TTAACCCTAT	TTCTCCCATC	AATGAGAAC'A	ACCCARCART.
12201	TATTAATOOC	TTANAACCAC	TTECATEACE	AACCATTCTC	COTCCCTCC	TAATCATTTC
1.3.341	ALLARIAU A		IIGUAIDAGU	ARGUALISIU		AAIUAIII.
13381	AAGCATTACC	CCCCTCAAGA	CCCUIGIJAI	ATCCATACCC	CCUTTAUTCA	AATTAGCIGC
13441	CCTCGGAGTT	ACAATTACAG	GATTATTTAT	TGCCCTC3AG	CTGGCAACAT	TGACCAATAA
13501	ACAGTACAAA	ATCACCCCTA	ATCTAGTTAC	TCACCACTTC	TCCAACATAC	TAGGCTTTTT
13561	CCCCTCCATC	ATTCACCGAT	TTATACCOAA	δητλοδτήτο	GTTTTAGGAC	ABCACTTON
10/01	- CCCCICOAIC	ATTCACCOAL	ammalamlal	ALANGERGOE	dodryyddyy	TOUCHCIICC
13021	CAGUCAACIA	ATTGACCAAA	CIIGACIAGA	GAAAGIIJGI	UJUAAAGUAA	ICIUIICAIC
13681	AAATATTCCC	TIGATTACAA	CAACAAGTAA	TACCCAACAA	GGAATAATTA	AAACATATCT
13741	CACCCTATTC	CTTCTCACCT	TAACCCTCGC	TGCCTTATTA	TTTACCCGTT	AAACTGCCCG
13801	ABAGCCCC	CRACTTAGTC	CACGAGTTAA	стеста астаста	AUAAATAAGG	ТААССАССАА
12041	A A CCC A CCCCC	CTANCTACCA	CTATACCTCC	COOTANTON	TACIATTA ACC	CAACCCCCCC
10001	AAUUUAUSUA	CIAAGIAUCA	BIAIACUICU NANGERENER	CUCTAAISAA	TACATTAACG	CAACUUUIUU
13921	AATATCGCCT	CUCAATACAG	AGAGCTCACT	AAGUTCATCA	GCCGGCACCC	AIGAGGACTC
13981	ATATCACCCC	CCTCAAAATA	CACTAGAAGC	CACCCCCACC	CCTACTAGGT	ACATCAACAT
14041	ATCACCTACA	ACAGGACCAC	TCACCCAACT	CTCCGGATAC	GGCTCAGCGG	CAAGCGCCGC
1/101	CIGAATATICA	2 2 C' 2 C' 2 2 C' T 2		CARATAAATC	222220000000	CTAATGATAG
1 4 4 4 4 4	JINGGGGGGGG		DIDIDUCUCU DOMESTICACIÓN			COALUSIAJ ATNOTANOCO
14101	RAAGGGICCC	CCATGACCIA	CCAATAUICU	GUAUUUUAIG	UUUUGUUAUAA	UTACTAAUUU

(Figure 3. Continued)
14221	CAAGGCAGCA	AAGTAAGGAG	AAGGGTTTGA	AGCAACT3CA	ACCAACCCTA	GAACTAATCC
14281	ΑΑΤΤΑΑΑΑΑΤ	AAAGACATAA	TATAAGTCAT	AATTCCTGCC	AGGACTTTAA	CCAGAACCAA
14341	TGGCTTGAAA	AACCACCGTT	GTTATTCAAC	TACAAGAACC	CGCTAATGGC	AAGTCTACGA
14401	AAGACACACC	CTCTCCTCAA	AATCGCAAAC	AATGCCCTAG	TTGACCTACC	CICCCCCTCA
14461	AATATTTCAG	TGTGATGAAA	CTTCGGATCT	CTCCTAGGAC	TCTGCTTAAT	TATCCAGATC
14521	CTCACAGGAC	TATTTTTAGC	CATACACTAT	ACCTCTGATA	TTGCTACGGC	TITTTCTCC
14581	GTTGCCCACA	TTTGTCGGGA	CGTAAATTAT	GGCTGACTCA	TCCGAAACCT	TCACGCTAAC
14641	GGTGCATCCT	TCTTCTTTGT	ATGTATTTAT	GCCCATATTG	GCCGCGGACT	TTACTACGGC
14701	TCATACCTCT	ATAAAGAAAC	ATGAAACATC	GGAGTAGTCC	TTCTACTTTT	AGTTATAATA
14761	ACTGCTTTCG	TTGGCTATGT	ACTACCCTGA	GGCCAAATGT	CCTTTTGAGG	TGCAACCGTT
14821	ATTACTAACC	TACTTTCCGC	AGTACCCTAC	GTAGGTAGCT	CTTTAGTCCA	ATGAATCTGA
14881	GGTGGGTTCT	CAGTAGACAA	TGCAACCETT	ACCCGATTCT	TTGCCTTCCA	CTTCCTATTC
14941	CCATTCGTAA	TTGCAGGAGC	AAGCATGGTA	CACCTTCTTT	TECTTEATEA	AACAGGATCA
15001	AATAATCCCC	TCGGCCTAAA	TTCAGACGCA	GATAAAATAA	GCTTCCACCC	CTATTTCTCA
15061	TATAAAGACT	TACTGGGGTT	CGCAGTACTT	GTCATTGCCC	TCACATGTCT	AGCTCTATTT
15121	TCGCCCAACC	TACTGGGAGA	CCCAGACAAC	TTCACCCCCG	CCAATCCACT	AGTTACCCCT
15181	CCCCACATCA	AACCAGAATG	ATATTTCCTG	TTCGCGTATG	CAATTCTACG	CTCCATCCCC
15241	AATAAACTAG	GAGGAGTATT	GGCCCTACTG	GCCTCAATCC	TCATTCTGAT	GCTCGTACCA
15301	TTTCTACACA	CGTCTAAACA	ACGAAGTCTT	ACCTTCC3GC	CGCTTACACA	ATTCTTGTTT
15361	TGAACCCTAA	TTGCAGATGT	TATTATTCTC	ACCTGAATTG	GAGGTATGCC	TGTATCACAC
15421	CCGTTTGTTA	TTATTGGACA	AATTGCGTCC	TTTTTTATACT	TTTTCCTCTT	CCTAGTCCTT
15481	ACACCATTAG	CAGGTTATGC	AGAGGATAAA	GCACTTGAAT	GAACTTGCAT	TAGTAGCTCA
15541	GCGTCAGAGC	CCTGGTCTTG	TAAACCAGAT	GTCGGAG3TT	AAAATCCTCC	CTACTGCTCA
15601	AAGAAAGJAG	ATTTTAACTC	CCACCCCTGG	CTCCCAAAGC	CAGGATTCTT	ААТТАААСТА
15661	TTCTTTGTAG	TGTATGTACA	ATAATTTTAA	ATACATATAT	GTATTATCAA	CATTAATTTA
15721	TATTAACCAT	ATCATAGGGC	ATTCAAGTAC	ATATATGTTT	TATCACCATA	TCTAGGGTTA
15781	CACCATTCAG	GAATTACATG	ATACGAAAAT	TTTACATAAA	GCATGATAAT	AATAATAAC
15841	AAGTACTTAT	AAACACCAGG	CGAAATCTAA	GACCTAACAC	AAAAACCCAT	AGGTTAAGTT
15901	ATACCTTTAC	TCAAAATCCC	GCCAAACTCA	AATATTTAAT	GTAGTAAGAG	CCGACCAACA
15961	AGTCCATTTC	TTAATGCCAA	CGTTTATTGA	AGGTGAGGGA	CAAAAATTGT	GGGGGTTTCA
16021	CAGAGTGATT	TATTCCCGGC	CATTGGGTTC	TTTTTCA3GG	CCCCAAATTG	TAAACCTCCC
16081	CCATAAATTA	TTTCCCCGGG	CATAGGGTAA	TGGGGGGATA	CAAAAGCGGG	AGCGGCCCCC
16141	CTGCCGG3CC	TTTTTTTTAT	AGGGCCTTGG	GCTTTTTTTT	TTTCTTTTTC	TTTTCAATAG
16201	AC'ATTTC'AC'A	GTGCACGCAA	TCTAGTTAAC	AAGGTGGGAG	TCATCATAGG	AAGCATGGAA
16261	ATAGTATJAG	TGATGAAAAG	TCTTTACTAA	AGAATTACAT	ΑΤΑΑΑΑΑΤΑΤ	CAAGGACATA
16321	ATATAGTJAA	ATTTAGTCGG	AAGATATCTA	TATGACCCCC	TTTGGCTTTT	TEGEGTAAAC
16381	CCCCCTACCC	CCCTAAACTC	CTGAGATAAC	TAACGCTCCT	GTAAACCCCC	CGGAAACAGG
16441	AAAAC					

Table 2. Summary of the mitochondrial genome of S. longispinis

\*H and L denote heavy and light strands, respectively.



Figure 4. A gene arrangement in the complete mitogenome Sebastes longispinis.

황볼락의 미토게놈은 13개의 단백질 암호화 영역, 2개의 리보좀, 22개의 tRNA 및 1개의 조절영역으로 이루어져 있었고, 전체 크기는 16,465 bp 이었다(GenBank Accession number KJ834063)(Figure 5). 황볼락의 미토콘드리아 DNA의 염기 조성비는 27.7% (A), 26.6% (T), 17.3% (G) 및 28.4% (C)이었고, AT 비율이 약간 높았음 (54.3%), COX1 (GTG)을 제외한 모든 단백질 암호영역의 개시코돈은 ATG이었고, 종결코돈은 TAA가 주를 이루었으나 ND1과 ND3는 TAG, ND4는 AGA였으며 Cytb는 종결코돈이 불완전(T)하였다(Table 3). tRNA형태는 tRNA<sup>Ser(AGY)</sup>(lacked DHU-arm 구조)을 제외하고 모두 전형적인 cloverleaf 구조로 되어 있었다(Figure 6).

1	GCTAGCGTAG	σττααττααα	ссать всъст.	GAAGATGTTA	ABATERACCC	тарааарсее
 	dadaadaaa.	ANGGGTTGGT	COTALACACI	CHACATOIIA	TAGAIOGAGUU	
10	LGUUUGUAUA	AAGGUIIGGI	CUIGACITIA	LIAICAACII	TAGUCAAATT	TACACATGCA
121	AGTATECGEC	CCCCTGTGAG	AATGCCUTAC	AGCTCCCTGC	CCGGGGAGCAA	GJAGCTGGTA
181	TCAGGCACAC	ATCTGTAAGC	CCATGACACC	TTGCTTAGCC	ACACCCTCAA	GGGAACTCAG
241	CAGTGATAAA	CATTAAGCCA	TAAGTGAAAA	CTTGACTTAG.	TTAAAGCTAA	GAGGGCCGGT
201	A A COTOTO	CHARCHARCE	CERCITATACE	AGAGACCCAA	GTTGATACCA	TTORGOGTAS
001	AAACCICUIU	THEADLCALLO	ADALENNAC		UTIONIACCA.	ATAGGATCA
-301	AGAGIGGIIA	1.5GAAAATAA	AGAULAAAGU	CGCACACCII	CAAABCIGII	ATAUGUATUU
421	GAAGGCTAGA	AGATCAACCA	CGAAGGTAGC	TTTACAACCC	CIGACCCCAC	GAAAGCTCTG
481	GCACAAACTG	GGATTAGATA	CCCCACTATG	CCTAGCCCTA	AACCTTGGTA	ATATATCACA
541	CACCCTACCC	GCCTGGGAAC	TACGAGCACC	AGCTTAAAAC	CCAAAGGACT	TGGCGGTGCT
601	TTAGACCCC	CTAGAGGAGC	CTGTTCTAGA	ACCGATAACC	CEEGTTEAAC	CTCACCCTTC
661	CTTCTTTATC	COCCCTATAT	Accessered	TCACCTTACC	CTCTCAACCC	СТААААСТАА
001	CITCITATC	CLOULAIAI	ACCOLOTO	CABCITACC	CIGIGARGGU	CIARABURA
121	GUAUAAUIGG	CAUAAUUUAA	AACGICASGI	CISABISTISTAIS	CSCATSSAAS	GSGAAGAAAI
-781	GGGCTACATT	CCCTACATTA	GGGAACACGA	ACGGCGCACT	GAAATACGCG	CUTGAAGGAG
841	GATTTAGTAG	TAAGCGGAAA	ATAGCGTGTT	CCGCTGAAAT	CGGCCCTGAA	GCGCGCACAC
901	ACCECCCETC	ACTCTCCCCA	AGCCTATCAC	TTTAAATAAT	TAAAAACCCA	AAAATCGCGG
Q61	AGGGGAGGCA	AGTOGTAACA	TEGTAAGGET	ACCEGAAGET	GCACTTGGTA	ATATCAGAGT
1001	CTACTTAAAA	ТАРААТААСА	CTTCCCTTAC	ACTENNENCA	CACCCCTCCA	AATCCCATCA
1.001		INDARIARLA GINCIOCENCI	CITCLUIIAC GOOLOLAADA	ALIGARGAGA ALIGARGAGA	CAUCUSISCA	TRAICUSCAICA
1081	CULIGATOU	CAACAGUIAG	CUCACAAACA		CAAUCAITAI	TIATAACUUS
1141	AAACGCACGA	GIGTTTTAAT	TAAACAAACC	ATTTTTCCCC	TITAGTATGG	GCGACAGAAA
1201	AAGGACTTAG	GAGCAATAGA	GAAAGTACCG	CAAGGGATCG	CTGAAAGAGA	AATGAAATAA
1261	CCCAGTGAAG	CTAAGTAAAG	CAGAGATTTA	TTCTCGTACC	TTTTGCATCA	TGATTTAGCC
1321	AGCGTGACCC	AAGCAAAGAG	TGCTTTAGTT	TGACACCCCG	AAACTAGGGG	AGCTACTCCA
1201	ACACACCETA	TTTATACCCC	CAACCCCTCT	CTETECCAAA	ACACTCCAAT	CACCTTTCAC
1444	TAGACAGCUIR	A A A COT A COC	AACCUUTT	TIGIGGCARA	AGAGIGGAAI	GAGCIIIGAS GALTAGAACT
1441	TAGAGGIGAT	ARALLIALUG	AACCIAGIIA	TAGUIGGIIG	CUUGGGAAAI	GJATAGAAGI
1501	TCAGCCTCTC	AGATTTTTTA	TTCACCTCAG	TATTACCCCA	CCTGATACCA	CAAGAAACTG
1561	TGAGAGTTAT	TCAAAGGGGGG	TACAGCCCCT	TTGAAACAAG	ATACAACTTT	TCCGGGAGGA
1621	AAAAGATCAT	ΑΑΤΤΑΑΑΤΑΑ	AGGTAAGTAT	TTGGGTG3GC	CTAAAAGCAG	CCACCCCAGT
1681	AGAAAGCGTT	ATAGCTCAAA	TAC'ATC'AC'TA	CCCCTCTCTA	TCCTGATCAT	TAATTCTTAC
17/1	TOCOOCTTO		CATCOCATCC	AAACATCCCA	CORACCOTEC.	TAATATCACT
1001		CONACCOURCE	CTCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	ACATORIA	CHEARCEAAC	
1001	AATAAGAGAG	CCAAGUUIUI	CIUCIIGUAU	ACAIGIAAII	COBAACGAAC	CUGUALUBAS
1861	CATTAACGGC	CCCAAACGAA	GAGGGACCTG	AACAACAACC	CAAACAACCA	GAAAAAAACT
1921	CAAACATAAA	CCGTTAACCC	TACACAGGTG	TGCACCTAGG	GAAAGACTAA	AAGAAAGAGA
1981	AGGAACTCGG	CAAACAAATC	AAGCCTCGCC	TGTTTACCAA	AAACATCGCC	TCTTGCAAAG
2041	CTAAAGAATA	AGAGGTCCCG	COTGCCCTGT	GACTATTAGT	TTAACGGCCG	CGGTATTTTG
2101	ACCETECAAA	GETAGCECAA	TCACTTGTCT	TTTAAATGAA	GACCTGTATG	AATGGCACAA
2161	CCACCCTTA	ACTETOTOCT	CTTTCAACTC	AATCAAATTC	ATCTCCCCCT	COVER VGCES
2101	CORDOCTIA	ACIGICICU	A A A A A A A A A A A A A A A A A A A	COACCETTA	ALCICCUUT	ACAGAAGCOG
2221	GGATATAAAC	ATAAGACGAG	AAGAJUUTAT	GGAGCITIAG	ACAULAAAGA	ASAIUUIGIC.
ZZ81	AAGTAACCCC	TTATAAGGGC	CIGAACTAAT	GGAACCUTTC	CUTAAIGICI	TIGGTIGGG3
2341	CGACCGCGGG	GAAACAAAAA	ACCCCCACGT	GGAAAGGGAG	CACCCCCTCC	TACAACTAAG
2401	AGCCGCAGCT	CTAATTAACA	GAATATCTGA	CCAATAAGAT	CCGGCAACGC	CGATCAACGG
2461	ACCGAGTTAC	CCTAGGGATA	ACAGCGCAAT	CCCCTTTTAG	AGCCCATATC	GACAAGGGGG
2521	TTTACCACCT	CGATGTTGGA	TCAGGACATC	CTANGETCC	AGCCGCTATT	ABGGTCCGT
2521	TTETTCAACE	CTTANACTOC:	TACCTCATCT	CARTTCARAC	CURACTANTC	CACCTCACTT
0644	THUTTCAACO		TACOLORICE	ANACCARCE ANACCE	ANALIAATC.	
2041	ICTATUTATO	GIGIGUIUII	TICINGIACE	AAAGGAUUGA	ANAGAAGAGG	
2701	AAGCAAGCCI	CAUCUCUACU	TAGI JAAJAC	AACTAAAGTA	GUCAAGAGGG	CATACCCCCA
2761	ATGCCTGAGA	GAACGGCATG	TTGG3GT3GC	AGAGCCC3GT	GAATGCAAAA	GACCTAAGCC
2821	CTTTTTACAG	AGGTTCAAGT	CCTCTCCTTA	ACTATGATCT	CAGTGCTTAT	TACCCATATT
2881	CTTAACCCCT	TEGCCTTCAT	TGTCCCCGTC	CTCTTAGCCG.	TEGECTTEET	CACACTTCTA
29/1	GAACGTAAAG	ТАСТАВВАТА	CATACAACTA	CGAAAGGETC	CANATATTET	ABRECETTAC
2001	CCCCTATTAC	ACCENTRO	TOATOCTOTA	A A ACTOTTA	TTANACACCC	Tattadadada
5001 50C4	THELE	ABUULATIOU ABUULATIOU	IUNIUUIUIA ATTTOTOTOTA	AAACICIIIA Oggagaaaga	TIANAGAGUU	
SUD1	ICCACTICCT Notacimental C	CLUCULGIACT	ATTICIUCIU	GUUUUAUIGU	I CIGUACIUAC	BUILIBUULTA
3121	ACCCTCTGAG	CCCCCATGCC	CCTCCCTTAC	CCAGTCATCG	ACITAAACCT	IJGGATCCTA
3181	TTTATTTTAG	CCCTATCAAG	CCTCGCTGTC	TACTCCATCC	TGGGCTCAGG	CTGAGCATCA
3241	AATTCAAAAT	ATGCTCTCAT	CGGGGGCCCTT	CGGGCTGTAG	CCCAAACCAT	TTCATATGAA
3301	GTTAGTCTGG	GCCTAATCCT	ATTAAGTACT	ATCATTTTTA	CAGGGGGTTT	TACACTACAA
3361		TTGCCCAACA	AAGCIGTCITCA	ATACTTCTCC	CARCOTRACC	ATTAGCCGCA
0001 0101	ATATOOTATA	TIGULUHADA	TOCACION	ANCOURCE	COTTONCE	TILLOUCOCK
0421	AIAIGUIAIA	TITURAUUUI		AACUUIUCAU	CUITIGAUUI	
3481	GAATCIGGAAC	TAGTCTCTGG	CLICAATGIC	GAATATGCAG	GIGGCCCGTT	I SCIUTATTI
3541	TTCTTAGCCG	AATATGCTAA	TATICTECTT	ATAAATACAC	TITCCGCCAC	CUTUTTCTTA

Figure 5. Sequences of the mitogenome of S. owstoni (GenBank Accession No. KJ834063)

3601	GGGGCCTCTC	ACTTTCCAAT	ACTACCTGAA	CTCACCGCAG	TAAACCTGAT	AACCAAAGCA
3661	GCCCTTCTAT	CTGTCCTATT	CCTATGAGTT	CGAGCCTCTT	ACCCACGATT	CCGATACGAC
3721	CAGCTCATGC	ATCTAATTTG	AAAAAC'TTC'	CTCCCGCTTA	CACTAGECT	AGTTATCTGA
3781	CACCTARCCC	TOCCOATTGC	ATTTGCTGGC	CTREERCOC	AGCTATAGAT	AAGAAGCCIGT
0001	COOTENNET	A ADDOCTACT	TTCATACACT	CACTTATICC	COTTONANTO	COCOCCCT
0041	GUUIGAAJIA	ARGOGULAUI	I I GATAGAGI Nggagagaga	CACITATOGO	DOTICARATC	
1.065		GUUGA	ACCULUEUCIA	AGGAGAGCAA	AACICCIEGI	GUICUUACIA
3961	CACTATTICC	TAGTAAAGTC	AGCTAATICT	AAGCTCTTGG	TCCCATACCC	CAAACACGAA
4021	GGTTAAAATC	CCTCCTTTAC	TAATGAACCC	TTACATCTTA	ACCGCCCTGC	TATTTGGTAT
4081	TGGTTTAGGC	ACTACTACCA	CCTTCGCAAG	CTCCCACTGA	CTACTAGCCT	GAATGGGCCT
4141	GGAGATAAAT	ACTCTCGCCA	TCATTCCCCT	AATAGCTCAA	CACCATCACC	CCCGAGCAGT
4201	TGAAGCAGCC	ACTAAATATT	TCTTGATTCA	AGCTGCC3GG	GCAGCTATGC	TACTCTTTGC
4261	CARCACCACC	AACGCTTGAT	TAACTGGACA	ATGAGACTTC	TTACAAATTG	COCACCCTT
1201	CCCAACTICT	CTTGTCACTT	TERCOTTOR	ACTABASTG	GRACTTECAC	CTGTACACTC
1001		CANCEACTE	AACCOCTACA	CCTAACTACA	COACT TOCAC	
4001	AIGACIACCI	CARGIACTIC	COTTACT	CURACIACA	TOTOCONICA	CONCORT
4441		- GUUULAITIG	CULTAITAGE	ACCARACULUU	- ISIGULAALA	LCAULTIII
4501	AGITATICIA	GGGCTTACCT	CAACCATIGI	AGGAGGUIGA	GJGGGGACICA	ACCAAACCUA
4561	ACTTCGCAAA	AICCITECCT	ACTECTECAT	CGCACATCTA	GGCTGGATAG	TAATTGTACT
4621	ACAATTCTCC	CCCTCCCTAA	CTATTTTAAC	ATTATTTACA	TACTTCATTA	TAACCTTCTC
4681	AGCATTTCTT	ATGTTTAAGC	TTAATAAAGC	AACCAGCATT	AATGCTCTAG	CAACCTCATG
4741	GGCAAAAACC	CCCGCCCTAA	CCGCCCTTGC	ACCCCTACTA	TTATTATCCT	TAGGGGGCCT
4801	CCCCCCACTG	ACAGGCTTTA	TGCCAAAGTG	ACTTATTCTT	CAAGAACTTA	CTAAGCAAGA
4861	COTTGOCOCC	GOTECAACAC	TGGCGGCGAT	AACCGCCCTC	CTTAGCCTAT	ATTTTTATCT
1001	GCGACTATCA	ТАСІССІААТАС	САСТААСТАТ	TTCACCCAAC		CANTTOCCO
4021	ATCACCCCTC	COUTCOTTAC	AACTAACIAL	ACCACCERSE	ACTICACCO	TACCTACCCC
4201 5044	COTOCUTO	COULCUITE		ACCACILICE	ACCOUNTING	CONCERCOUL
5041	BUIGUIIUIA		ULGUUGUAAI max xx aaaam	AGCACIANIA Nomelogiam	ACCULITANG	GSACTIAGGI GNAMAGOMAN
5101	TAAAACAAGA	CCAAGGGGCUI	ICAAAGUCUI	AAGIGAGGGI	6.5AAGICCCC	CAGICCUIGA
5161	TAAGGCTIGC	GGGACACTAC	CCCACATOIC	CIGIAIGCAA	AACAGGTACT	TTAATTAAGC
5221	TAAAGCCTTC	CTAGAAGGGC	AGGCCTCGAT	CCTGCAAGAT	CTTAGTTAAC	AGCTAAGCGC
5281	TC'AAAC'C'AGC'	GAGCATCCAT	CTACITTICC	CCCGCCTGAC	GGGCGGGCGG	AGGCGGGGGA
5341	AAGTCCCGGC	AGACGACTAA	CCTGCATCTT	CAGATTTGCA	ATCTGATATG	TATAACACCT
5401	CAAGACTTCT	GGTAAGAAGA	GGATTCAAAC	CTCTGTTTGT	GGGGCTACAA	TCCATCGCTT
5461	AAAAACTCAG	CCATCCTACC	TGTGGCCATC	ACACGTTGAT	TTTTCTCCAC	ΤΑΑΤΟΑΟΆΑΑ.
5521	GACATOGGCA	CCCTTTATCT	AGTATTTGGT	GCCTGAGCCG	GTATAGTAGG	CACAGCCCTC
5581	AGCCTACTCA	TTCGAGCAGA		CORRECTO	TOOTTGGAGA	CIGACICIAAATT
56/1	ТАТААТСТСА	ТОСКОСКОК	ACATICCER	CTAATCATTT	TOUTTEATACT	AATCCCAATT
57041		CTTTTCCAAA	CTCATTATC	CCCCTANTCA	TOPINIAGO	AGATATACCA
D701 E701	MICALICSAC TTALATIONAL		A A A A A A A A A A A A A A A A A A A	CULUTAAIGA ATTATAAAAA		AGENCENCE.
D/01	- TICCUTCSTA	TAAATAACAT	AAGITTUTGA	UTICIACCCC.		GCIACIACII
5821	GCCTCCTCTG	GAGTAGAAGC	AGGTGCUGGA	ACCEGETEAA	CAGIGIACCC	ACCCUTGGCT
5881	GGTAATTTAG	CCCACGCAGG	AGCATCAGTC	GACCTGACAA	TETTTEACT	TCACCTGGCA
5941	GGTATTTCCT	CAATTCTTGG	GGCAATCAAT	TTTATTACCA	CAATTATTAA	TATGAAACCT
6001	CCAGCCATCT	CTCAATACCA	AACACCCCTG	TTTGTGTGAG	CCGTCCTAAT	TACCGCTGTT
6061	CTTCTCCTTC	TETEETTACE	AGTCCTTGCT	GCCGGCATCA	CAATGCTCCT	TACCGACCGA
6121	AACCTTAATA	CCACCTTCTT	TGACCCGGCC	GGAGGAGGAG	ATCCAATCCT	TTACCAGCAC
6181	TTATTCTGGT	TETTTGGACA	CCCGGAAGTA	TATATTCTCA	TTCTGCCTGG	CTTTGGTATG
6241		TOTTOGICTA	TTACTCTGGC			TATGGGTATA
6301	GTGTGAGGAA	ТААТСССТАТ	TERCOTOTTA	CCCTTTATTC	TATEAGCTCA	COMPAREMENT
6261	ACACTTCCCA	TARIOUCIAL	CACACCTUR	TATTTACCT	CTECEACIAN	AATCATCCCA
6401	ACAGIIGSCA	ATATTA ANAT	ATTALGISCI		TIGULALAAI	ARICALCOCK ATTOTATTAAA
64ZI	ALICCCACCE	GIGIIAAAGI	ATTIAGUIGA	UTIGUAACUU TTTTTTTTTTT	TACATGGGGG	UTUTATTAAA
6481	TGAGAAACAC	CCCTTCTAIG	GGCCCTTGGC	TITATITICT	TATTTACAGT	AGGCGGGCTT
6541	ACAGGCATTG	TICTGGCCAA	TTCATCTCTA	GATATTGTAC	TCCACGATAC	CTATTATGTA
6601	GTAGCCCACT	TCCACTACGT	ATTATCTATG	GGGGCCGTAT	TTGCCATTGT	CGCCGCCTTT
6661	GTGCACTGAT	TCCCACTATT	CTCAGGCTAC	ACACTTCACA	GTACTTGAAC	GAAAATCCAC
6721	TTCGGTATTA	TGTTCTTAGG	GGTAAACTTA	ACCTTCTTCC	CACAACACTT	CCTCGGATTA
6781	GCCGGAATGC	CCCGACGATA	CTCTGACTAC	CCTGACGCCT	ATACCCTATG	AAATACAGTC
6841	TCCTCAATCG	GATCACTTAT	CTCCTTAGTA	GCTGTTATTA	TGTTCTTATT	TATTATTTGA
6901	GAGGCATTCG	CCGCCAAACG	TGAAGTTOTA	GCAACAGATT	ТААСААСААС	CAATGTAGAA
6961	ТСАСТАСАТС	GTTGCCTCC	CCCATACCAC	ACATTORAGE		ТЗТАСАЛАТА
7021	CARCINCATO		GGAGGAGTCC	A A C C C C C A T A	CTCCCTOCCTT CTCCCCTTTC	A ACCCCACCA
7041	CARDUSCAUL					AUCCUBLER
1081	CATAACCGCI	CIGULACIII	CITIATAAGA	CACIAGIAAA	RUABBABBAB	RURUUBUUII

7141	GTCAAGGCGG	AAGTGTGGGT	TAGACCCCCG	CGTGTCTTGC	TTTTAATGGC	CCATCCGTCA
7201	CAGCTTGGAT	TTCAAGATGC	AGCTTCACCT	GTTATAGAAG	AACTTCTTCA	TTTTCACGAC
7261	CATGCTTTAA	TAATCGTCTT	CCTAATTAGC	ACACTTGTGC	TTTACATTAT	TCTTGCTATA
7321	GTTACCACTA	AATTAACGAA	CAAATATATT	TTAGATTCAC	AAGAGATTGA	AATTATCTGA
7381	ACAATTCTCC	CAGCTATTAT	TCTAATTCTA	ATTGCACTCC	CCTCCCTCCG	CATCCTCTAC
7441	CTTATAGATG	AGATTAATAA	CCCCTTATTG	ACAATTAAAG	CCGTTGGCCA	CCAATGATAC
7501	TGAAGCTATG	AGTACACTGA	CTACGAAGAC	CTTGGCTTTG	ACTCATACAT	AATCCCACC
7561	CAAGACCTAA	COCCTERACIA	ΔΤΤΟΡΩΟΡΤΔ	TTAGAAGTOG		GGTTATTCCA
7621	CTTCAATCCC	CTATCCGACT	CTTATTATCT	CCACACCATC	TACTCCATTC	ATCACCACTC
7681	CCACCCTCC	CIRICOMOL	CINCIPATA	CCACCACCAC	TTAATCAAAC	ACCTTCATC
7001	CCAUCCCIOU	CACCCATATT	CTACICACIA	Тастсталлл	TINGICARAC	AAATCACACC
7001	TTTATACCTA	TTETACTECA	ACCARTICC	CTACAACACT	TTEANAACTE	ATCATCTCC
7061	ATACTTCAAC	ACCOUNT OF A	ACCASILLU	ATACCCCATA	COOTTACO	TTTAICIUGE TTTAACCTAA
7001				CATCOCCATA		
7941	TOCTATION	CICULAAULA	CANTOCTOTT	CAIGUUUUAA	A TTOOODOTA	VACUIIGAII
7961 0044			GAALSGIGII GTAATGTGGA	ULIBBUUSII	ATTUUUGUTA	
8041	CUACACUTIC	CCAAATACCC	CIAAICIGCA	AAGUGUAJAA	AAAUUUAAAA	CAGACCCCTG
8101	AACCIGACCA	TGACACTAAG	CITITIAC	CAGTITATAA	GCCCCACCTA	TOTIGGAATO
8161	CCACTAATAG	CCCTTGCCCT	TACCITACCC	TGACTCCTTT	ACCCCACACA	TACAACTCGA
8221	TGATTAAATA	ACCGATTCCT	CGCGCTTCAG	GGTTGATTTA	TCAACCGTTT	TACTCAACAG
8281	CTCCTCCTCC	CCTTAAATAT	TGGAGGCCAT	AAGTGAGCCG	CCCTCCTGAC	CTCATTAATG
8341	ATCTTTTTAA	TTACCCTAAA	TATGTTAGGA	CTTCTTCCCT	ATACTTTTAC	CCCCACCACC
8401	CAACTATCAC	TAAATTTAGG	GCTTGCGGTA	CCTCTCTGAT	TAGCAACTGT	TATTATTGGC
8461	ATGCGAAACC	AACCAACCCA	TGCCCTAGGA	CACCTCCTAC	CAGAAGGCAC	ACCTGGCCCC
8521	CTCATCCCCG	TACTTATCAT	TATCGAAACA	ATTAGCCTCT	TTATTCGCCC	CCTTGCCCTA
8581	GGAGTACGAC	TAACGGCCAA	TTCAACAGGT	GGCCACCTCT	TAATTCAAGT	AAATGGTAC'A
8641	GGTGCATTCG	TACTTCTCCC	CTTAATACCA	ACCGTCGCAA	TCATCACAAC	AACAGTACTG
8701	GTTCTCCTCA	CCCTATTAGA	AGTTGCTGTA	GCAATAATTC	AAGCCTACGT	CTTCGTTCTC
8761	CTACTAACAC	TATACCTACA	AGAAAACGTC	TAATGGCCCA	TCAAGCACAC	CCTTACCACA
8821	TAGTTGACCC	CAGCCCTTGA	CCCCTAACAG	GGGCAATTGC	TGCCCTCCTG	ATAAC'ATC'AG
8881	GCCTCGCAAC	CTGATTTCAT	TTTCGCTCAA	CAACCTTAAT	AACCTTAGGA	ACAGCCCTAC
8941	TGCTTCTTAC	AATATATCAG	TGATGACGAG	ACATCGTACG	AGAGGGTACA	TTCCAAGGAC
9001	ACCACACGCC	CCCCGTACAA	AAAGGTCTTC	GATACGGAAT	GATTCTTTTC	ATTACCTCCG
9061	AAGTATTCTT	TTTCCTAGGA	TTCTTCTGAG	CCTTTTACCA	CGCGAGCCTC	GCCCCTACTC
9121	CTGAGCTAGG	GGGCTGCTGA	CCTCCCACAG	GCATTACAAC	TETTGACCCA	TTTGAAGTCC
9181	CCCTCCTTAA	TACAGCTGTC	CTGCTTGCTT	CTEEAGTAAC	GGTCACCTGA	GCCCACCACA
9741	GCATTATGGA	AGETGAACGA		TCCAATCACT		ATTCTTCTAG
9301	GCTTTTATT	TACATTTCTT	CAAGCCCTGG	AATACTATGA	AGCCCCCTTT	ACAATTGCAG
9361	ATGGCGTATA	CGGCTCTACA	TTTTTCGTAG	CCACTEGATT	CTACGGACTA	CACGTTATTA
0/21	TTGGCTCCAC	ATTTTTAGCT	GTCTGCCTCC	TACGACAAAT	CTAATACCAC	TTTACATCCS
Q181		CGGGTTCGAA	GCACCTCCCT.	GATACTGACA	TTTCGTAGAC	GTTGTCTGAT
9541	ТАТТССТАТА	TATCTCTATC	TACTGATGAG	GCTCTTAATC	TTTCTAGTAT	TAAAACTAGT
9541 9601	ATAAGTGACT	TOCATCACC	CERTITIET	TAAAATCCAA	GIANGATAA	TRAMETAC
Q661	TATACOTORCI		CTATTTTCCT	TTCCGTAGTC	CTERCCATTE	TATCCTTCTC
Q721				GOTOTOCA	TATEAATETE	GTTTCGACCC
0791	CTTCCCCCCA	CTCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	CATTTTCCC	CORCICICCON	CTACTOCCA	TTOTOTOT
02/1	COTTOUGATER	ΤΤΛΩΛΛΑΤΤΩ	CONTRACT:	CCCUTTCIC	TARICOLLA	AATTAACCTC
0001	COUTTEATE	ACACTCTTCT	CACCITCICCI CACCCCTCCC	COMPACTOR	CTTCTTACCC	TTEECCTART
9901	TTACCACTCA	ATTCAACCAC	CATTACAATC	ACCCCANTAC	CITCITACCC	TAACAAAAAT
10021	ATTTCATTC	COCTCAAGGAG	CTTAGARIG	ABUUDAAIAD	TETETATIA	ACTOCOCOCTO
10021	ATTIGATIC	GUILAAAAG atalitaaaaa	TTAIGGIIA	AAGICCAIAA alamalalaa	COTACONTRA	ASICCUGUIS
10081	ACTICUTIT	UIUAIUGGUU TOOTOTTTT	TITALULIAG	GALIGALAGG		ATTCHARLES
10201	ACCILCICIC: COATTECARD	TOUTUIIIA COTOBACTA	TGUIIABAAB	OUTRICATION OF THE OUTPIT OUTPIT OF THE OUTPIT OUTPIT OF THE OUTPIT OUTP	CIUIIIAIII	AIIBBAUIII Aradrammaa
TUZUL 10061	CUATTIGAAC	ULIULAAUIA	AGGGLCCACAA	GITTICIGC TROCTTREACT	GUILUIAIA	CIUCIALIGS
10261		TIGIGAAGCA	AGCGCAGGGC	TIGUITTACT	GUINANAN	GCIUGUAUGC
10321	AIGGTTCAGA	TUGUUTTUAA	ACCITAAACC	ICITACAATG	CIAAAAATTC	TAATTCCCAC
10381	CCTAATGCTT		CONGACTIGC		I JATIATGAC	
10441	ATCCCACAGC	CTAGTTATTG	CGCTGGCCAG	COTCACTIGA	CTAAAGAATA	CATCTGAAAC
10501	AGGUIGAICT	I JCUTCACGC	CETTEATAGE	CACAGATECC	CIUTUAACAC	CICICCITUT
10561	TETTACCIGC	TJACTACTIC	CTCTTATAAT	TTTGGCAAGC	CAAAGCCACA	CAGCACTAGA
10621	ACCTATTAAC	CGCCAACGAA	CCTACATTAG	CCTGTTAACG	TCTCTGCAAG	TATTCCTTAT

10/201	TATACIATTC	COTOCTACTO	a a concuertra m	OTTTTATOTT	3 T 3 T T T T 3 3 G	CIT & CITCITT & T
TODST	TATAGCATIC	6516CIACI5	AAUILUIIAI	GIIIIAIJII	ATATTIJAAG	CIACICIIAI
10741	CCCCCACACTA	ATTATTATTA	CTCG3TG3GG	CAACCAGGCA	GAACGCCTTA	ATGCAGGAGT
1.0801	ATATTTTTG	TTTTATACCC	TAGCAGECTC	TCTCCCATTA	CTAGTIGCCC	TOTTGOTTOT
10061	THANADOAT	Adaddadad	TATAAATATT	- X 200 X 200 X 200 X X	TATACTACCT	OTACCOPTOT
TOODI	TCAAAAGGAI	ACABBUILLE		AAUUAIUUAA	TATACTASCT	CIACULUICI
10921	TTCATCTTAT	GCTGATAAAC	TTTTGATGAGC	AGGTTGCCTA	A FTGCATTTT	TAGTAAAAAT
10981	ACCCTTATAT	GGAGCACATC	TETESETACE	AAAAGCACAT	GTAGAAGCCC	CAGTTECAES
11041	CTCANTOTT	CTACCTCCAC	TTOTTOTAAA	ACTACCCCCC	TACCCTATCA	TOODAATAAT
11041	CICAAIDSII	CIABUIDUAD	TICTICIANA	ALLI ALGIGIGIGIGI	TAUGGIAIGA	I.J.GAAIAAI
11101	AGTCATATTA	GAACCTCTCA	CCAAGGAATT	AAGCTATCCC	TTTATTGTCC	TAGCGCTCTG
11161	AGGTGTAATT	ATAACTIGGCT	CCACCTRCCT	TCIGCICIAAACIA	GATETTAAAT	COOTCATORO
11001	OT A TTCA TCC	CTAACCONTA	Tadadaraar	COTTOCATOO	ATTOTTATOO	AGACIACICCE
	CIAIICAICC	GIAAGULAIA	1555001351	Clat RalaAlalata	ALICITATEC	ASACAUCUIS
11281	AGGTCTTGCC	GGCGCTGTAA	TCCTTATGAT	TGCACACGGC	CIGACGICCI	CGGCCCCCTCTT
11341	CTECTTEECC	AATACAAACT	ATGAACGCCT	CCATAGCCGA	ACAATACTAT	TAGCCCGGGG
11401	ATTACACA	OTCOTTOCAC	TCATAACAAC	ATCATCATT	ATTRACTACIÓN	TOCONNOTT
11401	ATTALAGATA	GIGUIIUUAU	ICATAACAAC	AIGAIGAIII	ATTIGULABUL	I.J.GUARAU I I
11461	AGCCCTTCCC	CUTCTGCCTA	ACCICATGGG	GGAACTETTA	ATTATTACCT	CATTATTTGG
11521	C'TGATC'ATGA	TGAACTCTCG	TACTCACAGG	GACAGGGACC	CITATIACCG.	CGAGCTATTC
11501	ACTTTATATC	TTCCTCATCA	COCANCERCE	TOCOTOCA	CCACATATTA	TTARCETAAA
11001	ACTITATATO	INCOLORIDA	LCCCAACOSOG	ICCCCICCCA	GUACATATIA	TIAGUCIAAA
11641	CCCCTCCTAT	ACCCGGGGAAC	ACCTAGTTAT	AGCCCTTCAC	CICCICCCGC	TACITCIACI
11701	TGTTTTAAAG	CCCGAATTAG	TATGAGGCTG	AACCACCTGT	AGATATAGTT	TAACAAAAAT
11761	ATTACATTCT	CATTOTAAAC	ACACACCTTA	AATCCCTT	ATCCACCEAC	ABARCETCA
1101		ONTICIANO	areagedia	THATCOCCIT	agayagaga	Nonoocicoc
118ZI	CAGCAACGAA	GACIGCIAAI	CICUSCBACU	TIGGIIGJAC	CCCAPPECIC	ACIUGGCUIS
11881	CTCCTAAAGG	ATAACAGCTC	ATCCATTGGT	CTTAGGAACC	AAAAACTCTT	GGTGCAAATC
110/1	CAAGTAGCAG	CTATECACTC	CTCATCACTT	δΤΤΔΤΔΤΎΔΤ	COAGCTTAGT	CATTATOTT
10001	TTACTACTAC	CIATATOCACIC	CICHICHOIT ATTENADAAT	ATTAINIATOAT	dedeedaa M	CATINICITI
TZÖÖT	TTACTATIAG	CATAICUTAI	CITTACGACI	CIGGAGECIE	GUUUUUGAAA	CCCIGAAIGS
12061	GCCGTTTCAC	ATGTTAAGAC	AGCGGTAGCC	CTGGCCTTCT	TEGTEAGEET	AATCCCCCTA
12121	TTTCTCTTC	TTAACGAAGG	CGCAGAAGCA	ATCATCACCT	CATGAAATTG	AATGAATACA
10101	CTANCOTTO	ACCTCANCAT	тастттала	TTTCATCATT	ACTCACTTAT	CTTTCTCCCC
12101	UTRACUTIUG	ACGIGAACAI	INGITTIANA	TITOATCAIL	ACICAGITAT	
12241	ATTGCCCTCT	ACGTCACTTG	ATCTATTCTA	GAGTTTGCAT	CATGATATAT	ACACGCAGAC
12301	CCATACATAA	ACCGATTCTT	TAAATATCTC	CTAATTTTCC.	TTATTGCCAT	AATCATTCTT
12261	GTCACACCA A	<u>Α ΤΑ ΑΤΟ</u> ΤΑΤΤ	CCACCTTTTC	ATTORTION	ABCCACTACC	саттатетса
12.001					100001000	
12421	TITCIACICA	LUGGUUGALG	ALAUGIGAUGA	GUGGATGUUA	ACACABCBBC	CLITTUAGGUU
12481	GTTGTGTATA	ATCGGGTCGG	AGACATTIGA	CTGCTATTCA	CAATAGCATG	AATAGCAACC
12541	AACGCTAACT	CCTGAGAGCT	ACAACAAATT	TTTGTGGCAA	CAAAAGACCT	GGATCTTACT
10601	TTACCCTAC	TAPPOPTAT	TATTOCCOCT	ACACCOACT	COPPOPERANT	TOCTOTOCAC
12001	COMPOSITION	CONCILAT	1AII3CC3CI	ACAGGCARGI	GEGGGGGGAN	13GICICCAC
17001	CCITGACITC	CULUIGUIAI	AGAGGGICCI	ACAUCGGICI	CIGUUCIACI	ACALICAAGC
12721	ACCATGGTTG	TIGCCGGCAT	TTTTCTCTTG	GTACGAACAA	GCCCCCCCCC	AGAGAATAAT
12781	CAAACTGCCC	TOACTACCTG	CCTATGCCTA	GGTGCCCTAA	CRACCCTATT	TACACCCACC
10011	TATAGATIN	GOGLARATO	TATUANAAAA	ATCOTOCCCIAN	TORCOLLAR	AAAAAAAA
12841	IGIGUUIIAA	CUCAAAAIGA	TAICAAAAAA	AICGIAGCAI	ISIULAUAIU	AAGUUAAUII
12901	GGCCTAATAA	TAGTTACCAT	CGGCTTAAAT	CAACCTCAAC	TAGCCTTCCT	CCACATCTGC
12961	ACCCATCCCT	TOTTCAAAGC	AATGTTGTTC	CTCTGTTTCCG	GCTCAATTAT	CCACAGCCTC
13021	AACGACGAAC	ABRATATCCC	AAAAATAGGG	GCCATACACC	ACCTTACCCC	CTTTACATCO
10001	TACOACOARC	ATATTACCO ATATTACTA	TTTACCOTC	JOUCHINCHCC		A GAREOTEC
13081	TUUTGUUTUA	CIAIIGGIAG	TITAGUUUTU	ACAGGIACIC	COLLCUIAGE	ASGAILCIIC
13141	TCCAAAGATG	CCATTATTGA	GGCACTAAAC	ACATCCCACC	TAAACGCCTG	AGCCCTAGTC
13201	CTAACCCTTC	TARCCACCTC	ATTEACABOO	ATCTACABOC	THERET	<b>ATTTTTTGT</b>
10061	TCANTACCITC	ACCACCAC	TARCACE	TOTOTATON		COOMPONDER
13/01		ACCURUITATI	TAACOUTATI	TUTUUTATUA	AILTADAALAA	
13321	ATTAACCCCT	TAAAGCGACT	TGCATGAGGA	AGCATTGICG	CIEECCICCT	AATCATTTCA
13381	AGCATTACCC	CCCTCAAGAC	CCCTGTGATA	TCTATACCCC	CCTTGCTCAA	ACTAGCTGCC.
13//1	CTCCCACTTA	CANTTACCCC	CCTACTCATC	CCCCTCCAAC	TACCACACT	<u>A ACCANTAAA</u>
10501	CICOGAGIIA	CANTIACOOD	agmigmmigg	GUUCICONNC.		ACCARIAAA
13201	CAGIACAAAA	TTACCCCCAA	CCTASTIACU	CACCACITCI	CCAACAIGCI	AGGUITCITC
13561	CCCTCGATTA	TTCACCGATT	TACCCCTAAA	CTAAATCTAG	TCTTAGGGCA	GACATTTGCC
13621	AGTCAACTAA	TTGACCAAAC	TTGATTAGAG	AAAGTCIGGCC	CCAAAGCAAT	C'TC'TTC'ATC'A
12601	A A CIATOCOCC	TANTTACAAC	AACAACCAAC		CANTANTTAN	CACATACCTC
10001	AACAIUUUUU	TARTIACAAC	AACAAGCAAC	ACACAACAAD	GAATAATTAA BEN GAGGEE	DACATACULU
13741	ACCUTATICC	TICTCACCCT	GACCENTIGCT	GCCCTATIGT	TTACCCGTTA	AACTECCCBA
13801	AGAGCCCCCC	GACTTAGTCC	TCGAGTTAAC	TCCAACACAA	CAAATAAAGT	GAGAAGCAGA
13861	ACCEACECAC	ТААСТАССАА	TATCOTOC	CCTAGTGAGT	ATTARCAC	AACCCCTCCA
10001	ATATACOCAC	TUNOINCON	ANCOURCE	AGGEGINGIONOI	UQQQQ3QQQ3	THEFT
10941	ATAICGUEIU	GCAGAACAGA	AAGUICACIA	AGUICAICAG	CLUBBLALLUA	IJAAGACICA
13981	TATUACUUCC	CICAAGGCAC	ATTACAAGCC	ACCCCCACCC	CCATGAAGTA	CATCAACATG
14041	TCACTTACAA	CAGGACCAAG	CACTCAACTA	TCAGGATAAG	GATCAGCAGC	AAGTGCCGCC
14101	GAGTATOTAA	<u>ממדמממי</u> קמייה <u>מ</u>	TATACCACCC	AAGTAAATCA	אאאריאאריארי	ТАНТНАТАНА
1 / 1 / 1 / 1	JANGETCOM	- ANTONACINA	- THIRDOROCO	- ANOIMANICA - ANOIMANICA		TACTORIACE
14101	RAAGGIUGUU	CATGAGUAAU	CANTAUTOCA	CAUCUUATGU	CORCEAURAC	TACTAACCCC
14221	AAGGCAGCAA	AGTAGGGAGA	AGGGITAAAG	GCAACTGCAA	CCAACCCTAA	AACTAATCCA

4.4004	1000111111111	3 3 3 3 4 3 7 3 7 3 7	OTALOTO		a a a a a a a a a a a a a a a a a a a	
14281		AAGACATAAT	GIAASICAIA mmammaaaam	ATTUCTUCCA	GGACITIAAC	
14341	GGUTIGAAAA	ACCACCETTE	TTATICAACT	ACAAGAACCC	ACTAATGGCA	AGICIACGAA
14401	AGACACACCC	TCTCCTCAAA	ATCGCAAACA	ATGCCCTAGT	TGACCTACCC	GCCCCCTCAA
14461	ATATITCAGT	GIGATGAAAC	TTCGGATCTC	TCTTGGGACT	CIGCITAATT	ATTCAAATCC
14521	TCACAGGACT	ATTITITAGCC	ATGCACTACA	COTOTGATAT	TECTACEECC	TETTCETCC3
14581	TTGCTCATAT	CIGCCGGGAC	GTAAATTACG	GCTGACTCAT	CCGAAACCTC	CACGCCAACG
14641	GTGCATCCTT	CTTCTTTGTA	TGCATCTATG	CCCACATTGG	CCGCGGACTT	TACTACGGCT
14701	CATACCTCTA	TAAAGAGACA	TGAAACATTG	GAGTAGTTCT	GCTACTTCTA	GTTATAATAA
14761	CTGCTTTCGT	CJGCTATGTA	CTACCCTGAG	GCCAAATGTC	CTTTTGAGGT	GCAACCGTTA
14821	TTACTAACCT	ACTCTCTGCA	GTACCCTACG	TAGGTGGCTC	TCTGGTTCAA	TGAATTTGAG
14881	GTGGATTCTC	AGTAGACAAT	GCAACCCTTA	CCCGATTCTT	TGCCTTCCAC	TICCTATICC
14941	CCTTTGTAAT	TGCAGGCGCA	ACCATAGTCC	ACCTTCTTTT	CCTTCACCAA	ACAGGGTCAA
15001	ATAATCCCCT	C GGCCTTAAT	TCAGACGCAG	ATAAAATAAG	CTTCCACCCC	TATTTTTCAT
15061	ACAAAGACCT	ATTAGGGTTT	GCAGTACTTG	TCATTGCCCT	CACATGTTTA	GCCTTATTCT
15121	CACCCAACCT	GCTAGGAGAC	CCAGACAACT	TCACCCCCGC	CAATCCGCTA	GTCACTCCTC
15181	CCCACATTAA	ACCAGAATGA	TACTTCCTGT	TCGCGTACGC	AATTCTACGC	TCCATCCCCA
15241	ATAAACT3GG	GGGAGTCTTA	GCCCTCCTGG	CCTCAATCCT	TATTCTGATA	CTCGTACCAT
15301	TTCTACACAC	GTCTAAACAA	CGAAGCCTCA	CTTTCCGACC	ACTTACACAA	TICTIGTIT
15361	GAACCCTAAT	CGCAGACGTT	ATTATTCTCA	CTTGAATTGG	AGGAATGCCC	GTGTCACACC
15421	CATTCGTCAT	TATTGGACAA	ATTGCATCCT	TTTTATACTT	TITCCTCTTC	CTAGTCCTTA
15481	CGCCACTAGC	AGGCTATGCA	GAAGATAAAG	CACTTGAATG	AGCTTGCATT	AGTAGCTCAG
15541	CGCCAGAGCC	CTGGTCTTGT	AAACCAGATG	TCGGAGGTTA	AAATCCTCCC	TAATGCTCAA
15601	AGAAAGGAGA	TITTAACTCC	CACCCCTGGC	TCCCAAAGCC	AGGATTCTTA	GTTAAACTAT
15661	TCTTTGTATT	ATATGTATAA	ΤΑΑΤΤΤΤΑΑΑ	TACATATATG	TATTATCAAC	ΑΤΤΑΑΤΤΤΑΤ
15721	ATTAACCATA	TCATATAGCA	TTCAAGTACA	TACATGTATT	ATCACCATAT	CTAGGATTTA
15781	ACCATTCAAG	AATTACACGA	AACGAAAAAT	CTTACGCAAA	ACATAACAAC	AAAAATCAAT
15841	AAACACTTAA	AAATACCAGG	CGACATTTAA	AACCTAACAC	AAACCCCCAT	GAGTCAAGTT
15901	ATACCTTTAC	TCAAATTCCC	GTCAAACTCA	ΑΑΤΑΤΤΤΑΑΤ	GTAGTAAGAG	CCGACCAACA
15961	AGTCCATTTC	TTAATGCCAA	CGGTTATTGA	AGGTGAGGGA	CAAAAACTGT	GGGGGTTTCA
16021	CACGGTGATC	TATTCCTGGC	ATTTGGTTCC	TATTTCAGGG	CCATAAATTG	TAAACATCCC
16081	CATAACTTAT	TCTAAAAGGC	ATAAGTTAAT	GGTGGAGAAC	AATAGCGGGA	GCGGCCACCA
16141	TGCCGAGCGT	TOTTTOCATA	GGGCATTTAG	CTCTTTTTTT	TTTTTTCCT	TTTCATAGA
16201	CATTTCACAG	TGCACGCAAT	CTAGTTAACA	AGGTGGGAAT	AATCCTAGGA	AGCAAGGAAA
16261	TAGTATECGT	GATGAAAAGT	СТТААСТААА	GAATTACATA	TAGAACTTTC	AAGGACATAA
16321	GATAGTGAAA	TTTAGTCGGA	AGATATCTAT	ATTACCCCCCT	TTTGGCTTTT	TEGEGTTAAA
16381	CCCCCCTACC	CCCCTABACT	CCTGAGATAA	CTAACGCTCC	TGTAAACCCC	CCGGAAACAG
16441	GAAAACCTCG	AGTCGTTTTT	ATGGT	OTUROOTOO	10111110000	000011110110
TLLOT	Ormmice 100	1010011111	111001			

\*H and L denote heavy and light strands, respectively.



Figure 6. A gene arrangement in the complete mitogenome Sebastes owstoni.

노랑볼락의 미토게놈은 13개의 단백질 암호화 영역, 2개의 리보좀, 22개의 tRNA 및 1개의 조절영역으로 이루어져 있었고, 전체 크기는 16,450 bp 이었다(GenBank Accession number KJ834060)(Figure 7). 노랑볼락의 미토콘드리아 DNA의 염기 조성비는 27.7% (A), 26.6% (T), 17.3% (G) 및 28.4% (C)이었고, AT 비율이 약간 높았음 (54.3%), COX1 (GTG)을 제외한 모든 단백질 암호영역의 개시코돈은 ATG이었고, 종결코돈은 TAA가 주를 이루었으나 ND1과 ND3는 TAG, ND4는 AGA였으며 Cytb는 종결코돈이 불완전(T)하였다(Table 4). tRNA형태는 tRNA<sup>Ser(AGY)</sup>(lacked DHU-arm 구조)을 제외하고 모두 전형적인 cloverleaf 구조로 되어 있었다(Figure 8).

1	GCTAGCGTAG	CTTAATTAAA	GC'ATAAC'AC'T	GAAGATGTTA	AGATGGACCC	TAGAAAGTCC
61	CGCCCGCACA	AAGGCTTGGT	CCTGACTITA	CTATCAACTT	TAGCCAAATT	TACACATGCA
121	AGTATCCGCC	CCCCTGTGAG	AATGCCCTAC	AGCTCCCTGC	CCGGGAGCAA	GGAGCTGGTA
181	TCAGGCACAC	ATTTGTAAGC	CCATGACACC	TTGCTTAGCC	ACACCCTCAA	GGGAACTCAG
241	CAGTGATAAA	CATTAAGCCA	TAAGTGAAAA	CTTGACTTAG	TTAAAGCTAA	GAGGGCCGGT
301	AAAACTCGTG	CCAGCCACCG	CGGCTATACG	AGAGACICICAA	GTTGATACC'A	TTCGGCGTAA
361	AGAGTGGTTA	TIGAAAATAA	AGAC'TAAAGC'	CGCACACCTT	CAAAGCTGTT	ATACGCATCC
421	GAAGGCTAGA	AGATCAACCA	CGAAGGTAGC	TTTACAACCC	CTGACCCCAC	GAAAGCTCTG
181	GUACIALACTE	GGATTAGATA	CCCCACTATE		A ACCTTEGTA	
5/1	TACCCTACC	CONTRACAIN	TACCACINIC	ACCTTANAAC	CONARCACT	TERRETERT
601	TTAGACCCCC	CITAGAGGAGC	CTGTTCTAGA	ACCIGATAACC	COCGTTCAAC	CTCACCCTTC
661	CTTCTTTATC	CARCANOLOC	ACCEPTEINCE	TCACCTTACC	CTCTCAACC	СТААЛАСТАА.
721	CITOTIATC	CACAACCCAA	ACCOLOTICO	CONCEPTION	CICIONAGO	CIARAGUAR
701	CCCCCACARCIGG	CHURACUURE	CCCAACACCA		COCATOGAGO	COTCARCARAT
0/1	CATTACTAC	TARCOLATATIA	ATACCCTTTT	COOCTONNA	CARACECOLO	COTORAGORO
041	ACCOCCCCCCC	ACTOTOCOCA	ALAGUGIJII	TTTAATAAT	TAAAAACCCA	AAAATCCCCC
901	ACCOLLET	ACTUICUUA	TOOTALCAC	ACCCC A ACCC	CONCEPTION A	ARAAIUGUGG ATATCACACT
1001	AUGUUAUJUA	AGILGIAACA	IGGIAAGOGI ATTOCOTTIA	ACCEBAASEI	CLACIIGDIA	ATAICAGAGI
1.001		TABAATAALA GINAGAGENG	CITULLIIAU	AL ISAABABA	CAULUSIBLA	ARIUSSAIUA TTATA ACCCC
1141	A A TOOLAGUU	CHACAGUIAG		ATTTTTTCCC	TTTACTATION	TIATAAUUUU
1141	AAAIGUAJGA	GIGITITAAT		ATTTTUCCC	TITAGIAIGG	GUGAUAGAAA
1201	AAGGACTIAG	GAGUAATAGA	GAAASIACUS	CAAGGGAICG	UIGAAAGAGA	AAIGAAACAA TAATTAAGG
1201		CIAAGIAAAG		TICICGIACC.	TITIGUATUA	I GATTIAGUU
1321	AGUGIGACCC	AAGCAAAGAG	IGCITIAGII	IGACACCCCG	AAACTAGGGG	AGCIACICUA
1381	AGACAGCCIA	TITATAGGGC	GAACCEGICT	CIGIGGCAAA	AGAGIGGAAT	GAGCITIGAG
1441	TAGAGGTGAT	AAACCTACCG	AACCTAGTTA	TAGCIGGIIG	CCCGGGAAAT	GGATAGAAGT
1501	TCAGCCTCTC	AGATTCTTTA	TTCACCTCAG	TATTACCCCA	CCTGATACCA	CAAGATAACT
1561	GTGAGAGITA	TICAAAGGGG	GTACAGCCCC	TTTGAAACAA	GATACAACTT	TTCCGGGAGG
1621	AAAAAGATCA	TAATTAAATA	AAGGTAAGTA	TTTGGGTGGG	CCTAAAAGCA	GCCATCCCAG
1681	TAGAAAGOGT	TATAGCTCAA	ATACATCACT	ACCCCTCTCT	ATCCTGATCG	TTAATTCTTA
1741	CTCCCCCCTT	CCCTACCGGG	CCATCCCATG	CAGACATGGG	AGGGACCCTG	CTAATATGAG
1801	TAATAAGAGA	GCCAAGCCTC	TCTCCTTGCA	TACGIGIAAT	TCGGAACGAA	CCCGCACCGA
1861	GCATTAACGA	CCCCAAACGA	AGAGGGACCT	GAACAACAAC	CCAAACAACC	AGAAAAAAT
1921	TCAAACATAA	ACCETTAACC	CTACACAGGT	ATGCACCTCA	GGAAAGACTA	AAAGAAAGAG
1981	AAGGAACTCG	GCAAACAAAT	CAAGCCTCGC	CTGTTTACCA	AAAACATCGC	CTCTTGCAAA
2041	GCTAAAGAAT	AAGAGGTCCC	GCCTGCCCTG	TGACTATTAG	TTTAACGGCC	GCGGTATTTT
2101	GACCGTGCAA	AGGTAGCGCA	ATCACTTGTC	TTTTAAATGA	AGACCTGTAT	GAATGGCACA
2161	ACGAGGGCTT	AACTGTCTCC	TCTTTCAAGT	CAATGAAATT	GATCTCCCCG	TGCAGAAGCG
2221	GGGATATAAA	CATAAGACGA	GAAGACCCTA	TGGAGCTTTA	GACACCAAAG	AAGATCCTGT
2281	CAAGTAACCC	CTTATAAGGG	CCTGAACTAA	TGGAATCCTT	CCCTAATGTC	TTTGGTTGGG
2341	GCGACCGCGG	GGAAACAAAA	AACCCCCACG	TGGAAAG3GA	GCACCCACTC	CTACAACTAA
2401	GAGCCGCAGC	TCTAATTAAC	AGAATATCTG	ACCAATAAGA	TCCGGCAATG	CCGATCAACG
2461	GACCGAGTTA	CCCTAGGGAT	AACAGCGCAA	TCCCCTTTTA	GAGCCCATAT	CGACAAGGGG
2521	GTTTACGACC	TCGATGTTGG	ATCAGGACAT	CCTAATG3TG	CAGCCGCTAT	TAAGGGTCCG
2581	TTTGTTCAAC	GGTTAAAGTC	CTACGTGATC	TGAGTTCAGA	CCGGAGTAAT	CCAGGTCAGT
2641	TTCTATCTAT	GGTGTGCTCT	TTTCTAGTAC	GAAAGGACCG	AAAAGAAGAG	GCCCCTGCTC
2701	TAAGC'AAGC'C'	TCACCCCCAC	CTAGTGAAGA	CAACTAAAGT	AGGCAAGAGG	GCATACCCCC
2761	CGTGCCTJAG	AGAACGGCAT	GTTG3GGTGG	CAGAGCCCGG	TAAATGCAAA	AGACCTAAGC
2821	CCTTTTTACA	GAGGTTCAAG	TCCTCTCCTT	AACTATGATT	TCAGTCCTTA	TTACCCATAT
2881	TCTTAACCCC	TIGGCCTICA	TCGTCCCCGT	CCTCTTA3CC	GICGCCTICC	TCACGCTTCT
2941	AGAACGTAAG	GTACTAGGGT	ATATACAACT	ACGAAAGGGT	CCAAATATTG	TAGGACCTTA
3001	CGGGCTGTTA	CAGCCTATCG	CCGATGGTGT	GAAGCTCTTT	ATTAAGGAGC	CIGIICGCCC
3061	CTCCACTTCC	TETECETAC	TTTTCCTCCT	CGCCCCCCTA	CTCGCACTTA	CGCTTGCCTT
3121	GACCCTTTGA	GCCCCCATGC	CTCTCCCATA	CCCAGTCATT	GACTTGAACC	TIGGGATTCT
3181	ATTTATTTTG	GCCCTATCAA	GCCTCGCTGT	CTACTCTATT	CTAGGCTCAG	GCTGAGCATC
3241	CAATTCAAAA	TATGCCCTCA	TCGGGGGCCCT	TCGGGCTJTA	GCCCAAACCA	TCTCATATGA
3301	GGTTAGTCTA	GGCCTAATCC	TATTAAGTAC	TATTATTTTT	ACAGGGGGTT	TTACCCTACA
3361	AACCTTCAAC	ATTGCTCAAG	AGAGCGTCTG	AATACTACTC	CCAGCTTGAC	CACTAGCCGC
3421	AATATGATAT	ATTTCAACCC	TTGCGGAGAC	AAACCGTGCA	CCTTTTGACC	TTACTGAAGG
3481	CGAATCCGAA	CTAGTCTCTG	GCTTCAATGT	CGAATATGCA	GGTGGCCCAT	TEGECETATT
3541	TTTCTTG3CC	GAATATGCTA	ATATICTACT	TATAAATACG	CTTTCCGCCA	CCCTCTTCTT

Figure 7. Sequences of the mitogenome of S. steindachneri (GenBank Accession No. KJ834060)

0.004	1000000000		~	100001	9733 J J 010177 J J	<b>—</b> • • • • • • • • • • • • • • • • • • •
3601	AGGGGCCTCT	CATTTTCCAA	CACTACUTGA	ACTCACCGCA	GTAAACCTAA	TAGTCAAAGC
3661	GGCCCTTCTG	TCTGTCTTAT	TTTTATGAGT	TCGAGCCTCT	TACCCACGAT	TCCGCTACGA
3721	ТСААСТСАТА	CATCTAATTT	GAAAAACTT	COTCCCGCTT	ACACTGGCCC	TEGTTATTE
2791	ACACCTACC	CTCCCATTC	CATTICCICC	CTTCCCACCC	CARCTATARA	TAACAACCCC
0001		ANACCCCATTO		TOCCACCO	CAUCIALAUA	
3841	IGUUIGAMGI	ANAGGGUUAU	1116AIA5AG	IGACITATES	GGGIICAAAI	
3901	TCTTAGAAAA	GEGEGACTCG	AACCCCGCCT	AAGGAGAGCA	AAACTCCTGG	TECTCCCACT
3961	ACACTATTTC	CTAGTAAAGT	CAGCTAATTC	TAAGCTCTTG	GTCCCATACC	CCAAACACGA
4021	AGGTTAAAAT	CCCTCCTTTG	CTAATGAACC'	CTTACATCTT	AACCGCCCTG	CTATTTGGTA
4081	TTGGTTTAGG	CACTACTACC	ACCTTCGCAA	GCTCCCACTG.	ACTACTCCCC	TGAATAGGCC
1111	TGGAAATAAA	ТАСТСТТВССС	ΔΤΤΔΤΤΩΎΤΩ	ТААТАВСТСА	ACACCATCAC	CHECKBERGEAR
4201	TTCAACCACC	CACTADATAT	TTATICT	Macracia	ACACCHICHC	CTACTATTT
4201		CAUTAAATAT	TICIIGATIC	AAGUIGUUGG	RUCAUCIAIA	CIACIALITS
4201	CUABUAULAU	CAAUGUIIGA	TIGAJIGGGU	AAIGGGAGUI	IIIGCAAAII ggglommggl	GUUCAUUUTT
4.321	TCCCAACTGT	TCITIGICACT	TIGECCUTCE	CACTAAAGGT	GEGACTIECA	CULETACACT
4381	CATEGCTACC	TGAAGTACTT	CAAGGCCTAG	ACCTAACCAC	AGGACTTATT	TIGTCGACCT
4441	GACAAAAACT	TJCCCCATTT	GCCTTATTAG	TCCAAACTCC	CTGTGCCAAC	ACCACCCTTT
4501	TAATTATCCT	CGGACTTACC	TCAACCATTG	TAGGAGGCTG	AGGAGGTCTC	AACCAAACCC
4561	AGCTTCGCAA	GATTCTTGCC	TACTOCTOCA	TCGCACACCT	CGGCTGGATA	GTAATTGTAC.
4621	TACAATTOTO	TOCOTOCTTG	ACTATTCTAA	CACTACTCAC	ATATTTATT	ATAACATTTT
4601	CARCATTECT	TATETTAAA	CTTAATAAAC	CANCEARENT	TAATCCTCTA	COARCOTCAT
4001	CAGCALLICI		AGGGGGGTTG		ATTECT	TENAGAGGGG
4/41	BUBUAAAJAU	ULULUBUUUIA Magaaamm	ACCERCETTE	CALULTICE	ALIGITATUU	I I AbbAbbbb
4801	TCCCCCCACT	TACAGGATIT	AIGCCAAAGT	GACITATICT	TCAAGAGCTT	GUTAAACAAG
4861	ACCTTGCCCC	CGCCGCAACA	CTGGCGGCAA	TAACCGCCCT	CCTTAGCCTA	TATTTTTATC
4921	TACGACTATC	ATACGCAATG	GCACTAACTA	TTTCGCCAAA	TAACCTAACC	GCAATTTCCC.
4981	CATGACGCCT	CCCCTCTTTA	CAACTAACAA	TGCCACTTGC	TACCTCAGCC	ATAGCTACGC
5041	TAATGCTTCT	ACCCCTAACA	CCCGCCGCAA	TARCACTAAT	ΑΑΓΓΓΓΤΤΤΑΑ	GEGACTTAGE
5101	TTAAAACAAG	ACCAAGGGCC	TTCAAAGCCC	TAAGTGAGGG	TEGAAGTCCC	CCAGTCCCTT
5161		CERCACACEA	CCCCACATCT	CCTCTATICA	AAACACCTAC	TTTAATTAAC
5101		COUCHCIA	CACCONCRICT	TOCTOCALOCA	TOTTAOTTA	CACCEARCCE
5004			TOTATION			
5281	UILAAAL.Ab	LOADLAILLA	TUTATUTITU	CLUBLEIGA	ISISISALISISISI IS	
5341	AAGICCCGGC	AGACGACTAG	ICIGCAICII	CAGATITISCA	AICIGAIAIG	TAAAACACCI
5401	CAAGACTTCT	GGTAAGAAGA	GGACTCAAAC	CTCTGTTTGT	GGGGCTACAA	TCCATCGCTT
5461	AAAAACTCAG	CCATCCTACC	TGTGGCCATC	ACACGTTGAT	TTTTCTCCAC	TAATCACAAA
5521	GACATCG3CA	CCCTTTATCT	AGTATTTGGT	GCCTGAGCCG.	GTATAGTAGG.	CACAGCCCTC
5581	AGCCTACTCA	TTCGAGCAGA	ACTAAGCCAA	CCEGECECTC	TOCTTGGAGA	CRACICIAAATT
5641	ТАТААТСТАА	TTGTTACAGC	ACATICTTC	GTAATGATTT.	TOTTTATACT	AATGCCAATT
57041	ATAATTOTAA	CTTTTCCAAA	OTCATTAATT.	COCCTATO	TTCCACCCC	ADATATACCA
		TOTITOGAAA	VIGALIANTI Martrata			AGAIAIAGUA
D/01	- TITUUTUSTA	1 SAATAACAT	AAGITTUTGA			ACTACTACTT
5821	GCCTCTTCTG	GAGTAGAAGC	GGGTGCCGGA	ACCEGETEAA	CAGIGIACCC	GCCCCIGGCI
5881	GGTAACTTAG	CCCACGCAGG	AGCATCAGTC	GACCTGACAA	TCTTTTCACT	TCACCTAGCA
5941	GGTATTTCCT	CAATCCTAGG	GGCAATCAAT	TTTATTACCA	CAATTATTAA	TATGAAGCCC
6001	CCGGCTATCT	CTCAGTACCA	GACACCCCTA	TTTGTGTGAG	CTGTCCTAAT	TACCGCTGTT
6061	CTTCTCCTTC	TCTCTCTACC	AGTTCTCGCT	GCCGGCATCA	CAATGCTCCT	TACCGACCGA
6121	ΑΤΓΎΤΤΑΑΤΑ	CCACCTTCT	TGACCCGGCA	GGAGGAGGGG	ATCCAATCCT	TTATCAACAC
6181	TTATTOTICT	TTTTTGGACA	CCCCGGAAGTA	ΤΔΤΔΤΤΩΤΤΔ	TTCTGCCTGG	CTTTEGTATS
6041	ATTTCACACA	TITIOUACA	TTATTOTOO	ANNANGANC		TATACCCATA
0241	ATTICACACA GENERADAGAN	TUBIUSUUIA	TIAIILISU.		ULTITUUTA TATANAGTA	TATABBLATA
6301	GIAIGAGCAA	TAAIGGUIAI	IGGCCICCIA	GGCTTTATIG	TATGAGUTUA	ICACATATIT
6361	ACAGTTGGCA	TGGACGTAGA	CACGCGTGCT	TATTICACAT	CIGCCACAAT	AATCATCGCA
6421	ATTCCCACCG	GTGTTAAAGT	ATTTAGCTGA	CTTGCGACCC	TACATGGGGG	CTCTATTAAA
6481	TGAGAAACAC	CCCTTTTTATG	AGCCCTTGGC	TTTATTTTCC	TGTTTACAGT	AGGAGGGCTT
6541	ACAGGCATTG	TTCTGGCCAA	TTCATCTCTA	GATATTGTAC	TCCACGATAC	CTATTATGTA
6601	GTAGCCCACT	TECACTACGT	ACTATCTATE	GGGGCCGTGT	TTGCCATTGT	CBCCBCCTTC
6661	GTGCACTGAT	TCCCGCTATT	TTCAGGCTAC	ACIGCITTCIACA	GCACTTGAAC	AAAAATCCAC
6721	TTCGGTATTA	TETTOTTAGE	GGTAAACTTA	ACCTTCTTCC	CACAACATTT	COTOGGATTA
6721	COCCATIN	COUCANCATA	CTCCCACTAC	COTELOC	ATACCCTATIC ATACCCTATC	AATACACTC
6041	TOOTONATOO	COUCCECTER	OTCOMOTAC.	COTOTONICOL	TATACCULATO	TATTACTOL
-0641 -6044	- ICUICAAIU5	COTCACTIAL	TANATTAT	OCTOTOALCA	TOTICITALL	INTINITOR
0901	BAGBCAIICG	CLUCCAAACG	IGAASIICIA GOOLTROOM	GUAAUAGAII	I RAURACAAC	CAAIGIAGAA
D901	IGACIACAIG	CONCOUNCE	CCCATACCAC	ACATICGAGG	AGUUIGUUIT	I JIACAAGTA
7021	CAABCABACT	AAUGAGAAAG	GGAGGAGTCG	AAUUUUUUATA	GJTCAGTTTC	AAGUUGACCA
7081	CATAACCGCT	CIGCCACTTT	CTTTATAAGA	CACTAGTAAA	AGAGTACATT	ACACCGCCTT

7141	GTCAAGGCGG	AAGTGTGGGT	TAGACCCCCG	CGTGTCTTGC	TTTTAATGGC	CCATCCGTCA
7201	CAGCTTGGAT	TTCAAGATGC	AGCTTCACCT	GTTATAGAAG	AACTTCTTCA	TITTCACGAC
7261	CATGCTTTAA	TAATCGTCTT	CCTGATTAGC	ACACTAGTGC	TTTATATTAT	TCTTGCTATA
7321	GTTACCACTA	AATTAACGAA	CAAATATATT	TTAGATTCAC	AAGAGATTGA	AATTATCTGA
7381	AC'AATTCTCC	CAGCTATCAT	TTTAATTCTG	ATTGCACTCC	CTTCCCTTCG	CATCCTCTAT
7441	CTTATAGATG	ΔΙΆΔΤΤΔΔΓΊΔΔ	CCCTTTATTA	ACATTAAAG	CHETTERCCA	CONGTEGEAC
7501	Таллестата	AATATACTCA	CTACZAAZAT	CTTCCCTTTC	ATTCATATAT	AATTCCCACC
7561	CANCARCINIC	COCCEPCION	ATTOOCOTA	TTACAACCOC	ACCATCCALL	COTTATTOCACO
7001			ATTUGUETA		TACTICUTC	JTGAGGAGTG
7021	GIIGAAIJUU	CLAILUGAGI		GLAGACGAIG	TAUTULAUTU	AIGAGEMAGIC
7081	CUBBUUUIBB	GAGIAAAAAI	GGAUSUASIU	CUABBUUBUU Maamamalal	TAAAUUAAAU	ASCULLATE
//41	GCATCUUGAU	CAGGUGIAII	CIACGGACAA	IGUICIGAGA	TCIGUGGAGU	AAATCACAGC
7801	TTTATACCTA	TIGTAGTGGA	AGCAGTTCCC	CTAGAACACT	TIGAAAACIG	ATCATCTCGA
7861	ATACTIGAAG	ACGCCTCGCT	AGGAAGCTAA	ATAGGGTATA	GCGTTAGCCT	TTTAAGCTAA
7921	AGATTGGTGG	CTCCCAACCA	CCCCTAACGA	CATECCCCAA	CTCAACCCCG	CACCTTGATT
7981	TGCTATTTTA	GTCTTTTCGT	GAATGGTCTT	CCTGGCCGTT	ATTCCCGCTA	AAGTTACAGC
8041	CCACACTTTC	CCAAAATACCC	CTAATCTGCA	AAGCGCAGAA	AAAGCCAAAA	CAGACCCCTG
8101	AACTTGACCA	TJACACTAAG	CTTTTTTGAC	CAGTTTATAA	GCCCCACCTA	TETTGGGATE
8161	CCATTAATAG	CCCTTGCCCT	TACCCTACCC	TGACTCCTTT	ACCCCACACC	TAC'AACTC'GA
8221	TGATTAAATA	ACCGATTCCT	CGCGCTTCAA	GGTTGATTTA	TTAACCGTTT	TACTCAACAG
8281	CTTCTCCTCC	CETTAAATAT	TGGAGGTCAT	AAGTGAGCTG	CCCTCCTAAC	CTCATTAATG
8341	ATCTTTTTAA	ТТАСССТААА	TATATTAGGA	CTTCTTCCCT	ACACTTTTAC	TOCTACCACO
8401	CAATTGTCAC	TAAATTTAGG	ACTEGRETA		TAGCAACTGT	TATTATTGGC
8461	ATACCARACC	AACCAACCCA	TECCOTACEA	CACCTCCTAC	CAGAAGGCAC	Addadadaa
0401	CTTATOCAR	TOCTTATCAT	TATCOANACA	ATTACCTO	TTATTOCOC	TOTTOCOCCE
0.021	CITAICCUCG	TOUTRICAL	TATUSAAACA	COTOLOCICI	TRICCOCCC	AATTCCCTA
0001	COCOCCETTO	TACTTCC		ACCORDINA	TARTICARCI	ANTIGUTACA
0041		COOTATTACA	AGTTANIACCA	ACCOLOCAA ACAATAATTA	A DOMESSION	ARCAGIACIS
- 8701 - 0704	GILLULIA	TATIATIANA TATAGANA	AGIIACIGIA	BUAATAATIU	ARGUATAUST TOMOGRAPHIC	
0/01 0004			AGAAAAUGIU	TAATGGUUUUA		CUITACCACA
8821	TAGETIGACCC	CAGCCCTTGA	CCCCTAACAG	GUGCAATIGU	I GUCUTACIG	ALAACATCAS
8881	GCCTCGCGAC	CIGATITCAT	TITCGCTCAA	CAACCITAAT	AACCITAGGA	ACAGUTOTAC
8941	TGCTTCTTAC	AATATACCAA	TGATGACGAG	ATATCGTACG	AGAAGGTACA	TICCAAGGAC
9001	ATCATAC3CC	CCCCGTACAA	AAAGGTCTTC	GATACGGAAT	AATTCTCTTC	ATTACCTCC3
9061	AAGTATTTTT	TTTCCTAGGA	TTCTTCTGAG	CCTTTTACCA	CGCAAGCCTA	GCCCCCACTC
9121	CTGAGCTAGG	GGGCTGCTGA	CCTCCAACGG	GCATTACAAC	TCTTGACCCA	TTTGAAGTCC
9181	CCCTCCTTAA	TACAGCTGTC	CTACTICCCT	CCGGGGGTGAC	GGTTACCTGA	GCCCACCACA
9241	GCATTAT3GA	AGGTGAACGA	AAACAAACCA	TTCAATCGCT	AGCCTTAACT	ATCCTTCTAG
9301	GCTTTTATTT	CACATTTCTT	CAAGGCCTGG	AATACTATGA	GGCCCCCTTT	ACAATTGCAG
9361	ATGGCGTATA	CGGCTCTACC	TTTTTCGTAG	CCACTGGCTT	TCACGGACTA	CATGTTATTA
9421	TTGGCTCCAC	ATTTTTAGCT	GTTTGCCTCC	TACGACAAAT	CCAATACCAC	TTTACATCCG
9481	AGCACCACTT	CEGGTTEGAA	GCAGCTGCCT	GATACTGACA	TTTCGTAGAC	GITGTGTGAT
9541	TATTCCTATA	TATCTCTATC	TACTGATGAG	GCTCTTAATC	TTTCTAGTAT	TAAAACTAGT
9601	ATAAGTGACT	TCCAATCACC	CGGTCTTGGT	TAAAATCCAA	GGAAAGATAA	TGAACGTAGC
9661	AATAGCTGTA	ATTACCATCA	CTATTTGCT	TTCCGTAATC	CTERCETTE	TATCCTTCTG
9721	GCTCCCCCAA	ATGACCCCCG	ATCATGAAAA	GETETEE	TATGAATGTG	GTTTCGACCC
9781	CTTAGGATCA	GTELECETAE	CATTTTCCCT	CCGCTTCTTC	CTARTCRCCA	TTOTTTTOOT
02/1	CCTTTTCCAT	TTAGAAATTG	CCCTTCTCCT	CCCACTCCCT	TRACCC	A ATTA ACTTC
0001	CCCCTTACTC	ACACTOTTOT	GAGCCGTGGC	CGTGCTTATT	CTTCTTACCC	TTGGCTTGGT
0061	ТТАССАСТСА	ATTCAACCAC	CTTTACAATC	ACCCANTAC	COAATTACTT	ТААСАААААТ
10021	ATTACOADION	COCTCANNAC	OTTINUATO CTTATCCTTA	AACTCCATAA	TTOTOTANTO	ACTOCOCCE
10021	ATTIGATIC ACTTCCCTTT	CTCLARAAG	TTTAIGUIA	CACTCALAR	COTACONTRO	CATCUUGELE
10144	ACTICUTI	THATCHER	TITALULIAU	CONTONERO	CUTABUALLU	A TOO A ACT
10201	AULIULIUL	ISUIUIIIA ASTALAACTA	TGUTTAGAAG	UGAIGAIJUI	CICILIAIII	AICGGACIII
10201 10264	CURTINGAAC	ULICCAACIA	ACCALCUACAA	RITICICIEC Traderersor	Gamagedaga	CIUCIALIES acradazada
10201	UTITICAGC	TIGIGAAGCA	AGUGUAGGGU	TIGUIIIACT	GETAGUUACA	BUIUGUAUGU
10321	ATGGTTCAGA	TCGCCTTCAA	ACCTTAAACC	ICTTACAATG	CTAAAAATCC	TAATTCCCAC
10381	CCTAATGCTT	CITCCCACAG	COTEGOTIEC	CCCTGCCAAA	IJATTATGAC	CTACTACCCT
10441	CIUCCACAGC	CTAGTCATTG	CACTGGCCAG	CCTCACTIGA	TTAAAAAATA	CATUTGAAAC
10501	AGGCTGATCT	TGCCTCACGC	CCTTCATAGC	CACAGACCCC	CICTCAACAC	CCCTCCTTGT
10561	TCTTACCTGC	TJACTACTCC	CTCTTATAAT	TTTGGCAAGC	CAAAGCCACA	CAGCACTAGA
10621	ACCTATTAAC	CGCCAACGAA	CCTACATTAG	CCTATTAACG	TETETGEAAG	TATTCCTTAT
10681	TATAGCATTC	GGTGCCACTG	AACTCCTTAT	GTTTTAT3TT	ATATTTGAAG	CTACTCTCAT

(Figure 7. Continued)

10741	CCCCACACTA	λΤΤΛΤΤΛΤΤΛ	erectored and	TANTCACCA	CARCCCTTA	ATCCACCACT
10741	UCCUALACIA	ATTATTATTA	0100310300	TAATUAGSUA	GRAUBUUTTA	AIGUADGAGI
10801	ATATTTTTG	TITTATACCC	TAGCAGGCTC	TCTCCCGTTA	CTAGTTGCCC	TELLECTICL
10861	TCAAAAGGAT	ACAGGATCCC	TCTCCCTCTT	AACCATCCAA	TATACTAGCT	CTACCCCTCT
1.0921	TTCATCTTAT	GCTGACAAAC	TTTGATGAGC	AGGCTGCCTA	ATTGCATTTT	TAGTAAAAAT
10021	ACCTTTATAT	CONCINENT	TOTCATTACC	AAAAQCACAT	GTAGAAGCCC	CARTTECARZ
44044		GURGERCATE.	TUTURITACU	NOTVOOLAGE	TRUBARUUUU.	TOURGE
11041	CICAAIGJII	CIAGUIGUAG	TICIICIAAA	ACTAGGAGGC	TACGGIAIAA	ICCGAAIGAI
11101	AGTTATATTA	GAACCTCTCA	CCAAGGAATT	AAGCTATCCC	TITATIGTCC	TAGCCCTCTG
11161	AGGTGTAATT	ATAACTGGCT	CCACCTGCCT	TCGCCAAACA	GATCTTAAAT	CCCTCATCGC
11001	CTATTCATCC	GTAACCCATA	TRECOGRET	COTTORADOT	ATTOTTATOO	A A ACIACICICITIZ
11001	AGEATERICA	OINNOCCHIN	TOODSCIDDI	Tagyayagaa		ADDACACCETT
11281	AGGICIIGCC	BalbligIAA	IUUIIAISAI	ISCACACSSC.	UIGAUGIUUI	
11341	CIECTIESCC	AATACAAACT	ATGAACGCCT	CCATAGUUGA	ACAATACTAT	TAGCCCGGGGG
11401	ATTACAGATA	GTGCTTCCAC	TCATAGCAAC	ATGATGATTT	ATTGCCAGCC	TCGCAAACTT
11461	AGCCCTTCCC	CCTCTGCCTA	ACCTCATGGG	AGAACTTTTA	ATTATTACCT	CATTATTTGG
11521	TTGATCATGA	TGAACTCTCG	TACTCACAGE	GGCGGGGGACC	CTTATTACCG	CGAGCTATTC
11521	ΔΟΤΤΤΑΤΑΤΑ	TTCCTCATCA		TOCOUTOCA	CTACATATTA	TTACCCTCAA
11/21	ACTITATATA AGGGTGGTAG	ACCORDANCE	ACCENCUSUU	ACCOUNTER	araaraaaa	TIAUCCICAN
11041	CULUTUUTAU	ALUUGGGAAU	ACCIAGITAT	AGUULTICAU		IJUIIUIAUI
11701	TGTATTAAAG	CCCGAATTAG	TATGAGGCTG	AACCACCIGT	AGATATAGIT	TAACAAAAAT
11761	ATTAGATTGT	GATTCTAAAG	ACAGAGGTTA	AAATCCCCTT	ATCCACCGAG	AGAGGCTCGC
11821	CAGCAACGAA	GACTGCTAAT	CTCCGTGACC	TTGGTTGGAC	CCCAGGGCTC	ACTOGGCCTG
11881	CTECTAABGG	ATAACABCTC	ATCCATTGGT	CTTAGGAACC	<u>888886'TC'TT</u>	GETECALATE
11041	CAACTACCAC	CTATCCACTC	CTCATCACTT.	ATTATOTOAT	COACCTTACT	CATTATO
11941	UAADIAGAG	CIAIGUAUIU	UTCATCACIT	ATTAIGLAT	CCABUITAGI	CATTAICTIT
12001	TIACIATIAG	CATAICUTAI	CITTACGACU	CIGGACUCIC	GUUUUUGAAA	CUCIGACIGS
12061	GCCGTTTCCC	ATGTTAAGAC	AGCGGTCGCC	CTGGCCTTCT	TCGTTAGCCT	AATCCCTTTA
12121	TTTCTCTTTTC	TTAACGAGGG	GGCGGAAGCA	ATCATCACCT	CATGAAATTG	AATGAATAC'A
12181	CTAACCTTCG	ACGTGAATAT	TAGCTTCAAG	TTTGACCACT	ACTCAGTTAT	CTTTGTACCC.
122/1	ATTRCCCTCT	ACIETCIACITTE	ATCTATTCTA	CACTTECAT	CATCCTATAT	ACACACACAC
10001	COATACATAA	ACCIENCIES	TAATACOTC	CTACTTOCAL	TEATOOIAIAI	
12301	ULATACATAA OTGUGUGUA	ACCOALICIT	IAAAIAUGIU GOLLOTTE	UTABITITUU	1 TATIGULAT	AATTATICT
12361	GTCACAGCAA	ACAAICTAIT	CCAACITITIC	ATTGGTT3GG	AGGGAGTAGG	CATTATATCA
12421	TTTCTACTCA	TTGGCTGATG	ATACGGACGA	GCGGATGCCA	ACACAGCGGC	COTTCAGGCO
12481	GTCGTGTATA	ATCGGGTAGG	AGACATTGGA	TTGCTATTC'A	CAATAGCTTG	AATAGCAACC
12541	AACGCTAACT	COTGAGAGTT	ACIAACIAAATT	TTTGTAGCAA	CTAAGGACCT	CGATETTACC
12601	CTACCCCTAC	TACCCCTCAT	TOTTOCOCO	ACACCCAACT	COCCOANT	TOCTOTICS
10001	CINCLUINC	COTOTON	AAAAAAAAA	ACAGGCAAGI	CJUCCAAII GEOGOGEAGE	ISGICITCAC
12001	LUTIGACIUU	CULUIGUIAI	AGAGGGICUI	ACAUCEBIAI	CIGUUIAUI	GUAIIUGAGU
12721	ACCATAGICG	TIGCCGGTAT	TITITCICCIA	GTACGAACAA	GICCCCTCCT	GJAAAATAAT
12781	CAAACTGCCC	TCACCACCTG	CCTATGCCTA	GGTGCCCTAA	CGACACTATT	TACAGCCACC
12841	TGTGCCCTAA	CCCAAAATGA	TATCAAAAAG	ATCGTAGCAT	TCTCCACATC	AAGTCAACTA
12901	GGTCTAATAA	TAGTTACTAT	TGGCTTAAAT	CAACCTCAAC	TAGCCTTTCT	CHACATTER
12061	ACCONTRACT	TOTTANCCO	AATACTATTC	CTCTCTCTC	COTCAATTAT	TOACACCCTO
10001	MODIAN	A DATATOOC	ANIACIATIC	CICIGITCIG	ACCENTIAL	TEACAGECIE
130ZL	AAUGAUGAAU	AAGAIAIUUG	AAAAAIAJGA	GUIAIACAIC	AULIIGUUU	
13081	TCCTGCCTTA	CTATTGGTAG	TTTA3CCCTC	ACAGGCACCC	CCITCCTAGC	AGGATTCTTC
13141	TCCAAAGATG	CCATTATTGA	GGCACTAAAC	ACATCTCACC	TAAACGCCTG	AGCCCTAGTC
13201	CTAACCCTTC	TAGCCACCTC	ATTCACCGCC	ATCTACAGTC	TCCGCGTAGT	GTATTTTGTT
13261	TCAATGGGCC	A CEE A CEETT	таасстатт	тесесатез	δηθασατα	CHEAGEGETT
13371	ATTAACCCCT	TAAACCICACT	TECATEACEA		CTRECCTCCT	AATTATCTCA
10001	ALLARCULL Addamerraddd			TOTALIGICS		ATTAICICA
13301	AGUATIAJUU	CULTIAAGAU	UUUUSIGAIA	ICIAIACUIC	CUITGUILAA	ACTAGUIGUI
13441	CITGTAGITA	CAATTATAGG	ATTACTCATT	GCCCTCGAGC	TAGCAACACT	CACCAATAAA
13501	CAGTACAAAG	TTACCCCTAA	TCTAGTTACC	CATCACTTCT	CCAACATGTT	ACGCTTTTTC
13561	CCCTCGATTG	TICACCGATT	TACCCCCAAA	CTAAATCTAG	TCTTAGGACA	GACACTTGCC
13621	AGCCAACTGA	TIGACCAAAC	TTGACTAGAG	AAAGTTGGTC	CCAAAGCAAT	CTCTTCATCA
13681	ATATTCCC	TAATTACAAC	A AC'A ACC'A AC	ACACAACAAC	CAATAATTAA	CACCTACCTC
10001	ANIALICUCC	TARTIACAAC	AACAAGCAAC	ACHCARCARG ACHCARCARG	TTACCOTTA	BACGIACCIC.
15/41	ACCUTATICC	TICTIACUUT	GAULETIGUI	GUULIAITAI	TIALUUGIIA	AACIGCCCGA
13801	AGGGTCCCCC	GACTAAGTCC	CCGAGTTAAC	TCCAGCACAA	CAAACAAGGT	GAGAAGC'AAG
13861	ACCCACGCAC	TAAGTACTAA	TAACCCTCCC	CCTAATGAGT	ACATTAACGC	AACCCCTCCC
13921	ATATCGCCTC	GCAAGACAGA	GAGCTCACTA	AGCTCATCAG	CCGGCACCCA	TGAAAACTCA
13981	TATCACCCCC	CTCAAAATAT	ACTAGAAGCC	ACCCCCACCC	CTACTAAGTA	TATCAACATC
14041	ТСАССТАСАА	CARRACCACT	TACCOACCTT	TCCGGGTAAG	GALABORE	
1/1/04	CACCIACAA	AGATOACCACT	TACCONUCTI	A A A T A A T C A	AYAYGYGGYG	CACTOLOCO
14101	BABIAUBCAA	ACAIGACIAG	INTECCACUL	AAATAAAICA	AAAALAGLAU	CAGIGAIAGA
14161	AAGGGTCCCC	CATGICCAAC	CAATACTCCA	CAUCCCATGC	CUGUUACGAC	TACTAACCCT
14221	AAGGCAGCAA	AGTAAGGAGA	AGGGTTAGAG	GCAACTGCAA	CCAACCCTAG	AACTAACCCA
14281	ATTAAAAATA	AAGACATAAT	GTAAGTCATA	ATTCCTGCCA	GJACTTTAAC	CAGAACTAAT

14341	GGCTTGAAAA	ACCACCGTTG	TTATTCAACT	AC'AAGAAC'C'C'	ACTAATGGC'A	AGTCTACGAA
14401	AGACACACCC	TCTCCTCAAA	ATCGCAAACA	ATGCCCTAGT	TGACCTACCC	GCCCCCTCAA
14461	ATATTTCAGT	GTGATGAAAC	TTCGGCTCTC	TCTTGGGACT	CTGCTTAATT	ATTCAAATCC
14521	TCACGGGACT	ATTTTTAGCC	ATACACTATA	CCTCTGATAT	TGCTACAGCT	TITTCTTCCG
14581	TTGCTCATAT	TTGCCGAGAC	GTAAATTACG	GGTGATTCAT	CCGAAACCTT	CACGCCAACG
14641	GTGCATCCTT	CTTCTTTGTA	TGCATCTATG	CCCACATTGG	CCGCGGACTT	TACTACGGCT
14701	CATACCTTTA	TAAAGAGAC'A	TGAAACATCG	GAGTAGTCCT	ACTACTECTA	GTTATAATAA
14761	CTGCTTTCGT	CGGTTATGTG	TTACCCTGAG	GCCAAATGTC	CTTTTGAGGT	GCCACCGTTA
14821	TTACCAACCT	ACTCTCCGCA	GTACCCTACG	TAGGTAACGC	CCTCGTTCAG	TGAATTTGAG
14881	GTGGATTCTC	AGTAGACAAT	GCAACCCTTA	CCCGATTCTT	CICCTTCCAC	TITTTATTCC
14941	CCTTTGTAAT	TGCAGGCGCG	ACCATAGTCC	ACCTCCTTTT	CCTTCATCAG	ACAGGATCAA
15001	ATAATCCCCT	CGGCCTAAAT	TCAGATGCAG	ATAAAATAAG	CTTCCACCCC	TACTTCTCAT
15061	ACAAAGACTT	ATTAGGGTTT	GCAGTACTTG	TCATTGCCCT	TACATGTCTA	GCTTTATTT
15121	CACCCAACCT	GCTAGGAGAC	CCAGACAACT	TCACCCCCGC	CAATCCACTA	GTTACTCCAC
15181	CCCACATCAA	GCCAGAGTGA	TATTTTCTGT	TCGCATACGC	AATTCTACGC	TCCATTCCCA
15241	ATAAACTAGG	GGGAGTTTTA	GCCCTCCTAG	CTTCAATCCT	TATTCTTATG	CTAGTGCCCT
15301	TTCTACACAC	GTCTAAACAA	CGAAGCCTCA	CTTTCCGACC	ACTCACACAA	TTCTTGTTTT
15361	GAACCCTGAT	CGCAGACGTT	ATTATTCTCA	CCTGAATCGG	GGGAATGCCC	GTATCGC'ACC
15421	CGTTCGTCAT	TATCGGACAA	GTCGCATCCT	TTTTATACTT	TTTCCTCTTT	CTAGTCCTTA
15481	CACCACTAGC	AGGCTATGCA	GAGGACAAAG	CACTTGAATG	AGCTTGCACT	AGTAGCTCAG
15541	CGTCAGAGCC	CTGGTCTTGT	AAACCAGATG	TCGGAGGTTA	GAGTCCTCCC	TACTGCTCAA
15601	AGAAAGGAGA	TTTTAACTCC	CACCCCTGGC	TCCCAAAGCC	AGGATTCTTA	GTTAAACTAT
15661	TCTTTGTAAT	ATATGTACAA	TAATTTTATA	TACATATATG	TATTATCACC	ATTAATTTAT
15721	ATTAACCATA	TCATATAGCA	TTCAAGTACA	TATATGTATT	ATCACCATAT	CTAGGGTTTA
15781	ACCATTCAAG	TGTTATATTA	AACGAATAAT	TTACATAAAG	CAGAATAATA	AAAAACAATA
15841	AACACTTATA	AATACCGGGC	GAAACTTAAG	ACCTATCACA	ATAACTCATA	AGTCAAGTTA
15901	TACCTTTACT	CAAAATCCCG	CCAAACTCAA	ATATTTAATG	TAGTAAGAGC	CGACCAACAA
15961	GTCCATTICT	TAATGCCAAC	GGTTATTGAA	GGTGAGGGAC	AACTATTGTG	GGGGTTTCAC
16021	ACAGTGA I''I'''	ATTCCTGGCA	THIGHTCOT	ATTICAGGCC	CACATATTGT	AAACCCCCCC
16081	ATACGTCTAT	CGTAGAAAGC	ATAAGTTAAT	GGTGGAAAAC	AATAGCIGGGA	GCGGCCACCA
16141	TGCCGAGCGT	TCTTTCCATA	GGGCATTAAG		TITTTTCCT	TTTCAATGGA
16201	CATTTCACAG	TGCACGCGAT	CTGATTAACA	AGGTGGGAAT	AATCTTAGGA	AGCAAGGAAA
16261	TAGTATGAGT	GGTGAAAGGT	CTTTACAAAA	GAATTACATA	TAAGGATTTC	AAGGACATAA
16321	AGTAGTGAAA	TITAGTCGGA	AGATATUTAT	ATTACCCCCT	TTTGGCTTTT	TOGOGITTAAA
16381	CCCCCCTACC	CCCCTAAACT	CGTGAGATAA	CTAACGCTCC	TGTAAACCCC	COGGCCTTAG
16441	GAAAACTTCG					

Table 4. Summary of the mitochondrial genome of S. steindachneri

\*H and L denote heavy and light strands, respectively.



Figure 8. A gene arrangement in the complete mitogenome Sebastes steindachneri.

# 3. 결론 및 토의

동해서식 볼락류의 자원특성을 조사하고 유전특성을 파악하기 위하여 다양한 체색 패턴을 가진 볼락류를 대상으로 연구를 수행하였다. 우럭볼락, 흰꼬리볼락, 황볼락, 노랑볼락의 미토게놈 정보를 분석하여 염기서열 정보를 세계유전자은행인 GenBank에 등록하였고 분석정보는 논문으로 출판하였다. 본 연구를 통하여 확보한 4종의 볼락류 미토게놈 정보는 해양어류의 유전정보 뱅크화에 직접 사용이 가능하다.

# 제2절 해양어류자원의 기능유전체 정보 이용기술 탐색

## 1. 연구 재료 및 방법

### 가. 어류시료 확보

볼락류의 체색관련 기능유전체 분석을 위하여 체색이 다양한 개볼락을 임원항에서 구입하였다. 개볼락 시료는 반점이나 마크를 기준으로 color type과 wild type으로 구분 하였으며, 체색관련 기능유전체 분석 시료로 활용하였다.

# 나. 개볼락 체색타입별 total RNA 추출

개볼락의 근육조직을 각각 100mg씩 취하여 호모게나이저를 사용하여 파쇄하였다. TRIzol<sup>®</sup> Reagent(Invitrogen, USA) 1ml을 첨가하여 상온에서 5분간 방치하면서 완전히 섞어 준 다음, chloroform을 0.2ml 넣는다. 상온에 3분간 두었다가 원심분리(12,000x*g*, 4℃, 15분) 하였다. 상층액에 isopropyl alcohol을 첨가하여 상온에 10분간 두었다가 원심분리(12,000x*g*, 4℃, 10분)하여 RNA pellet만 취했다. 75% 에탄올로 세정 후에 RNA pellet을 건조시킨 후, RNase-free water를 첨가하여 충분히 녹인다.

추출한 total RNA는 Agilent사의 2100 BioAnalyzer를 이용하여 QC 과정을 거쳤으며, BAConcentration(ng/ul), RIN value, 28s/18s ratio 값을 활용하였다.

## 다. 개볼락 근육조직 RNA-seq 분석

개볼락 체색 타입별 근육조직 각각의 total RNA 2µg을 oligo(dT)를 사용하여 mRNA를 분리·정제하였다. Library는 Paired-end 100 bp로 진행되었으며, Illumina사의 TruSeq RNA Sample Prep Kit을 이용하여 Library 준비를 진행하였다. 분리·정제된 mRNA는 fragmentation 단계를 거쳐 random hexamer primer를 사용하여 single-stranded cDNA로 합성하였다. 이를 주형으로 하여 double stranded cDNA를

합성하였다. Blunt-end를 만들기 위한 End Repair, Adapter를 붙이기 위한 A-tailing, Adapter ligation 과정을 순차적으로 거친 후 PCR(Polymerase Chain Reaction)을 이용 증폭하였다. PCR 최종 산물은 2100 BioAnalyzer를 사용하여 확인하였으며, 제작된 library는 KAPA library quantification kit을 이용하여 정량한 후 cluster generation하여 Hiseq 2500을 이용하여 염기서열 해독을 진행하였다(Figure 9).



Figure 9. Workflow of amplicon library construction and Hiseq2500 sequencing.

라. 염기서열 조립 및 유전자 확보

전사체 서열을 얻기 위한 조립과정에 앞서 서열 해독과정에서 생긴 잘못된 서열을 가진 리드(Read)나 혹은 어댑터 서열을 제거하고, 낮은 품질의 서열로 인한 조립 오류를 줄이기 위한 서열 필터링 작업을 진행하였다.

서열 조립은 전체 샘플의 서열 데이터를 하나로 합친 뒤, de Bruijin graph (DBG) 알고리즘을 이용하여 RNA-seq de novo assembly를 수행하는 대표적인 RNA Assmebler인 Trinity (Grabherr et al., 2011; Hass et al., 2013) 프로그램을 이용하여 서열을 조립하였다. Trinity의 서열 조립과정은 3단계(Inchworm, Chrysalis, Butterfly)로 이루어져 있고, 첫 번째 단계인 Inchworm에서는 메모리의 효율적인 사용을 위하여 각 서열을 서로 연결하여 서브그룹으로 나누는 작업을 하였다. Inchworm 단계에서는 contig를 구성하였고, Chrysalis 단계에서는 Inchworm에서 구성된 contig를 이용하여 리드에 대한 집단화 (clustering)를 진행하였다.

조립된 전사체의 집단화는 주어진 서열을 1대 1로 비교하여 서열간의 유사도를 계산하고, CAP3 (Huang and Madan, 1999)으로 각 집단에 속한 서열을 재조립하여 집단을 대표하는 서열을 만들어 내는 TGICL (Pertea et al., 2003) 프로그램을 이용하여 수행하였다.

Unigene의 기능을 알기 위하여 단백질 발현 부분인 coding sequence (CDS)를 예측하였다. CDS 예측과정은 TransDecoder (Haas et al., 2013)를 이용하였으며, 주어진 서열에서 가능성 있는 모든 CDS 서열을 예측하였고, GeneID(Blanco et al., 2007)로 가능성을 검증하여 log-likelihood score가 0보다 큰 것을 선택하였다.

### 마. 유전자 기능 분석

조립된 unigene과 이를 기반으로 예측된 CDS의 기능을 파악하기 위하여 DNA 서열 및 단백질 서열의 상동성 검색을 진행하였다. DNA 서열의 상동성은 BLAST 검색을 하였고, 단백질 서열의 상동성은 InterProScan을 이용하였다.

 DNA
 서열에
 기반한
 상동성
 검색은
 NCBI
 Blast
 2.2.28+와
 NCBI

 Non-redundant(nr)
 database(2013/07/17)가
 사용되었고, Blastx를
 사용하여
 unigene

 서열에
 대한
 모든
 가능한
 단백질
 서열을
 non-redundant
 database를
 검색하여서
 그

 기능을
 예측하였다.
 검색
 과정에서
 상동성의
 유의성은
 E-value
 < 1e<sup>-5</sup>로
 설정하였다.

 단백질
 서열
 기반의
 상동성
 검색은
 Hidden
 Markov
 Model을
 이용하여
 단백질

 기능의
 단위인
 도메인
 수준에서
 유사성을
 검색하여
 기능을
 예측하는
 InterProScan

(v5)을 이용하였다. ProDom, PfamA,Panther, SMART, SuperFamily, Gene3d의 총 6개의 단백질 데이터베이스를 E-value < 1e<sup>-5</sup>를 기준으로 검색하였다.

바. 유전자 발현량 측정 및 차등발현유전자 기능정보 분석

개볼락 체색타입별 근육조직 발현유전자의 발현량 분석은 유전자의 발현량을 측정하는 도구인 RSEM (Li and Dewey, 2011) 프로그램을 이용하였으며, Bowtie를 사용하여 리드를 전사체에 정렬한 뒤 directed graph model로 전사체의 발현량을 계산 하였다.

조직별 발현량을 이용하여 체색타입별로 발현량 차이를 나타내는 차등발현유전자 (Differentially Expressed Gene, DEG)를 탐색하였다. DEG 분석은 TCC (Sun et al., 2013) 프로그램의 DEGES/DESeq 방법을 활용하였다. DEGES/DESeq 방법은 Negative-binomial 분포를 이용하여 DEG 분석을 수행하는 DESeq(Anders and Huber, 2010) 방법에 기반한 분석 방법으로 초기 두 시료 사이의 normalization 과정을 반복 수행함에 따라 DEG 선정의 정확도를 높여준다(Kadota et al., 2012). 본 연구에서는 normalization 과정을 3번 반복 수행하였으며, 기준값을 q-value < 0.05 이하로 설정하여 DEG를 선정하였고, 차등발현유전자의 기능 정보는 Gene Ontology (GO) enrichment 분석으로 확인하였다. 2. 연구결과

가. 볼락류의 체색관련 기능유전체 분석을 위한 시료 확보

체색이 다양한 개볼락 시료를 확보하여 반점이나 mark가 있는 color type과 반점이나 mark가 없는 wild type으로 구분하여 체색관련 기능유전체 분석 시료로 활용하였다 (Figure 10).



Figure 10. The skeletal muscles of S. pachycephalus for RNA-seq analysis.

나. 개볼락 체색타입별 근육조직 total RNA 추출

Color type과 wild type 개볼락의 근육조직 total RNA를 추출하여 전기영동으로 확인한 후, RNA-seq 분석에 사용하기 위하여 순도를 조사하였다(Figure 11).



Figure 11. The result of quality assessment of total RNA.

다. 볼락류의 체색관련 기능유전체 확보를 위한 RNA-seq analysis

(1) 개볼락 체색타입별 근육조직 전사유전체 서열 조립 및 unigene 확보

개볼락 체색타입별 근육조직 각각의 cDNA 라이브러리를 제작한 후, 전사체 정보를 분석하기 위하여 염기서열을 해독하였다. 염기서열 정보 중 N으로 나타난 염기의 비율이 전체 서열의 10% 이상 포함되어 있거나, Q20 미만의 염기가 20% 이상인 리드, 평균 quality가 Q20 이하인 리드 및 서열의 양끝에 존재하는 Q20 미만의 염기를 제거하여 전사유전체 서열 조립을 위한 필터링을 수행하여(Table 5) unigene을 확보 하였다.

Sample	Ra	W	Cle	- Low Quality reads		
name	Reads	Basepair	Reads	Basepair		
Bbrf-Wt	78,199,118	7,819,911,800	73,392,934 (93.9%)	7,316,256,602 (93.6%)	4,806,184 ( 6.1%)	
Bbrf-Ct	77,946,600	7,794,660,000	73,454,304 (94.2%)	7,323,430,566 (94.0%)	4,492,296 ( 5.8%)	

Table 5. Summary of short-read data from skeletal muscle of S. pachycephalus

서열 조립은 개볼락 체색 타입별 근육시료의 서열 데이터를 하나로 합친 뒤 Trinity 프로그램을 이용하여 Inchworm, Chrysalis, Butterfly 3단계를 거쳐 진행하였 고, 조립된 전사체 서열의 집단화(Clustering)는 TGICL 프로그램을 이용하였으며, 클 러스터링을 위한 서열 유사성 기준은 0.94로 설정하여 101,112개의 unigene(평균길이 655bp)을 확보하였다(Figure 12).



Figure 12. Results of the blass bloched rockfish (S. pachycephalus) de novo assemly

확보한 101,112개 unigene의 DNA 서열을 BLASTn 검색한 결과 67,419개(66.7%) 유 전자는 상동성이 확인되지 않았고, 상동성이 확인된 33,693개(33.3%)의 유전자를 annotation 하였다(Figure 13). Unigene의 기능은 아미노산 서열을 기반으로 InterProScan v5을 이용하여 예측하였으며, 6개의 단백질 데이터베이스(ProDom, PfamA, Panther, SMART, SuperFamily, Gene3d)를 검색하였으며, 기준은 e-value < 1e<sup>-5</sup>로 정하였다(Figure 14). 개볼락 체색타입별 근육조직 전사체로부터 확보한 unigene을 검색하여 유전자의 기능이 밝혀진 35,482개와 기능이 밝혀지지 않은 65,630개 유전자의 염기서열 정보 및 아미노산 서열정보를 뱅크화 하였다(Figure 15, Table 6, Figure 16, Figure 17).



Figure 13. Frequancy of the blass bloched rockfish unigenes BLAST annotation

35k									Predicte Valide	d d
	3A	1 I 9 Ø	4	8	9		-0			
Number of CDS of unigene	s 1	2	3	4	5	6	7	8	9	10
Predicted	39710	6829	1229	234	39	9	2	0	0	1
Validated	29363	797	30	4	0	0	0	0	0	0
CDS type	SUM	Has ho	mologou	5			No Hor	nolog	ous	
Total	58,254	31,06	3	53.3%			27,191	1		46.7%
Complete	16,941	7,35	51 43.49		9,59		9,590			56.6%
5prime_partial	14,882	7,93	7,932 53		6,950		í Í		46.7%	
3prime_partial	9,756	3 <mark>,</mark> 64	3	37.3%	.3%		6,113			62.7%
Internal	16.675	12 13	7	72.8%			1 5 2 0			27.20

Figure 14. Predict of CDS of unigenes by InterProScan



Figure 15. Annotation consensus between Blastx and InterProScan of the unigenes

1.1	A	В	C	D	E	F	G	Н	1	J	K	L	M N	4	0	P C	2	R S T U V W X Y
1 Ge	neOrder Ge	neID	GeneAcc	GeneName	Chr	Start	End N	Transcript Tra	nOrder	TranID	TranAcc	TranName	Chr Star	irt Er	nd St	trand Leng	th N_E	ixons
2	186 TBI 715 TBI	U000186	TBI_c10/54_g1_i2	Dgkb	Un	1	201	2	203	TBIM000203 TBIM000804	m.239/4	-	Un	2 1	143 -		183	0 Diacylglycerol kinase beta
4	871 TBI	U000871	TBI_c12458_g1_i1	BAZ2B	Un	1	201	1	974	TBIM000974	m.28915	-	Un	3 18	877 +	1	875	0 Bromodomain adjacent to zinc finger domain protein 2B
5	1190 TBI	U001190	TBI_c13334_g1_i1	Sox4	Un	1	201	1	1322	TBIM001322	m.31394	-	Un	3 (	617 +		615	0 Transcription factor SOX-4
6	1722 TBI	U001722	TBI_c14888_g1_i1	R3HDM2	Un	1	201	1	1913	TBIM001913	m.35452	20 00	Un	20 1	462 +		462	0 R3H domain-containing protein 2
8	2483 TBI	U002483	TBI_c17084_q1_i1	KATNAL2	Un	1	201	1	2765	TBIM002011 TBIM002765	m.41054	-	Un	3	185 +	-	183	0 Katanin p60 ATPase-containing subunit A-like 2
9	2617 TBI	U002617	TBI_c17418_g1_i1	meox2	Un	1	201	1	2913	TBIM002913	m.41909	5	Un 4	133 7	762 +		330	0 Homeobox protein MOX-2
10	2792 TBI	U002792	TBI_c17998_g1_i1	Dgat1	Un	1	201	1	3107	TBIM003107	m.43178	-1	Un	3 3	308 +		306	0 Diacylglycerol O-acyltransferase 1
12	2814 TBI 2817 TBI	10002814	TBL_c1807_c1_i2	PLCG1 mfed5	Un	1	201	1	3131	TBIM003131 TBIM003134	m.43328	*: -	Un 1	3 4	440 + 283 ±		438	U 1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase gamma-1 0 Molybdate-anion transporter
13	2929 TBI	U002929	TBI c1850 g1 i1	Pygm	Un	1	201	1	3256	TBIM003256	m.3792	-	Un 1.	123 9	986 +		864	0 Glycogen phosphorylase, muscle form
14	3209 TBI	U003209	TBI_c19695_g1_i1	DIP2C	Un	1	201	1	3573	TBIM003573	m.46117	8	Un	3 1	257 +		255	0 Disco-interacting protein 2 homolog C
15	3372 TBI	U003372	TBI_c20285_g1_i1	DSEL	Un	1	201	1	3757	TBIM003757	m.47312	-	Un	1 4	471 +	-	471	0 Dermatan-sulfate epimerase-like protein
10	3482 TBI 2402 TBI	10003482	TBL_C20647_g1_i2	- Enh4111	Un	1	201	1	3886	TRIM003886	m.48085	-	Un	2 4	421 +		420	0 - 0 Rand 41-like protein 1
18	3630 TBI	U003630	TBI c21222 a1 i1	Pagr3	Un	1	201	1	4043	TBIM004043	m.49156	-	Un	3 3	287 +		285	0 Progestin and adipoO receptor family member 3
19	3787 TBI	U003787	TBI_c21760_g1_i1	hmox	Un	1	201	1	4222	TBIM004222	m.50496	-	Un 1	101 6	697 +		597	0 Heme oxygenase
20	3842 TBI	U003842	TBI_c21874_g1_i1	kbtbd8	Un	1	201	1	4282	TBIM004282	m.50985	e) .	Un	3 5	554 +		552	0 Kelch repeat and BTB domain-containing protein 8
21	3911 TBI 4602 TBI	10003911	TBL_c22051_g1_i1	- PIGG	Un	1	201	1	4361	TBIM004361 TBIM005157	m.51618	2) 21	Un	95 4	235 +		141	0 - 0 GPI ethanolamine phosphate transferase 2
23	4606 TBI	U004606	TBI_c23667_q1_i1	-	Un	1	201	2	5163	TBIM005163	m.57386	-	Un 4	130 5	579 -		150	0 -
24	5085 TBI	U005085	TBI_c24876_g2_i1	FMN1	Un	1	201	1	5707	TBIM005707	m.61215	~	Un	3	245 +		243	0 Formin-1
25	5286 TBI	U005286	TBI_c25322_g1_i1	-	Un	1	201	1	5943	TBIM005943	m.63049	-	Un	2 1	208 -		207	0 -
26	5311 IBI	0005311	1BL_c2538_g1_i1	-	Un	1	201	1	59/0	TRIM005970	m.2608	-	Un	1 1	144 -		0.70	0-
28	5430 TBI	U005430	TBI c25655 g1 i1	Cxxc5	Un	1	201	1	6112	TBIM000031	m.64387	-	Un	3 3	338 +		336	0 CXXC-type zinc finger protein 5
29	5476 TBI	U005476	TBI_c25759_g1_i1	-	Un	1	201	1	6169	TBIM006169	m.64795	5	Un	58	237 +		180	0-
30	5529 TBI	U005529	TBI_c25904_g1_i1	mboat2	Un	1	201	1	6230	TBIM006230	m.65357	2	Un	2 2	277 +		276	0 Membrane-bound O-acyltransferase domain-containing protein 2
31	5550 TBI	U005550	TBL_c25946_g1_i1	EDF1	Un	1	201	2	6255	TBIM006255	m.65525	•	Un 2.	227 6	673 +		447	0 Endothelial differentiation-related factor 1 homolog
33	5717 TBI	U005602	TBL c26275 g1_11	ndafra	Un	1	201	1	6456	TBIM006456	m.66973		Un 1-	3 6	495 +	-	348	0 Platelet-derived growth factor receptor alpha
34	6137 TBI	U006137	TBI_c27160_q1_i1	dip2ba	Un	1	201	1	6937	TBIM006937	m.70894	-5	Un	2 1	295 +		294	0 Disco-interacting protein 2 homolog B-A
35	6677 TBI	U006677	TBI_c28359_g1_i1	babam1	Un	1	201	1	7546	TBIM007546	m.75115	-	Un 1	162 3	311 -		150	0 BRISC and BRCA1-A complex member 1
36	6784 TBI	U006784	TBI_c2858_g1_i1	che-1	Un	1	201	1	7673	TBIM007673	m.6488	C1	Un	2 1	196 +		195	0 Transcription factor che-1
2019 2020 2021	A 21118 TE 22351 TE 22430 TE	B BIU021118 BIU022351 BIU022430	C TBL_c46960_g2_l4 TBL_c47782_g1_l2 TBL_c477830_g1_l5	D TLK2 Golga2 Angptl3	E Un Un Un	F 1 1	G 215 215 215	H 1 3 2	I 25897 27600 27715	J TBIM025897 TBIM027600 TBIM027715	K m.22680 m.24234 i m.24349	L 9 - 8 - 1 -	M N Un Un Un	N 2 1 3 2	O 268 + 069 + 991 +	P (	267 069 990	R         S         T         U         V         W         X         Y           0]Serine/threenine-protein kinase tousied-like 2         0.Golgin subfamily A member 2         0.Golgin subfamily A member 2         0.Golgin subfamily A member 3         0.Golgin Subfamily A member 3
2022	22672 TE	BIU022672	TBI_c47985_g1_i1	Dgkb	Un	1	215	2	28061	TBIM028061	m.24663	1 -	Un 8	802 2	559 +	1	758	0 Diacylglycerol kinase beta
2023	23588 TE	BIU023588	TBI_c48458_g1_i5	Bag6	Un	1	215	3	29355	TBIM029355	i m.25805	- C	Un 9	993 4	004 +	3	012	0 Large proline-rich protein BAG6
2024	23808 TE	BIU023808	TBI_c48558_g2_i2	Hectd2	Un	1	215	2	29672	TBIM029672	m.26110	8 -	Un 14	2	298 +		297	0 Probable E3 ubiquitin-protein ligase HECTD2
2025	24341 TE 24524 TE	BIU024341 BIU024524	TBI_c48811_g1_i8	- SPAG9	Un	1	215	1	30441	TBIM030441	m.2686/	5 -	Un 14	454 1	567 +		567	0 - 0 C-lun-amino-terminal kinase-interacting protein 4
2027	24983 TE	BIU024983	TBI_c49107_q2_i4	Cog1	Un	1	215	1	31354	TBIM031354	m.27800	6 -	Un 4	412	975 +		564	0 Conserved oligomeric Golgi complex subunit 1
2028	25172 TE	BIU025172	TBI_c49175_g1_i1	saal1	Un	1	215	2	31623	TBIM031623	m.28102	5 -	Un 6	680 1	378 +		699	0 Protein saal1
2029	25174 TE	BIU025174	TBI_c49176_g1_i1	EII3	Un	1	215	2	31626	TBIM031626	m.28104	8 -	Un 3	350	565 +		216	0 RNA polymerase II elongation factor ELL3
2030	25290 TE	BIU025290	TBI_c49219_g3_i2	ABCB7	Un	1	215	1	31801	TBIM031801	m.28280	1 -	Un	3	653 +		651	0 ATP-binding cassette sub-family B member 7, mitochondrial
2031	20425 10	BIU020425 BIU027361	TBI_c49921_g1_i3	PIEZO2	Un	1	215	1	34734	TBIM034734	m.31154	1 -	Un	100	621 +		621	0 Piezo-type mechanosensitive ion channel component 2
2033	27931 TE	BIU027931	TBI_c50097_g3_i6	RABGGTA	Un	1	215	1	35588	TBIM035588	m.32029	5 -	Un 1	124 1	536 +	1	413	0 Geranylgeranyl transferase type-2 subunit alpha
											•							
4	A	В	С	D	E	F	G	Н	I	J	К	L	M I	N	0	P (	2	R S T U V W X Y
48016	40740 T	BIU040740	TBL_c75882_g1_i1	- ST5	Un		4934	1	5030	TBIM050305	5 m.42577	3 -	Un	3	239 +		237	0 - 0 Suppression of tumoridenicity 5 protein
48018	42484 T	BIU042484	TBI_c80163_q1 i1	-	Un	1	4939	1	5221	B TBIM057578	8 m.43714	6 -	Un	1	192 +	-	192	0-
48019	47261 T	BIU047261	TBI_c92353_g1_i1	-	Un		L 4944	1	5740	7 TBIM057407	7 m.47189	1 -	Un 4	409	528 -		120	0-
48020	47104 T	BIU047104	TBI_c91897_g1_i1	-	Un		4983	1	5724	2 TBIM057242	2 m.46866	7 -	Un	39	251 +	-	213	0-
48021	40819 T	BIU040819	TBL_c76091_g1_i1	-	Un	1	L 5007	2	5039	I TRIMOSOSO	2 m.42627	7 -	Un 2	247	372 +		126	0-
48022	40919 T 45044 T	BIU040919	TBI c86728 a1 i1	-	Un		L 5007	1	5499	1 TBIM054991	1 m.4268/	7 -	Un	2 792	+12 + 917 +		126	0-
48024	47641 T	BIU047641	TBI_c933_g1_l1	Strip1	Un		5019	1	5781	1 TBIM05781	1 m.1826	-	Un	2	223 +	-	222	0 Striatin-interacting protein 1
48025	35377 T	BIU035377	TBI_c62614_g1_i1	SMAD7	Un	8	L 5070	1	4442	8 TBIM044428	8 m.38954	0 -	Un	3	200 +		198	0 Mothers against decapentaplegic homolog 7
48026	47390 T	BIU047390	TBI_c9271_g1_i1	-	Un	1	5071	1	5754	4 TBIM057544	4 m.21452	-	Un 1	101	235 +	-	135	0-
48027	47642 T	BIU037082	TBL c69255 g1 i1	-	Un		5163	1	4727	TBIM057812	2 m.47208	4 -	Un	3	242 +	-	138	0-
48029	46298 T	BIU046298	TBI_c89753_q1 i1	CLDN4	Un		1 5163	1	5636	TBIM056360	0 m.46325	4 -	Un	16	228 -		213	0 Claudin-4
48030	47452 T	BIU047452	TBI_c92873_g1_i1	-	Un	3	L 5193	1	5761	1 TBIM05761	1 m.47092	1 -	Un	2	211 +	-	210	0 -
48031	45716 T	BIU045716	TBI_c88309_g1_i1	atp1b3	Un	1	L 5208	1	5572	9 TBIM055729	9 m.45918	4 -	Un	2	304 +		303	0 Sodium/potassium-transporting ATPase subunit beta-3
48032	46297 T	BIU0148297	TBL c41290 c1 i1	Dcun1d2	Un		5256	1	56359 1746	TBIM056359	m.46325	9 -	Un	2	499 +		498	0 DCN1-like protein 3
48034	43058 T	BIU043058	TBI_c81424_q1 i1	-	Un	1	5285	1	5284	5 TBIM052845	5 m.44090	2 -	Un 1	179	340 -	8	162	0-
48035	47259 T	BIU047259	TBI_c92346_g1_i1	-	Un		5318	1	5740	5 TBIM057405	5 m.46972	2 -	Un	3	104 -		102	0-
48036	21177 1	BIU021177	TBI_c46998_g1_i3	Myole	Un	1	L 5410	1	2598	TBIM025980	0 m.22735	0 -	Un	31	1886 +		.884	0 Unconventional myosin-le
48037	44505 T	BIU044505	(BL_c85313_g1_i1 TBL_c9136_c1_i1	- cxcl12	Un	1 2	5422	1	5440	TBIM054405	m.45077	4 -	Un	2	313 +		312	0 Stromal cell-derived factor 1
48039	40697 T	BIU040697	TBI_c75778 a1 i1	GPHN	Un	1	5520	1	5025	BIM050258	8 m.42550	2 -	Un	2	358 +	-	357	0 Gephyrin
48040	44947 T	BIU044947	TBI_c86500_g1_i1	-	Un		5541	2	5488	TBIM054880	0 m.45360	4 -	Un	3	395 +	-	393	0-
48041	46723 T	BIU046723	TBI_c90827_g1_i1	PapIn	Un	8	L 5546	1	5682	7 TBIM05682	7 m.46603	8 -	Un	74	286 +	-	213	0 Papilin
48042	45529 T	BILI020100	IBL c87873_g1_i1	-	Un	1	5550	1	5552	2 (BIM05552)	2 m.45787	4 -	Un	3	275 +	-	273	0 Cutoplasmic dynain 2 heavy chain 1
48044	45362 T	BIU045362	TBI_c87474_a1_i1	Arhaef10	Un		5608	1	5533	7 TBIM05533	7 m.45665	1 -	Un	3	302 +		300	0 Rho guanine nucleotide exchange factor 10
49045	47451 7	DTI 10474E1	TOI -02071 -1 11	CCC214	Um	1	5620	1	5761	TRIMOS7610	a m 47001	4	Un	2	250 .		257	0 Protoin transport protoin Soc214

# Table 6. Unigenes transcripts detected in this study

el protein type

ha B

Couter New •	101112 total and	pencer		
Mode Select / Skile	Selector 0 Fination (2)	Sequence Mark None Numbers Mark None	Sur D	
	The DOLLAR STATE OF STATE	CITAL THE CONTRACTOR	south and a second s	
W. I D T D M co 4	. 后面的时间 私國家 供	STACK	i meet door 🗛 📲 tot	
		10 20	20 30 40 50 60 70 80 90 100 110 120 130 140	150 2
TBIU048853 TBI c20 c	1 11	TROCCCAGIASTGAATCOTOT	TOTTETCAARCARACTARTARTAGETETESAAAAAACAAACAAACEAACETTACGGGGCTGTGTGTGTGTGTGTGTGTGCGACACACAGAGAACGGCGACAACTGAGAACCETTCAGAGAGCGCACAGAGAGCGCGGGGGGGGGG	PCCAAACAGAAAAT
TRIUDGBUSG TRI CT Q1	at it	AAAOTOOCTGAGAAAOTOOCTG	THAD BAY THAT DAN DATE OF THE DAMA OF THE DAMA OF THE DAMA THAT OF THACK AND ACCOUNT AT COMPACT THAT THAT THAT THAT THAT THAT THAT T	ECCTPTGGCCGCEC
7810003854 781 c18 ;	1 il cep89 Centrosonal p	FUE ACCTGAGGTCATCTOCAGGCAG	icasocassocratorearerearerearerearerearerearerearer	ATTOTCCARACAGG
Thioddensa Thi cl234	q1 11	OTOTTATTICCTOTTATCAT TGLCLAGTTLTACAGLCTOTCLAG		PEATPTAAACTOCT
TBIUD46395 TBI c8 g1	il ARRGAP20 Eho OTFase-	AGE CARCOACAACCAGATGAACGCC	1900TTCAACCTOPCC0TCTGCATC9CTCCCAGCATGCTTTG5GCTCC00GCCCAAGCACCCCCCGAGATG3AGGCGAGGGCACTAAGAAGGTGTGTGAGACGGTTC5CCCCCCTCATAGAGAAGTGCA	AGRAGEGECCEGGG
TBIU048058 VBI c13 c	111	COATCIGTGTGTGTATAGTGTOTA	KOTADTOT DEADEUTUTOTA IFUTUTOT DE TATATOT DE TRANSPORTA EL CONTRA COSOCETA HADOUT UTUTOT CADARUCADA AUTOUT DE ACUADA ALA CADA ALA C	SCATCATATITOTT
7810048009 TB1 C19 c	1 11	ACTOCOGTTTTTGGATTGATG		CTGACATGCACAA
7810000808 781 c123	q1 i1 = -	CACACCOUCTCOACAOTTCAOT	aproxyada caraara triatectro cart to an economic and a criteric constantia da da ana tricer triatector internation of the constantian of the constant of the const	ADDTARATORA
Thi0040061 Thi cl234	41,11	TPCCGATGTTAAGCTCGTTTG		CATTTAUTURADAT
TRIUD48063 TRI c30 c	1 11	GTGATGTAGTGCACAACACGCO	SOCK-CCCARGEOGETEN-TOTOGTOTTAACAGGCAGCTOAGCAGCTOAGAAGATGGTTATTTTGGTTTAGTCTPTTTTGGTTTAGCATCCAGAAGGCTOCCCAAACCATCTTAAATTA	CATACAGTATACTA
TEIU000841 TEI c123	gl il zic4 Zinc finger	Pro CADAGGTGAGAAGCCCTTCAAJ	IAATYFOADTTTOLOGGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	TOODTORAGOTCOCT
2810048064 081 01240	di il	GAGCATGATATGOGACCTTTAL	TRADUCT AGUARACOURO ARCANTOCCCARGOUT CACTANTITI AVAILARY TATUAT CAN AN ANTITI AUA CANTUR CARGOTTALAR TOTAL TOTAL CARGOTTALAR TOTAL CARGOTTAL AND TATUAT	PTTRACAAOTTTOC
2810048065 781 c36 c	1 11	ACCEASATCTGOTOGAACAAA	urreger and the second s	200A00TTA7A0T0
TBI0048066 TBI 037 (		GAATATCCTTOAOTTOTTSAA TCTGAAGAGTGAATTAAAGTT		SANGAACTTACATT
TRIU019834 TRI C45	1 11 MFSD12 Major facili	TAT OGARGACATOGCATTTAGTTGO	$\label{eq:constraint} Transformed the transformed a constraint of the constraint of the transformed transformed that the transformed the transformed transformed to the transformed transformed to the transformed tra$	29996999999869
TBI0021103 TBI c46 c	1 11 ARIDIN A7-rich inte	TAC CATCOTCCGGAGCCTGTCCTTT	TTATE DOGATED CTEATATE CONDUCT DOT CONTOUR DOGATO TO TOTE TO A CARCACTERATING ALCANT OCCUPATION OF A DAGAC	JAGCASCASGASCA
2810048068 TB1 c50 c	1 il	00CTCRASASASCC00TACO		CCTGOCTOTCTGAG
7810048068 781 c51 c	1 11	BOCARATTOCTCARCETCAR	CATCEAGAAAADEACCTIGATCETCOUTOACCATCIGATCEADECOUCTOUCTOUTCACTURCTCURCEADUGUTCUAACATCATOCTUTOTECOUCTOUTCAAC	SCREDCECCAGOC
TRIU0333/6 TRI 05/ 0 TRIU033994 TRI 058 0	1 11 SEMLAR Semanhorin-6	a CACSCASATCACTSTCCATCS	THE REPORT OF A MARANEE STATATION TO THE ATTOC A REAL ASSOCIATE A TABLE OF A TABLE AND A DATA AND A DATA AND A	receasesseses
TRIU035509 TRI c62 c	1 il LMAN2L VIP36-like p	TOT SCTATTTTCTTCACCOTGCTOT	ITOTTTCCAFGCT0007T0CATCTCCTCAFCGTCAT000CCT000FCTACAGCCACT00AA5GAGAGCCGACCGACCCACT0TACT0AGAGAAAGAACAAAAATACAGAFGAAFFT0AFGAFFT0AFGAFGAFGAAF	107GACGACTOCOO
TEIU035510 TEI C62 C	2 11 IMANZL VIPLE-like p	GCATGGATAGOGGTAGOCTATTA		ATAAOCATAOCCAG-
2810037534 TB1 c67 g	1 il GTF218D2 General tr	and GAACTORATOTORACCTACAGO	CAGADARAAN TONOTTTOTACA TARCA TATACAARA TOTGADAR COTTCARA TOTACO TO ACCATOT TO TO CAGADARCA TOTGADARCA ATT TO COCATA TARCA TARCA TOTACAA	AAGAGGCCCCCTGA
TBI0038237 TBI -69 (	1 11 + -	CACACACACACACTEACACA	INCATABY'S ADDITE OF TACTORIAN CACACITY OF THAACACITY OF TO TAADAD THAT AT TACT TACTADACIAN CACACITAT CACOUNT TUTO CAT DAD TO THE	ANOTTOTTOOTTOO
ThIU046072 TBI 075 0		ARTOTAAACACATTTTTATTAC	TACTUTT ON TURKE, AUTOMPTT CORAL ARGARAGUAR AUTOMPTOTATION ART TUTTTTTTT AUTACTARA TARA TARA TARTATTAR TUTUTT	2TGTGAA3GATTTT
TRI0041159 TRI c76 c	1 41	CCTGGAGACCTFTCGCTCFGFG	A TRADUCTURA DE LA CONTRACTOR DE LA CONTRACTOR DE LA CONTRACTÓRIA DE LA CONTRACTÓRIA DE LA CONTRACTOR DE LA CONTRACTÓRIA DE LA	ITTG0007TTCPTT
TRIUG48073 TRI CBC C		CCATATOSGGCATOCAGGCTTA	TRACTATION TARGET AND A TRACTATION AND TARGET	TOCCATATTTATA
2010044375 001 084	1 11	CACTOOPCCTCTACCTCACCTC	RETUTION ON OF CONTRACTOR OF CONTRACTOR CONTRACTOR CONTRACTOR OF CONTRACTOR CON	100ATTTCACTACA
7810044376 781 c84 c	2 il + - 1 il latal Decreis caras	TOCCCTOTTOCACCTOCTCASC	AND TECHNOLOGIUS, TEACTERISE NO. AS AN TITUTE AT TRADUCTOR AS A TRADAR AND COMPACT AT TRADAT SOCIETY AND	ADGCACACACADAT
7810045581 781 c87 c	1 11 code134 Coiled-coil	do CTCACACTGATOTCACAOGAGO	INSCRIMENTAL SUBJECT STATE TAACCOSTAC SCALE AGAGA SATE SIGT COLOR SUBJECT STATE STA	AAGAGAAGAGAGAAA
TBI0048075 TBI C88 (	141	AAATATTAATTTOCAACAAAA	алтатталттостисталаттасалтаттисалалтаттиссалалтасти балссалатисталтисталтисталтисталтисские жестосла сасылтальоттисскогла	ATANTAGTTTGCG
TBIU048077 TBI 053 0	1 11	OTTATOOTULCAOTATATOOL	NORMALIAN ANALASIAN DI UTUTUTUTUTU ANALASI ANTAN CAMBINATIAN CANTUN ANTAN ANTAN ANTAN ANTAN ANTAN ANTAN ANTAN A	SOCTORATAGOGTT
2010047976 001 095	1 11	0020020020020020020020	10707ACT400CCA64CHCCATCCATCCTTCT0CA6ACATCTTCTCCA6A6A0A686ACCTCA6CCTCCTTCTCTCTCTCTCTCCACCCACCCACC	JCT00CAGAAAGAG
TBIU048078 TBI 097 ( TBI0048079 TBI 0102	1 11 REEO1 RMA exonuclea	TOGACTOTTOAGAAAACAAM		CATTATOSACATSC
ThIU040040 TBI 0104	g1 11 = =	GAGAAACAATATGTCTGOGTAA	FEARTATOAASSASTCAAATACCATACTACTACTACTACTACTCACTC	AAATTTTTAGATGT
TRI0048081 TRI 0105	91 11	AATATACCTUTATTCACCCTCT COGEDOECETTCOCACCCTCT		ATCRARGENCE ATCRARGENCE
TRIU000380 TRI cili	di il pop-l Multidrug re	ISIS CTATCOTCASASCTOPSOTTOS	TO SALESCO CAROTALY ATA, TO SALESCO COCCASE, TO SALETO SA	SACCOTOGAGAASG
2210000418 221 cl12	ul 11 + +	CTCTCTOCTOCACAGOTOSTTC	ALGEBRAR DEBRECESSTAR ADECAR SECARES CACASATES AS A TECCARE SECARES AS	1907CAACCACAGE
		a stated		

Figure 16. DNA sequences of unigene transcripts



Figure 17. Amino acid sequences of unigene transcripts

(2) 개볼락 체색타입별 근육조직 전사유전체의 발현량 측정

RSEM 프로그램을 이용하여 개볼락 체색 타입별 근육조직 발현량을 분석하였으며 Bowtie를 이용하여 리드를 전사체에 정렬한 뒤 directed graph model로 전사체의 발현량을 측정하였다. 개볼락 근육조직에서 발현되는 유전자 개수가 체색타입별로 차이를 나타내었다. 노란색 반점을 가진 color type (Bbrf-Ct) 근육조직에서는 53,499개의 유전자가 발현되었으며, wild type (Bbrf-Wt) 근육조직에서는 55,508개의 유전자가 발현되었다(Figure 18).



Figure 18. Expressiom levels of unigene transcripts

(3) 개볼락 근육조직의 체색타입별 차등발현유전자 탐색

개볼락의 체색 타입에 따른 DEG를 탐색하고 분석한 결과, 164개의 유전자가 발현량 차이를 나타내었으며(Figure 19), 차등발현유전자중 wild type에서만 발현되는 68개의 유전자와 color type에서만 발현되는 51개의 유전자를 확보하였다(Table 7).

개볼락 color type 근육조직에서 차등발현하는 유전자 개수는 83개 중 51개 유전자는 wild type 근육조직에서는 발현되지않고 color type 근육조직에서만 발현되었으며, 35개는 기능이 밝혀진 유전자이었다(Table 8). 반면, 개볼락 wild type 근육조직에서 차등발현하는 유전자 개수는 81개이었고, 68개 유전자는 color type 근육조직에서는 발현되지않고 wild type 근육조직에서만 발현되었으며, 44개가 기능이 밝혀진 유전자 이었다(Table 9).



Figure 19. Differentially expressed genes (DEG) analysis of unigene transcripts

GeneId	GeneName	q-value	DEG:UP_DOWN	EXP:Bbrf_WT:FPKM	EXP:Bbrf_CT:FPKM
TBIU062084	-	0.000288592	DOWN	270.35	0
TBIU064175	-	7.42E-05	UP	0	72.53
TBIU064176	-	1.60E-05	DOWN	341.79	0
TBIU065291	-	0.00541945	DOWN	29.63	0
TBIU065468	-	0.000878531	DOWN	82.08	0.1
TBIU067415	rnaset2	0.00327647	DOWN	9.03	0
TBIU067654	-	0.00332399	UP	0	14.36
TBIU068429	-	0.000901415	UP	0.11	55.42
TBIU069081	-	0.00317218	DOWN	23.33	0
TBIU069579	-	0.00541945	UP	0	9.14
TBIU069971	-	0.000136337	UP	0.16	246.51
TBIU069985	-	0.00101459	UP	0	21.7
TBIU069986	-	0.00174934	DOWN	19.56	0
TBIU070060	-	0.00648428	DOWN	20.85	0
TBIU070198	-	0.00877295	DOWN	10.6	0
TBIU070559	-	0.000205273	DOWN	72.34	0
TBIU070668	-	0.00374817	DOWN	20.79	0.02
TBIU070994	-	0.000562856	UP	1.91	978.01
TBIU071086	-	0.00397836	DOWN	9.72	0
TBIU071087	-	0.00399712	UP	0	9.39
TBIU071155	-	0.00225047	DOWN	37.96	0
TBIU071992	-	0.00374817	UP	0	35.14
TBIU073514	-	0.0092286	UP	0	15.61
TBIU072198	-	0.00201045	DOWN	27.79	0
TBIU072278	-	0.00790906	UP	7.13	688.67
TBIU073257	-	0.00631178	UP	0	11.92
TBIU073667	Etnk1	0.000350436	DOWN	11.09	0
TBIU073754	-	3.84E-05	UP	0	158.4
TBIU073969	-	0.000581513	UP	0	15.98
TBIU074890	-	0.00572695	UP	0	19.16
TBIU075697	-	0.00545732	UP	0.14	56.99
TBIU075704	-	0.00397836	UP	0	18.8
TBIU019675	letm2	0.00182905	UP	0.04	14.88
TBIU024777	CUL1	5.88E-05	UP	0	58.69
TBIU011833	-	0.000443256	UP	0	48.18
TBIU027578	NMRK2	1.60E-05	DOWN	453.55	0.02
TBIU004254	Sds	0.00101915	UP	0.22	65.74
TBIU027579	NMRK2	1.60E-05	DOWN	520.22	0
TBIU022000	-	0.00313787	DOWN	122.01	0.46

Table 7. List of color pattern type-specific differentially expressed genes

GeneId	GeneName	q-value	DEG:UP_DOWN	EXP:Bbrf_WT:FPKM	EXP:Bbrf_CT:FPKM
TBIU005394	bcas2	0.00180366	DOWN	21.63	0
TBIU020370	-	0.00273889	UP	0	17.38
TBIU004387	TOMM7	5.94E-05	DOWN	147.43	0
TBIU025303	-	0.000903742	UP	0	13.79
TBIU027572	NMRK2	1.60E-05	DOWN	184.81	0
TBIU018571	LACTB	0.00349544	UP	0	11.22
TBIU026119	MST4	0.000831836	DOWN	18.56	0
TBIU027793	DCAF11	0.000640213	DOWN	14.4	0
TBIU011844	mrp63	0.000831836	UP	0	50.88
TBIU028221	murc	0.000382233	DOWN	17.89	0
TBIU028294	-	0.000500669	DOWN	41.26	0
TBIU027575	NMRK2	3.84E-05	DOWN	129.75	0
TBIU028890	S100a9	0.00371605	DOWN	24.49	0
TBIU010317	vma21	8.86E-05	DOWN	67.08	0
TBIU025670	MRPL18	0.000382233	DOWN	43.5	0
TBIU014060	MYH7	2.43E-05	UP	0	441.29
TBIU027577	NMRK2	0.00660113	DOWN	91.4	0.78
TBIU019691	-	0.000146252	UP	0	24.81
TBIU021247	ANXA6	0.000665209	UP	0	15.31
TBIU023744	Rnf213	0.000355721	DOWN	40.56	0
TBIU015845	SCEL	0.000471375	DOWN	19.74	0
TBIU022338	NBR1	0.00973217	UP	0.04	14.22
TBIU010365	-	0.000351338	UP	0.34	245.76
TBIU014932	MRPS14	0.00310118	UP	0	30.76
TBIU028775	samhd1	0.00131392	DOWN	18.96	0
TBIU010202	Wdr33	0.000382233	UP	0	40.17
TBIU013082	fos	0.00431954	DOWN	38.75	0.24
TBIU020844	YPEL5	0.00522832	DOWN	9.06	0
TBIU026224	ascc3	0.000562856	UP	0	9.65
TBIU023185	tiprl	0.00476572	UP	0	9.89
TBIU001129	C1QTNF3	0.00215564	UP	0.57	105.08
TBIU023347	GIMAP8	0.00973217	DOWN	100.51	1.05
TBIU028776	samhd1	0.000903742	UP	0	20.02
TBIU021383	Ppan	0.00141128	DOWN	27.74	0
TBIU008482	PFKM	0.000760019	UP	0.61	193.8
TBIU017755	isca1	9.91E-05	UP	0	67.85
TBIU008448	HRSP12	0.000382233	DOWN	54.72	C
TBIU028292	-	0.00101915	DOWN	70.09	0
TBIU024593	GIMAP7	0.00840863	DOWN	6.62	C
TBIU010127	SNTX subunit beta	0.000471375	DOWN	1331.32	0.43

GeneId	GeneName	q-value	DEG:UP_DOWN	EXP:Bbrf_WT:FPKM	EXP:Bbrf_CT:FPKM
TBIU016181	ZRSR2	0.00259161	DOWN	8.55	0
TBIU019960	Lgals9	0.000180366	UP	0	38.83
TBIU017754	isca1	9.91E-05	DOWN	68.02	0
TBIU009302	-	7.36E-05	UP	0	126.86
TBIU004388	TOMM7	6.02E-05	UP	0	163.5
TBIU023658	cmas	0.00028491	UP	0	23.93
TBIU012460	-	0.000894994	DOWN	20.32	0
TBIU020950	PCBP2	0.000279397	DOWN	23.78	0
TBIU030108	klhl20	0.00166919	UP	0.22	42.75
TBIU015715	IGFN1	0.00476572	UP	0	6.88
TBIU028527	Emilin1	0.000208004	UP	0	14.56
TBIU025158	-	0.000298688	UP	0	31.45
TBIU025851	TNS	0.000125783	DOWN	65.93	0.06
TBIU017534	ACADM	7.42E-05	DOWN	46.83	0
TBIU026028	slc38a2	0.000124019	DOWN	18.78	0
TBIU016140	unc45b	0.00572695	UP	0	5.72
TBIU021898	RBM12B	0.000740309	UP	0	8.94
TBIU016179	ZRSR2	0.000629851	UP	0	13.07
TBIU021897	RBM12B	0.000424178	DOWN	11.05	0
TBIU023497	SGCA	8.86E-05	DOWN	220.52	0.13
TBIU017010	LEG_ELEEL	0.00729205	UP	3.36	2635.33
TBIU018148	FEL	0.000640213	UP	0.51	160.02
TBIU005395	bcas2	0.00264582	UP	0	19.06
TBIU013080	fos	0.000333069	DOWN	25.34	0
TBIU025160	-	0.000148257	DOWN	39.9	0
TBIU008231	ARL3	0.00190558	UP	0	24.19
TBIU005548	-	0.0003629	DOWN	148.14	0
TBIU025389	-	0.000529899	DOWN	22.94	0
TBIU010126	SNTX subunit beta	0.000124019	UP	0.15	125.42
TBIU026370	-	0.00235698	DOWN	71.57	0
TBIU028859	Dhx8	0.00314893	DOWN	10.11	0
TBIU017395	-	0.00186158	DOWN	35.51	0
TBIU030393	MYOT	5.94E-05	UP	0	20.64
TBIU026051	Nfkb2	0.00374817	UP	0	8.85
TBIU023584	Bag6	0.00724879	DOWN	4.23	0
TBIU027220	CTL	8.86E-05	UP	0	71.44
TBIU015844	SCEL	0.00259161	UP	0.2	37.93
TBIU019692	-	0.00028491	DOWN	18.9	0
TBIU008479	PFKM	0.00374817	UP	0.51	113.64
TBIU022019	tmem66	0.00541772	DOWN	10.57	0

GeneId	GeneName	q-value	DEG:UP_DOWN	EXP:Bbrf_WT:FPKM	EXP:Bbrf_CT:FPKM
TBIU017325	Ube2l3	0.000529899	DOWN	26.21	0
TBIU023496	SGCA	0.000216366	UP	0.23	145.08
TBIU026697	fam73a	0.000476091	DOWN	11.1	0
TBIU025540	BAG4	0.000374894	DOWN	18.58	0
TBIU019961	LGALS9	0.000773707	DOWN	22.97	0
TBIU025183	CAC1S_CYPCA	0.000208004	UP	0	35.2
TBIU029397	Aloxe3	0.00313787	DOWN	15.58	0
TBIU028649	-	0.000443256	DOWN	14.43	0
TBIU031738	GPI	0.000903742	UP	0.69	598.34
TBIU010316	vma21	9.21E-05	UP	0	63.64
TBIU026225	ascc3	0.00141128	DOWN	7.24	0
TBIU026410	-	0.00180366	UP	0	8.23
TBIU029225	brd4	0.00407128	DOWN	5.67	0
TBIU038365	MYSS_CYPCA	0.00401887	UP	2.08	299.64
TBIU025184	CAC1S_CYPCA	0.000873524	DOWN	21.09	0
TBIU027136	CUZD1	0.00122484	UP	0.71	185.84
TBIU028904	ZNF106	5.88E-05	UP	0	18.34
TBIU029809	NEB	0.000119949	UP	0	18.82
TBIU028488	Acot11	0.00186827	UP	0	14.25
TBIU026001	-	0.0022221	DOWN	29.66	0
TBIU020166	-	0.00484085	DOWN	84.23	0
TBIU028706	-	0.00174934	DOWN	19.86	0
TBIU025069	ASL	0.00186825	UP	0	15.92
TBIU023901	V-MIL	0.00322677	UP	0	19.34
TBIU015572	egr1	0.00943071	DOWN	46.82	0.47
TBIU022079	Gm129	9.91E-05	DOWN	23.03	0
TBIU029492	Thrap3	0.0031625	UP	0.1	15.4
TBIU025757	GOT1	0.00837356	UP	0.59	53.97
TBIU020411	FAM57B	0.00440512	DOWN	12.89	0
TBIU026698	fam73a	5.94E-05	UP	0	28.19
TBIU008401	lyrm7	0.00886531	UP	0	21.49
TBIU020043	krt13	0.00652609	UP	2.93	867.66
TBIU020410	FAM57B	0.00724879	UP	0	10.64
TBIU024378	ACVR1	0.00474586	DOWN	16.87	0
TBIU025825	POL	0.00101915	DOWN	14.92	0.03
TBIU029849	TNNI3	0.00141128	UP	0.51	106.61
TBIU027072	SLC4A7	0.000471375	DOWN	20.99	0
TBIU033342	Apcs	0.00392642	UP	2.17	306.99
TBIU029784	b2m	0.000102214	UP	0	57.3
TBIU026562	FLNC	5.88E-05	DOWN	13.49	0
TBIU029975	-	3.84E-05	UP	0	46.64
TBIU024774	CUL1	0.000180366	DOWN	97.11	0.11
TBIU029786	b2m	5.88E-05	DOWN	66.42	0
TBIU045153	-	0.00584705	UP	3.61	398.03
TBIU042518	-	0.000562856	UP	0.33	173.1
		5.555552050	01	0.00	1, 3.1

No.	GeneId	GeneName	Description	EXP:Bbrf_WT:FPKM	EXP:Bbrf_CT:FPKM
1	TBIU024777	CUL1	Cullin-1	0	58.69
2	TBIU018571	LACTB	Serine beta-lactamase-like protein LACTB, mitochondrial	0	11.22
3	TBIU011844	mrp63	Ribosomal protein 63, mitochondrial	0	50.88
4	TBIU014060	MYH7	Myosin-7	0	441.29
5	TBIU021247	ANXA6	Annexin A6	0	15.31
6	TBIU014932	MRPS14	28S ribosomal protein S14, mitochondrial	0	30.76
7	TBIU010202	Wdr33	pre-mRNA 3' end processing protein WDR33	0	40.17
8	TBIU026224	ascc3	Activating signal cointegrator 1 complex subunit 3	0	9.65
9	TBIU023185	tiprl	TIP41-like protein	0	9.89
10	TBIU028776	samhd1	Deoxynucleoside triphosphate triphosphohydrolase SAMHD1	0	20.02
11	TBIU017755	isca1	Iron-sulfur cluster assembly 1 homolog, mitochondrial	0	67.85
12	TBIU019960	Lgals9	Galectin-9	0	38.83
13	TBIU004388	TOMM7	Mitochondrial import receptor subunit TOM7 homolog	0	163.5
14	TBIU023658	cmas	N-acylneuraminate cytidylyltransferase	0	23.93
15	TBIU015715	IGFN1	Immunoglobulin-like and fibronectin type III domain-containing protein 1	0	6.88
16	TBIU028527	Emilin1	EMILIN-1	0	14.56
17	TBIU016140	unc45b	Protein unc-45 homolog B	0	5.72
18	TBIU021898	RBM12B	RNA-binding protein 12B	0	8.94
19	TBIU016179	ZRSR2	U2 small nuclear ribonucleoprotein auxiliary factor 35 kDa subunit-related protein 2	0	13.07
20	TBIU005395	bcas2	Pre-mRNA-splicing factor SPF27	0	19.06
21	TBIU008231	ARL3	ADP-ribosylation factor-like protein 3	0	24.19
22	TBIU030393	MYOT	Myotilin	0	20.64
23	TBIU026051	Nfkb2	Nuclear factor NF-kappa-B p100 subunit	0	8.85
24	TBIU027220	CTL	Galactose-specific lectin nattectin	0	71.44
25	TBIU025183	CAC1S_CYP CA	Dihydropyridine-sensitive L-type skeletal muscle calcium channel subunit alpha-1	0	35.2
26	TBIU010316	vma21	Vacuolar ATPase assembly integral membrane protein vma21	0	63.64
27	TBIU028904	ZNF106	Zinc finger protein 106	0	18.34
28	TBIU029809	NEB	Nebulin	0	18.82
29	TBIU028488	Acot11	Acyl-coenzyme A thioesterase 11	0	14.25
30	TBIU025069	ASL	Argininosuccinate lyase	0	15.92
31	TBIU023901	V-MIL	Serine/threonine-protein kinase-transforming protein mil	0	19.34
32	TBIU026698	fam73a	Protein FAM73A	0	28.19
33	TBIU008401	lyrm7	Complex III assembly factor LYRM7	0	21.49
34	TBIU020410	FAM57B	Protein FAM57B	0	10.64
35	TBIU029784	b2m	Beta-2-microglobulin	0	57.3

Table 8. Known genes of color-type skeletal muscle specific expressed genes

No.	GeneId	GeneName	Description	EXP:Bbrf_WT:FPKM	EXP:Bbrf_CT:FPKM
1	TBIU067415	rnaset2	Ribonuclease T2	9.03	0
2	TBIU073667	Etnk1	Ethanolamine kinase 1	11.09	0
3	TBIU027579	NMRK2	Nicotinamide riboside kinase 2	520.22	0
4	TBIU005394	bcas2	Pre-mRNA-splicing factor SPF27	21.63	0
5	TBIU004387	TOMM7	Mitochondrial import receptor subunit TOM7 homolog	147.43	0
6	TBIU027572	NMRK2	Nicotinamide riboside kinase 2	184.81	0
7	TBIU026119	MST4	Serine/threonine-protein kinase MST4	18.56	0
8	TBIU027793	DCAF11	DDB1- and CUL4-associated factor 11	14.4	0
9	TBIU028221	murc	Muscle-related coiled-coil protein	17.89	0
10	TBIU027575	NMRK2	Nicotinamide riboside kinase 2	129.75	0
11	TBIU028890	S100a9	Protein S100-A9	24.49	0
12	TBIU010317	vma21	Vacuolar ATPase assembly integral membrane protein vma21	67.08	0
13	TBIU025670	MRPL18	39S ribosomal protein L18, mitochondrial	43.5	0
14	TBIU023744	Rnf213	E3 ubiquitin-protein ligase RNF213	40.56	0
15	TBIU015845	SCEL	Sciellin	19.74	0
16	TBIU028775	samhd1	Deoxynucleoside triphosphate triphosphohydrolase SAMHD1	18.96	0
17	TBIU020844	YPEL5	Protein yippee-like 5	9.06	0
18	TBIU021383	Ppan	Suppressor of SWI4 1 homolog	27.74	0
19	TBIU008448	HRSP12	Ribonuclease UK114	54.72	0
20	TBIU024593	GIMAP7	GTPase IMAP family member 7	6.62	0
21	TBIU016181	ZRSR2	U2 small nuclear ribonucleoprotein auxiliary factor 35 kDa subunit-related protein 2	8.55	0
22	TBIU017754	isca1	Iron-sulfur cluster assembly 1 homolog, mitochondrial	68.02	0
23	TBIU020950	PCBP2	Poly(rC)-binding protein 2	23.78	0
24	TBIU017534	ACADM	Medium-chain specific acyl-CoA dehydrogenase, mitochondrial	46.83	0
25	TBIU026028	slc38a2	Sodium-coupled neutral amino acid transporter 2	18.78	0
26	TBIU021897	RBM12B	RNA-binding protein 12B	11.05	0
27	TBIU013080	fos	Proto-oncogene c-Fos	25.34	0
28	TBIU028859	Dhx8	ATP-dependent RNA helicase DHX8	10.11	0
29	TBIU023584	Bag6	Large proline-rich protein BAG6	4.23	0
30	TBIU022019	tmem66	Store-operated calcium entry-associated regulatory factor	10.57	0
31	TBIU017325	Ube2l3	Ubiquitin-conjugating enzyme E2 L3	26.21	0
32	TBIU026697	fam73a	Protein FAM73A	11.1	0
33	TBIU025540	BAG4	BAG family molecular chaperone regulator 4	18.58	0
34	TBIU019961	LGALS9	Galectin-9	22.97	0

Table 9. Known genes of wild color-type skeletal muscle specific expressed genes
No.	GeneId	GeneName	Description	EXP:Bbrf_WT:FPKM	EXP:Bbrf_CT:FPKM
35	TBIU029397	Aloxe3	Hydroperoxide isomerase ALOXE3	15.58	0
36	TBIU026225	ascc3	Activating signal cointegrator 1 complex subunit 3	7.24	0
37	TBIU029225	brd4	Bromodomain-containing protein 4	5.67	0
38	TBIU025184	CACIS_CYPCA	Dihydropyridine-sensitive L-type skeletal muscle calcium channel subunit alpha-1	21.09	0
39	TBIU022079	Gm129	Uncharacterized protein C1orf51 homolog	23.03	0
40	TBIU020411	FAM57B	Protein FAM57B	12.89	0
41	TBIU024378	ACVR1	Activin receptor type-1	16.87	0
42	TBIU027072	SLC4A7	Sodium bicarbonate cotransporter 3	20.99	0
43	TBIU026562	FLNC	Filamin-C	13.49	0
44	TBIU029786	b2m	Beta-2-microglobulin	66.42	0

개볼락 체색 타입별 근육조직 전사체 분석에서 확보한 차등발현하는 유전자의 기 능정보를 Gene Ontology(GO)enrichment 분석으로 확인한 결과, 164개의 DEG가 1,613개의 GO ID에 해당하고(Figure 20), DEG 2개는 molecular function, 46개는 biological process, 6개는 cellular component 기능그룹에 해당하였다(Table 10, Table11, Table 12).



Figure 20. Histogram of the Gene Ontology classification

No.	ACC	Category	# Transcripts	Definition
1	GO:0050262	phosphotransferase activity, alcohol group as acceptor	5	ribosylnicotinamide kinase activity
2	GO:0030628	organic cyclic compound binding	2	pre-mRNA 3'-splice site binding

Table 10. Molecular function group of the differentially expressed genes

Table 11. Celluar component group of the differentially expressed genes

No.	ACC	Category	# Transcripts	Definition
1	GO:0030016	cytoplasmic part	10	myofibril
2	GO:0043292	cytoplasmic part	10	contractile fiber
3	GO:0016011	cell part	2	dystroglycan complex
4	GO:0016012	cell part	2	sarcoglycan complex
5	GO:0005742	mitochondrial membrane part	2	mitochondrial outer membrane translocase complex
6	GO:0030017	contractile fiber part	8	sarcomere

No.	ACC	Category	# Transcript s	Definition
1	GO:0045662	regulation of muscle tissue development	5	negative regulation of myoblast differentiation
2	GO:0009435	organonitrogen compound biosynthetic process	5	NAD biosynthetic process
3	GO:0019359	organonitrogen compound biosynthetic process	5	nicotinamide nucleotide biosynthetic process
4	GO:0019363	organonitrogen compound biosynthetic process	5	pyridine nucleotide biosynthetic process
5	GO:0072525	organonitrogen compound biosynthetic process	5	pyridine-containing compound biosynthetic process
6	GO:0019674	organic cyclic compound metabolic process	5	NAD metabolic process
7	GO:0051148	muscle structure development	5	negative regulation of muscle cell differentiation
8	GO:0045661	regulation of muscle tissue development	5	regulation of myoblast differentiation
9	GO:0006094	organic substance metabolic process	5	gluconeogenesis
10	GO:0048742	regulation of muscle tissue development	5	regulation of skeletal muscle fiber development
11	GO:0070070	protein complex subunit organization	2	proton-transporting V-type ATPase complex assembly
12	GO:0070072	protein complex subunit organization	2	vacuolar proton-transporting V-type ATPase complex assembly
13	GO:0019362	organic cyclic compound metabolic process	5	pyridine nucleotide metabolic process
14	GO:0046496	organic cyclic compound metabolic process	5	nicotinamide nucleotide metabolic process
15	GO:0019319	organic substance metabolic process	5	hexose biosynthetic process
16	GO:0072524	organic cyclic compound metabolic process	5	pyridine-containing compound metabolic process
17	GO:0046364	organic substance metabolic process	5	monosaccharide biosynthetic process
18	GO:0045445	cellular developmental process	5	myoblast differentiation
19	GO:0051146	striated muscle cell differentiation	9	striated muscle cell differentiation
20	GO:0006733	cellular metabolic process	5	oxidoreduction coenzyme metabolic process
21	GO:0010830	muscle structure development	5	regulation of myotube differentiation
22	GO:0048641	regulation of muscle tissue development	5	regulation of skeletal muscle tissue development
23	GO:0051188	cofactor metabolic process	7	cofactor biosynthetic process

Table 12. Biological process group of the differentially expressed genes

No.	ACC	Category	# Transcripts	Definition
24	GO:0048741	skeletal muscle organ development	5	skeletal muscle fiber development
25	GO:0051153	muscle structure development	5	regulation of striated muscle cell differentiation
26	GO:0014904	striated muscle cell development	5	myotube cell development
27	GO:0007519	skeletal muscle organ development	7	skeletal muscle tissue development
28	GO:0048747	striated muscle cell development	6	muscle fiber development
29	GO:0060538	skeletal muscle organ development	7	skeletal muscle organ development
30	GO:0070071	protein complex subunit organization	2	proton-transporting two-sector ATPase complex assembly
31	GO:0016202	regulation of muscle tissue development	5	regulation of striated muscle tissue development
32	GO:1901861	biological regulation	5	regulation of muscle tissue development
33	GO:0061061	anatomical structure development	12	muscle structure development
34	GO:0048634	biological regulation	5	regulation of muscle organ development
35	GO:0055002	striated muscle cell development	7	striated muscle cell development
36	GO:0016051	organic substance metabolic process	6	carbohydrate biosynthetic process
37	GO:0007517	anatomical structure development	9	muscle organ development
38	GO:0014706	anatomical structure development	9	striated muscle tissue development
39	GO:0006006	single-organism carbohydrate metabolic process	7	glucose metabolic process
40	GO:0042692	cellular developmental process	9	muscle cell differentiation
41	GO:0051147	muscle structure development	5	regulation of muscle cell differentiation
42	GO:0030150	establishment of protein localization to organelle	2	protein import into mitochondrial matrix
43	GO:0060537	anatomical structure development	9	muscle tissue development
44	GO:0055001	muscle cell development	7	muscle cell development
45	GO:0014902	striated muscle cell differentiation	5	myotube differentiation
46	GO:0051186	cellular metabolic process	7	cofactor metabolic process

# 제 4 장 연구개발 목표 달성도 및 대외기여도

## 제1절 연구개발 목표 달성도

연구목표 및 달성도 평가	달성도 (%)
<ul> <li>아유용 어류자원 탐색 및 볼락류 유전자 정보 탐색</li> <li>- 동해서식 볼락류(우럭볼락, 흰꼬리볼락, 황볼락, 노랑볼락)의 미토콘드리 염기서열을 분석하였고, 서열정보는 GenBank에 등록하였음</li> </ul>	리아 100
<ul> <li>볼락류의 체색 관련 기능유전체 정보 탐색</li> <li>개볼락 체색 타입별 근육조직 전사체 정보를 분석하였음</li> <li>개볼락 체색 타입별 차등발현 또는 특이발현 유전자의 염기서열 아미노산 서열 정보를 뱅크화하였고, 기능을 분석하였음</li> </ul>	100 및

제2절 대외 기여도

1. 주요 연구성과의 우수성

동해서식 볼락류의 유전자 정보를 분석하였고, 체색과 관련된 기능유전체를 탐색 하여 다음과 같은 결과를 도출하였다.

가. 동해서식 볼락류의 유전자 정보 등록

- 우럭볼락 미토게놈 염기서열(16,453 bp) 정보를 GenBank에 등록

-76-

- 흰꼬리볼락 미토게놈 염기서열 (16,445 bp) 정보를 GenBank에 등록
- 황볼락 미토게놈 염기서열(16,465 bp) 정보를 GenBank에 등록
- 노랑볼락 미토게놈 염기서열(16,450) 정보를 GenBank에 등록
- 나. 동해서식 개볼락의 근육조직 발현 전사체 정보 뱅크화
  - 개볼락 체색 타입별 근육조직 전사체의 unigene을 확보(101,112개)하여 35,482개의
     기능이 밝혀진 전사체와 65,630개의 기능이 밝혀지지 않은 전사체의 염기서열 및
     아미노산 서열 정보를 뱅크화하였음
- 다. 동해서식 개볼락의 체색타입별 근육조직에서 차등발현 유전자 확보
  - 개볼락 color type 근육조직에서 발현되는 53,499개 유전자의 발현량 계산
  - 개볼락 wild type 근육조직에서 발현되는 55,508개 유전자의 발현량 계산
  - 개볼락 체색 타입별 근육조직에서 차등발현하는 유전자 164개 확보
  - 개볼락 color type 근육조직에서 차등발현하는 유전자 83개와 color type 근육 조직에서만 특이적으로 발현되는 51개의 유전자 확보
  - 개볼락 wild type 근육조직에서 차등발현하는 유전자 81개와 wild type 근육 조직에서만 특이적으로 발현되는 68개의 유전자 확보
- 라. 동해서식 개볼락의 체색관련 차등발현 유전자의 기능 분석
  - Molecular function에 해당하는 2개의 차등발현 유전자 확보
  - Biological process에 해당하는 46개의 차등발현 유전자 확보
  - Celluar component에 해당하는 6개의 차등발현 유전자 확보

#### 2. 연구실적

구 분			최종	(2014)		
			목표	목표	실적	
논문 게재	SCI(E)		2	2	4	
	기타 국외학술지					
	국내 학술지					
학회	국외					
발표	국내					
	국외	출원				
트쿼		루				
= 9	국내	출원				
		루				
저서						
세미나(워크샵), 전문가활용 등						

가. SCI 논문

- 1. Complete mitochondrial genome of Sebastes hubbsi (Scorpaenidae, Scorpaeniformes) from the East Sea, Korea 학술지명 / doi : Mitochondrial DNA / 10.3109/19401736.2014.971311
- 2. Complete mitochondrial genome of Sebastes longispinis (Scorpaenidae, Scorpaeniformes) from the East Sea, Korea 학술지명 / doi : Mitochondrial DNA / 10.3109/19401736.2014.987240
- 3. Complete mitochondrial genome of Sebastes owstoni (Scorpaenidae, Scorpaeniformes) from the East Sea, Korea 학술지명 / doi : Mitochondrial DNA / 10.3109/19401736.2014.989501
- 4. Complete mitochondrial genome of Sebastes steindachneri (Scorpaenidae, Scorpaeniformes) from the East Sea, Korea 학술지명 / doi : Mitochondrial DNA / 10.3109/19401736.2014.1003834

## 제 5 장 연구개발 결과의 활용계획

동해서식 볼락류의 자원특성 조사 및 유전자 정보 탐색 분야에서 확인한 볼락류의 미토게놈 정보는 유용어류자원의 정보뱅크화 작업에 필요한 기초자료이다. 해양어류의 미토게놈 정보뱅크는 어류의 서식환경과 관련된 유전특성을 파악하고, 개체군 수준에서 서식지 적응 또는 이동에 따른 개체군 수준의 유전자 변이 파악에 이용될 수 있다.

본 연구사업 수행으로 확보한 개볼락의 체색타입별 근육조직 전사체 정보 및 차등 발현유전자는 기능유전체 정보 이용 기술 개발 연구를 위한 기초자료로 활용할 계획 이다. 또한, 어류의 체색 패턴과 유전현상을 이해하고, 기능유전체 정보를 이용한 유 전자마커 등 분자유전학적 도구 개발에도 이용할 예정이다. 향후 연구에서는 어류의 체색 결정 또는 생명현상 유지에 필요한 유전자의 기능 및 작용기작을 밝히고, 다양 한 생물의 유전체 빅데이터를 융복합한 목표지향적인 기능유전체 활용 연구와 유전체 정보의 산업화를 위한 이용기술을 개발하고자 한다.

### 제 6 장 참고문헌

- Agilent Technologies. (2002) Quantitation comparison of total RNA using the Agilent 2100 bioanalyzer, ribogreen analysis, and UV spectrometry. Agilent Application Note, Publication Number 5988-7650EN.
- Anders S. and Huber W. (2010) Differential expression analysis for sequence count data. Genome Biol. 11(10):R106.
- Blanco E., Parra G. Guigó R. (2007) Using geneid to identify genes. Curr Protoc Bioinformatics. Chapter 4:Unit 4.3.
- Colihueque N. and Araneda C. (2014) Appearance traits in fish farming: progress from classical genetics to genomics, providing insight into current and potential genetic improvement. Front Genet. 5:Article 251. doi: 10.3389/fgene.2014.00251
- Grabherr M.G., Haas B.J., Yassour M., Levin J.Z., Thompson D.A., Amit I., Adiconis X., Fan L., Raychowdhury R., Zeng Q., Chen Z., Mauceli E., Hacohen N., Gnirke A., Rhind N., di Palma F., Birren B.W., Nusbaum C., Lindblad-Toh K., Friedman N., Regev A. (2011) Nat Biotechnol. 29(7):644–652.
- Greenwood A.K., Cech J.N., Peichel C.L. (2012) Molecular and developmental contributions to divergent pigment patterns in marine and freshwater sticklebacks. 14(4):351–362.
- Greenwood A.K., Jones F.C., Chan Y.F., Brady S.D., Absher D.M., Grimwood J., Schmutz J., Myers R.M., Kingsley D.M., Peichel C.L. (2011) The genetic basis of divergent pigment patterns in juvenile threespine sticklebacks. 107(2):155–166
- Haas B.J., Papanicolaou A., Yassour M., Grabherr M., Blood P.D., Bowden J., Couger M.B., Eccles D., Li B., Lieber M., Macmanes M.D., Ott M., Orvis J., Pochet N., Strozzi F., Weeks N., Westerman R., William T., Dewey C.N., Henschel R., Leduc R.D., Friedman N., Regev A. (2013) De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc. 8(8):1494–1512.

- Harms L., Frickenhaus S., Schiffer M., Mark F.C., Storch D., Held C., Pörtner H.O., Lucassen M. Gene expression profiling in gills of the great spider crab Hyas araneus in response to ocean acidification and warming. (2014) BMC Genomics. 12;15:789. doi: 10.1186/1471-2164-15-789.
- Henning F., Jones J.C., Franchini P., Meyer A. (2013) Transcriptomics of morphological color change in polychromatic Midas cichlids. BMC Genomics. 13;14:171. doi: 10.1186/1471-2164-14-171.
- Huang X. and Madan A. (1999) CAP3: A DNA sequence assembly program. Genome Res. 9, 868–877.
- Jiang Y., Zhang S., Xu J., Feng J., Mahboob S., Al-Ghanim K.A., Sun X., Xu P. (2014) Comparative transcriptome analysis reveals the genetic basis of skin color variation in common carp. PLoS One. 25;9(9):e108200. doi: 10.1371/journal.pone.0108200.
- Kadota K., Nishiyama T., Shimizu K. (2012) A normalization strategy for comparing tag count data. Algorithms Mol Biol. 7(1):5. doi: 10.1186/1748-7188-7-5.
- Kim H.A., Lim C.J., Kim S., Choe J.K., Jo S.H., Baek N., Kwon S.Y. (2014) High-throughput sequencing and de novo assembly of Brassica oleracea var. Capitata L. for transcriptome analysis. PLoS One. 28;9(3):e92087. doi: 10.1371/journal.pone.0092087.
- Kuraku S, Meyer A. (2008) Genomic analysis of cichlid fish 'natural mutants'. Curr Opin Genet Dev. 18(6):551–558
- Li B. and Dewey C.N. (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics, 4;12:323. doi: 10.1186/1471-2105-12-323.
- Li H. and Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform, Bioinformatics. 25(14):1754–1760.
- Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G., Durbin R; 1000 Genome Project Data Processing Subgroup. (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics.

25(16):2078-2079.

- Lister J.A., Robertson C.P., Lepage T., Johnson S.L., Raible D.W. (1999) nacre encodes a zebrafish microphthalmia-related protein that regulates neural-crest-derived pigment cell fate. Development. 126(17):3757-3767.
- Manfred S., Susanne K., Brigitta W., Toni W., Christiaan V. H., Herman P. S., Svenja M. (2012) Conserved Expression Signatures between Medaka and Human Pigment Cell Tumors. 7(5):e37880. doi: 10.1371/journal.pone.0037880
- Martin J.A. and Wang Z. (2011) Next-generation transcriptome assembly. Nat Rev Genet. 12(10):671-682.
- Metz H.C., Manceau M., Hoekstra H.E. (2011) Turing patterns: how the fish got its spots. Pigment Cell Melanoma Res. 24(1):12-14.
- Parichy D.M., Mellgren E.M., Rawls J.F., Lopes S.S., Kelsh R.N., Johnson S.L. (2000) Mutational analysis of endothelin receptor b1 (rose) during neural crest and pigment pattern development in the zebrafish Danio rerio. Dev Biol. 15;227(2):294–306.
- Parichy D.M., Rawls J.F., Pratt S.J., Whitfield T.T., Johnson S.L. (1999) Zebrafish sparse corresponds to an orthologue of c-kit and is required for the morphogenesis of a subpopulation of melanocytes, but is not essential for hematopoiesis or primordial germ cell development. Development. 126(15):3425-3346.
- Pertea G., Huang X., Liang F., Antonescu V., Sultana R., Karamycheva S., Lee Y., White J., Cheung F., Parvizi B., Tsai J., Quackenbush J. (2003) TIGR Gene Indices clustering tools (TGICL): a software system for fast clustering of large EST datasets. Bioinformatics, 19(5):651–652.
- Qian X., Ba Y., Zhuang Q., Zhong G. (2014) RNA-Seq technology and its application in fish transcriptomics. OMICS. 18(2):98-110
- Santos M.E., Braasch I., Boileau N., Meyer B.S., Sauteur L., Böhne A., Belting H.G., Affolter M., Salzburger W. (2014) The evolution of cichlid fish egg-spots is linked with a cis-regulatory change. Nat Commun. 9(5):5149. doi:

10.1038/ncomms6149.

- Sun J., Nishiyama T., Shimizu K., Kadota K. (2013) TCC: an R package for comparing tag count data with robust normalization strategies, BMC Bioinformatics, 14:219. doi: 10.1186/1471-2105-14-219.
- Wang C., Wachholtz M., Wang J., Liao X., Lu G. (2014) Analysis of the skin transcriptome in two oujiang color varieties of common carp. PLoS One. 6;9(3):e90074. doi: 10.1371/journal.pone.0090074.
- Watanabe M., Iwashita M., Ishii M., Kurachi Y., Kawakami A., Kondo S., Okada N. (2006) Spot pattern of leopard Danio is caused by mutation in the zebrafish connexin41.8 gene. EMBO Rep. 7(9):893–897.
- Wolf J.B. Principles of transcriptome analysis and gene expression quantification: an RNA-seq tutorial. (2013) Mol Ecol Resour. 13(4):559-572
- Yan B., Liu B., Zhu C.D., Li K.L., Yue L.J., Zhao J.L., Gong X.L., Wang C.H. (2013) microRNA regulation of skin pigmentation in fish. J Cell Sci. 1;126(Pt 15):3401–3408.
- Yang Y. and Smith S.A. (2013) Optimizing de novo assembly of short-read RNA-seq data for phylogenomics. BMC Genomics. 14:328.

주 의 1. 이 보고서는 한국해양과학기술원에서 수행한 주요사업의 연구결과보 고서입니다. 2. 이 보고서 내용을 발표할 때에는 반드시 한국해양과학기술원에서 수 행한 주요사업의 연구결과임을 밝혀야 합니다. 3. 국가과학기술 기밀유지에 필요한 내용은 대외적으로 발표 또는 공개하여서는 안됩니다.