

THESIS

Course code: BI309F

Candidate number:

Reproductive genes of *Fucus distichus* and
Fucus serratus.

Date: 15.02.2016

Total number of pages: 89

Content

Abstract	ii
1. Introduction	1
2. Materials and methods.....	9
2.1 Sampling	9
2.2 Laboratory work.....	9
2.3 Processing of data.....	12
3. Results	16
3.1 Laboratory work.....	16
3.2 Results of data processing	16
4. Discussion.....	32
1. Genes under positive selection in transcriptomes of <i>F. serratus</i> male and female	32
2. Reproductive genes of <i>F. serratus</i> male transcriptome.	33
2.1 carbohydrate metabolism.....	33
2.2 Genes of detoxification.....	33
2.3. Flagella-associated genes	33
3. Reproductive genes of <i>F. serratus</i> female transcriptome.	34
3.1 Cell-wall genes	34
3.2. Genes of vesicle transport.....	34
3.3 Signaling molecules.....	35
3.4 Male-biased genes (according to the literature) found in transcriptome of <i>F. serratus</i> female....	35
4 Unknown proteins within both transcriptomes of <i>F. serratus</i> male and female	36
5 Reproductive genes of <i>F. distichus</i>	36
5. Conclusion.....	38
6. References	39
7. Acknowledgments	41
8. Supplementary materials	42

Abstract

Speciation is the process of appearance of new species. One important aspect bound with speciation in reproductive isolation (inability to mate or to give viable and fertile progeny). Genomic studies suggest that reproductive isolation results from small number of genes under selection, called "speciation genes". The genus *Fucus* is an ideal to study mating system, reproductive isolation and speciation genes, as it consist of two lineages, both with a hermaphroditic and a dioecious species, hybridizing in some part of their range. Furthermore the genus *Fucus* has been the subject of numerous research on ecology, phylogeography, phylogenetics and population genetics. The main aim of presented thesis to identify putative speciation genes that may be involved in the reproduction isolation between two species of the genus *Fucus*: *F. distichus* and *F. serratus* from zone of sympatry for 10000 years.

We compared the transcriptomes of *F. serratus* (male, female and vegetative tissues) and *F. distichus* (reproductive and vegetative). In total, we identified 750 genes linked to reproduction processes for *F. serratus* male, 3213 for *F. serratus* female and their homologues in *F. distichus*. Among these genes 434 genes between *F. serratus* female vs. *F. distichus*, and 14 genes between *F. serratus* male vs. *F. distichus* had a dN/dS ratio >1, suggesting divergent selection. These genes are thus candidate genes for speciation but their role in the hybrid zones will need to be confirmed by future studies

F. serratus female transcriptome was characterized by processes of cell wall recovery and maintenance. We also observed high number of vesicle transport genes that are essentials for gamete formation. But the most important findings in *F. serratus* female transcriptome were genes of signaling molecules providing egg-sperm reaction of fusion and fertilization (Rab6). A large part of *F. serratus* male transcriptome were flagella-associated genes that refers to important function of flagella movement involved in gamete fusion and fertilization. Also male genes provides biosynthesis of polysaccharides and cell detoxification. *F. distichus* transcriptome was characterized by a high number of flagella-associated and ribosomal genes. But we have found also gene 14-3-3 protein from a family of regulation proteins involved in transport of K^+ , Ca^{2+} and Mg^{2+} that might impact on the reproduction by regulation of sperm movements.

1. Introduction

Speciation is a central topic of evolutionary science (Butlin *et al.* 2011). The balance between speciation and extinction determines species diversity (Butlin *et al.* 2011). The most common definition for speciation is the process of appearance of new species (Martin *et al.* 2008). Four different types of speciation can be characterized by the presence of a geographic barrier and the exchange of genes (Futuyma 2013). There are allo-, peri-, para- and sympatric speciation (Figure 1) (Ridley 2004). Allopatric speciation is characterized by the occurrence of a geographical barrier, which prevents gene exchange between populations (Futuyma 2013). Peri- and parapatric speciation are both characterized by appearance of a new niche within a single population (Ridley 2004). While peripatric speciation describes the isolation of a single small colony in an isolated niche, parapatric speciation describes reproductive isolation of one population in two adjacent niches. (Futuyma 2013). During sympatric speciation, reproductive barriers are evolving within habitat of its ancestor (Futuyma 2013, Ridley 2004).

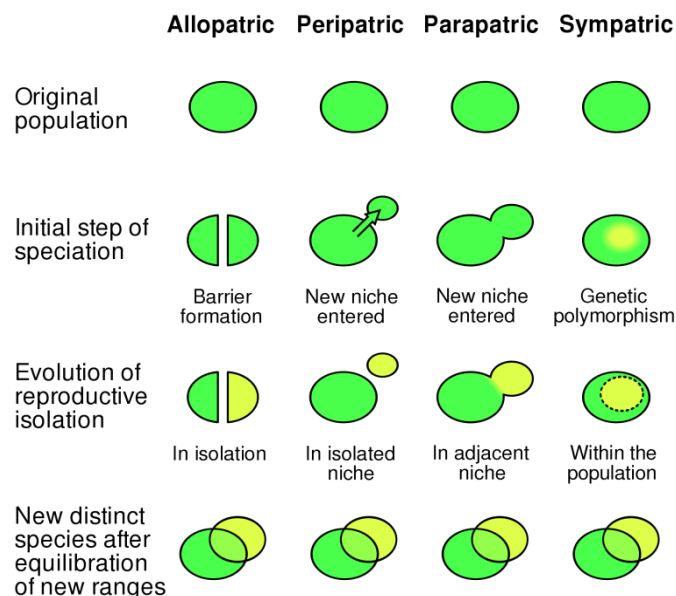


Figure 1 Types of speciation

http://upload.wikimedia.org/wikipedia/commons/thumb/5/53/Speciation_modes.svg/2000px-Speciation_modes.svg.png

One important aspect bound with speciation is reproductive isolation (inability to mate or give viable and fertile progeny) (Butlin *et al.* 2011). Reproductive isolation can occur

before (prezygotic) or after (postzygotic) hybrid formation (Butlin *et al.* 2011). Prezygotic isolation can result from gamete incompatibility, mechanism of protection from polyspermy or unsuccessful gamete recognition and sympatry describes absence of geographical barriers (Ridley). Postzygotic isolation is characterized by unfit hybrids (Coyne *et al.* 1998). Presumably due to incompatibility alleles brought from different species in one organism they are also genes called «complementary genes» (Coyne *et al.* 1998). These genes can lead to decreasing hybrid viability even if they have only weak effect on hybrid fitness and are only slightly incompatible between organisms (Coyne *et al.* 1998).

The origin of species is now better understood than it was in 1859 when Darwin's book was published (Darwin 1859). Darwin's idea that speciation is driven by natural selection was accepted in a broad sense (Coyne *et al.* 2004). Nowadays we know that drivers of speciation are both natural and sexual selection (Coyne *et al.* 1998), but also sexual selection by itself (Lande 1981; 1982; Anderson *et al.* 2006) and also sexual conflict (Parker *et al.* 1998; Gavrilets 2000). In some cases, complete reproductive isolation evolved because of changes in female pheromones, genes of male attraction and female perception, or in other terms, male traits and female preferences (Coyne *et al.* 1998). For example guppy *Poecilia reticulata* local population shows female preference to mate with local males more than with foreign (Endler *et al.* 1995). Recent investigations support the idea that sexual and natural selection in combination may be able to initiate and complete speciation (Butlin *et al.* 2011).

Some studies indicate that sexual selection might not have enough power to create a complete reproductive barrier between two populations (Maan *et al.* 2011), but sexual selection is not confined to sexual and morphological features acting before fertilization but could include sperm selection by females and sperm competition between males within an impregnated female (Coyne *et al.* 1998). It presumably provides a selective pressure that induces the rapid evolution of reproductive proteins (Swanson *et al.* 2002). Sperm competition could be intense due to sperm density that is released by males, for example in the case of sea urchins there are billions of sperm cells released during spawning period by one individual (Swanson *et al.* 2002).

There is a mechanism, which contributes to speciation: reinforcement. It is the process by which selection against hybrids increases reproductive isolation between emerging species (Howard 1993; Servedio *et al.* 2003). The role of reinforcement in speciation has been underestimated (Servedio *et al.*; 2003), because increased reproductive isolation, the main predicted pattern of the process could occur due to for example gene flow or because of the process of ecological character displacement (Butlin *et al.* 2011). Several models of reinforcement have shown that mating preferences diverge mainly due to natural selection (Servedio *et al.* 1997; Cain *et al.* 1999; Kirkpatrick *et al.* 1999; Servedio 2000) and only in some cases due to sexual selection (Kirkpatrick *et al.* 1999; Kirkpatrick *et al.* 2002). Some studies revealed that unfit or unviable hybrids may be caused by malevolent interactions between genes on sex chromosomes and genes on autosomes (Pantazidis *et al.* 1993; Turelli *et al.* 1995; Lamnissou *et al.* 1996; Noor *et al.* 2001; Orr *et al.* 2001).

Complete reproductive isolation between two species results from several factors (Coyne *et al.* 1998). Genomic studies suggest that reproductive isolation can result from a small number of genes under selection (Turner *et al.* 2005; Hohenlohe *et al.* 2010), called "speciation genes" (Nosil *et al.* 2011). Speciation genes are difficult to recognize, especially if they take part in the initial phase of reproductive isolation (Butlin *et al.* 2011). The numbers of speciation genes was suggested to control the strength of pre-mating isolation (Liou *et al.* 1994), but the position of contributing genes on the chromosomes may play a more important role for the evolution of reproductive isolation (Servedio *et al.* 2003). Speciation genes on sex chromosome increase the probability of reinforcement as compared with speciation genes on autosomes (Kelly *et al.* 1996) because sex-linked genes initiate pre-mating isolation and subsequently reinforcement due to the immediate expression of recessive alleles in the heterogametic sex (Servedio *et al.* 2003).

Speciation genes have been well-studied in sea urchins where the egg carbohydrate outer coat (known also as jelly coat) activates sperm, insufficient activation of the sperm during interspecific fertilization can lead to failure of gamete fusion (Biermann 1998). More common is failure of sperm proteins (bindin) of one species to attach to the egg surface receptor of another species (Metz *et al.* 1996). It varies within *Echinometra* sp. and

Strongylocentrotus sp. (Biermann 1998). Between *Echinometra* and *Strongylocentrotus* increased dN/dS ratio suggested positive selection in bindin regions that may be involved in adhesion of sperm to the egg (Swanson *et al.* 2002).

An entirely different system of gamete recognition was found in the snail genera *Haliotis* (abalone) and *Tegula* (turbin snail). Here: the sperm protein lysin when released from the acrosome interacts with the protein VERL (Vitelline Envelop Receptor for Lysin) on the egg surface forms a pore in the outer coat of the egg, allowing sperm penetration (Swanson *et al.* 1998; Swanson *et al.* 2002). Different forms of VERL were found in related species of abalone, which could be a reason of lysine divergence (Swanson *et al.* 2002). Lysines vary between species, indicating adaptive evolution to the environment, which evolved presumably to reinforce reproductive isolation (Metz *et al.* 1998; Hellberg *et al.* 1999; Yang *et al.* 2000b). Protein sp 18, a mediator in the sexual process that is released by abalone's sperm can evolve very rapidly between species (Metz *et al.* 1998). Rapid evolution was demonstrated by a comparing 20 times faster diverging rate of exons as compared with introns (Metz *et al.* 1998).

The genus *Fucus* is an ideal system to study mating system, reproductive isolation and speciation genes, as it consist of two lineages, both with a hermaphroditic and a dioecious species, hybridizing in some part of their range. Furthermore the genus *Fucus* has been the subject of numerous research on ecology (Sideman *et al.* 1983; Arrontes 1993; Pearson *et al.* 1996; Mathieson *et al.* 2006), phylogeography (Canovas *et al.* 2011; Coyer *et al.* 2011), phylogenetics (Alice *et al.* 1999; Coyer *et al.* 2006) and population genetics (Coyer *et al.* 2002; Hoarau *et al.* 2007). External fertilization in *Fucus* sp. involves chemical attractants (pheromones) and free-swimming gametes (Mohr *et al.* 2012). The mating system is easier to analyze in comparison with for example *Drosophila* sp., because of the lack of sexual behaviors such as courtship dance (Lessios 2007).

The genus *Fucus* originated in the North Pacific from where it spread into the North Atlantic after the opening of the Bering Strait 5,54 Myr ago, and diverged into two lineages: the first one includes *F. distichus* and *F. serratus* and the second one –*F. spiralis* and *F.*

vesiculosus (Figure 2) (Serrao *et al.* 1999; Gladenkova *et al.* 2002; Coyer *et al.* 2006; Coyer *et al.* 2006; Hoarau *et al.* 2007; Canovas *et al.* 2011; Coyer *et al.* 2011). Each lineage includes one dioecious (*F. serratus*, *F. vesiculosus*) and one hermaphroditic species (*F. distichus*, *F. spiralis*) (Martins *et al.* 2013; Hoarau *et al.* 2015).

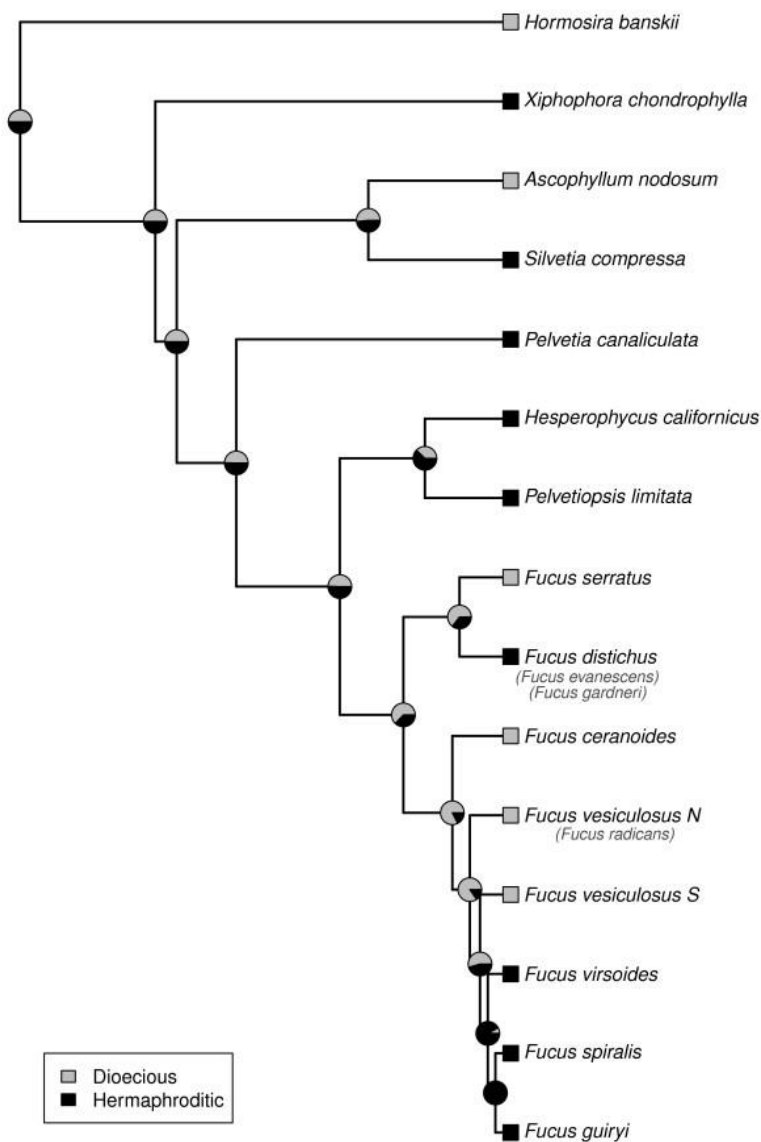


Figure 2 *Fucus* sp. phylogeny (Canovas *et al.* 2011).

Recent molecular work has shown that hybridization is prevalent among sister taxa and therefore, always involves one hermaphroditic and one dioecious parental species e.g. *F. distichus* and *F. serratus* (Coyer *et al.* 2002, Coyer *et al.* 2006), *F. spiralis* and *F. vesiculosus* (Wallace *et al.* 2004; Billard *et al.* 2005; Engel *et al.* 2005; Coyer *et al.* 2006; Mathieson *et al.* 2006).

Our investigation continues and expands research on hybridization with different time of coexistence (sympatry) between *F. serratus* and *F. distichus* (Figure 3) in four hybrid zones with different times of coexistence (sympatry) (Hoarau *et al.* 2015):

- 1 and 2: Tromsø and Kirkenes (Northern Norway), recolonized by *F. serratus* after the Last Glacial Maximum (18000 – 20000 years ago (Coyer *et al.* 2006) from Ireland (10000 years of sympatry) (Hoarau *et al.* 2007).
- 3: Iceland (Heimaey), where *F. serratus* was presumably introduced from the central part of Norway (Oslofjord) 100 years of sympatry (Coyer *et al.* 2006).
- 4: The Kattegat sea (Blushø) where *F. distichus* was introduced to an endemic *F. serratus* population (100 years of sympatry) (Schueller *et al.* 1994; Wikström *et al.* 2002).



Figure 3 Locations of presence for *F. serratus* (light grey line) and *F. distichus* (black line and dark grey in zone of sympatry). T – Tromsø, K – Kirkenes (Northern Norway), H – Heimaey (Iceland) and B - Blushø (Denmark) (Hoarau *et al.* 2015).

Reinforcement is more likely to occur in the older zones of sympatry (10 000 years), because they are characterized by fewer hybrids (2.1-3.1 %) in comparison with younger

zones (12.9-23.9 %) (Table 1) (Hoarau *et al.* 2015). Furthermore, F1 hybrids (first generation hybrids) were not found in older zones while F1 hybrids represented 3.7-8.3 % of individuals in the younger zones (Hoarau *et al.* 2015).

In-vitro crosses between the two species were significantly less successful for gametes collected from Tromsø (0-1.1 %, Kirkenes was not observed) in comparison with gametes from Iceland and Denmark (9.6- 43.1 %) (Table 1) (Hoarau *et al.* 2015).

Table 1 Characteristics for old and young contact zones between *F. serratus* and *F. distichus*.

Location	Number of years of sympatry	Number of hybrids (%)	Interspecific fertilization success (%)
Kirkenes (Norway)	~10 000	2.1	-
Tromsø (Norway)	~10 000	3.1	0-1.1
Heimaey (Iceland)	~100	23.9	23.7-38.3
Blushøy (Denmark)	~100	12.9	9.6-43.1

Individuals from the two older sympatry zones carry approximately 3% of alleles from other species (Hoarau *et al.* 2015). A single individual in Tromsø carried the chloroplast of *F. distichus* and the nuclear material from *F. serratus* indicating that hybridization occurred between the two species but decreased with increasing time of coexistence (Hoarau *et al.* 2015).

The low number of hybrids and unsuccessful in-vitro crosses between the two species in the older zones of sympatry was suggested to result from increasing gamete

incompatibility with the time of contact (Hoarau *et al.* 2015). Incompatibility is likely based on reduced gamete attraction and recognition, favoring prezygotic mechanism of isolation and presumably reinforcement (Hoarau *et al.* 2015).

The main aim of the present thesis was to identify putative speciation genes that may be involved in the reproductive isolation between two species of the genus *Fucus*: *F. distichus* and *F. serratus*. We aimed to: 1) identify reproductive genes in the transcriptome of *F. distichus* and *F. serratus*, by comparing the transcriptomes derived from vegetative and non-vegetative tissues, 2) to identify only those reproductive genes that are potentially under diverging selection.

2. Materials and methods

2.1 Sampling

Samples were collected in Tromsø (10,000 years old sympatry zone) along the shoreline (69°37'56.6"N 18°57'22.9 "E) (Figure 4) during low tide (<http://vannstand.no/index.php/nb/english-section/tide-table>) on the 17th of June, a time of maximum spawning regulated by the lunar cycle (Pearson *et al.* 1998; Pearson *et al.* 2006b).

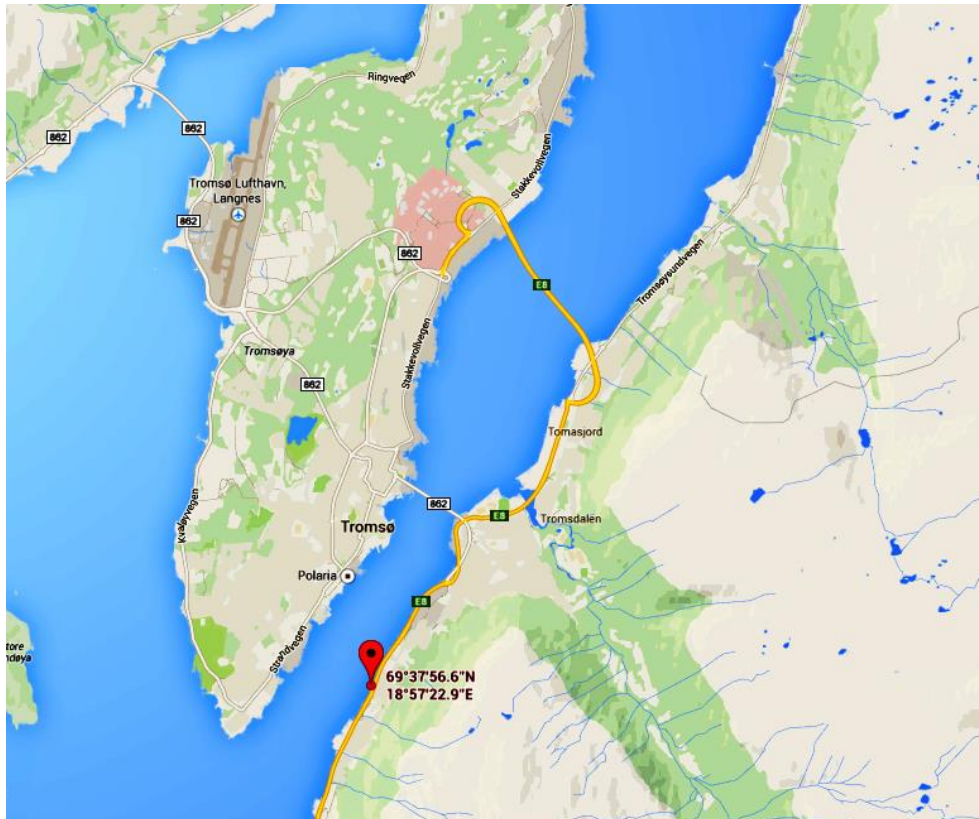


Figure 4 Sampling site. Location of sampling marked in red.

<https://www.google.com/maps/place/69%C2%B037'56.6%22N+18%C2%B057'22.9%22E/@69.6595664,18.9998712,12z/data=!4m2!3m1!1s0x0:0x0>

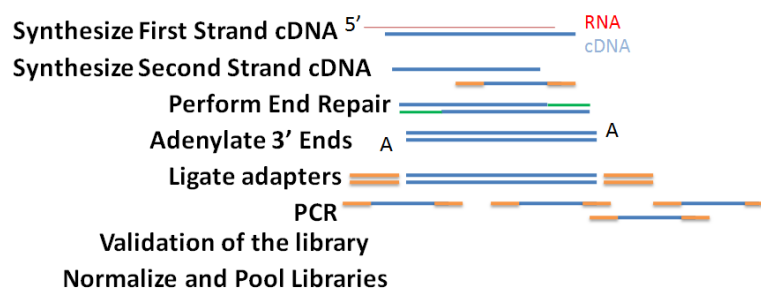
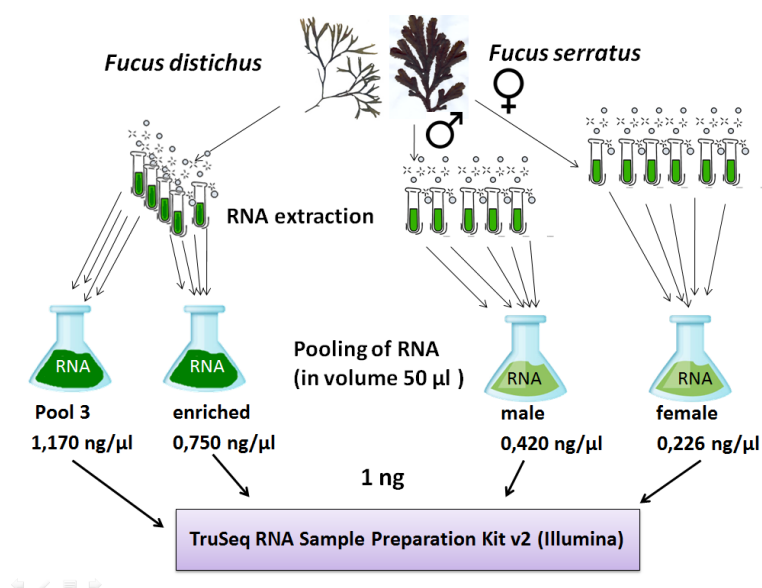
Reproductive tissues were collected from 15 individuals of the monoecious *Fucus distichus* and 40 individuals of the dioecious *Fucus serratus*. A higher numbers of samples *F. serratus* individuals were collected as it is difficult to determine the sex in the field and we targeted a minimum number of 10 males and 10 females.

2.2 Laboratory work

The algal tissues placed into cooler box and transported to the laboratory of the University of Nordland (Nord University since 01.01.2016), where the sex of each sample

was determined using a light microscope. Samples were frozen in liquid nitrogen and then freeze-dried using BenchTop «K» Series Freeze Dryer at -77.7°C and 140 mBar for 3 days. Samples were subsequently stored at -80°C .

RNA was extracted using a protocol developed specifically for *Fucus* sp. (Figure 5) (Pearson *et al.* 2006a). RNA extraction from seaweed can be problematic due to polysaccharides and polyphenols (Cheong-Xin *et al.* 2004). Poly T magnetic beads (Ambion) were used to isolate mRNA. For *F. distichus*, an enriched pool of mRNA was generated using another round of Poly A selection (Figure 1S, 2S).



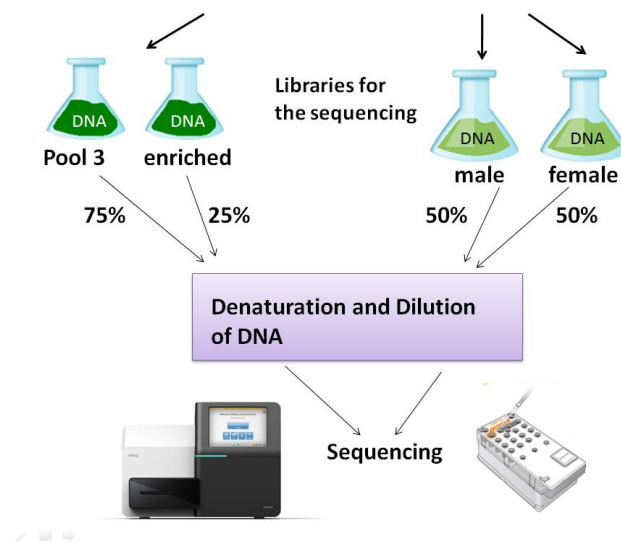


Figure 5 Processing of algal tissue. 1. The figure shows steps of RNA extraction, pooling extracted RNA for *Fucus distichus* pool 3 and enriched one and *Fucus serratus* male and female, and loading the pools to TruSeq RNA Sample Preparation Kit v2. 2. Steps of reverse transcription of RNA to DNA libraries. 3. Denaturation and dilution of DNA using Preparing Libraries for Sequencing on the MiSeq® and loading to the sequencing machine.

RNA quality was checked using agarose gel electrophoresis (Sambrook *et al.* 2001) and quantities were measured using Qubit 2.0 fluorometer (Life technologies) with the Qubit® RNA HS Assay Kit. Quality control of the RNA was done using the Agilent 2200 TapeStation System (Agilent Technologies), using the Agilent RNA 6000 Nano Kit.

The OneStep-96™ PCR Inhibitor Removal Kit (Zymo research) was used to remove remaining tannin and polysaccharides, PerfeCta® DNase I (Quanta biosciences) was used to remove DNA. The RNA was concentrated to 50 µl ZR-96 RNA Clean & Concentrator™ (Zymo research). We created 4 pools of 10 individuals: 1) *F. distichus* enriched, 2) *F. distichus* pool 3, 3) *F. serratus* male, and 4) *F. serratus* female (Figure 5).

The cleaned RNA (1 ng in 50 µl of each pool) was loaded to the TruSeq RNA Sample Preparation Kit2 and reversely transcribed to the first and second cDNA strands (Figure 5). The four pools were barcoded with four illumina TruSeq adapters (Table 2).

ATCT3

TruSeq Indexed Adapter (for the reverse strand) (Istvan 2014).

5 GATCGGAAGAGCACACGTCTGAACTCCAGTCAC-NNN(place for special barcode adapter)NNN-ATCTCGTATGCCGTCTTCTGCTTG 3.

Table 2 List of 6 nucleotide adapters for each library

Library	Number of adapter	Nucleotides
<i>Fucus serratus</i> female	Adapter 12	CTTGTA
<i>Fucus serratus</i> male	Adapter 6	CCAATG
<i>Fucus distichus</i> Pool 3	Adapter 5	ACAGTG
<i>Fucus distichus</i> enriched	Adapter 19	GTGAAA

After qPCR quantification and normalization according to analyzing profile comparison of every pool (Figure 2S) we decided to load 75% of *F. distichus* pool 3 and 25% of enriched one (Figure 5), because *F. distichus* enriched consist on shorter (Figure 2S). As sample concentration of *F. serratus* male and female pools were almost identical (Figure 2S), 50% of each pool was loaded in a single chip for the sequencing.

cDNA was denatured and diluted according to the MiSeq® protocol as a final preparation for sequencing (Figure 5). Pools of *F. distichus*, *F. serratus* female and male, with concentrations within the range of 0,1-1 ng, were loaded to the sequencing machine (Figure 5). Sequencing has been performed using the MiSeq (Illumina), using TruSeq RNA Sample Preparation Kit v2 which provides up to 2 * 250 bp or 500 cycles of sequencing (Figure 5).

2.3 Processing of data

Fucus serratus transcriptome, derived from non-reproductive tissues, using IonTorrent (life technologies) was used as negative control. The data were available from an earlier project (Jueterbock unpublished).

Data from IonTorrent and Illumina did not need to be demultiplexed, because platforms do it automatically using barcodes on adapters (Deagle *et al.* 2013; Wong *et al.* 2013). To

analyze the data we used the steps, showed on figure 6. After transcriptome assembly we merged the two *F. distichus* pools. We used De Novo Assembly because no reference genome is yet available for *Fucus* sp.

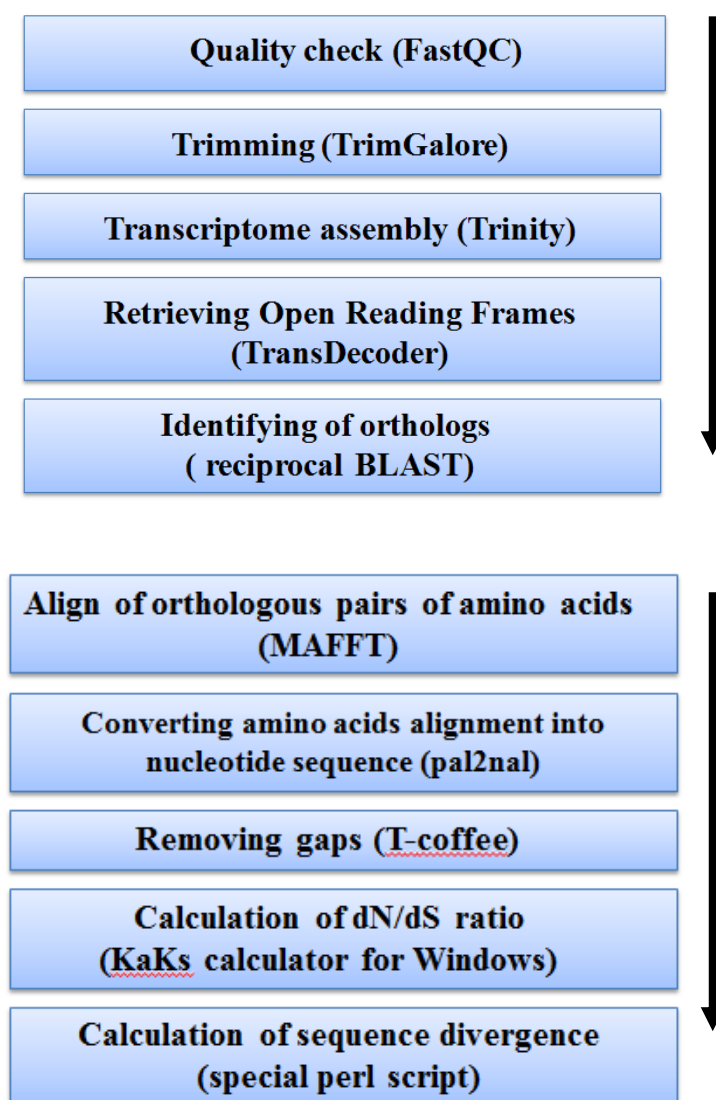


Figure 6 Raw data

References for programs: FastQC (Andrews 2010), TrimGalore! (Krueger 2012) (Table 1S), Trinity (Grabherr *et al.* 2011), TransDecoder (<http://transdecoder.github.io/>), Proteinortho (Lechner *et al.* 2011), MAFFT (Kato *et al.* 2013), pal2nal.pl (Suyama *et al.* 2006), T-coffee (Notredame *et al.* 2000), PAML (Yang 1997), KaKs calculator (Zhang *et al.* 2006a).

Reciprocal BLAST searches were performed according to Table 3. First we compared non-reproductive *F. serratus* IonTorrent and reproductive *F. serratus* male and female transcriptomes (green1.1, 1.2 cells on the Table 3) to filter out genes that were expressed in

non-reproductive tissue and which likely are not involved in reproduction. For the same purpose we compared somatic *F. serratus* IonTorrent transcriptome with reproductive *F. distichus* (blue 2.1 cell on the Table 3). We compared *F. distichus* and *F. serratus* male and female (orange 3.1, 3.2 cells on the Table 3) transcriptomes to identify differences in reproductive genes between a monoecious and dioecious species.

Table 3 Steps of reciprocal BLAST. Green cells: comparison of somatic *F. serratus* IonTorrent vs. reproductive *F. serratus* male and female to exclude non-reproductive genes. Blue cell: comparison of *F. serratus* IonTorrent vs. *F. distichus* to exclude non-reproductive genes and compare monoecious and dioecious species. Orange cells: comparison of reproductive transcriptomes of *F. distichus* vs. *F. serratus* male and female to perform BLAST and compare sex-biased genes of the two sexes in within one species between two sexes of *F. serratus*. *F. d* – *Fucus distichus*, *F. s* – *Fucus serratus* somatic (IonTorrent), *F. s. f*, *F. s. m* – *F. serratus* male and female

	<i>F. d</i>	<i>F. s s</i>	<i>F. s f</i>	<i>F. s m</i>
<i>F. d</i>			3.1	3.2
<i>F. s s</i>	2.1		1.1	1.2

All of the procedures (sequencing RNA, De Novo assembly, receiving ORF, preparing alignments to dN/dS ratio calculation (Figure 6)) were performed to prepare raw data to comparing transcriptomes with each other and to calculating dN/dS ratio. We calculated dN/dS ratio as the ratio of nonsynonymous to synonymous nucleotide substitutions (Yang *et al.* 2000a). Synonymous (silent) substitutions are changes that does not change the protein sequence, whereas nonsynonymous substitutions result in amino acid changes of the protein (Kimchi-Sarfaty *et al.* 2007). It was supposed that if the dN/dS ratio of a gene is higher than one, this presumably was under natural selection and if it is below than one it means purifying (stabilizing) selection (Hurst 2002).

The sequence divergence has been calculated using a perl script (Figure 6) (Johnson *et al.* 2013). The sequence divergence was defined as the number of different nucleotides divided by the length of the sequence which shows genetic variance of each orthologs.

The dN/dS ratio was calculated using KaKs calculator for windows (Zhang *et al.* 2006b), which estimates the dN/dS ratio based on 9 methods: NG (Nei *et al.* 1986), LWL (Li *et al.* 1985), MLWL(Modified LWL), LPB (Li 1993), MLPB (Shiu *et al.* 2004), YN (Yang *et al.* 1999), MYN (Zhang *et al.* 2006a), GY (Goldman *et al.* 1994), MS and MA (Posada 2003) (Figure 6). We chose 9 methods because the results can vary between them. We defined a dN/dS ratio significantly high only if at least 1 method revealed p-value higher than 0,05.

3. Results

3.1 Laboratory work

Extracted RNA concentration was lower for *F. distichus* (7,19 ng/μl in average for one individual) than *F. serratus* (17,13 ng/μl in average for one individual). As a material for the TruSeq RNA Sample preparation Kit (Illumina) we used pools of *F. distichus* pool 3, enriched and *F. serratus* male, female (Figure 2S).

For reverse transcription of RNA to cDNA, we used two different pools of *F. distichus* – pool 3 (with concentration of peak 1,170 ng/μl) (Figure 3S) and enriched one (0,750 ng/μl) (Figure 4S), for the libraries of *F. serratus* we used one pool of *F. serratus* male (with concentration of peak 420 pg/μl) (Figure 5S) and one of female (226 pg/μl) (Figure 6S, Figure 5).

The library pools of *F. distichus* enriched were characterized by sequences of 400 bp on average and a lower concentration than pool 3 and *F. distichus* pool 3 has 300 bp on average (Figure 3S, 4S). Lower concentration of *F. distichus* enriched explains only 25 % of this pool loaded in cheap for sequencing, the rest 75 % were complemented by *F. distichus* pool 3.

The library pools of *F. serratus* male and female were characterized by fragments of ~ 350 bp and concentrations of 0,420 ng/μl and 0,226 ng/μl respectively (Figure 5S, 6S).

Sequencing of the *F. distichus* transcriptome was successful (Figure 7S). According to figure 8S, sequencing of *F. serratus* work well just with for the forward sequence, but after 301 cycle intensity of sequencing was decreasing.

3.2 Results of data processing

Transcriptome of *F. distichus* had a higher number of one strand reads after sequencing – 5853302 (Table 4) as compared with the *F. serratus* male (140000) and female transcriptome (119948) due to failure of sequencing the reverse strands.

Table 4 Results of data processing. *F. d* – *Fucus distichus*

Transcriptome		<i>F. serratus</i> male		<i>F. serratus</i> female		<i>F. d</i> pool 3		<i>F. d</i> enriched		<i>F. serratus</i> IonTorrent	
Number of reads after sequencing (one strand)		140 000		119 948		810 531		5 853 302		2 864 279	
Number of reads after trimming (one strand)		134 158		118 380		391 223		808 564		2 610 349	
Parameters after De novo assembly	Total number of contigs	12 075		10 795		67 221		1 227			
	Minimum length	200		200		200		200			
	Maximum length	10 795		10 795		10 795		1 678			
	Average length	327.3		437.3		420.6		295			
Number of Open Reading Frames	Filetype*	.pep	.mRNA	.pep	.mRNA	.pep	.mRNA	.pep	.mRNA		
	Total	1 823		7 546		44 048		768			
	Minimum length	49	200	49	200	49	200	49	200		
	Maximum length	1 062	1 0795	1 062	10 795	1 062	10 795	231	1 678		
	Average length	104.3	419.3	102.4	429.2	108.5	484.6	75	328		
Number of orthologs vs. <i>F. d</i> transcriptome		2 936		8 234		-		-			
Number of orthologs vs. <i>F. s</i> IonTorrent		12		51		67					
Number of reproductive genes		750		3 213		40					
Candidate genes with dN/dS>1		14		434		-		-			

After cutting adaptors and barcodes the quality of the libraries was checked with the program FastQC (Table 4) (Andrews 2010). Using TrimGalore! were removed reads with length less than 40 (Table 1S) and usually using that program sequences lose 0.1-0.2 % of reads (Krueger 2012). Transcriptome of *F. distichus* enriched get rid of 86,18% of low

quality reads, *F. distichus* pool 3 – 51,73%, *F. serratus* IonTorrent – 8,86%, *F. serratus* male 4,17% and male 1,30% (Table 4).

Most libraries have increased its quality by trimming that was shown on the table 5. We compared several criteria using FasQC: per base sequence quality, per sequence quality scores, per base sequence content, per sequence GC content, sequence duplication levels and overrepresented sequences (Table 5). Quality of *F. serratus* female have not increased its quality, but *F. serratus* male and *F. distichus* pool 3 and enriched improved their quality and *F. serratus* male characterized by high quality in comparison with other libraries.

Table 5 Comparison of sequence quality using FastQC before and after using TrimGalore! “-” – means low quality, ”+” - high quality, -/+ medium quality,

Parameters	<i>F. distichus</i> pool 3 and enriched		<i>F. serratus</i> female		<i>F. serratus</i> male	
	before	after	before	after	before	after
Per base sequence quality	-	+/-	-	-	-	+
Per sequence quality scores	-	+/-	+	+	-	+
Per base sequence content	-	-	-	-	-	-
Per sequence GC content	+/-	+/-	+	+	+/-	+/-
Sequence duplication Levels	-	-	+	+	+	+
Overrepresented sequences	+/-	+//-	+	+	+	+

We merged together transcriptomes of *F. distichus* pool 3 and enriched one using special script and the result was *F. distichus* mixed. The number of assembled contigs (overlapping DNA fragments) was about six times higher in the *F. distichus* mixed transcriptome (67 221) (Table 4) than in the *F. serratus* female transcriptome 10 795 or *F. serratus* male transcriptome (12 075).

In Table 4 there is a results of Open Reading Frame (ORF) prediction. The highest number of candidate coding regions has been found within *F. distichus* mixed transcriptome (44 048) and the lowest one - in *F. serratus* male (1 823). Reciprocal BLAST identified 2 936 orthologs between the transcriptomes of *F. distichus* and *F. serratus* male and 8 234 orthologs between *F. distichus* and *F. serratus* female (Table 4, Figure 4, 5).

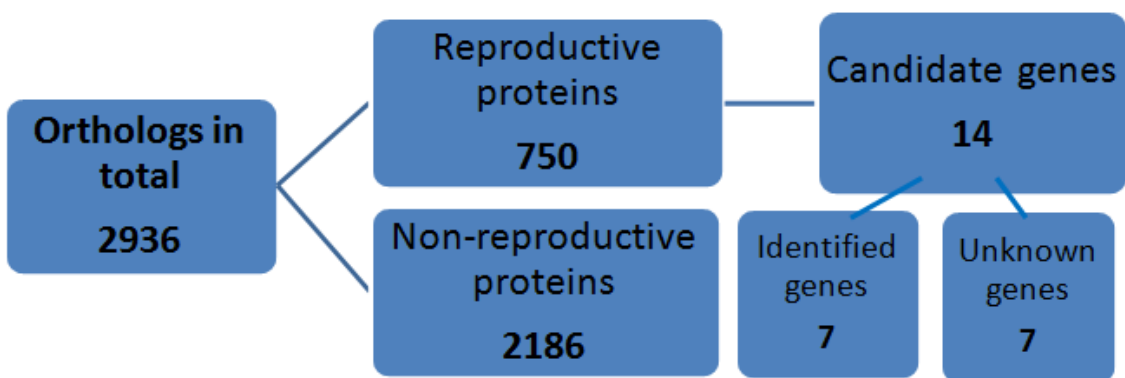


Figure 4 Analysis of orthologs found between *F. serratus* male vs. *F. distichus*.

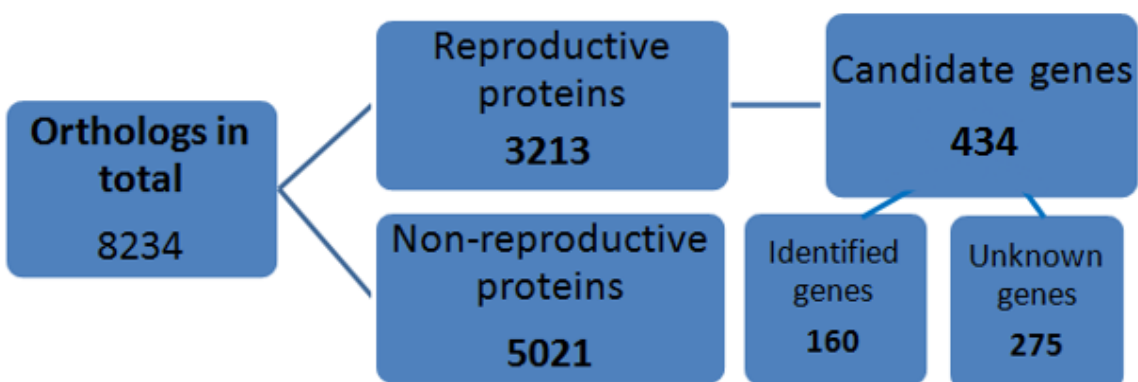


Figure 5 Analysis of orthologs found between *F. serratus* female vs. *F. distichus*.

Orthologs between the somatic *F. serratus* IonTorrent and reproductive transcriptomes were filtered out to select only the genes that are expressed during reproduction. In total, we identified 750 genes linked to reproduction processes for *F. serratus* male and 3213 for *F. serratus* female and with *F. distichus* orthologues (Figure 4, 5). Those genes that are reproduction-linked genes that had a dN/dS ratio >1, responded putatively to divergent selection were identified: 434 genes between *F. serratus* female vs. *F. distichus*, and 14 genes between *F. serratus* male vs. *F. distichus* (Figure 4, 5). Candidate genes are visualized in figure 6, 7 by red rectangles.

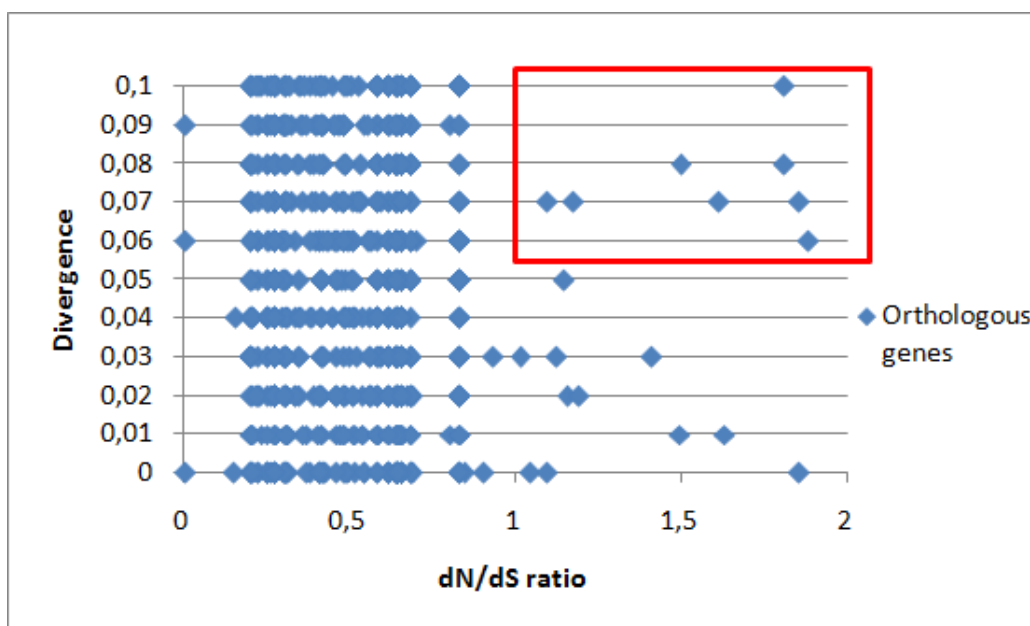


Figure 6 dN/dS ratio between pairwise comparison of orthologs, found between *F. serratus* female and *F. distichus* transcriptomes. The red square highlights putative speciation genes with dN/dS ratio > 1, indicating impact of natural selection and high genetic divergence. Approximately 1000 of orthologs are present in the figure.

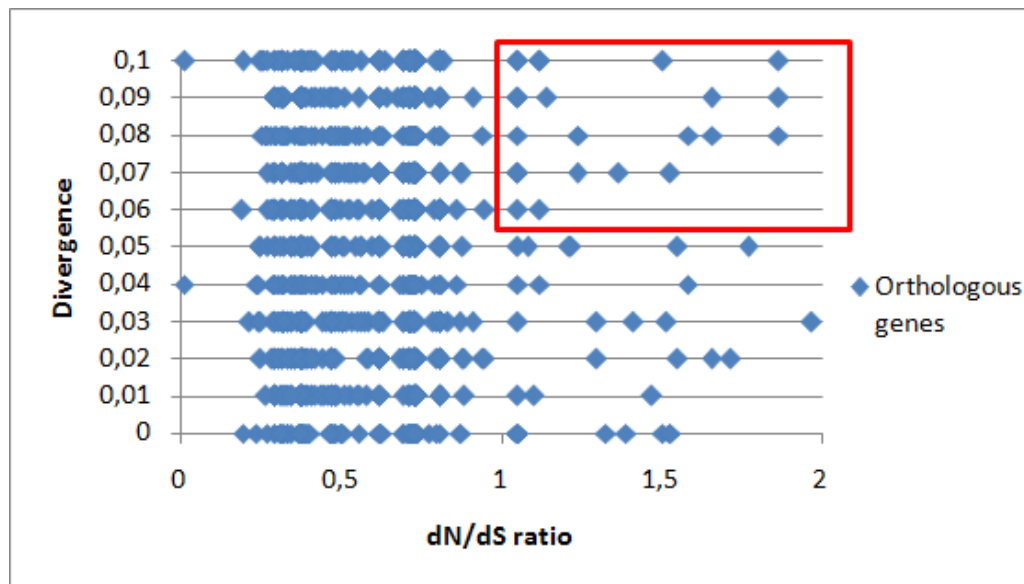


Figure 7 dN/dS ratio between pairwise comparison of orthologs, found between *F. serratus* male and *F. distichus* transcriptomes. The red square highlights putative speciation genes with dN/dS ratio > 1, indicating impact of natural selection and high genetic divergence. Approximately 1000 of orthologs are present in the figure.

The most of identified reproductive orthologous proteins found between *F. distichus* vs. *F. serratus* male (70%) matches to transcriptome of *Ectocarpus siliculosus*, 3% matching to brown algae and 1% to *F. distichus* and *F. vesiculosus* (Figure 8). In transcriptome of *F. serratus* male were identified genes that can possibly be associated with reproduction process (Table 6). Genes taking part in biosynthesis of amino acids translation and transcription present in table 7, 8. The majority of reproduction genes refer to biosynthesis of carbohydrate (Gdp-d-mannose 4,6-dehydratase, Glutamate dehydrogenase 2) and genes of cell detoxification (Glutathioneperoxidase, luthione S- transferase 3, 4), genes associated with flagella (dynein, kinesin) (Table 6). There was identified male-biased genes that was found also in female genome – ankyrin and beta-tubulin. In the male transcriptome were found genes that seems to be female-biased (Martins *et al.* 2013, Lipinska *et al.* 2013). There were observed cell wall gene Mannuronan C-5-epimerase and gene of vesicle transport SNAP (Soluble NSFAttachment Protein) receptor. In transcriptome of *F. serratus* male we found 11 types of reproductive genes in total (Table 6).

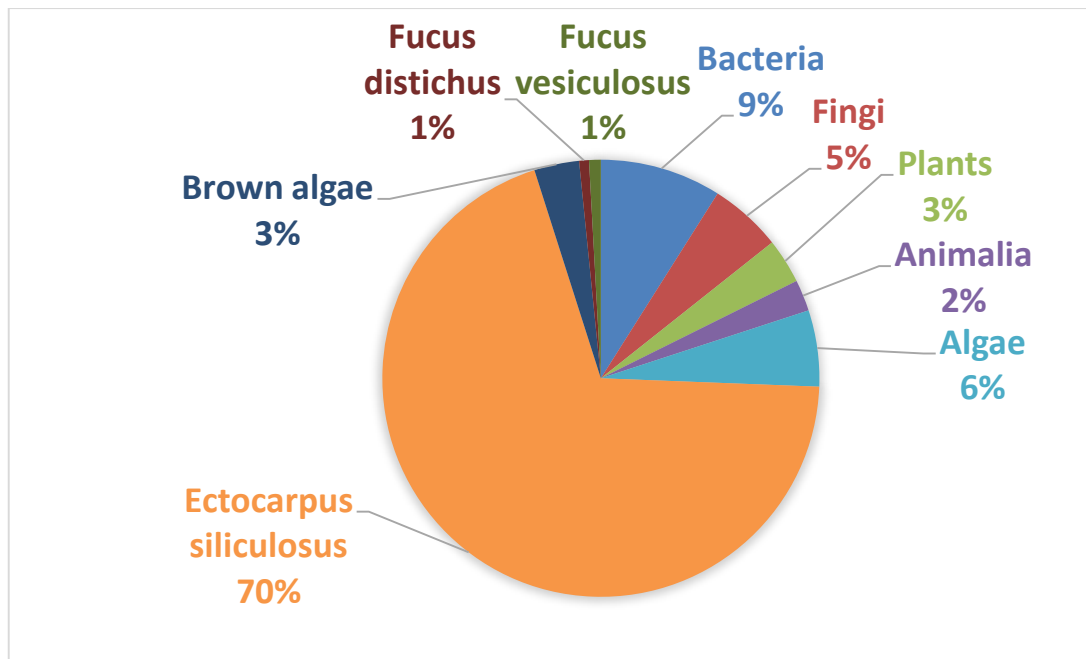


Figure 8 Percentage of the best BLAST hit of identified reproductive orthologous genes found between *F. distichus* vs. *F. serratus* male transcriptomes to different groups of organisms: animalia, plants, fungi, bacteria, algae, brown algae, *F. distichus* and *F. vesiculosus*.

Table 6 Reproductive genes of *F. serratus* male, female and *F. distichus* transcriptomes. According to the literature genes were defined as male-, female-biased and a common genes observed in both male and female transcriptomes (Lipinska *et al.* 2013; Martins *et al.* 2013).

Male-biased genes (according to the literature)			
Group name		Function	Transcriptome
Biosynthesis of polysaccharides			
1	Gdp-d-mannose 4,6-dehydratase	fructose and mannose metabolism	<i>F. serratus</i> male
2	Glutamate dehydrogenase 2		<i>F. serratus</i> male
Genes of detoxification			
3	Glutathioneperoxidase	peroxidase detoxification in cells	<i>F. serratus</i> male
4	lutathione S-transferase 3, 4	detoxification of xenobiotics	<i>F. serratus</i> male
Microtubules			
5	dynein	converts chemical energy into mechanical	<i>F. serratus</i> male
6	kinesin	maintain processes of cell division and transport	<i>F. serratus</i> male

7	alpha-tubulin	component of flagella	<i>F. distichus</i>
8	beta-tubulin	component of flagella	<i>F. serratus</i> male, female
9	D1LIC (Dynein light intermediate chain)	converts chemical energy into mechanical	<i>F. serratus</i> female
10	cytoplasmic dynein 1 light intermediate chain	converts chemical energy into mechanical	<i>F. serratus</i> female
11	Actin-related protein 2/3 complex subunit 3	component of flagella	<i>F. distichus</i>
12	Actin	component of flagella	<i>F. distichus</i>
13	Tubulin beta-5 chain	component of flagella	<i>F. distichus</i>
14	Beta-5-tubulin	component of flagella	<i>F. distichus</i>
15	Flagennar basal body protein	component of flagella	<i>F. distichus</i>
	Signalling molecules		
16	phosphatidylinositol 4-kinase	intermediate in a number of signaling pathways	<i>F. serratus</i> female
17	rho family GTPase of <i>Fucus distichus</i>	initial sperm-egg fusion	<i>F. serratus</i> female
	Protein-protein interactions		
18	ankyrin	provides protein-protein interuptions	<i>F. serratus</i> male, female
	Female-biased genes (according to the literature)		
	Cell-wall genes		
1	cellulose synthase	synthetase of cellulose	<i>F. serratus</i> female
2	aquaporin	controls water movements between extra-cellular matrix and the cell, maintaining the osmotic pressure	<i>F. serratus</i> female
3	sulfotransferase	growth and development of the cells, communication of the cells and defense	<i>F. serratus</i> female
4	Mannuronan C-5-epimerase	catalyzes the last step of biosynthesis of alginates	<i>F. serratus</i> male, female
	Vesicle transport		
5	arf1, ARF family GTPase	microtubule-dependent functions	<i>F. serratus</i> female
6	clathrin	formation of coated vesicles	<i>F. serratus</i> female
7	coatomer protein complex, beta sub-unit	Golgi budding and vesicular trafficking	<i>F. serratus</i> female
8	Vesicle coat complex COPI, beta subunit	Vesicles transporting proteins	<i>F. serratus</i> female

9	SNAP (Soluble NSF Attachment Protein) REceptor	mediated fusion of the vesicles	<i>F. serratus</i> male, female
Signalling molecules			
10	Rab6, RAB family GTPase	growth, differentiation and survival of the cell	<i>F. serratus</i> female
Genes that refers to male and female transcriptomes (according to the literature)			
1	14-3-3-like protein	participants in reactions between signaling proteins (kinases, phosphatases, and transmembrane receptors)	<i>F. distichus</i>
2	30S and 23S ribosomal	The transcription and translation processes	<i>F. serratus</i> male
3	40S small subunit	The transcription and translation processes	<i>F. distichus</i>
4	component of cytosolic 80S ribosome a	The transcription and translation processes	<i>F. distichus</i>
5	Polyadenylate-binding protein	Binding with poly-A tail of the mRNA	<i>F. distichus</i>
6	ribosomal protein S23	The transcription and translation processes	<i>F. distichus</i>
7	Ribosomal protein S27	The transcription and translation processes	<i>F. distichus</i>

Table 7 Reproductive genes among orthologs between *F. serratus* male vs. *F. distichus* taking part in biosynthesis of amino acids.

№	Gene ID	Gene name	Amino Acid
1	CBJ33413.1	Arginine biosynthesis bifunctional protein ArgJ	Arginine
2	CBJ25592.1	Arginyl-tRNA Synthetase	Arginine
3	CBJ29662.1	Asparagine synthase (glutamine-hydrolyzing)	Asparagine
4	CBN75233.1	Cysteine desulfurase	Cysteine
5	CBJ29069.1	Glycyl-tRNA Synthetase	Glycine
6	CBN79120.1	Homoserine dehydrogenase	Serine

7	CBN75596.1	Serine hydroxymethyltransferase 2	Serine
8	CBN78666.1	serine/threonine-protein kinase	Serine
9	CBN74675.1	Leucine rich repeat protein	Leucine
10	CBN79710.1	Lysyl-tRNA ligase	Lysine
11	CBN75077.1	Ornithine cyclodeaminase	Ornithine

Table 8 Reproductive genes among orthologs between *F. serratus* male vs. *F. distichus* taking part in the processes of translation and transcription.

№	Gene ID	Gene name
1	CBJ25582.1	elongation factor EF-3
2	CBN78248.1	Eukaryotic initiation factor 3 d subunit
3	ADG35721.1	eukaryotic translation initiation factor 3 subunit 6 interacting protein
4	ADG35721.1	eukaryotic translation initiation factor 3 subunit 6 interacting protein
5	CBN73980.1	eukaryotic translation initiation factor 4E-like 1
6	CBJ32071.1	global transcription factor group

Analysis of the genes with $dN/dS > 1$ shows that between reproductive genes of *F. distichus* vs. *F. serratus* male the highest number of accessions belongs to *Ectocarpus siliculosus* (64,3%) duo to presence of annotated *E. siliculosus* transcriptome on the second place is *Eimeria mitis* (28,6%) (Table 9). In total 14 genes have dN/dS ratio significantly more than 1 proved by p -value $< 0,05$ (Table 9) and by two methods in case of 11 genes and

by one method with the rest 3 genes. 50% of genes were identified (Figure 4, 5) for the rest unknown seven genes sequences were submitted in table 5S to explore more details. Among seven identified genes two refer indirectly to reproduction process: 30S ribosomal protein S24 of *Saprolegnia diclina* and ribosomal protein S23, component of cytosolic 80S ribosome and 40S small subunit *Ectocarpus siliculosus*.

Table 9 Identified reproductive genes of *F. distichus* vs. *F. serratus* male transcriptomes, which have dN/dS ratio more than 1. ID number refers to transcriptome of *F. serratus* male. Some sequences includes 2 methods of dN/dS ratio calculation. Genes that refer to the processes of reproduction highlight with bold font.

Nº	ID number in transcriptome	Method	dN/dS	P-value	Name of gene	Length (Amino acids)
1	TR1255	LWL	1,41269	0,001476	protein disulfide isomerase [Ectocarpus siliculosus]	92
		MLPB	1,49696	0,000122		
2	TR1377	LWL	1,41269	0,001476	asn/thr-rich large protein family protein [Ectocarpus siliculosus]	136
		MLPB	1,49696	0,000122		
3	TR1387	LWL	1,41269	0,001476	30S ribosomal protein S24e [Saprolegnia diclina VS20]	61
		MLPB	1,49696	0,000122		
4	TR865	LWL	1,41269	0,001476	ornithine cyclodeaminase [Vibrio xuii]	60
		MLPB	1,49696	0,000122		
5	TR1374	MLWL	2,05153	0,002916	Light harvesting complex protein [Ectocarpus siliculosus]	123
6	TR1487	MLWL	1,61323	0,000191	dihydrofolate reductase dfrA [Mycobacterium tuberculosis]	140
7	TR1621	MLWL	1,61323	0,000191	Ribosomal protein S23, component of cytosolic 80S ribosome and 40S small subunit [Ectocarpus siliculosus]	145

As in the case of genes identified between *F. distichus* vs. *F. serratus* female the most part of identified genes between *F. distichus* vs. *F. serratus* male (71%) concerns to genes of *Ectocarpus siliculosus*, 4% attributes to algae, 3% to brown algae and 1% to *Fucus* sp. (Figure 9). Genes taking part in biosynthesis of amino acids (Table 10) and translation and

elongation might have impact on the reproduction (Table 11). Transcriptome of *F. serratus* female had a majority of cell wall and transport genes (Table 6). To the cell wall genes refers cellulose synthase, aquaporin, aquaporin and Mannuronan C-5-epimerase. Group of vesicle transport in female transcriptome consist on arf1, ARF family GTPase, clathrin, coatomer protein complex, beta sub-unit, Vesicle coat complex COPI, beta subunit SNAP. Rab6, RAB family GTPase as a signaling molecule was found among reproduction genes in female transcriptome. According to the literature phosphatidylinositol 4-kinase and rho family GTPase of *Fucus distichus* usually refers to the male transcriptome (Martins *et al.* 2013, Lipinska *et al.* 2013), but was identified in *F. serratus* female transcriptome. We identified also male-biased genes associated with flagella (DILIC (Dynein light intermediate chain) and cytoplasmic dynein 1 light intermediate chain). 30S and 23S ribosomal gene that do not refers to male or female transcriptome (Martins *et al.* 2013, Lipinska *et al.* 2013) was observed in *F. serratus* female transcriptome. In total in *F. serratus* transcriptome we found 16 types of genes (Table 6).

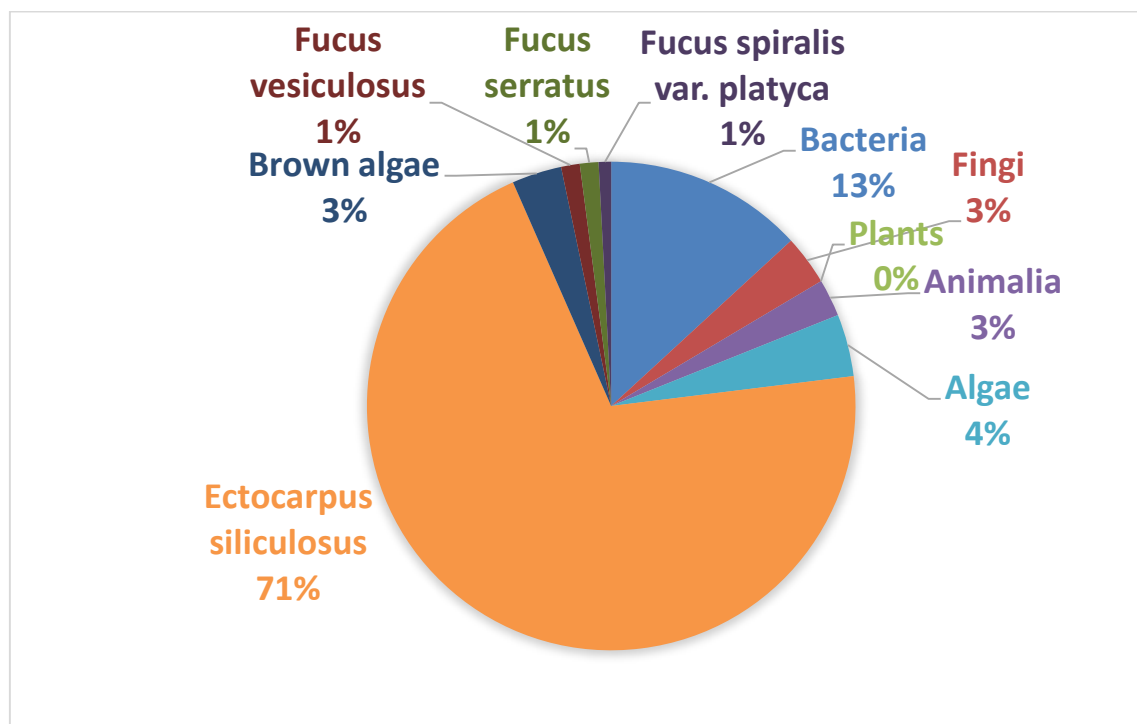


Figure 9 Percentage of the best BLAST hit of identified reproductive orthologous genes found between *F. distichus* vs. *F. serratus* female transcriptomes to different groups of organisms: animalia, plants, fungi, bacteria, algae, brown algae, *F. distichus* and *F. vesiculosus*.

Table 10 Reproductive genes among orthologs between *F. serratus* female vs. *F. distichus* taking part in biosynthesis of amino acids.

№	Gene ID	Gene name	Amino Acid
1.	CBN77982.1	Argininosuccinate lyase	Arginine
2.	CBN78379.1	Argininosuccinate synthetase	Arginine
3.	CBJ30598.1	Argonaute 1	Arginine
4.	CBN78830.1	3-Deoxy-D-arabinoheptulosonate 7-phosphate (DAHP)	Phenylalanine Tyrosine Tryptophan
5.	CAC36148.1	ALA dehydratase	Alanine
6.	CBN74554.1	Cystathionine beta-synthase	Cysteine
7.	CBN75996.1	Cystathionine gamma-lyase	Cysteine
8.	CBN75233.1	Cysteine desulfurase	Cysteine
9.	NP_199566.1	Cysteine proteinase inhibitor 5	Cysteine
10.	CBJ33850.1	Cysteine synthase	Cysteine
11.	CBJ48668.1	Glycine cleavage system H protein	Glycine
12.	CBJ29069.1	Glycyl-tRNA Synthetase	Glycine
13.	CBN75879.1	Leucine rich repeat protein	Leucine
14.	CBN75077.1	Ornithine cyclodeaminase	Ornithine Proline
15.	CBN76803.1	Serine Carboxypeptidase	Serine

16.	CBN75444.1	Serine hydroxymethyltransferase 2	Serine
17.	CBN78358.1	Serine O-acetyltransferase, incomplete	Serine
18.	CBN80294.1	Thiamine monophosphate synthase	Thiamine
19.	CBJ33029.1	Tyrosinase	Tyrosine
20.	CBN79225.1	Tyrosinase family member (tyr-4)	Tyrosine
21.	CBN79231.1	Tyrosinase-like protein 2	Tyrosine

Table 11 Reproductive genes among orthologs between *F. serratus* female vs. *F. distichus* taking part in the processes of translation and transcription.

№	Gene ID	Gene name
1.	CBJ27261.1	DnaJ-like/ Sec63 translocase subunit
2.	EEF1A2	eukaryotic translation elongation factor 1 alpha
3.	CBJ32863.1	EF2, translation elongation factor 2
4.	CBN74782.1	EIF2S3, gamma subunit of the eukaryotic translation initiation factor 2
5.	CBN77232.1	EIF3C EUKARYOTIC TRANSLATION INITIATION FACTOR 3 subunit c
6.	ETL93330.1	elongation factor 1-alpha
7.	CBJ25582.1	elongation factor EF-3
8.	CBN77382.1	eukaryotic elongation factor-1 B beta
9.	CBJ29821.1	Eukaryotic initiation factor 1

10.	CBN77297.1	eukaryotic translation initiation factor 2B alpha subunit
11.	CBJ32071.1	global transcription factor group
12.	CBN76113.1	HBS1, eRF3-like GTPase involved in mRNA
13.	CBN80083.1	high mobility group protein
14.	NP_593314.1	High-mobility group non-histone chromatin protein
15.	CBJ26585.1	Mef1, mitochondrial translation elongation factor EF-G
16.	CBN73983.1	putative initiation factor eIF3 g subunit
17.	CBN78637.1	putative Translation Initiation Inhibitor
18.	CBJ26618.1	similar to activating transcription factor 6
19.	EDN00410.1	transcriptional regulator, ArsR family
20.	CBN74067.1	translation initiation factor eIF3 f subunit

According to our analysis we found 434 genes with $dN/dS > 1$ (p-value $< 0,05$ in the most cases of one method of dN/dS calculation) (Figure 5, 6, Table 5, Table 6S). One of these genes shows significant dN/dS ratio >1 with 5 methods and 23 were confirmed by 2 methods (Table 6S). Genes under positive selection were divided into 8 group (Table 6S). The majority of genes belongs to the group of metabolically genes (68,5 % from total amount of genes under positive selection) and genes of translation and transcription (11, 5%) (Table 6S). The other genes refer to the groups of photosynthesis (7 %), flagella-associated proteins (4,2 %), vacuolar genes (2,7 %), signaling molecules (0,007 %), hypothetical and conserved proteins (2 %) and viral genes (3,2 %).

We found 60 % of reproductive genes between transcriptomes of *F. distichus* vs. non-reproductive *F. serratus* IonTorrent by comparing them with reproductive transcriptomes of *F. serratus* male and female (Table 4S). 27 % them were unknown putative proteins, but

the rest of genes were identified. The majority were male-biased flagella-associated genes: actin-related protein 2/3 complex subunit, actin, tubulin beta-5 chain, alpha-, beta-5-tubulin and flagellar basal body protein. There were found high number of a ribosome genes: 40S small subunit, component of cytosolic 80S ribosome, ribosomal protein S23 and ribosomal protein S27. 14-3-3-like protein providing reactions between signaling molecules (kinases, receptors of membranes) and polyadenylate-binding protein that attach poly-A tail to mRNA (Kahvejian *et al.* 2005). We found 12 types of genes in *F. distichus* transcriptome.

4. Discussion

We have identified genes linked with reproduction within the transcriptomes of *F. serratus* male, female and *F. distichus*. Analysis of the genes shows that *F. serratus* female transcriptome involve higher number of genes in comparison with *F. serratus* male. The most part of *F. serratus* female are normally female-biased (usually found in *F. vesiculosus* female transcriptome (Martins *et al.* 2013), but it was also found male-biased genes. According to the literature female-biased genes (specific for female transcriptome) were more numerous than male-biased in transcriptome of *Ectocarpus siliculosus* (Martins *et al.* 2013). Transcriptome of *F. distichus* as *F. serratus* male have less number of genes than female 11 in comparison with 16. However, technical issues have resulted in limited depth of sequencing for the transcriptomes of *F. serratus*, so we should take our results with caution.

1. Genes under positive selection in transcriptomes of *F. serratus* male and female

Under the hypothesis of reinforcement we would expected that female respond stronger for selection against hybrids (Lipinska *et al.* 2013; Martins *et al.* 2013; Perry *et al.* 2014; Sharma *et al.* 2014). Indeed, we have found 434 genes under positive selection within female transcriptome in comparison with 14 genes in the male transcriptome. The majority of genes refers to the group of metabolically genes (68,5 %) and seems not to have impact on reproduction processes before fertilization. These genes presumably may provide and maintain the cell processes after fertilization (Xin *et al.* 2011): metabolism, transport of chemical compound within the cells that provide genes of vesicle transport (3 %), food supply process by genes of photosynthesis (7 %). Group of translation and transcription genes (11,5 %) may impact directly on reproduction by synthesis of proteins that presumably refer to gamete attraction and interaction. Gamete recognition and chemo sensation presumably are provided by flagella-associated genes (5 %) (Muller *et al.* 1988; Schmid 1993) that may provide gamete interaction. The most important found gene was Rab6 from RAB family GTPase that refers to the processes of growth, cell differentiation and maintain cell vitality and it presumably initiates fusion of egg and sperm (Lipinska *et*

al. 2013). These genes are suggested to a role in gamete incompatibility (Lipinska *et al.* 2013). Incompatibility is likely based on reduced gamete attraction and recognition, favoring prezygotic mechanism of isolation and presumably reinforcement (Hoarau *et al.* 2015). Among 7 genes of *F. serratus* male transcriptome under positive selection we have not found genes that may impact on gamete recognition and attraction. The majority of found genes refer to metabolically pathways, components of ribosome and genes of photosynthesis that may impact on post fertilization processes, but according to the literature this role mostly plays female (Xin *et al.* 2011).

2. Reproductive genes of *F. serratus* male transcriptome.

2.1 carbohydrate metabolism

According to the recent investigations high number of genes of carbohydrate metabolism can be found in the reproductive tissue of brown algae (Martins *et al.* 2013, Lipinska *et al.* 2013). *F. serratus* male transcriptome had gdp-d-mannose 4,6-dehydratase that refers to the fructose and mannose synthesis (Mulichak *et al.* 2002). Genes of carbohydrate metabolism were female-biased genes (according to the literature (Martins *et al.* 2013, Lipinska *et al.* 2013)) that was explained by polysaccharides ability to form cell wall that is an essential reproduction function for females (Martins *et al.* 2013).

2.2 Genes of detoxification

We found several genes of cell-defense in *F. serratus* male transcriptome. Glutathioneperoxidase, which participates in detoxification of the cell from peroxides (Diane *et al.* 2009) . Lutathione S-transferase 3 and glutathione S-transferase 4 detoxifying xenobiotics were observed (Josephy *et al.* 2010). They are found mostly in male transcriptome of brown algae and were assigned to the male-biased genes (Martins *et al.* 2013, Lipinska *et al.* 2013). These genes are might be possibly important for reproduction process by maintenance of sperm and other cells vital activity (Diane *et al.* 2009).

2.3. Flagella-associated genes

F. serratus male transcriptome had a high number of flagella-associated genes: kinesin maintain some cellular reactions, such as cell division and transport moving along

microtubule (Vale 2003). Dynein that bound to flagella converts chemical energy encompassed within ATP molecule into energy of movement (mechanical) (Karp *et al.* 2005) and presumably initiated and maintained sperm motility (Martins *et al.* 2013). Repeats of ankyrin providing protein-protein interactions that might be essential for gamete recognition because it mediated interaction between protein and cell wall (Michaely *et al.* 2002). In *F. serratus* male transcriptome we found male-biased beta-tubulin as a part of the flagella. This gene was observed in *F. serratus* female, but usually it has low expression in female transcriptome (Martins *et al.* 2013)³. Genes of flagella might be involved in acrosome reaction of the sperm during penetration to the egg (Gilbert *et al.* 2000).

3. Reproductive genes of *F. serratus* female transcriptome.

3.1 Cell-wall genes

It was observed that genes, which take part in biosynthesis of carbohydrates and cell wall genes, were over expressed in females, because after fertilization cell wall needs to be restored after penetration of sperm (Callow *et al.* 1978b). One of the minor female-biased gene is cellulose synthase that was observed within genes of *F. serratus* female transcriptome. Aquaporin, sulfotransferase and mannuronan C-5-epimerase are the other female-based genes observed in *F. serratus* female transcriptome. The first gene maintaining the osmotic pressure and possibly controls movements of water between extra-cellular matrix and cells, which assists in gametangia extrusion (Agre 2006; Martins *et al.* 2013). Sulfotransferase catalyzes reaction components of what provides interaction, development, growth and defense of the cell (Negishi *et al.* 2001). Mannuronan C-5-epimerase catalyzes the last step of alginates biosynthesis (Nyvall *et al.* 2003; Martins *et al.* 2013). Alginates are the essential forming components of the algae cell (Andresen *et al.* 1977). These genes are also do not provide gamete recognition by itself, but they might impact indirectly.

3.2. Genes of vesicle transport

In our data group of vesicle transport genes contains arf1, ARF family GTPase that might participates in membrane traffic and functions associated with microtubes (Kahn *et al.* 2005). Clathrin is a female-biased gene found within *F. serratus* transcriptome that

plays a role in formatting of coated vesicles involved in cell interaction (Pearse 1976). The other two genes of vesicle transport are coatamer protein complex coating vesicles without clathrin coat (De Baere *et al.* 1999) and vesicle coat complex COPI is also covering vesicles of the complex Golgi (Serafini *et al.* 1991). Active Golgi bodies plays an important role in formation of the gamete and spores according to photographs of *Fucus* eggs, which was not fertilized, showing presence of vesicles under the plasma membrane (Baker *et al.* 1973; Callow *et al.* 1978b; Maier 1997). These evidences corresponds to over expression of clathrin coat proteins (move vesicles from the Golgi apparatus towards to the plasma membrane) (Kirchhausen 2000) and dynamins (incision of vesicles) (Kirchhausen 2000). The components of the Golgi apparatus play major role in formation of important in reproduction polysaccharides such as phlorotannins, alginates and sulfated fucans, which usually transported towards to the cell membrane using vesicles (Callow *et al.* 1978a; Schoenwaelder 2000). It was assumed that biogenesis of cell wall could be maintained by vesicular transport of cell wall components and formation of cellulose, which supported by over expression of genes responsible for synthesis of these chemical compound (Lipinska *et al.* 2013). There was a hypothesis that these substances take a part in gamete adhesion as a part of cell membrane, which could effect on reproduction (Charrier *et al.* 2008). For example, sulfated fucans and galactans response for inducing of acrosome reaction of sea urchins by adhesion to the sperm REJ (Receptor for Egg Jelly) (Alves *et al.* 1997; Mourão 2007).

3.3 Signaling molecules

There were observed male-biased (according to the literature) signaling molecules: phosphatidylinositol 4-kinase mediated signaling between gametes and participate in the most important reaction of fertilization – breakdown induces by sperm by phospholipase C activating (Sillers *et al.* 1985; Halet *et al.* 2002).

3.4 Male-biased genes (according to the literature) found in transcriptome of *F. serratus* female

In our data in *F. serratus* female transcriptome were found male-biased (according to the literature) genes (Martins *et al.* 2013) of flagella: D1LIC (Dynein light intermediate

chain), cytoplasmic dynein 1 light intermediate chain, similar to cytoplasmic dynein intermediate chain, beta-tubulin and ankyrin. SNARE protein was found within transcriptomes of *F. serratus* male and female. SNARE proteins - soluble NSF (N-ethylmaleimide-sensitive factor) attachment receptor protein mediated fusion of the vesicles (Malsam *et al.* 2011).

4 Unknown proteins within both transcriptomes of *F. serratus* male and female

More than 50% (54% in female and 51% in male transcriptome) of the genes we found were coding for unknown proteins. Similar results have been found in *Ectocarpus* and *Fucus vesiculosus* transcriptomes (Martins *et al.* 2013, Lipinska *et al.* 2013). There is a suggestion that some of these genes could be involved in gamete incompatibility and subsequently reproductive isolation, but their function and structure needs to be investigated (Martins *et al.* 2013, Lipinska 2013).

5 Reproductive genes of *F. distichus*

Presence both male- and female-biased genes was expected in transcriptome of monoecious *F. distichus*, but we have not found any female-associated genes. The most expressed male-related genes were flagella forming genes and genes of ribosome protein. As was mentioned previously genes of flagella have an indirect impact on reproduction, because it effect on the sperm movement and these cytoskeletal proteins putatively realize ultimate steps of fertilization, but full function remains to be not completely understudied (Dvoráková *et al.* 2005). Transcriptome of *F. distichus* characterized by appearance of genes that was not found previously in reproductive transcriptome of *F. vesiculosus* (João *et al.* 2013) or *Ectocarpus siliculosus* (Lipinska 2013): polyadenylate-binding genes and 14-3-3 protein. These genes might taking part in the process of reproduction, because the first one participates in translation by presumable indirect impact on reproductive protein when they synthesized by ribosome during translation. 14-3-3 protein inhibited P-type H⁺ ATPases (Jahn *et al.* 2002) that involved in transport of K⁺, Ca²⁺ and Mg²⁺ (Stokes *et al.* 2003; Lenoir *et al.* 2007; Lopez-Marques *et al.* 2010). Activity of that kind of potassium channels might be connected with chemotaxis of the sperm (the same principle which induces motility of

sea urchin sperm) (Neill *et al.* 2004; Darszon *et al.* 2008). Hyperpolarization and penetration of Na⁺ and Ca²⁺ ions in sperm induced by dynamics of K⁺ ion and activity of cyclic nucleotide messengers (Hansbrough *et al.* 1980; Lee *et al.* 1986; Babcock *et al.* 1992; Galindo *et al.* 2000; Granados-Gonzalez *et al.* 2005; Strünker *et al.* 2006). Temporary increase in concentration of Ca²⁺ ions leads to chemotaxis orientation of the sperm using a fast Ca²⁺ imaging system it could have direct influence on the reproduction process (Shiba K 2008; Darszon *et al.* 2011; Collingridge *et al.* 2012). We found 12 types of reproduction genes that more than in *F. serratus* male (11), but less in comparison with female transcriptome (16). It may provide theory that female organism contribute more in after fertilization processes (Xin *et al.* 2011).

In our investigation we have not found important for reproduction process genes of the sexually induced gene family – Sig1 and Sig2-like genes that were referred as male-specific and first were described within *Thalassiosira* genome (Armbrust 1999). It was suggested that these genes may have gamete recognition function due to existence of different forms of Sig family proteins between closely related species (Sorhannus *et al.* 2006) under positive selection (Armbrust 1999). We did not found either g protein–coupled receptors (GPCR) that are male-specific genes participated in acrosome reaction in the middle of egg-sperm fusion during the process of fertilization or there are a lot of pheromones, which could bound to GPCR receptor (Ward *et al.* 1992; Ohta *et al.* 2000; Spehr *et al.* 2004).

5. Conclusion

Our study compares reproductive genes in transcriptomes of *F. serratus* male; female and *F. distichus*. Analysis of reproductive protein shows that *F. serratus* female has higher types of reproductive genes – 16 in comparison with 11 in *F. serratus* male and 12 in *F. distichus*.

The most part of *F. serratus* male transcriptome genes are flagella-associated genes that refers to important function of flagella movement needs for the process of gamete fusion and fertilization. Also male genes provides biosynthesis of polysaccharides and cell detoxification.

F. serratus female transcriptomes characterized by involvement in processes of cell wall recovery and maintenance that linked with reproduction function of female organism to support the new appeared cells. There were observed high number of vesicle transport genes that are essentials for gamete formation. But the most important findings in *F. serratus* female genome related with genes of signaling molecules providing egg-sperm reaction of fusion and fertilization (Rab6, RAB family GTPase)

F. distichus characterized by high number of flagella-associated and ribosomal genes. But we have found also gene 14-3-3 protein involved in transport of K^+ , Ca^{2+} and Mg^{2+} that might impact on chemotaxis of the sperm.

Our aim was to identify candidate genes under positive selection with potential impact on reproduction in transcriptomes of *F. serratus* male; female and *F. distichus*. Some candidate genes was revealed but it needs to measure dN/dS in younger zones of sympatry and find out whether genes there are under positive selection. Then these genes needs to be compared with our findings.

6. References

- P. Agre (2006). "The aquaporin water channels." Proc Am Thorac Soc **1**: 5-13.
- L. A. Alice, S. H. Brawley, E. A. Serrao (1999). "Evolution of the Fucaceae (Phaeophyceae) inferred from nrDNA-ITS. ." J. Phycol. **35**: 382-394.
- A.-P. Alves, B. Mulloy, J. A. Diniz, P. A. S. Mourão (1997). "Sulfated polysaccharides from the egg jelly layer are species-specific Inducers of acrosomal reaction in sperms of sea urchins. ." J Biol Chem **272**: 6965–6971.
- M. J. Anderson, J. Nyholt, A. F. Dixson (2006). "Sperm competition affects the structure of the mammalian vas deferens." J. Zool. **264**: 97-103.
- I. L. Andresen, O. Skipnes, O. Smidsrød, K. H. Estgaard *et al.* (1977). "Some biological functions of matrix components in benthic algae in relation to their chemistry and the composition of seawater." ACS Symp. Ser. **48**: 361-381.
- S. Andrews (2010). "FastQC." from <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
- E. V. Armbrust (1999). "Identification of a New Gene Family Expressed during the Onset of Sexual Reproduction in the Centric Diatom *Thalassiosira weissflogii*." Appl Environ Microbiol. **65**(7): 3121–3128.
- E. V. Armbrust, H. M. Galindo (2001). "Rapid evolution of a sexual reproduction gene in centric diatoms of the genus *Thalassiosira*." Appl Environ Microbiol. **67**(8): 3501-3513.
- J. Arrontes (1993). "Nature of the distributional boundary of *Fucus serratus* on the north shore of Spain." Mar. Ecol. **93**: 183–193.
- D. F. Babcock, M. M. Bosma, D. E. Battaglia, A. Darszon (1992). " Early persistent activation of sperm K + channels by the egg peptide speract." Proc Natl Acad Sci U.S.A. **89**: 6001–6005.
- J. R. J. Baker, L. V. Evans (1973). "The ship fouling alga *Ectocarpus*. ." Protoplasma **77**: 1-13.
- C. Biermann (1998). "The molecular evolution of sperm bindin in six species of sea urchins (Echinoida:Strongylocentrotidae)." Mol Biol Evolution **15**: 1761-1771.
- R. Butlin, A. Debelle, C. Kerth, R. R. Snook *et al.* (2011). "What do we need to know about speciation?" Trends Ecol Evol. **27**(1): 27-39.

- M. I. Cain, V. Andreasen, D. Howard (1999). "Reinforcing selection is effective under a relatively broad set of conditions in a mozaic hybrid zone." Evolution **53**: 1343-1353.
- M. E. Callow, S. J. Coughlan, L. V. Evans (1978a). "The role of Golgi bodies in polysaccharide sulphation in *Fucus* zygotes. ." J Cell Sci **32**: 337–356.
- M. E. Callow, L. V. Evans, G. P. Bolwell, J. A. Callow (1978b). "Fertilization in brown algae. I. sem and other observations on *Fucus serratus*." J. Cell Set. **32**: 45-54.
- F. Canovas, C. Mota, E. Serrão, G. A. Pearson (2011). "Driving south: a multi-gene phylogeny of the brown algal family Fucaceae reveals relationships and recent drivers of a marine radiation. ." BMC Evol. Biol. **11**: 371–376.
- B. Charrier, S. M. Coelho, A. Le Bail, T. Tonon *et al.* (2008). "Development and physiology of the brown alga *Ectocarpus siliculosus*: two centuries of research. ." New Phytol **177**: 319–332.
- C. Cheong-Xin, T. Swee-Sen, H. Chai-Ling, R. Y. Othman *et al.* (2004). "Optimisation of RNA Extraction from *Gracilaria changii* (*Gracilariales, Rhodophyta*)." Journal of Applied Phycology **16**(4): 297-3001.
- P. W. Collingridge, G. Wheeler, C. Brownlee (2012). "The molecular mechanisms of calcium signalling in *Chlamydomonas* flagella." Cilias **1**: 19.
- J. Coyer, H. G., M. Skage, T. W. Stam *et al.* (2006). "Origin of *Fucus serratus* (Heterokontophyta; Fucaceae) populations in Iceland and the Faroes: a microsatellite-based assessment." Eur. J. Phycol **41**(2): 235–246.
- J. A. Coyer, G. Hoarau, J. F. Costa, B. Hogerdijk *et al.* (2011). "Evolution and diversification within the intertidal brown macroalgae *Fucus spiralis*/*F. vesiculosus* species complex in the North Atlantic. ." Mol. Phylogenet. Evol. **58**: 283–296.
- J. A. Coyer, G. Hoarau, M. P. Oudot-Le Secq, W. T. Stam *et al.* (2006). "A mtDNA-based phylogeny of the brown algal genus *Fucus* (Heterokontophyta; Phaeophyta)." Molecular phylogenetics and evolution **39**(1): 209-222.
- J. A. Coyer, A. F. Peters, G. Hoarau, W. T. Stam *et al.* (2002). "Inheritance patterns of ITS1, chloroplasts and mitochondria in artificial hybrids of the seaweeds *Fucus serratus* and *F. evanescens* (*Phaeophyceae*)." Eur. J. Phycol **37**: 173-178.
- J. A. Coyne, H. A. Orr (1998). "The evolutionary genetics of speciation." Philos Trans R Soc Lond B Biol Sci. **353**(1366): 287–305.
- J. A. Coyne, H. A. Orr (2004). Speciation.

- A. Darszon, A. Guerrero, B. E. Galindo, T. Nishigaki *et al.* (2008). "Sperm-activating peptides in the regulation of ion fluxes, signal transduction and motility." Int J Dev Biol **52**: 595-606.
- A. Darszon, T. Nishigaki, C. Beltran, C. L. Treviño (2011). "Calcium channels in the development, maturation, and function of spermatozoa." Physiol Rev. **91**(4): 1305-1355.
- C. Darwin, Ed. (1859). On The Origin of Species by Means of Natural Selection, or The Preservation of Favoured Races in the Struggle for Life.
- E. De Baere, F. Speleman, N. Van Roy, A. De Paepe *et al.* (1999). "Assignment of the cellular retinol-binding protein 1 gene (RBP1) and of the coatomer beta subunit gene (COPB2) to human chromosome band 3q23 by in situ hybridization." Cytogenet Cell Genet **82**(3-4): 226–227.
- B. E. Deagle, A. C. Thomas, A. K. Shaffer, A. W. Trites *et al.* (2013). "Quantifying sequence proportions in a DNA-based diet study using Ion Torrent amplicon sequencing: which counts count?" Molecular Ecology Resources **13**: 620–633.
- E. H. Diane, L. Edith, Y. Yi, D. John *et al.* (2009). "Glutathione Peroxidase-1 Regulates Mitochondrial Function to Modulate Redox-dependent Cellular Responses." J Biol Chem. **284**(18).
- K. Dvoráková, H. D. Moore, N. Sebková, J. Palecek (2005). "Cytoskeleton localization in the sperm head prior to fertilization." Reproduction **130**(1): 61-69.
- R. H. E. Martin, Ed. (2008). A Dictionary of Biology.
- J. A. Endler, A. E. Houde (1995). "Geographic variation in female preferences for male traits in *Poecilia reticulata* " Evolution **49**: 456-468.
- D. J. Futuyma (2013). Evolution, Sinauer Associates.
- B. E. Galindo, C. Beltrán, E. J. Cragoe, A. Darszon (2000). "Participation of a K(+) channel modulated directly by cGMP in the speract-induced signaling cascade of stronglylocentrotus purpuratus sea urchin sperm. ." Dev Biol **221**: 285–294.
- S. Gavrilets (2000). "Rapid evolution of reproductive barriers driven by sexual conflict." Nature **403**: 886-889.
- S. F. Gilbert, M. A. Sunderland (2000). "Recognition of Egg and Sperm." Developmental Biology **6**.

- A. Y. Gladenkova, A. E. Oleinikb, L. J. Marincovich, K. B. Barinovd (2002). "A refined age for the earliest opening of Bering Strait." Palaeogeography, Palaeoclimatology, Palaeoecology **183**(3-4): 321–328.
- N. Goldman, Z. Yang (1994). "A codon-based model of nucleotide substitution for protein-coding DNA sequences." Mol Biol Evol. **11**(5): 725-736.
- M. G. Grabherr, B. J. Haas, M. Yassour, J. Z. Levin *et al.* (2011). "Full-length transcriptome assembly from RNA-seq data without a reference genome." Nat Biotechnol. **29**(7): 644-652.
- G. Granados-Gonzalez, I. Mendoza-Lujambio, E. Rodriguez, B. E. Galindo *et al.* (2005). "Identification of voltage-dependent Ca²⁺ channels in sea urchin sperm. ." FEBS Lett. **579**: 6667–6672.
- G. Halet, R. Tunwell, T. Balla, K. Swann *et al.* (2002). "The dynamics of plasma membrane PtdIns(4,5)P(2) at fertilization of mouse eggs." J Cell Sci **115**: 2139–2149.
- J. R. Hansbrough, G. S. Kopf, D. L. Garbers (1980). "The stimulation of sperm metabolism by a factor associated with eggs and by 8-bromo-guanosine 3',5'-monophosphate. ." Biochim Biophys Acta **630**: 82–91.
- M. E. Hellberg, V. D. Vacquier (1999). "Rapid evolution of fertilization selectivity and lysin cDNA sequences in teguline gastropods." Mol Biol Evolution **16**: 839-848.
- G. Hoarau, J. Coyer, M. C. W. G. Giesbers, A. Jueterbock *et al.* (2015). "Pre-zygotic isolation in the macroalgal genus *Fucus* from four contact zones spanning 100–10 000 years: a tale of reinforcement?" Royal society open science.
- G. Hoarau, J. A. Coyer, J. H. Veldsink, W. T. Stam *et al.* (2007). "Glacial refugia and recolonization patterns in the brown seaweed *Fucus serratus*." **16**: 3606–3616. .
- P. A. Hohenlohe, P. C. Phillips, W. A. Cresko (2010). "Using population genomics to detect selection in natural populations: key concepts and methodological considerations." Int J Plant Sci. **171**(9): 1059–1071.
- D. J. Howard (1993). Reinforcement: origin, dynamics and fate of an evolutionary hypothesis In Hybrid zones and the evolutionary process, Oxford, UK: Oxford University Press.
- L. Hurst (2002). "The Ka/Ks ratio: diagnosing the form of sequence evolution." Trends in Genetics **18**(9): 486–489.
- I. Istvan (2014). "A Short Guide to Illumina Sequencing."

T. P. Jahn, A. Schulz, J. Taipalensuu, M. G. Palmgren (2002). "Post-translational modification of plant plasma membrane H(+)-ATPase as a requirement for functional complementation of a yeast transport mutant." The Journal of Biological Chemistry **277**(8): 6353–6358.

K. P. Johnson, J. M. Allen, B. P. Olds, L. Mugisha *et al.* (2013). "Rates of genomic divergence in humans, chimpanzees and their lice." Proc Biol Sci. **281**(1777).

P. D. Josephy, J. D. Hayes, J. U. Flanagan, I. R. Jowsey (2010). "Genetic variations in human glutathione transferase enzymes: significance for pharmacology and toxicology." Hum Genomics Proteomics **45**: 51–88.

R. A. Kahn, L. Volpicelli-Daley, B. Bowzard, P. Shrivastava-Ranjan *et al.* (2005). "Arf family GTPases: roles in membrane traffic and microtubule dynamics." Biochem Soc Trans. **33**(6): 1269-1272.

A. Kahvejian, Y. V. Svitkin, R. Sukarieh, M. N. M'Boutchou *et al.* (2005). "Mammalian poly(A)-binding protein is a eukaryotic translation initiation factor, which acts via multiple mechanisms." Genes Dev. **19**(1): 104–113.

G. Karp, K. Beginnen, S. Vogel, S. Kuhlmann-Krieg, Eds. (2005). Molekulare Zellbiologie (in French). Springer.

K. Katoh, D. M. Standley (2013). "MAFFT multiple sequence alignment software version 7: improvements in performance and usability." Mol Biol Evolution **30**(4): 772-800.

J. K. Kelly, M. A. F. Noor (1996). "Speciation by reinforcement: A model derived from studies of *Drosophila*." Genetics **143**: 1485-1497.

C. Kimchi-Sarfaty, J. M. Oh, I.-W. Kim, Z. E. Sauna *et al.* (2007). "A "Silent" Polymorphism in the MDR1 Gene Changes Substrate Specificity." Science **315**(5811): 525-528.

T. Kirchhausen (2000). "Three ways to make a vesicle. ." Nat Rev Mol Cell Biol **1**: 187–198.

M. Kirkpatrick, V. Ravigne (2002). "Speciation by natural and sexual selection: models and experiments." Am. Nat. **159**(22-35).

M. Kirkpatrick, M. R. Servedio (1999). "The reinforcement of mating preferences " Proc. Natl. Acad. Sci. **94**: 1282-1286.

F. Krueger (2012). "Trim Galore!". from http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/.

- K. Lamnissou, M. Loukas, E. Zouros (1996). "Incompatibilities between Y chromosome and autosomes are responsible for male hybrid sterility in crosses between *Drosophila virilis* and *Drosophila texana*." *Heredity* **26**(603-9).
- R. Lande (1981). "Models of speciation by sexual selection on polygenic traits." *Proc Natl Acad Sci USA* **78**: 3721-3725.
- R. Lande (1982). "Rapid origin of sexual isolation and character divergence in a cline." *Evolution* **36**: 213-223.
- M. Lechner, S. Findeiß, L. Steiner, M. Marz *et al.* (2011). "Proteinortho: Detection of (Co-)orthologs in large-scale analysis." *BMC Bioinformatics* **12**(124).
- H. C. Lee, D. L. Garbers (1986). "Modulation of the voltage-sensitive Na⁺/H⁺ exchange in sea urchin spermatozoa through membrane potential changes induced by the egg peptide speract. ." *J Biol Chem* **261**: 16026–16032.
- G. Lenoir, P. Williamson, J. C. Holthuis (2007). "On the origin of lipid asymmetry: the flip side of ion transport." *Curr Opin Chem Biol* **11**(6): 654–661. .
- H. A. Lessios (2007). "Reproductive isolation between species of sea urchins." *BULLETIN OF MARINE SCIENCE* **81**(2): 191-208.
- W. H. Li (1993). "Unbiased estimation of the rates of synonymous and nonsynonymous substitution." *J Mol Evol* **36**(1): 96-99.
- W. H. Li, C. I. Wu, C. C. Luo (1985). "A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes." *Mol Biol Evolution* **2**(2): 150-174.
- L. W. Liou, T. D. Price (1994). "Speciation by reinforcement of premating isolation." *Evolution* **48**: 1451-1459.
- A. P. Lipinska, S. D'hondt, E. J. Van Damme, O. D. Clerch (2013). "Uncovering the genetic basis for early isogamete differentiation: a case study of *Ectocarpus siliculosus*." *BMC Genomics* **14**(909).
- R. L. Lopez-Marques, L. R. Poulsen, S. Hanisch, K. Meffert *et al.* (2010). "Intracellular targeting signals and lipid specificity determinants of the ALA/ALIS P4-ATPase complex reside in the catalytic ALA alpha-subunit." *Mol Biol Cell* **21**(5): 791–801.

- M. E. Maan, O. Seehausen (2011). "Ecology, sexual selection and speciation." Ecol. Lett. **14**: 591-602.
- I. Maier (1997). "The fine structure of the male gamete of *Ectocarpus siliculosus* (Ectocarpales, Phaeophyceae). 2. The flagellar apparatus. ." Eur J Phycol(32): :255–266.
- J. Malsam, T. H. Söllner (2011). "Organization of SNAREs within the Golgi stack." Cold Spring Harbor perspectives in biology **3**(10).
- E. Martin, R. Hine (2008). *A Dictionary of Biology*. O. U. Press.
- M. J. F. Martins, C. F. Mota, G. A. Pearson (2013). "Sex-biased gene expression in the brown alga *Fucus vesiculosus*." BMC Genomics **14**(294).
- A. C. Mathieson, C. J. Dawes, A. Wallace, A. S. Klein (2006). "Distribution, morphology, and genetic affinities of dwarf embedded *Fucus* populations from the northwest Atlantic Ocean." Bot. Mar. **49**: 283–303.
- E. C. Metz, S. R. Palumbi (1996). "Positive selection and sequence rearrangements generate extensive polymorphism in the gamete recognition protein bindin." Mol Biol Evolution **13**: 397-406.
- E. C. Metz, R. Robles-Sisisaka, V. D. Vacquier (1998). "Nonsynonymous substitutions in abalone sperm fertilization genes exceeds substitution in introns and mitochondrial DNA." Proc Natl Acad Sci USA **95**: 10676-10681.
- P. Michaely, D. R. Tomchick, M. Machius, R. G. Anderson (2002). "Crystal structure of a 12 ANK repeat stack from human ANK1." EMBO J. **21**(23): 6387–6396.
- H. Mohr, P. Schopfer, Eds. (2012). Plant Physiology, Springer Science & Business Media.
- P. A. S. Mourão (2007). "A carbohydrate-based mechanism of species recognition in sea urchin fertilization. ." Braz J Med Biol Res **40**: 5–17.
- A. M. Mulichak, C. P. Bonin, W. D. Reiter, R. M. Garavito (2002). "Structure of the MUR1 GDP-mannose 4,6-dehydratase from *Arabidopsis thaliana*: implications for ligand binding and specificity." Biochemistry **41**(52).
- D. G. Muller, C. E. Schmidt (1988). "Qualitative and quantitative determination of pheromone secretion in female gametes of *Ectocarpus siliculosus*." Biol Chem H-S **396**: 647-653.

- M. Negishi, L. G. Pedersen, E. Petrotchenko, S. Shevtsov *et al.* (2001). "Structure and function of sulfotransferases." Arch Biochem Biophys **390**(2): 149-157.
- M. Nei, T. Gojobori (1986). "Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions." Mol Biol Evolution **3**(5): 418-426.
- A. T. Neill, V. D. Vacquier (2004). "Ligands and receptors mediating signal transduction in sea urchin spermatozoa." Reproduction **127**: 141-149.
- M. A. F. Noor, K. L. Grams, L. A. Bertucci, Y. Almendarez *et al.* (2001). "The genetics of reproductive isolation and the potential for gene exchange between *Drosophila pseudoobscura* and *D. persimilis* via backcross hybrid males." Evolution **55**(3): 512-521.
- P. Nosil, D. Schluter (2011). "The genes underlying the process of speciation." Trends in ecology & evolution **26**(4): 160-167.
- C. Notredame, D. G. Higgins, J. Heringa (2000). "T-coffee: a novel method for fast and accurate multiple sequence alignment." Journal of Molecular Biology **302**(1): 205–217.
- P. Nyvall, E. Corre, C. Boisset, T. Barbeyron *et al.* (2003). "Characterization of Mannuronan C-5-Epimerase Genes from the Brown Alga *Laminaria digitata*." Plant Physiol **133**(2): 726–735.
- K. Ohta, C. Sato, T. Matsuda, M. Toriyama *et al.* (2000). "Co-localization of receptor and transducer proteins in the glycosphingolipid-enriched, low density, detergent-insoluble membrane fraction of sea urchin sperm. ." Glycoconjugate J **17**: 205–214.
- H. A. Orr, S. Irving (2001). "Complex epistasis and the genetic basis of hybrid sterility in the *Drosophila pseudoobscura* Bogota-USA hybridization." Genetics **158**(3): 1089-1100.
- A. C. Pantazidis, V. K. Galanopoulos, E. Zouros (1993). "An autosomal factor from *Drosophila arizonae* restores spermatogenesis in *Drosophila mojavensis* males carrying the *D.arizonae* Y chromosome." Genetics **134**: 309-318.
- G. A. Parker, L. Partridge (1998). "Sexual conflict and speciation." Philos. Trans. R. Soc. **353**: 261-274.
- B. M. Pearse (1976). "Clathrin: a unique protein associated with intracellular transfer of membrane by coated vesicles." Proceedings of the National Academy of Sciences of the United States of America **73**(4): 1255–1259.
- G. Pearson, A. Lago-lesten, M. Valente, E. Serra'o (2006a). "Simple and rapid RNA extraction from freeze-dried tissue of brown algae and seagrasses." Journal of Phycology **41**: 97-104.

- G. A. Pearson, S. H. Brawley (1996). "Reproductive ecology of *Fucus distichus* (Phaeophyceae): an intertidal alga with successful external fertilization." Mar Ecol Prog Ser **143**: 211-223.
- G. A. Pearson, E. A. Serrão (2006b). "Revisiting synchronous gamete release by furoid algae in the intertidal zone: fertilization success and beyond?" Integr Comp Biol **46**: 587-597.
- G. A. Pearson, E. A. Serrão, S. H. Brawley (1998). "Control of gamete release in furoid algae: sensing hydrodynamic conditions via carbon acquisition." 79(1725-1739).
- J. C. Perry, P. W. Harrison, J. E. Mank (2014). "The ontogeny and evolution of sex-biased gene expression in *Drosophila melanogaster*. ." Mol Biol Evol. **31**(5): 1206-1219.
- D. Posada (2003). "Using Modeltest and PAUP* to select a model of nucleotide substitution. ." Current Protocols in Bioinformatics **6,5**: 1-6.
- M. Ridley (2004). Evolution, Wiley-Blackwell.
- J. Sambrook, D. Russell, Eds. (2001). Molecular Cloning - A Laboratory Manual, CSHL Press.
- C. E. Schmid (1993). "Cell-cell-recognition during fertilization in *Ectocarpus siliculosus* (Phaeophyceae)." Hydrobiologia **260/261**(437-443).
- M. E. A. Schoenwaelder (2000). "Phenolic compounds in the embryo development of several Northern Hemisphere Furoids. ." Plant Biol **2**: 24–33.
- G. H. Schueller, A. F. Peters (1994). "Arrival of *Fucus evanescens* (Phaeophyceae) in Kiel Bight (Western Baltic)." Bot. Mar. **37**(471-477).
- T. Serafini, L. Orci, M. Amherdt, M. Brunner *et al.* (1991). "ADP-ribosylation factor is a subunit of the coat of Golgi-derived COP-coated vesicles: a novel role for a GTP-binding protein." Cell **67**(2): 239–253. .
- E. A. Serrao, L. A. Alice, S. H. Brawley (1999). "Evolution of the Fucaceae (phaeophyceae) inferred from nrDNA-ITS1." Phycol. **35**: 382–394.
- M. R. Servedio (2000). "Reinforcement and the genetics of nonrandom mating." Evolution **54**: 21-29.
- M. R. Servedio, M. Kirkpatrick (1997). "The effects of gene flow on reinforcement." Evolution **55**(1909-20).

- M. R. Servedio, M. Noor (2003). "The role of reinforcement in speciation: theory and data." Annu. Rev. Ecol., Evol. Syst. **34**(339-364).
- E. Sharma, A. Künstner, B. A. Fraser, G. Zipprich *et al.* (2014). "Transcriptome assemblies for studying sex-biased gene expression in the guppy, *Poecilia reticulata*." BMC Genomics **15**(400): 1471-2164.
- B. S. Shiba K, Inoue T, Yoshida M (2008). "Ca²⁺ bursts occur around a local minimal concentration of attractant and trigger sperm chemotactic response. ." PNAS **105**: 19312–19317.
- S. H. Shiu, W. M. Karlowski, R. Pan, Y. H. Tzeng *et al.* (2004). "Comparative analysis of the receptor-like kinase family in Arabidopsis and rice." Mol. Biol. Evol. **21**(2290-2298).
- E. J. Sideman, A. C. Mathieson (1983). " Ecological and genealogical distinctions of a high intertidal, dwarf form of *Fucus distichus*." J. Exp. Mar. Biol. Ecol. **72**: 171–188.
- P. J. Sillers, A. Forer (1985). "Ca²⁺ in fertilization and mitosis: the phosphatidylinositol cycle in sea urchin gametes and zygotes is involved in control of fertilization and mitosis. ." Cell Biol Int Rep **9**: 275–282.
- U. Sorhannus, S. L. Kosakovsky Pond (2006). "Evidence for positive selection on a sexual reproduction gene in the diatom genus *Thalassiosira* (Bacillariophyta)." J Mol Evol. **63**(2): 231-239. .
- M. Spehr, K. Schwane, J. A. Riffell, J. Barbour *et al.* (2004). "Particulate adenylyl cyclase plays a key role in human sperm olfactory receptor-mediated chemotaxis." J Biol Chem **279**: 40194–40203.
- D. L. Stokes, N. M. Green (2003). "Structure and function of the calcium pump." Annu Rev Biophys Biomol Struct **32**: 445–468.
- T. Strünker, I. Weyand, W. Böningk, Q. Van *et al.* (2006). "AK⁺ –selective cGMP-gated ion channel controls chemosensation of sperm. ." Nat Cell Biol, **8**: 1149–1154.
- M. Suyama, D. Torrents, P. Bork (2006). "PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments." Nucleic Acids Res **34**: 609-612.
- W. Swanson, V. D. Vacquier (1998). "Concerted evolution in an egg receptor for a rapidly evolving abalone sperm protein." Science **281**: 710-712.
- W. J. Swanson, V. D. Vacquier (2002). "The rapid evolution of reproductive proteins." Nat Rev Genet. **3**(2): 137-144.

- M. Turelli, H. A. Orr (1995). "The dominance theory of Haldain's Rule " Genetics **140**: 389-402.
- T. L. Turner , M. W. Hahn, N. S. V. (2005). "Genomic Islands of Speciation in *Anopheles gambiae*." PLoS Biol. **9**.
- R. D. Vale (2003). "The molecular motor toolbox for intracellular transport " Cell **112**(4): 467–480.
- C. R. Ward, B. T. Storey, G. S. Kopf (1992). "Activation of a G protein in mouse sperm membranes by solubilized proteins of the zona pellucida, the egg's extracellular matrix. ." J Biol Chem **267**: 14061-14067.
- S. A. Wikström, T. Wachenfeldt, L. Kautsky (2002). "Establishment of the exotic species *Fucus evanescens* C. Ag. (Phaeophyceae) in Öresund, Southern Sweden." Bot. Mar. **45**: 510-517.
- K. H. Wong, Y. Jin, Z. Moqtaderi (2013). "Multiplex Illumina sequencing using DNA barcoding." Curr Protoc Mol Biol. **7**(7.11).
- H. P. Xin, X. B. Peng, J. Ning, T. T. Yan *et al.* (2011). "Expressed sequence-tag analysis of tobacco sperm cells reveals a unique transcriptional profile and selective persistence of paternal transcripts after fertilization." Sex Plant Reprod. **24**(1): 37-46.
- Z. Yang (1997). "PAML: a program package for phylogenetic analysis by maximum likelihood." Comput Appl Biosci. **13**(5): 555-556.
- Z. Yang, J. P. Bielawski (2000a). "Statistical methods for detecting molecular adaptation." Trends in ecology & evolution **15**(12): 496–503.
- Z. Yang, R. Nielsen (1999). "Estimating Synonymous and Nonsynonymous Substitution Rates Under Realistic Evolutionary Models." Molecular Biology and Evolution **17**(1): 32-43.
- Z. Yang, W. J. Swanson, V. D. Vacquier (2000b). "Maximum likelihood analysis of molecular adaptation in abalone sperm lysin reveals variable selective pressures among lineages and sites." Mol Biol Evolution **17**: 1446-1455.
- Z. Zhang, L. Jun, Y. Jun (2006a). "Computing Ka and Ks with a consideration of unequal transitional substitutions." BMC Evolutionary Biology **6**: 44.
- Z. Zhang, J. Li, X. Q. Zhao, J. Wang *et al.* (2006b). "KaKs Calculator: Calculating Ka and Ks through model selection and model averaging." Genomics Proteomics Bioinformatics **4**: 259-263.

7. Acknowledgments

I sincerely acknowledge my main supervisor Professor Galice Hourau for the exciting Fucus project, for inspiration, opportunity to try myself in the solution of unexpected challenges and for the astonishing ideas during the project and correction of versions of this manuscript. I cordially appreciate my co-supervisor Alexander Jüterbock for his guidance, support; patience and politeness during some of the stage of the project (collect of the samples, some laboratory work, processing of the data using software). For the company and enthusiastic and amazing guidance through laboratory work I want to acknowledge Martina Kopp. Thanks to faculty staff members and students for creating a supportive and positive environment.

The study was conducted at the Faculty of Bioscience and Aquaculture, University of Nordland (Nord University since 01.01.2016) and was supported by The Research Council of Norway.

I gratefully acknowledge my family and boyfriend for providing a great support during my education.

8. Supplementary materials

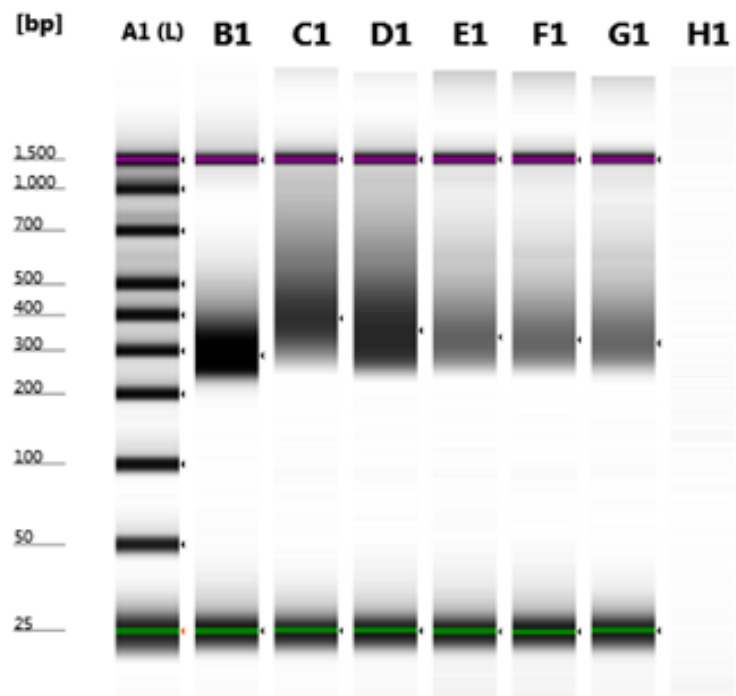
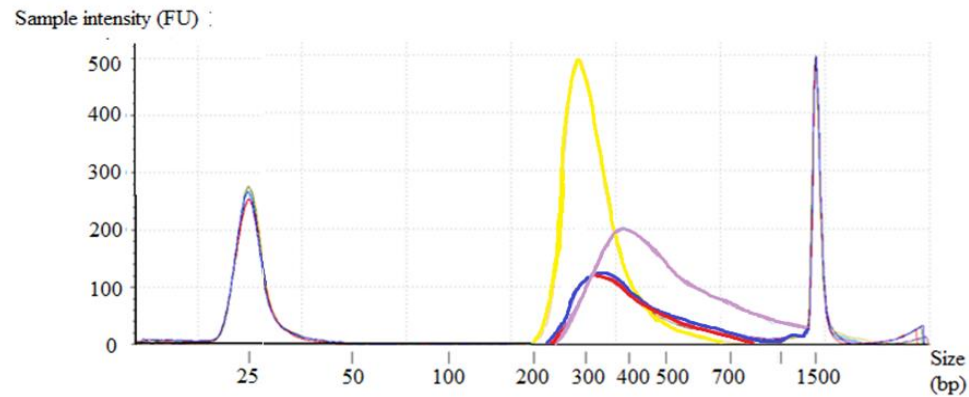


Figure 1S Gel picture from the Tape station.



Indicator	ID	Name of the sample
	E1	<i>Fucus serratus</i> male
	G1	<i>Fucus serratus</i> female
	B1	<i>Fucus distichus</i> enriched
	C1	<i>Fucus distichus</i> pool 3

Figure 2S Profile comparisons between samples. Sample intensity equals concentration. The higher peak – the more concentration of fragments with precise length contain in the sample.

It visualizes concentration of samples – the darker strand the higher the concentration. Additionally, it shows the length of fragments which contains in samples and their concentration, for example sample B contains more fragments with length 300 bp. A1 – ladder, B1 *F. distichus* pool3, C1 – *F. distichus* enriched, D1 – *Fucus distichus* pool 10nM, E1 – *F. serratus* male, F1 – *F. serratus* female, G1 - *F. serratus* male + female pool 10 nM, H1 – empty.

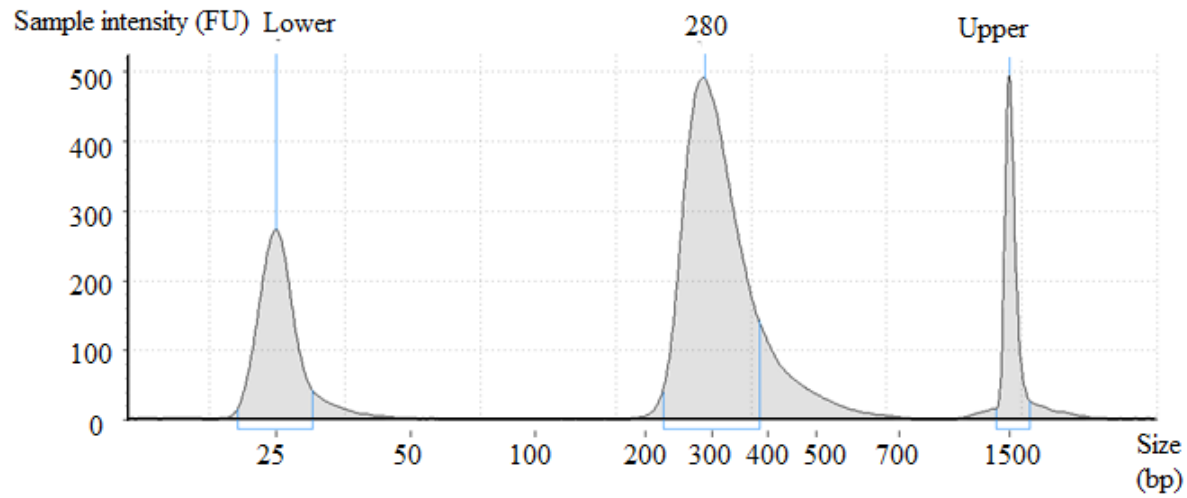


Figure 3S Intensity of *F. distichus* pool 3. Sample intensity shows concentration of sample in direct ratio; the higher Sample intensity the higher concentration of the RNA peaks. X-axis shows size of base pairs.

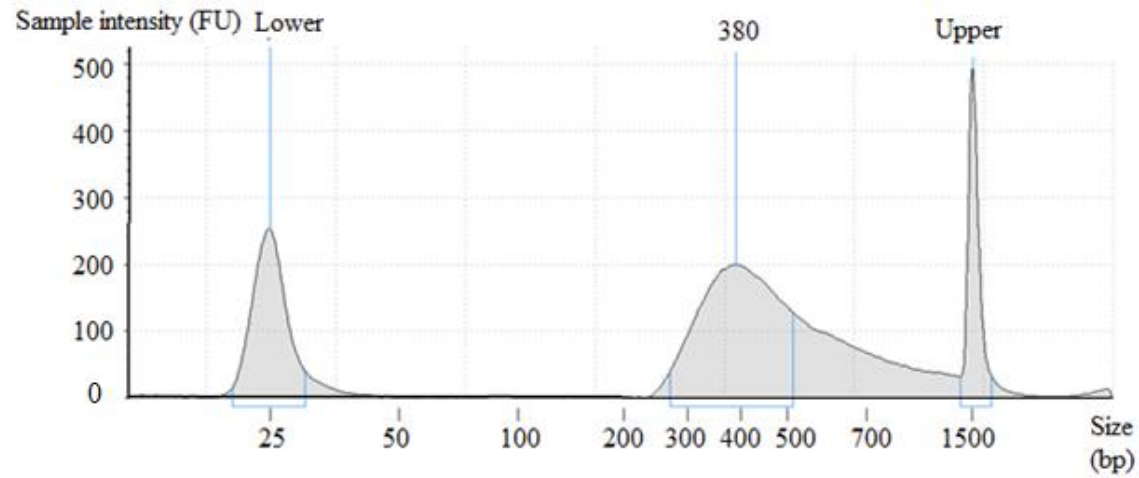


Figure 4S Intensity of *F. distichus* enriched pool. Sample intensity shows concentration of sample in direct ratio; the higher Sample intensity the higher concentration of the RNA peaks. X-axis shows size of base pairs.

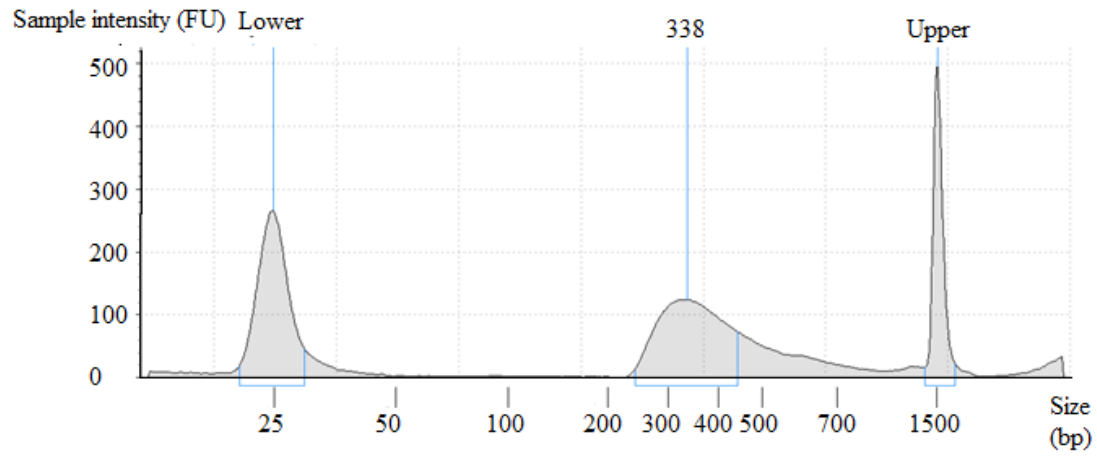


Figure 5S Intensity of *F. serratus* male pool. Sample intensity shows concentration of sample in direct ratio; the higher Sample intensity the higher concentration of the RNA peaks. X-axis shows size of base pairs.

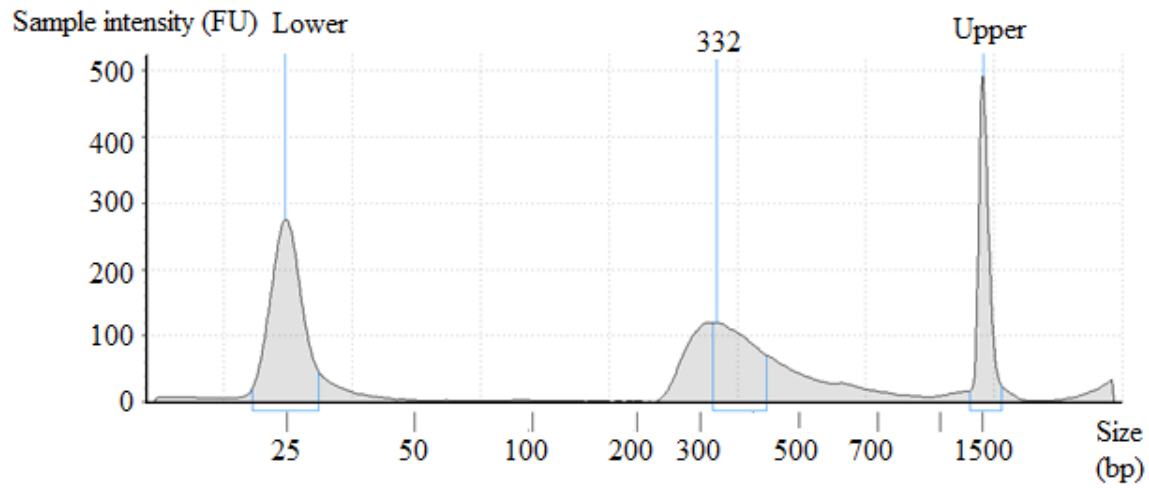


Figure 6S Intensity of *F. serratus* female pool. Sample intensity shows concentration of sample in direct ratio; the higher Sample intensity the higher concentration of the RNA peaks. X-axis shows size of base pairs.

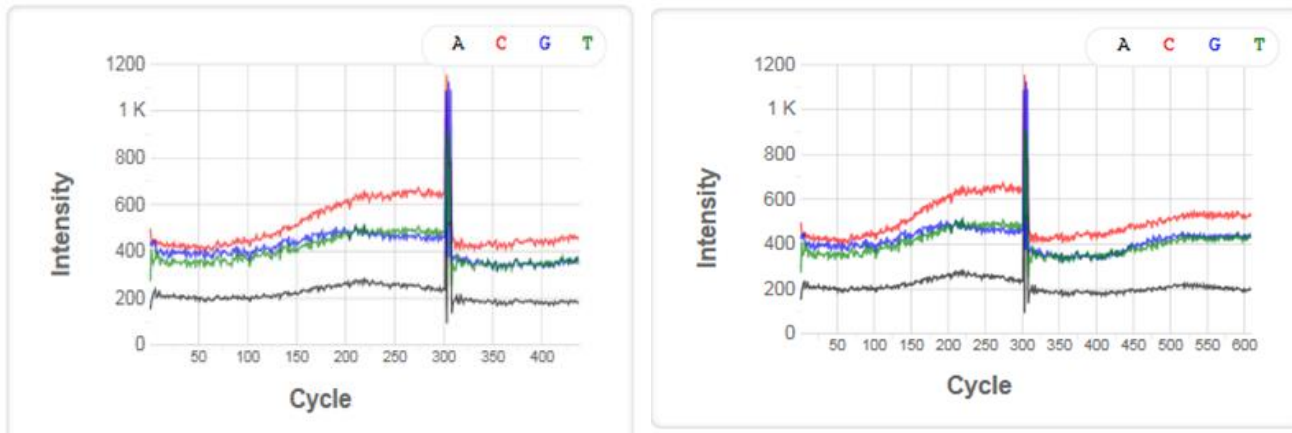


Figure 7S Sequencing intensity of *F. distichus* poo3 and enriched one.

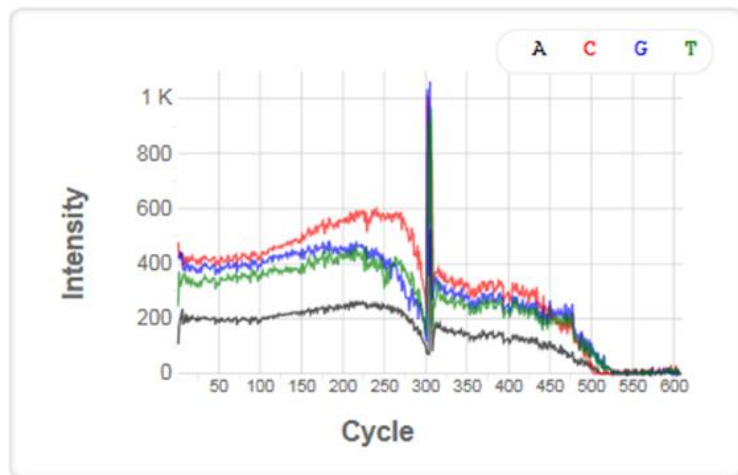


Figure 8S Sequencing intensity of *F. serratus* male and female. There is a drop of sequence intensity after 300 cycle.

Table 1S Example of code for trimming *F. serratus* female:

```
Trim_galore
--phred33
--fastqc
--stringency 5
--paired
--clip_R1 10
--clip_R2 10
--length 40
-a
AATGATACGGCGACCACCGAGATCTACACTCTTTCC
CTACACGACGCTCTTCCGATCT
-
a2GATCGGAAGAGCACACGTCTGAACTCCAGTCAC-
CTTGTA -ATCTCGTATGCCGTCTTCTGCTTG
-o /output directory path/
-q 20
/filepath/filename.fq
```

For *Fucus serratus* IonTorrent I use another code, because sequence data isn't paired:

```
trim_galore
--phred33
```

```
--fastqc
--dont_gzip
--stringency 5
-a
CCACTACGCCTCCGCTTTCCTCTCTATGGGCAGTCG
GTGAA
-o /output directory path/
-q 20
/filepath/filename.fq
I repeat trimming twice with F. serratus IonTorrent because
there are 2 adapters and there is no opportunity to cut them
out in one run:
trim_galore
--phred33
--fastqc
--dont_gzip
--stringency 5
-a CCATCTCATCCCTGCGTGTCTCCGACTCAG
--length 40
36
-o /output directory path/
-q 20 /filepath/filename.fq
```

Table 2S Orthologous genes between reproductive *F. serratus* male vs. non-reproductive *F. serratus* IonTorrent

№	Gene ID in genome	Gene name	Length (bp)
1.	TR1123 c0_g2_i1 m,3945	conserved unknown protein [<i>Ectocarpus siliculosus</i>]	54
2.	TR1688 c0_g5_i1 m,5799	imm-downregulated 8 [<i>Ectocarpus siliculosus</i>]	67
3.	TR1744 c0_g4_i2 m,5987	transcetolase [<i>Ectocarpus siliculosus</i>]	86
4.	TR600 c0_g2_i1 m,2116	hypothetical protein [<i>Bacillus azotoformans</i>]	79
5.	TR1487 c0_g3_i1 m,5135	poly(A) binding protein [<i>Ectocarpus siliculosus</i>]	121
6.	TR616 c0_g1_i1 m,2185	hypothetical protein, conserved [<i>Eimeria mitis</i>]	120
7.	TR393 c0_g1_i1 m,1325	heat shock protein 20 [<i>Fucus serratus</i>]	111
8.	TR1044 c0_g1_i1 m,3664	Light harvesting complex protein [<i>Ectocarpus siliculosus</i>]	100
9.	TR1044 c0_g2_i1 m,3671	Light harvesting complex protein [<i>Ectocarpus siliculosus</i>]	58
10.	TR679 c0_g1_i1 m,2401	Predicted uncharacterized protein LOC101741608 [<i>Bombyx mori</i>]	72
11.	TR1177 c0_g1_i1 m,4118	conserved unknown protein [<i>Ectocarpus siliculosus</i>]	69
12.	TR1744 c0_g4_i2 m,5987	transketolase [<i>Ectocarpus siliculosus</i>]	96

Table 3S Orthologous genes between reproductive *F. serratus* female vs. non-reproductive *F. serratus* IonTorrent

№	Gene ID in genome	Gene name	Length (bp)
1.	TR1549 c0_g9_i1 m,5289	hypothetical protein JCGZ_18639 [<i>Jatropha curcas</i>]	58
2.	TR3181 c0_g2_i1 m,11853	hypothetical protein, conserved [<i>Eimeria mitis</i>]	72
3.	TR3181 c0_g6_i1 m,11865	hypothetical protein, conserved [<i>Eimeria mitis</i>]	69
4.	TR1013 c0_g1_i1 m,3525	hypothetical protein, conserved [<i>Eimeria mitis</i>]	96
5.	TR1013 c0_g2_i1 m,3526	hypothetical protein, conserved [<i>Eimeria mitis</i>]	120
6.	TR1013 c0_g2_i2 m,3530	hypothetical protein, conserved [<i>Bacillus azotoformans</i>]	113
7.	TR1013 c0_g3_i1 m,3535	EsV-1-163 [<i>Ectocarpus siliculosus</i>]	106
8.	TR5318 c0_g1_i1 m,19025	DEAD box helicase [<i>Ectocarpus siliculosus</i>]	56
9.	TR276 c227_g1_i1 m,9924	ribosomal protein L34 [<i>Ectocarpus siliculosus</i>]	78
10.	TR26 c406_g1_i1 m,10229	conserved unknown protein [<i>Ectocarpus siliculosus</i>]	98
11.	TR3276 c0_g1_i1 m,12161	hypothetical protein, conserved [<i>Eimeria mitis</i>]	100
12.	TR5776 c0_g1_i1 m,20610	imm downregulated 8 [<i>Ectocarpus siliculosus</i>]	103
13.	TR4901 c0_g1_i1 m,17570	n/a [<i>Ectocarpus siliculosus</i>]	100
14.	TR2422 c0_g1_i1 m,8292	imm upregulated 3 [<i>Ectocarpus siliculosus</i>]	121
15.	TR2726 c69_g5_i1 m,9456	conserved unknown protein [<i>Ectocarpus siliculosus</i>]	67
16.	TR226 c210_g1_i1 m,9894	conserved unknown protein [<i>Ectocarpus siliculosus</i>]	69
17.	TR6010 c0_g3_i1 m,21452	metallothionein [<i>Fucus vesiculosus</i>]	86
18.	TR6010 c0_g4_i1 m,21455	metallothionein [<i>Fucus vesiculosus</i>]	59
19.	TR6010 c0_g5_i1 m,21457	metallothionein [<i>Fucus vesiculosus</i>]	105
20.	TR3086 c0_g1_i1 m,11537	Ribosomal protein S27 [<i>Ectocarpus siliculosus</i>]	132

21.	TR6357 c0_g7_i1 m,22788	actin [<i>Saccharina japonica</i>]	80
22.	TR1169 c0_g2_i1 m,4023	40S ribosomal protein S4 [<i>Ectocarpus siliculosus</i>]	78
23.	T2726 c260_g1_i1 m,9984	hypothetical protein THAOC_13056 [<i>Thalassiosira oceanica</i>]	89
24.	2726 c359_g1_i1 m,10152	EEF1A2, eukaryotic translation elongation factor 1 alpha [<i>Ectocarpus siliculosus</i>]	120
25.	TR6081 c0_g1_i1 m,21720	Heat shock protein 90 [<i>Ectocarpus siliculosus</i>]	113
26.	TR4608 c0_g1_i1 m,16597	conserved unknown protein [<i>Ectocarpus siliculosus</i>]	106
27.	TR4709 c0_g1_i1 m,16951	conserved unknown protein [<i>Ectocarpus siliculosus</i>]	87
28.	TR4071 c0_g5_i1 m,14870	hypothetical protein [<i>Bacillus azotoformans</i>]	80
29.	TR7070 c0_g4_i1 m,25065	conserved unknown protein [<i>Ectocarpus siliculosus</i>]	127
30.	TR3559 c0_g2_i2 m,13164	conserved unknown protein [<i>Ectocarpus siliculosus</i>]	63
31.	TR1549 c0_g1_i1 m,5264	Hypothetical protein GL50803_6525 [<i>Giardia lamblia ATCC 50803</i>]	64
32.	TR1549 c0_g5_i1 m,5276	conserved unknown protein [<i>Ectocarpus siliculosus</i>]	79
33.	TR382 c0_g2_i1 m,1455	resolvase domain containing protein [<i>Shewanella decolorationis S12</i>]	78
34.	TR2726 c21_g1_i1 m,9334	conserved unknown protein [<i>Ectocarpus siliculosus</i>]	64
35.	TR2726 c80_g2_i3 m,9563	hypothetical protein DAPPUDRAFT_37179 [<i>Daphnia pulex</i>]	97
36.	TR3026 c0_g2_i1 m,11192	hypothetical protein THAOC_21443 [<i>Thalassiosira oceanica</i>]	121
37.	TR2726 c37_g1_i1 m,9377	ribosomal protein L24 [<i>Ectocarpus siliculosus</i>]	111
38.	TR3998 c0_g1_i1 m,14666	heat shock protein 20 [<i>Fucus serratus</i>]	116

39.	TR3998 c0_g2_i1 m,14669	heat shock protein 20 [<i>Fucus serratus</i>]	108
40.	TR3111 c0_g3_i2 m,11614	hypothetical protein, conserved [<i>Eimeria mitis</i>]	121
41.	TR3111 c0_g5_i2 m,11625	hypothetical protein, conserved [<i>Eimeria mitis</i>]	103
42.	TR3111 c0_g5_i1 m,11622	hypothetical protein, conserved [<i>Eimeria mitis</i>]	107
43.	TR6157 c0_g3_i1 m,22032	conserved unknown protein [<i>Ectocarpus siliculosus</i>]	112
44.	TR5218 c1_g5_i1 m,18677	light harvesting protein lhcf4 [<i>Saccharina latissima</i>]	121
45.	TR5218 c1_g6_i1 m,18681	Light harvesting complex protein [<i>Ectocarpus siliculosus</i>]	70
46.	TR5776 c0_g1_i1 m,20610	imm downregulated 8 [<i>Ectocarpus siliculosus</i>]	73
47.	TR4340 c0_g1_i1 m,15736	heat shock protein 8 [<i>Ectocarpus siliculosus</i>]	68
48.	TR2462 c0_g2_i1 m,8426	hypothetical protein [<i>Shewanella decolorationis</i>]	89
49.	TR916 c0_g1_i2 m,3227	Predicted: antigen presenting glycoprotein CD 1 dl-like isoform X3 [<i>Ficedula albicollis</i>]	70
50.	TR6332 c0_g3_i1 m,22687	hypothetical protein [<i>Shewanella decolorationis</i>]	73
51.	TR6332 c0_g3_i1 m,22687	conserved unknown protein [<i>Ectocarpus siliculosus</i>]	79

Table 4S Orthologous genes between reproductive *F. distichus* mixed vs. non-reproductive *F. serratus* IonTorrent. With bold font highlighted reproduction genes.

№	Gene ID in genome	Gene name	Length (bp)
1.	TR12567 c1_g1_i1 m,38406	14-3-3-like protein [<i>Ectocarpus siliculosus</i>]	80
2.	TR8138 c0_g1_i1 m,25277	40S ribosomal protein S3 [<i>Aphanomyces invadans</i>]	85
3.	TR38397 c0_g3_i1 m,116568	40S ribosomal protein S4 [<i>Ectocarpus siliculosus</i>]	152
4.	TR34804 c0_g1_i1 m,105385	40S ribosomal protein-like protein [<i>Thalassiosira pseudonana</i> CCMP1335]	97
5.	TR9649 c1_g2_i1 m,29691	60S ribosomal protein L10A [<i>Ectocarpus siliculosus</i>]	159
6.	TR23025 c0_g1_i1 m,69974	60S ribosomal protein L2 [<i>Clavispora lusitaniae</i> ATCC 42720]	96
7.	TR44976 c0_g1_i1 m,136358	actin-1, putative [<i>Acanthamoeba castellanii</i> str, Neff]	135
8.	TR20855 c1_g3_i1 m,63277	Cell wall-associated hydrolase [<i>Blattabacterium</i> sp, (<i>Periplaneta americana</i>) str, BPLAN]	66
9.	TR34670 c0_g1_i1 m,104975	Cell wall-associated hydrolase [<i>Roseobacter</i> sp, AzwK-3b]	127
10.	TR13427 c0_g3_i1 m,41022	Conserved unknown protein [<i>Ectocarpus siliculosus</i>]	186

11.	TR13427 c0_g4_i1 m,41026	Conserved unknown protein [<i>Ectocarpus siliculosus</i>]	187
12.	TR42860 c0_g1_i1 m,130184	conserved unknown protein [<i>Ectocarpus siliculosus</i>]	128
13.	TR5133 c0_g1_i1 m,15854	conserved unknown protein [<i>Ectocarpus siliculosus</i>]	56
14.	TR21621 c1_g4_i1 m,65745	conserved unknown protein [<i>Ectocarpus siliculosus</i>]	103
15.	TR35809 c0_g1_i1 m,108486	conserved unknown protein [<i>Ectocarpus siliculosus</i>]	114
16.	TR222 c1_g1_i1 m,718	conserved unknown protein [<i>Ectocarpus siliculosus</i>]	207
17.	TR31139 c0_g1_i2 m,94571	conserved unknown protein [<i>Ectocarpus siliculosus</i>]	152
18.	TR3043 c0_g1_i1 m,9316	conserved unknown protein [<i>Ectocarpus siliculosus</i>]	50
19.	TR11851 c0_g1_i1 m,36344	conserved unknown protein [<i>Ectocarpus siliculosus</i>]	102
20.	TR11851 c0_g1_i1 m,36344	conserved unknown protein [<i>Ectocarpus siliculosus</i>]	102
21.	TR34464 c0_g6_i1 m,104361	conserved unknown protein [<i>Ectocarpus siliculosus</i>]	73
22.	TR15388 c0_g8_i1 m,46831	DEAD box helicase [<i>Ectocarpus siliculosus</i>]	81
23.	TR17454 c0_g2_i2 m,53037	DEAD box helicase [<i>Ectocarpus siliculosus</i>]	138
24.	TR13453 c0_g3_i2 m,41120	EsV-1-163 [<i>Ectocarpus siliculosus</i>]	160

25.	TR5638 c0_g1_i1 m,17445	glutathionyl-hydroquinone reductase YqjG [<i>Pseudomonas syringae</i>]	77
26.	TR15379 c0_g2_i1 m,46775	heat shock protein 90 [<i>Bicyclus anynana</i>]	80
27.	TR29172 c0_g1_i1 m,88154	heat shock protein 90 Hsp90 [<i>Helicosporidium</i> sp, ATCC 50920]	96
28.	TR6601 c0_g1_i1 m,20421	heat shock protein 90kDa [<i>Heliconius cydno chioneus</i>]	144
29.	TR14570 c0_g1_i1 m,44464	heat shock protein hsp82 [<i>Oryza sativa</i>]	106
30.	TR40319 c0_g1_i1 m,122448	hypothetical protein [<i>Aeromonas salmonicida</i>]	69
31.	TR40319 c0_g1_i1 m,122448	hypothetical protein [<i>Aeromonas salmonicida</i>]	69
32.	TR20212 c0_g1_i1 m,61055	hypothetical protein [<i>Bacillus licheniformis</i>]	63
33.	TR33901 c8_g25_i1 m,102673	hypothetical protein [<i>Shewanella decolorationis</i>]	130
34.	TR33901 c8_g68_i1 m,102787	hypothetical protein [<i>Shewanella decolorationis</i>]	132
35.	TR833 c0_g1_i1 m,2499	hypothetical protein [uncultured <i>Verrucomicrobiales</i> bacterium HF0200_39L05]	98
36.	TR20212 c0_g2_i1 m,61056	hypothetical protein [<i>Vibrio vulnificus</i>]	63

37.	TR4298 c0_g1_i1 m,13291	hypothetical protein AURANDRAFT_32917 [<i>Aureococcus anophagefferens</i>]	168
38.	TR4298 c0_g6_i1 m,13306	hypothetical protein AURANDRAFT_32917 [<i>Aureococcus anophagefferens</i>]	127
39.	TR36048 c0_g1_i1 m,109526	hypothetical protein FSCG_01459 [<i>Fusobacterium nucleatum subsp, vincentii</i> 4_1_13]	79
40.	TR10939 c0_g1_i1 m,33465	hypothetical protein GUITHDRAFT_84932 [<i>Guillardia theta</i> CCMP2712]	68
41.	TR18342 c0_g1_i1 m,55551	hypothetical protein SAMD00019534_123350 [<i>Acytostelium subglobosum</i> LB1]	162
42.	TR45491 c0_g3_i2 m,137953	hypothetical protein SHD_0204 [<i>Shewanella decolorationis</i> S12]	107
43.	TR20740 c6_g36_i1 m,62695	imm upregulated 3 [<i>Ectocarpus siliculosus</i>]	116
44.	TR20740 c6_g66_i1 m,62788	imm upregulated 3 [<i>Ectocarpus siliculosus</i>]	108
45.	TR20740 c6_g68_i1 m,62791	imm upregulated 3 [<i>Ectocarpus siliculosus</i>]	121
46.	TR20740 c6_g74_i1 m,62808	imm upregulated 3 [<i>Ectocarpus siliculosus</i>]	103
47.	TR20740 c6_g85_i1 m,62841	imm upregulated 3 [<i>Ectocarpus siliculosus</i>]	107

48.	TR20740 c6_g87_i1 m,62847	imm upregulated 3 [<i>Ectocarpus siliculosus</i>]	112
49.	TR20740 c6_g91_i1 m,62858	imm upregulated 3 [<i>Ectocarpus siliculosus</i>]	121
50.	TR20740 c6_g26_i1 m,62667	imm upregulated 3 [<i>Ectocarpus siliculosus</i>]	70
51.	TR39375 c0_g1_i1 m,119471	metallothionein [<i>Fucus vesiculosus</i>]	101
52.	TR3703 c0_g1_i1 m,11580	phosphonate ABC transporter ATPase [<i>Salmonella enterica subsp, enterica serovar Agona str, SH11G1113</i>]	100
53.	TR40626 c0_g1_i1 m,123474	putative elongation factor 1-alpha [<i>Diaphorina citri</i>]	85
54.	TR40762 c0_g1_i1 m,123874	QM protein [<i>Spodoptera litura</i>]	99
55.	TR35572 c0_g2_i1 m,107771	RecName: Full=Tubulin alpha-1 chain [<i>Pelvetia fastigiata</i>]	141
56.	TR31888 c0_g2_i1 m,96602	ribosomal protein L35 [<i>Ectocarpus siliculosus</i>]	68
57.	TR45802 c0_g1_i1 m,139172	Ribosomal protein S27 [<i>Ectocarpus siliculosus</i>]	89
58.	TR45491 c0_g1_i1 m,137940	RTX toxin [<i>Neisseria meningitidis</i> LNP27256]	96
59.	TR6492 c0_g1_i3 m,20097	unknown [<i>Medicago truncatula</i>]	136
60.	TR33456 c0_g1_i1 m,101391	unnamed protein product [<i>Chondrus crispus</i>]	110

61.	TR25671 c2_g157_i1 m,78106	vanadium-dependent bromoperoxidase 2 [<i>Ascophyllum nodosum</i>]	120
62.	TR25671 c2_g160_i1 m,78115	vanadium-dependent bromoperoxidase 2 [<i>Ascophyllum nodosum</i>]	113
63.	TR25671 c2_g162_i1 m,78119	vanadium-dependent bromoperoxidase 2 [<i>Ascophyllum nodosum</i>]	106
64.	TR25671 c2_g2_i1 m,77661	vanadium-dependent bromoperoxidase 2 [<i>Ascophyllum nodosum</i>]	87
65.	TR36498 c0_g4_i1 m,110954	vanadium-dependent bromoperoxidase 2 [<i>Ascophyllum nodosum</i>]	97
66.	TR36498 c0_g6_i1 m,110961	vanadium-dependent bromoperoxidase 2 [<i>Ascophyllum nodosum</i>]	93
67.	TR36498 c0_g8_i1 m,110968	vanadium-dependent bromoperoxidase 2 [<i>Ascophyllum nodosum</i>]	92

Table 5S Unknown genes of *F. distichus* vs. *F. serratus* male transcriptomes, which have dN/dS ratio more than 1. ID number refers to transcriptome of *F. serratus* male. Some sequences includes 2 methods of dN/dS ratio calculation, which shown in graph «Methods».

№	ID number in transcriptome	Methods	dN/dS	P-value	Protein sequence	Length (amino acids)
1	TR1111 c0_g1_i1	LWL	1,41269	0,001476	VFFFFQAEDGIRDIGVTGVQTCALPICMNASAFKISL	62
1	TR1111 c0_g1_i1	MLPB	1,49696	0,000122	QTYAMYDSPKSTVVPVSVSVDPPSM	
2	TR1448 c0_g1_i1	LWL	1,41269	0,001476	FFFSSRRRHTRYWRDWSDDVCSSDLVKLLSFYNFL	50
2	TR1448 c0_g1_i1	MLPB	1,49696	0,000122	SVVHQYLQLRSMCRS	
3	TR1525 c0_g1_i1	LWL	1,41269	0,001476	KPVKKKKIKREIKILQNLCGGTNIQLLDVVRDPQSKT	125
3	TR1525 c0_g1_i1	LWL	1,41269	0,001476	PSLVFEYINNTDFKVL YPTLTDYDIRYYIFELLKAL DYSHSNGTMHRDVKPHNV MIDHDHKQLRLIDWGLA EFYHPGREYNVRVASR	
4	TR1628 c0_g1_i1	LWL	1,41269	0,001476	FFQEEDGIRDYKVTGVQTCALPISGLIMSDNEYEGGD	103
4	TR1628 c0_g1_i1	MLPB	1,49696	0,000122	GGEGYDVFAQEDEPQMEEPTQEEADLEIMDE AEEGGEGLEADERITTKYMTKYERARILGTRAI	
5	TR1641 c0_g1_i1	LWL	1,41269	0,001476	FFFSSRRRHTRLQGDWSSDVCSSDLRTATERGVCHRR	56
5	TR1641 c0_g1_i1	MLPB	1,49696	0,000122	GSSRCDPCSWWNTPYGVLG	
6	TR933 c0_g1_i1	LWL	1,41269	0,001476	FFSSRRRHTRLQGDWSSDVCSSDLAGTRAPVGANLEL	80
6	TR933 c0_g1_i1	MLPB	1,49696	0,000122	QVDFLSLRDRQIPHLLHGARLSLCAPLLFLLLLL SCTSLRGTF	
7	TR874 c0_g1_i1	MLWL	1,17381	0,000228	FFFQAEDGIRDIGVTGVQTCALPISVIQAAIMFSRVAAT	191
7	TR874 c0_g1_i1	MLWL	1,17381	0,000228	GARAKGTL SARLATS GGGRGYRMPIIQPAFQTF HFGAPREQAVDKPAPSRATEIEANKNTPLFGENPR LQGFIDAGKLDGAKRLLPFALFLSVPALCREWYVI TEETMLMGCFFLFLGAAMDFGGEGIGKYFDDQ TAAIKNAQT TAEDNLIQ	

Table 6S Identified reproductive genes of *F. distichus* vs. *F. serratus* female transcriptomes, which have dN/dS ratio > 1. ID number refers to transcriptome of *F. serratus* male. Some sequences includes two methods of dN/dS ratio calculation. Genes refer to the processes of reproduction highlight with bold font.

	ID number in transcriptome	Method	dN/dS	P-value	Group/Name of gene	Length (amino acids)
					Metabolical genes (68,5 %)	
1.	TR3502 c0_g1_i1	GY-HKY	1,26071	0,00366	26S proteasome beta type 7 subunit	138
	TR3502 c0_g1_i1	MA	1,24944	0,00358		
	TR3502 c0_g1_i1	YN	2,33065	0,01806		
	TR3502 c0_g1_i1	MYN	2,33065	0,01806		
	TR3502 c0_g1_i1	MS	1,23321	0,00998		
2.	TR6537 c0_g2_i1	MLWL	1,23638	0,00585	26S proteasome regulatory subunit	71
3.	TR3838 c0_g2_i1	LPB	1,04853	0,03516	Aconitatehydratase	108
4.	TR6898 c0_g2_i1	LPB	1,04853	0,03516	acyl carrier protein	98
5.	TR3773 c0_g1_i1	MLWL	1,43407	0,00325	alcoholdehydrogenase	316
	TR3840 c0_g2_i1	LPB	1,04853	0,03516		
6.	TR2525 c0_g2_i1	LPB	1,04853	0,03516	Aminoalcoholphosphotransferase	73
7.	TR2740 c0_g2_i1	LPB	1,04853	0,03516	Argonaute 1	83
8.	TR5645 c0_g2_i1	LPB	1,04853	0,03516	aspartateaminotransferase	166
9.	TR1163 c0_g2_i1	LPB	1,04853	0,03516	Beta-galactosidase, family GH2	126
10.	TR3344 c0_g2_i1	LPB	1,04853	0,03516	biotinsynthase	76
11.	TR382 c0_g2_i1	LPB	1,04853	0,03516	C2H2 type zinc finger domain-containing protein	94
12.	TR6346 c0_g2_i1	LPB	1,04853	0,03516	carbonicanhydrase	117

13.	TR5765 c0_g2_i1	LPB	1,04853	0,03516	Casein kinase (serine/threonine/tyrosine protein kinase)	79
14.	TR3726 c0_g2_i1	LPB	1,04853	0,03516	co-chaperone grpE family protein	109
15.	TR1831 c0_g2_i1	LPB	1,04853	0,03516	copper/zinc superoxide dismutase	51
16.	TR1377 c0_g2_i1	LPB	1,04853	0,03516	COX4 neighbour	194
17.	TR4724 c0_g2_i1	LPB	1,04853	0,03516	Cytochrome b6-f complex iron-sulfur subunit	225
18.	TR4711 c0_g1_i1	LPB	1,33854	0,00108	Cytochrome c oxidase subunit Vb	131
19.	TR6794 c0_g2_i1	LPB	1,04853	0,03516	Cytochrome P450 67	67
20.	TR500 c0_g2_i1	LPB	1,04853	0,03516	cytosolic glutamine synthetase	99
21.	TR1980 c0_g2_i1	MLWL	1,23638	0,00585	DEAD box helicase	354
22.	TR3562 c0_g2_i1	LPB	1,04853	0,03516	DEAD-box ATP-dependent RNA helicase 15	115
23.	TR3059 c0_g2_i1	LPB	1,04853	0,03516	diguanylate cyclase	95
24.	TR6947 c0_g2_i1	LPB	1,04853	0,03516	DNA topoisomerase VI subunit B-like protein Topoisomerase 6 subunit B	117
25.	TR3083 c0_g2_i1	LPB	1,04853	0,03516	DRG2, developmentally regulated GTPase 2	121
26.	TR4552 c0_g2_i1	LWL	1,43872	0,00077	enolase	256
27.	TR5470 c0_g1_i1	LWL	1,20853	0,00061	Fe Superoxide dismutase, Destroys radicals which are normally produced within the cells and which ar	99
	TR5470 c0_g2_i1	LPB	1,04853	0,03516	Fe Superoxide dismutase, Destroys radicals which are normally produced within the cells and which ar	
28.	TR2133 c0_g2_i1	MLWL	1,23638	0,00585	FKBP-type peptidyl-prolyl cis-trans isomerase 11	113
29.	TR7067 c0_g2_i1	LPB	1,04853	0,03516	FKBP-type peptidyl-prolyl cis-trans isomerase 5	90
30.	TR1639 c0_g2_i1	LPB	1,04853	0,03516	Flavin-binding monooxygenase-like subfamily	68
31.	TR2163 c0_g1_i1	LPB	1,33854	0,00108	Gal-2,6-Sulfurylases I	581
32.	TR5067 c0_g2_i1	LPB	1,04853	0,03516	GatB/YqeY	90
33.	TR3443 c0_g2_i1	LPB	1,04853	0,03516	globin	353
34.	TR371 c0_g2_i1	LPB	1,04853	0,03516	Glutaredoxin	102
35.	TR2232 c0_g2_i1	LPB	1,04853	0,03516	glutathione S-transferase 3	94
36.	TR6162 c0_g2_i1	LPB	1,04853	0,03516	glutathione S-transferase 4	126

37.	TR5670 c0_g2_i1	LPB	1,04853	0,03516	Glycosyltransferase, family GT4	59
38.	TR4727 c0_g2_i1	LPB	1,04853	0,03516	GMP synthase	86
39.	TR6768 c0_g2_i1	LPB	1,04853	0,03516	Guanylate kinase (IC)	322
40.	TR3998 c0_g2_i1	LPB	1,04853	0,03516	heat shock protein 20	146
41.	TR2501 c0_g2_i1	LPB	1,04853	0,03516	heat shock protein 33	81
42.	TR4340 c1_g1_i1	MLWL	1,092	0,00015	Heat shock protein 70	115
	TR4420 c0_g2_i1	LPB	1,04853	0,03516	Heat shock protein 70	
43.	TR5841 c0_g2_i1	LPB	1,04853	0,03516	heat shock protein 70	113
44.	TR3127 c0_g1_i1	LPB	1,33854	0,00108	Heat shock protein 90	202
45.	TR3868 c0_g2_i1	LPB	1,04853	0,03516	high affinity phosphate transporter, probable	115
46.	TR597 c0_g2_i1	LPB	1,04853	0,03516	histone H2A (ISS)	132
47.	TR2940 c0_g2_i1	LPB	1,04853	0,03516	immdownregulated 23	139
48.	TR4407 c0_g2_i1	LPB	1,04853	0,03516	Indole-3-glycerol-phosphate synthase	90
49.	TR1637 c0_g5_i1	MYN	2,13898	0,00925	inorganicpyrophosphatase	166
50.	TR2321 c0_g2_i1	LPB	1,04853	0,03516	Inositol 2-dehydrogenase	76
51.	TR3790 c0_g2_i1	MLWL	1,23638	0,00585	L-ascorbate peroxidase	215
52.	TR319 c0_g2_i1	LPB	1,04853	0,03516	long chain acyl-coA synthetase	50
53.	TR1632 c0_g2_i1	LPB	1,04853	0,03516	Mannose-6-phosphate isomerase	186
54.	TR1471 c0_g2_i1	LPB	1,04853	0,03516	Molecular chaperones GRP78/BiP/KAR2, HSP70 superfamily	157
55.	TR6416 c0_g2_i1	LPB	1,04853	0,03516	Myosin light chain kinase (MLCK)	85
56.	TR5006 c0_g2_i1	LPB	1,04853	0,03516	N-acyl-D-amino-acid deacylase domain protein	88
57.	TR729 c0_g1_i1	LWL	1,20853	0,00061	NADH dehydrogenase (ubiquinone)	395
58.	TR6068 c0_g2_i1	LPB	1,04853	0,03516	nitrite transporter NAR1	209
59.	TR814 c0_g2_i1	LPB	1,04853	0,03516	NUOF homolog, NADH dehydrogenase (ubiquinone) subunit	60
60.	TR826 c0_g1_i1	LPB	1,33854	0,00108	OLA1, Obg-like ATPase 1 (YchF-related GTPase)	113
61.	TR1097 c0_g2_i1	LPB	1,04853	0,03516	Ornithine cyclodeaminase	68
62.	TR5868 c0_g2_i1	MLWL	1,23638	0,00585	oxidoreductase, NAD-binding, myo-inositol 2-dehydrogenase	94

63.	TR6165 c0_g2_i1	LPB	1,04853	0,03516	oxygen-evolving enhancer protein	324
64.	TR6996 c0_g2_i1	LPB	1,04853	0,03516	peptidase	94
65.	TR556 c0_g2_i1	MLWL	1,23638	0,00585	Phosphoadenylyl-sulfate reductase (thioredoxin)	320
66.	TR6344 c0_g2_i1	LPB	1,04853	0,03516	phosphoglyceratekinase	297
67.	TR6557 c0_g2_i1	LPB	1,04853	0,03516	PREDICTED: peptidyl-prolyl cis-trans isomerase Pin1	113
68.	TR5566 c0_g2_i1	LPB	1,04853	0,03516	Prohibitin complex subunit 1	87
69.	TR1941 c0_g2_i1	LPB	1,04853	0,03516	prolyl 4-hydroxylase alpha-1 subunit precursor-like protein	58
70.	TR3216 c0_g2_i1	LPB	1,04853	0,03516	protein disulfide isomerase	168
71.	TR707 c0_g2_i1	LPB	1,04853	0,03516	Protein trm112, putative	112
72.	TR260 c0_g2_i1	LPB	1,04853	0,03516	pyruvate dehydrogenase E1 component subunit alpha	91
73.	TR4848 c0_g2_i1	LPB	1,04853	0,03516	Calmodulin; Short=CaM	141
74.	TR185 c0_g2_i1	LPB	1,04853	0,03516	resolvase domain-containing protein	77
75.	TR4244 c0_g1_i1	LWL	1,20853	0,00061	Sec11 homolog, catalytic subunit of the Signal Peptidase Complex	139
76.	TR4399 c0_g2_i1	LPB	1,04853	0,03516	Serine/threonine protein kinase	95
77.	TR701 c0_g2_i1	LPB	1,04853	0,03516	SerineCarboxypeptidase	97
78.	TR3904 c0_g2_i1	LPB	1,04853	0,03516	similar to chromatin modifying protein 2B	76
79.	TR2626 c0_g2_i1	LPB	1,04853	0,03516	similar to Glyoxalase 1	90
80.	TR2997 c0_g2_i1	LPB	1,04853	0,03516	similar to Inositol polyphosphate multikinase (Inositol 1,3,4,6-tetrakisphosphate 5-kinase)	77
81.	TR6882 c0_g2_i1	LPB	1,04853	0,03516	similar to oral cancer overexpressed 1	111
82.	TR2629 c0_g2_i1	LPB	1,04853	0,03516	similar to thioredoxin-like protein	64
83.	TR6402 c0_g2_i1	LPB	1,04853	0,03516	small nuclear ribonucleoprotein D2	108
84.	TR5115 c0_g1_i1	LWL	1,20853	0,00061	Small nuclear ribonucleoprotein-associated protein D3	58
85.	TR3819 c0_g2_i1	LPB	1,04853	0,03516	small ubiquitin-like modifier	9
	TR3819 c0_g2_i1	MLWL	1,23638	0,00585	small ubiquitin-like modifier	
86.	TR1937 c0_g2_i1	LPB	1,04853	0,03516	Sterolmethyltransferase	76
87.	TR1686 c0_g2_i1	LPB	1,04853	0,03516	transposase	61

88.	TR3168 c0_g2_i1	LPB	1,04853	0,03516	transposase IS1016	122
89.	TR2639 c0_g2_i1	LPB	1,04853	0,03516	transposase, IS481 family	80
90.	TR387 c0_g5_i1	MYN	2,13898	0,00925	Triosephosphateisomerase	216
91.	TR7258 c0_g1_i1	LPB	1,33854	0,00108	type III polyketide synthase	184
92.	TR74 c0_g2_i1 m	MLWL	1,23638	0,00585	ubiquinol cytochrome c reductase cytochrome c1 prec (ISS)	108
93.	TR2572 c0_g1_i1	LPB	1,32096	0,00013	vanadium-dependent bromoperoxidase 2	58
94.	TR204 c0_g2_i1	LPB	1,04853	0,03516	xanthinedehydrogenase	81
95.	TR6529 c0_g2_i1	LPB	1,04853	0,03516	yellowcameleon 2.60	152
					Genes of translation and transcription (11,5 %)	
1.	TR7183 c0_g2_i1	LPB	1,04853	0,03516	translation initiation factor eIF3 f subunit	96
2.	TR5373 c0_g3_i5	MLWL	1,59516	0,00032	EEF1A2, eukaryotic translation elongation factor 1 alpha	281
3.	TR3157 c0_g2_i1	LPB	1,04853	0,03516	EF2, translation elongation factor 2	210
4.	TR6146 c0_g2_i1	LPB	1,04853	0,03516	high mobility group protein	114
5.	TR3719 c0_g4_i1	MLPB	1,38728	0,00031	C2H2 type zinc finger domain-containing protein	94
	TR3719 c0_g4_i1	LWL	1,21488	0,04772		
6.	TR1169 c0_g2_i1	LPB	1,04853	0,03516	40S ribosomal protein S4	91
7.	TR1200 c0_g2_i1	MLWL	1,23638	0,00585	ribosomal protein rpl30	89
8.	TR1733 c0_g7_i1	LWL	1,24124	0,00039	eukaryotic elongation factor-1 B beta	83
9.	TR1885 c0_g1_i1	LWL	1,20853	0,00061	Prohibitin complex subunit 2	168
10.	TR1943 c0_g4_i1	MLPB	1,38728	0,00031	60S ribosomal protein L38	86
11.	TR4347 c0_g2_i1	LPB	1,04853	0,03516	50S ribosomal protein L28	57
12.	TR4905 c0_g1_i1	LPB	1,32096	0,00013	ribosomal protein L24	147
13.	TR7239 c0_g2_i1	LPB	1,04853	0,03516	Ribosomal protein L3	115
14.	TR6956 c0_g1_i1	LPB	1,33854	0,00108	eukaryotic elongation factor-1 B beta	236
15.	TR3461 c0_g2_i1	LPB	1,04853	0,03516	Isoleucyl-tRNA Synthetase	73
16.	TR3177 c0_g2_i1	LPB	1,04853	0,03516	plastid thioredoxin M	125
					Genes of photosynthesis (7 %)	

1.	TR5426 c0_g4_i1	MLPB	1,38728	0,00031	transketolase	271
	TR5426 c0_g4_i1	LWL	1,21488	0,04772		
2.	TR5847 c0_g2_i1	LPB	1,04853	0,03516	Light harvesting complex protein	107
	TR5847 c0_g5_i1	MYN	2,13898	0,00925		
3.	TR2978 c0_g2_i1	LPB	1,04853	0,03516	Light harvesting complex protein	105
4.	TR5080 c0_g2_i1	LPB	1,04853	0,03516	Light harvesting complex protein	124
5.	TR5360 c0_g2_i1	LPB	1,04853	0,03516	Plastid ribosomal protein S1	112
6.	TR6270 c0_g2_i1	LPB	1,04853	0,03516	Light harvesting complex protein	121
7.	TR2772 c0_g2_i1	LPB	1,04853	0,03516	4-hydroxy-3-methylbut-2-enyl diphosphate reductase, chloroplast precursor	124
8.	TR630 c0_g2_i1	LPB	1,04853	0,03516	chloroplast clp protease P	53
9.	TR7272 c0_g2_i1	LPB	1,04853	0,03516	Mitochondrial Processing Peptidase alpha subunit	57
					Flagellar-associated proteins (4,2 %)	
1.	TR1434 c0_g2_i1	LPB	1,04853	0,03516	Actin-related protein 2/3 complex subunit ARPC4 or p20-Arc (Arp2/3 complex 20 kDa subunit)	71
2.	TR1460 c0_g2_i1	LPB	1,04853	0,03516	filamentous temperature sensitive Z	76
3.	TR2726 c66_g1_i	MLWL	1,14151	0,0005	alpha-tubulin	145
4.	TR2794 c0_g2_i1	LPB	1,04853	0,03516	Beta-tubulin-like protein	71
5.	TR3842 c0_g2_i1	LPB	1,04853	0,03516	Actin-related protein 2/3 complex subunit ARPC2 or p34-Arc (Arp2/3 complex 34 kDa subunit)	117
6.	TR28 c0_g2_i1 m	LPB	1,04853	0,03516	Beta-tubulin-like protein	76
					Vacuolar genes (2,7 %)	
1.	TR5174 c1_g1_i1	MLWL	1,092	0,00015	similar to Vacuolar protein sorting 29 (Vesicle protein sorting 29)	118
	TR5174 c1_g1_i1	MYN	2,38369	0,00059		
2.	TR1883 c0_g2_i1	LPB	1,04853	0,03516	RHD3a, RHD3/Sey1 family GTPase involved in the ER-to-Golgi traffic	71
3.	TR207 c0_g2_i1	LPB	1,04853	0,03516	Coatomer protein complex, gamma sub-unit	112

4.	TR5174 c0_g2_i1	LPB	1,04853	0,03516	similar to Vacuolar protein sorting 29 (Vesicle protein sorting 29)	118
					Signalling molecules (0,0007%)	
1.	TR5398 c0_g2_i1	LPB	1,04853	0,03516	Rab7, RAB family GTPase	107
					Hypothetical and conserved proteins (2 %)	
1.	TR1499 c0_g2_i1	LPB	1,04853	0,03516	hypothetical protein, conserved	132
2.	TR1681 c0_g2_i1	MLWL	1,23638	0,00585	Putative flagellar associated protein, transcriptional coactivator-like protein	139
3.	TR4091 c0_g2_i1	LPB	1,04853	0,03516	Putative: flagellar basal body protein	144
					Viral genes (3,2 %)	
1.	TR1013 c0_g2_i1	MLWL	1,23638	0,00585	EsV-1-163	230
2.	TR1360 c0_g2_i1	LPB	1,04853	0,03516	integrase	79
3.	TR4024 c0_g1_i1	MLPB	1,197	0,01219	related to retrotransposon HobShobase	124
4.	TR6644 c0_g1_i1	MLWL	1,22203	0,02711	retrovirus-related pol polyprotein from transposon tnt 1-94	119
5.	TR6278 c0_g2_i1	LPB	1,04853	0,03516	virulence factor BrkB domain protein	60

Table 7S Unknown genes of *F. distichus* vs. *F. serratus* female transcriptomes, which have dN/dS ratio more than 1. ID number refers to transcriptome of *F. serratus* male.

№	ID number in transcriptome	Amino acid sequence	Length (Amino acids)
1	TR1053 c0_g2_i1	YYLSTKKAFKMPLGDLIKPTNKEGYDETTLEDMPA YPGLTLTFHVLLKNIQVLSV VGSIVVGAWLRSPKKG FASLFLRK CPRYAGFGGAAGLLTASAMMFL KMPSIDLEGLEDRAYRIVLNKKVQALDFYSAIGGVTSLLVGVSMRSPVRGPK	153
2	TR277 c0_g4_i1	ALMFTSGSKSTGILCIFFFFSSRRRHTRLQGDWSSDVCSSDLAKEVPDLHRSARDGN VDGEMGVTEAHLIKVAL	74
3	TR2771 c0_g2_i1	FFFQAEDGIRDYKVTGVQTCALPICLVVHKRYCLPSAEAM SALAFSNASLVLPFASIL STAAVGGLRMRMSRFNSM	78
4	TR3111 c0_g4_i1	LDKLWSMLPEGTLEKAKGVDGKAPVIDVTESGYFKVLGKGKLPGVPIIVRAKIFSKL AEQKIRAAGGACLRSK	73
5	TR3165 c0_g4_i1	IHDTKMRSDCQVSIVLCFFFFSSRRRHTRLQGDWSSDVCSSDLERRALRR	50
6	TR4842 c0_g2_i1	VCSFFFSSRRRHTRLQGDWSSDVCSSDLSTMRVIHSNRTSRAILCSILLLSQHVIIVDAKG WVKSQPSRSTTLAVAVPGG RDRPLRDWKAFGRNAKKKIEESDLKKKIETKAEEAGEALEEVLPGGSTGDVEETVRAE AAAADPFLEKATPQSIALLG KVVVVTTVGLYVAKKAVAFTKAFRKARKDGKARRDAAADAFATASPTTGTGDGVGE VLSDLKDAVEE	226
7	TR5288 c1_g1_i1	SDVCSSDLGQVARDLGLFNPLAVQSMYIFKQPKIGGEVKPHQDGAFLYTEPQSVLGYW WALEDCTLSNACLWALPR SHSRVV	82
8	TR529 c0_g2_i1	TVKSMLSLFFFQAEDGIRDYKVTGVQTCALPIWGRLAATGKSKKRISDSQNETTNF HHRGNGDYGGGERSGCTHT	75

9	TR5397 c0_g2_i1	MGGGHDHGMGKATEGDMISAGLPMAFRDSCAHLIPLNKCRRGTMYPWECSHE RHIYEKCYLDFKERKGM MKEKKEES	81
10	TR5451 c0_g4_	QTLAQLSIFFFFSSRRRHTRLQGDWSSDVCSSDLRVWLGSFVATSSSFCVRLRLVTFL TDTSAVKT	66
11	TR5682 c0_g4_i1	PRKCSWTNRLLQAKDHSATQINVGKLDPVSGTFTGESDVYALAGYLRGRGRSEEHT SELQSPCNLVCRLLEKKKIDNI TSFTLYYFTNLLLA	93
12	TR1025 c0_g4_i1	LASLKAPRSRHPQLPQATPCFRNRESTEPRSESLPGSPQPPCRIMFIQKKNRVAVYSYL FKEGVLVAKKDFKPKHEEIE DVTNLEVLNLMKSMKSRGFVRETFNWQYFYWYLTVEGIEYLRAYLHLPEEIVPATL KKTASRAARPGPPPGEREGGR GGYGGDRFGDKKMGPGGDFNPEFQRG	183
13	TR594 c0_g4_i1	MCFFSSRRRHTRLQGDWSSDVCSSDLCKALRHQPSFLERRTRLFFSASTLEVSAIKSDH	60
14	TR6010 c0_g4_i1	LFFFFSSRRRHTRLQGDWSSDVCSSDLWECGLVLSPLQASWNSQMGQSTVSLGDER	56
15	TR6092 c0_g4_i1	FFSSRRRHTRLQGDWSSDVCSSDLPPHSLPGRGTSRHSPISTRVCVDFDCDSWRADG LHLILCPADMQVEA	72
16	TR690 c0_g2_i1	TMKQKGKRQMTKSPEFLVAFFSSRRRHTRLQGDWSSDVCSSDLFPSPCKQEELLFLN PRHPSCQSVVAPVPRGTSTA	78
17	TR4938 c0_g4_i1	FFSSRRRHTRLQGDWSSDVCSSDLGSGRRLTSQFVVGPAERLAVGGRRLAGEDN STAENVTLVDIVQYQLNLWT GVGLAFIAFLAVYATMTMDVQ PDSLLYAKFIADTSGGGLKTD	119
18	TR1437 c0_g1_i1	FFSSRRRHTRLQGDWSSDVCSSDLTAHYRLGLGSVLFLVFQLIYHLTGHAAADRGHVD NVCD	62
19	TR1041 c0_g2_i1	FSRVLFFFFSSRRRHTRLQGDWSSDVCSSDLSDARSKEGVSNMTRKRGSKKPTGGKVG NHSSARRRAMIQPAQDAK	92

		ALKRKKAAALSVA AEKE	
20	TR1157 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLVRVGTYPTRNFATLGPLLLRPPFTGASVISVVIN	59
21	TR1161 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLRPTCLRNWFRETGSAVQSRVSLFISILKLNMLTH GISPDFR	59
22	TR1355 c0_g2_i1	KDSKEHPECRYSSSTAKECSSNNGKYTCETIRRIFRSCPNTRPEKIYEKTTDRKSTRLNSS HLVISYAVFCLKKK	75
23	TR1356 c0_g2_i1	MDESGKLIAIIGDEDVTGFLLAGVGHRTVNTTNFLVVKNDTTVSQIEDAFNRLTAREDI AIVLINQHVAGEIRHVLSGY SKTIPTVLEIPSKDVPYDPEQDYIMQRINMLGGGQ	116
24	TR1375 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLERLPVACYGDLDLVLHCERNGKPWSNVRVTLTN VAVVPGVWFDLISFNQIET HPILLDKEGAHILGS	93
25	TR1451 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLGRRRVGLRLDRPLLFCFEDWDNRVGEKGGSGYGRE LGRRCPAGIMPLVIESSLS	78
26	TR1477 c0_g2_i1	HCSRGVDFFFSSRRRHTRLQGDWSSDVCSSDLTGLDYPRSRGTVTVGIVCAALQDGLL PNLHRGKADVPGFDYAAGADR	79
27	TR155 c0_g2_i1	SSWTCAELKASLRNMLQLADSAPLYVYCNSVFEPSPDQMSDLYKCFHVNRRALLMRD RKSTRLNSS	66
28	TR1558 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLSVERLHTRVSSSFLMRGKPTKIYL	49
29	TR16 c0_g2_i1 m	FFFSSRRRHTRLQGDWSSDVCSSDLTEMGRTEIGNTSDPDWTHQFVIDFLFNEIQYLKV TVYDDD	66
30	TR1606 c0_g2_i1	FFFFQAEDGIRDYKVTGVQTCALPISDTQSTTVSVNATNAASACDLFRPALSMISATTCG FLTILAVSSLVAFIMLIP	78
31	TR161 c0_g2_i1	FFFSNRRRHTRLQGDWSSDVCSSDLSLCTEMVRRRTLEKHAENVRLGQQRLHLLSPFLG GHVLQEHHHGLEVHL	74
32	TR1674 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLRNEPTARAIFTIQSALDSWVKLGTSEEQLRRIESKH DIARKSPEAKDNVHKKVLQAL ENFDVPITAWISLVRNSLESRRQNHRSHYPSRLPRGTENLCLAGHA	129
33	TR1676 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLKNGTLSTPGSFARQPCPKRLPSVLEHLKLSRRFGR VQRLPPTSGALDR	73
34	TR1702 c0_g5_i1	IMKFIAACALAGVASSQAFMAQPTMKTTGPSPDATSRADFAKIVTGAVAASFVAAPA	120

		VAKSGTAAKQNYFGVLGADQ NLGGGMSNYFPESETYSPYSGTPDKALYNEADPFMIKIKV	
35	TR1823 c0_g2_i1	MLRHDRVDQHLGSPLPLNGNISYPVSRDSRAQDGSARSEEHTSELQSPCNLVCRLLEKK	60
36	TR1836 c0_g2_i1	FFSSRRRHTRLQGDWSSDVCSSDLGCSIVSVVLGGVALKKMKSSPKIGFLKTVFCTCSR DYIPYDKS	68
37	TR1855 c0_g2_i1	TQEGIRDDEVTGFFFFSSRRRHTRLQGDWSSDVCSSDLAFSMAFMREDCSDVALSMRLC HSCEASRYSYRSSITSSLVSGS NSYSTIISSAPSPDLDIMGRVLIEVS	108
38	TR1899 c0_g2_i1	FFFQAEDGIRDYKVTGVQTCALPSIMTLVHADVDVLQGASTCFAEKVQETEETAEPDENG EVVIEVDEVFAECIQEVEDAD DTKDTAATADFLSKLIGSDSKICSCLPFVEKTPDCDMWMRYRHFAMHLEDLCEAVDSA VENAAMPGIVELEAANQTEN QDEAFEAAPQGNTKGSNGVMDEHPFGPCPPGQETHPEDPTKCQTSENQETYFNEHESSG GRGPPPEGQGSIFGALMS VVGVVVALAGVVGVLGGLWAWGKFHGGARQPGVRYMAANVDNDIGTELGPFG	287
39	TR1924 c0_g2_i1	FFSSRRRHTRLQGDWSSDVCSSDLSDTPLKFMRSSLLRKGFKIRIKGRNVMSIRQSSQRP YHRWPRSPSIMTTLQTP	80
40	TR1874 c0_g2_i1	FFSSRRRHTRLQGDWSSDVCSSDLVDGFILKVQKRAVEKRKEMDEDRESGLELSKEER MGPGGLDPVEVFPTLPIGMQE AFESKDTEMLKSALLALTPQDQKYHMKRCEDSGLWVAGAGGSMSVGDDDDDD	132
41	TR1942 c0_g1_i1	FFSSRRRHTRLQGDWSSDVCSSDLAWALRRPEATTPPLSCHKSRRKSMNAHRSCGF SATSTT	64
42	TR195 c0_g2_i1	FFSSRRRHTRLQGDWSSDVCSSDLGRHVVICLGHAKALMYRMFSLVPSVQGCFLHP WQMFLGE	64
43	TR1966 c0_g2_i1	FFSSRRRHTRLQGDWSSDVCSSDLAAFHAHEGTTLRSLPKQRFQGKYLRREQNVSLR RSSPKASDRKSSKH	72
44	TR1976 c0_g1_i1	HWSHLLAAVALGGSILSGISAIDTVDKLIVDKYLGRWYQAYGNAVARTITRDSTCIAA DYGLDKATGNITVINSREHIPTG ELNTITGYAYGVDQSKPGQLKVHLGQVPVDGDYWIVGLGPATFGPEKAYQWAIVSGP EAESLFILVRDVPAFEKKYAKDVL ALVKDLGFTGREKEPRRIPQRGCIPFDPMTIHAGKKEITEWVDYTYATLEDEGDRKSTR	241

		LNSSHLVISYAVFCLKKK	
45	TR2027 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLGKARSVLTGGILGAWRPHPPDRMCDAFGVREK TGTKGGHKSDGKTE	71
46	TR2037 c0_g2_i1	SFFFSSRRRHTRLQGDWSSDVCSSDLQYLKNLMLKYLGTDEGEAREHMERAIATVLQ FSEKERSELKEKRQAQTL	75
47	TR2051 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLPYTLIQKHSQTPTYTTLVSSSLSPAHTP	53
48	TR2060 c0_g2_i1	ILLYFWFFFSSRRRHTRLQGDWSSDVCSSDLRSAEGNKSTRLPWFRLSRALTSRTNAPM LWCRWKQK	67
49	TR2113 c0_g2_i1	SVFSLRLVCHFLCFFFSSRRRHTRLQGDWSSDVCSSDLYVLLSEGGQVWRIFGEWQGT TALNTCHAIYPRFERDEPGTPSH VVGSLTTF	90
50	TR2117 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLSQLRPETSLHGVVMGGVQPINKMNVVKKRKT FARHQSDRFMRIKNSGWRKPK GIDGRVRR	87
51	TR2125 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLDEIHKNATDISAASVSRANVKQPPEAVSASSYA	58
52	TR2166 c0_g2_i1	FFFQAEDGIRDYKVTGVQTCALPISGKIANDANSVLNVDTAGQRKMLEARQEGEKSK VPELTARSELYGESNPHVELDKE KLEEA VKQQAFLSGGEEKDDRKR	106
53	TR2227 c0_g1_i1	FFFSSRRRHTRLQGDWSSDVCSSDLHCMRFPGGGGSPFASAWATMSTSACHHHGLTL	57
54	TR2230 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLEGAMASTGSLLPQALIVAGAAVICSAPELWRGE LPLVAVIAIWMLRAVVVVGVAV AVFFYRAY	88
55	TR224 c0_g2_i1	FFSSRRRHTRLQGDWSSDVCSSDLPKPRQMLLKGVPVGGRLSASVWIGVETSATREV ATAIATAWPTFR	69
56	TR2297 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLASQCCTRLHPYPSLANVVGNGQENALPGFTEA SESRLAGVAITPEEGGSEICAQGMQ	83
57	TR2330 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLVWSVGFRRKNGEGKSKLNSKNQKEMKFF	53
58	TR2340 c0_g2_i1	MSGKGGKGGKGGKGGKAPTSSKAPQSRSSKAGLQFPVGRHRFLKSRVHSHQRVGA TSAVYTAAILYLTAEVLELAGNACK DLKVKRITPRHLQLAIRGDEELDTLIKATIAGGGVIPHIHKS LVNKTTKKRI	134
59	TR2352 c0_g2_i1	MNAGEEGAVVVGELLKTDRPEQGFNAATGEYCDLIAAGIIDPTKVVVRTALIDAQSV	99

		ALLMTAEAMVTDLPADERSEEHTS ELQSPCNLVCRLLEKKK	
60	TR2361 c0_g2_i1	FFFQAEDGIRDYKVTGVQTCALPIWQFGMYGKQNVGNWMGMWQEYVAQEIEAGSE PKGDDAEDNVIPMVGIVPGRH FSAN	80
61	TR2366 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLQDYNNPSRQDTESHESNEQCWYRMNR	51
62	TR241 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLGKGPVRTDPVLDTCVKYVKNVDMMLCLYCNWR	57
63	TR2451 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLSAPVQSTRGPLPKSPHFSTTSVVKGKK	52
64	TR2460 c0_g2_i1	FFFYSSRRRHTRLQGDWSSDVCSSDLASLSEKNGCCVSHSDATEFHLFITPPFATAFTHTP	60
65	TR248 c0_g4_i1	CIDPVQYIIKHRSKLLFSAFFQAEDGIRDYKVTGVQTCALPICRNGDMVVRPYTPTTTND TKGHFDLVIKIYPEGNVSQLHMA NLKIGDKVEVKGPFKIDYKPGMKKEIGMIAGGTGITPMLQ	124
66	TR2534 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLRQTDGNVQICCGAILPLQTTVRGPAPSAQPG QEDIVEEALYFFRANVLYRNYQIN GSADRTLIIYLTF	95
67	TR2538 c0_g2_i1	PQVRTMVMRSLPRAGGGHDDHPHLVFDTPTKGPFTKTQAGVVVGGGVFLAIAIVIG AVSHQQIKHGFWKK	71
68	TR2539 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLRQLGAQKHPYEQSSRPLLFHGGRANTPRLYSTG VSPLSIHTYTHNTQNTCN	76
69	TR2558 c0_g1_i1	FFFSSRRRHTRLQGDWSSDVCSSDLSTSLSGNAAASGRELKQKGEMLPNQIHRLEVKAV STTQPLQHNTVLPNISTSLPGNA AASGQRAAFLHKGISYGPVHDMAQ	107
70	TR2580 c0_g2_i1	IDCAWTVSFDPKFDPLLEAVREATNAGISHAGIDVPLNEVGAAIQEVMESYEVEINGTT YPIKSVRNLNEIGRAHV	76
71	TR2581 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLARCSLGTFCGFQFFSWQLSQLQEYSTNFWLSING VGNLWGWQPLS	70
72	TR2594 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLCVELKYATIRRGGAAPAYNRIVPVRINAQFPFPA NSSYAPKT	67
73	TR2698 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLSTVPKAVRVNNAMGGDISPPERLVRPLTRKDIPK SKSTVNLFKGIIGGKFSDRQQHC LAKDATLGPIYEAVMEECEQWADKPTRFSLCCSREVCVELKDFL	127

74	TR2702 c0_g2_i1	FFFQAEDGIRDYKVTGVQTCALPILTLTNKTVDALNKYLGYYLERPMVHKKEGIVCRST LINGQLPCSTKSHIYIPQNQRH AMAATTDNTK	94
75	TR2704 c0_g2_i1	FFSSRRRHTRLQGDWSSDVCSSDLGSRCSSSLTRVGSTKTTRPSSRK	49
76	TR2705 c0_g1_i1	FSSRRRHTRLQGDWSSDVCSSDLPGVPVDGKWVEERYCGLITWVAVSTGLVYSYCCP CDRRTVWISGGVKFNEKGAM TNDDL	84
77	TR2709 c0_g2_i1	KVAAAVSSNLIFFTPLQSSKVKSPIMKVLSEFFIVCVTAVGAFVAPAPRVSSVVRPAMSS TAMRMGLKEDIQKVATGVSTA LLTAAPALATEGTGEPLGIEDSKGLIPLFGLLLINLLFNNWAKDQPEGDFGEYDQRR	140
78	TR2717 c0_g2_i1	FFFQAEDGIRDYKVTGVQTCALPICLDLAPHACVLDDEKRLGTYPKQSVFDLVG MHDGDAAIFGAS	69
79	TR2751 c0_g2_i1	SCTVLCLFFFSSRRRHTRLQGDWSSDVCSSDLRKLSDVRDGIIFHGLGGDIFGRHVCGLD VARSPRLRPQVDFSSKKRASLR	81
80	TR2764 c0_g2_i1	SFFFSSRRRHTRLQGDWSSDVCSSDLESSTPSSPWGERKRKPTRPLSSLLSESTWNVLS SNIFRQAEGHLHFQTTTTILRTTSAL LLWRLSDHVKKTQKIKNGPSPLVEELLLTLFFHLIDKILQVRLEGLLHR	134
81	TR2903 c0_g2_i1	FFSSRRRHTRLQGDWSSDVCSSDLKYVAHEPSLRVRKDLHRFIRSDVQRHSTLRRI FGGKPLGASCRVLNL	74
82	TR2926 c0_g2_i1	FFSSRRRHTRLQGDWSSDVCSSDLVYSVSTDSTIRPTRKVPIHPFVQSDWLRDSAHP LRAKWRGRLSPRTSNATIKCHK KNSVLLTDKSLQR	94
83	TR2928 c0_g1_i1	FFFSSRRRHTRLQGDWSSDVCSSDLAKRSSVKSAGSKKTPGEIRIQKDIADLDGGSV AEITFPNPNDLTSFRVKVKPDSGF WKGASYNFIFTIPALYPHDPKVLCEKTIYHPNINLQGNVCLNILREDWKPVLINAV IYGLIYLFYEPNPEDPLNHEAAELY RENISQFERTVCR	178
84	TR299 c0_g2_i1	LEVRQAEEIRELQAELTVRDVRLHNQKLGVECMAKELDDSKKTIERFRRSEEHTSEL QSPCNIVCRLLEKK	72
85	TR3020 c0_g2_i1	LFSSRRRHTRLQGDWSSDVCSSDLGDTMWTLVATLVDEVVISETGDFYQYLSYIDA VNVIDYDPTCEDGTEAFTGA	78

86	TR3026 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLANPQPRERAWNNQRGKKTLLSLTLVRRCEMT YEV	59
87	TR3038 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDIKGIDDHVEQSNGSLDDPRLPHRLEKQLPLIRHG EAVVRSADIQHRVALGLVRQA VHQ	83
88	TR3108 c0_g2_i1	MRWKKKKWCDFFFSSRRRHTRLQGDWSSDVCSSDLENNVRHREVARTFELFLT CPS CIEARQLFLLAGRPGLYLLVRVLSI IHIIYIKKIVDART	94
89	TR3111 c0_g2_i1	MVGGKEVCCVFLFFSSRRRHTRLQGDWSSDVCSSDLAGSASQVTAMPSRLSK NRKKRGHVSAGHGVRGKHKRHP GGRGNAGGQHHRILFDKFHPGYFGKVGMRNFHLLRNRDHCPIVNLDKLWSM	130
90	TR3133 c0_g2_i1	FYLFFSSRRRHTRLQGDWSSDVCSSDLLLAVPAPRRVEFDHHVLPFFLHHIFEVR GGQHNNVLLVNVL	70
91	TR3201 c0_g2_i1	LAPIALVGRQEISLAAGLEGRHSVGLIHPLYRGPVSEVDAFILVGGIEHRPVSRSED HTSELQSPCNLVCRLLEKK	77
92	TR3297 c0_g2_i1	THDVERFCCLCVVPALVFICFFSSRRRHTRLQGDWSSDVCSSDLHNSRSTHSTMT KAEAAKAPKWYSGDDVKPKHAPQ KHKLTKLRSSITPGTILILSGRFR GKR VVFLKQLPSGTLVTGPFQINGVPLRRVNQAYVIATSM	145
93	TR3326 c0_g2_i1	LLIVVFFSSRRRHTRLQGDWSSDVCSSDLNALLKVNQIGSVSEIAAVKLAK QSGWGVMSHRSETEDNYIADLAV GLCTGQIKTGAPCRSERLSKYNQLLRIEELGDKAVYSGQNFRKPGWM	129
94	TR333 c0_g2_i1	FYFFSSRRRHTRLQGDWSSDVCSSDLIAAIPLDISHHNGLREMRVQAQQGHCS	54
95	TR3408 c0_g2_i1	FFFQAEDGIRDYKVTGVQTCALPICRNDKPRASHGESLDCNAGRRGGSHMRAS NDTYSRGRSAKKVDTNGK	73
96	TR3410 c0_g2_i1	ATGDIAAVTRKEALEVKLGLSLVSAVQSDELDPAYCKILRERIMRDKILALHLEGE KRDEEAADVMKRIRAMEEELAGVPE DQRGDRKSTRLNSSHLV	98
97	TR3411 c0_g2_i1	KRCDYIVLFLVFFSSRRRHTRLQGDWSSDVCSSDLGTDKSTRIPIGPENIFRCTAV SGGVW	62
98	TR3414 c0_g2_i1	MLSFLTDLQPIPRVIVIALIAAAGCTGIYFLVITGAILDWPLPFYPWQILTLVFWAISK	338

		SAIASGGRLDGGGAIGAEVPSGNKD QSSSES VHYFMPAGWAFAIWAPILLGETALAIYQALPTHGVAKAAEWLAALSPWL SAAFLQALWCLSFRKWAKDAG LLWFPTLLLIGTALCLDRAHHVLRVAEALGAVSLLQYFLVHLPVSLHFAWITCAVLV NLNKYIGGISRFSSTVKFGFSMASI VVAVLLAVGITGKTGDPVYAAVVAWALLAVAGKGAENMERPGVQDMVANEDA STDTPLMTSGGGVESSKLTEN LVARIGSGLCVVAAVLFFFF	
99	TR3438 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLKTSITREYARDPRLDSSYHTSCTQLQYIRFPK LI	61
100	TR3496 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLGAVDTHGAPWPLRHYILSLNKLK	49
101	TR35 c0_g2_i1 m	FFFFQAEDGIRDYKVTGVQTCALPISILSPHSDKSCSVVRNTMKLIILFLAFVAMASA FVPSTSCRGALTASRTRSSTPLSA EEEDGPKHVGS GGMADTRDPEP	105
102	TR3536 c0_g2_i1	MEQGGSPQISGYLQKKAKTGKWQKRWFETNAHYLTYYSKKMEKLLAALNLPQ	54
103	TR3544 c0_g2_i1	VVSVLFMIFFFQAEDGIRDYKVTGVQTCALPISTGVLTGSSEASVNPQRMHVCVNRIS MEEGSR LGPSLELSRALEATSLV VFAATPRDQDRFFWSPLHPKPDWSKAREDLGCALRCTDSEWGAVPLL VVLSVQDSV RSRGKKRRLGDATLREL VVQD AWEGLGLQEVAQRGVSVRVTVVGLHEFESKSKEAAASLEGLVFSGEAEEASSDDLA SYVAELSGQA	224
104	TR3558 c0_g2_i1	MIFFFFSSRRRHTRLQGDWSSDVCSSDLNKNNTYEMRYFSIVNNEAEADEE	51
105	TR3452 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLLRVRTHQSRKISHVHQGSEVVQSRYLQRAE	55
106	TR3568 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLANPQLLGWVLPVRQGTSAAVPTAYRFSESFPS MTARAPRAMNAS	69
107	TR3573 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLRDVVHLPLPPQELRKRLQRNRGWQPQPQHV PLLAQTHHEHQYHQHQLRSP TARDVSAEGDPA	67
108	TR3578 c0_g2_i1	VSIQLRLGLGNVRVSLAFLGLVLVALVVSRCRTVLGPVPHANPAKFPRSEEHTSEL QSPCNLVCRLLEKK	72
109	TR3647 c0_g2_i1	HAILRRETCAGLLVFFSSRRRHTRLQGDWSSDVCSSDLFTTDGRSESPAPAADFY	67

		ALLAASAVAS	
110	TR3649 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLGLIYYTKKRVIFAPKTFGVLTLPYFNITITTTVL	60
111	TR3661 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLGMISQMFGSDVGRDAYIERAVPLRCRHYLST LLFTFRTSPLDLYQVLE	73
112	TR3674 c0_g2_i1	SVVYVVFCAFFLFLVVFSSRRRHTRLQGDWSSDVCSSDLFDTTQPCIHQNGRLG LPPLLWRPLLSLTASR	73
113	TR3677 c0_g1_i1	DLGEMAVVNWVQSYAEDRGMVRAAKMLRKKKMASYLTAEMRGGSMNDTTVA IVWLQ	57
114	TR3679 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLLCDPKLIQSLVISDVYVIVLAGNVILRIVRGGACPPHYHPYMHIPVQAVLL	76
115	TR3701 c0_g2_i1	SKASLFAAGSSGLLTAFGYGGYLEYKRRGTVSKLWTIASLVVSNAIWIMMVKRISK KTGNLMPSAPV GALGLG MTLFYV YQMINPPTKKED	91
116	TR3831 c0_g2_i1	TFFLFNIGFFFSSRRRHTRLQGDWSSDVCSSDLSTRDIFNGTISRQAPANRTMKFTAIF LALWAMCASAFIAPSAVVKS RRFVTPEVRSTSSTVAPKTTMNMGLEQIVDMVPM	114
117	TR3911 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLTCS DGYEYKDCYEDIVCDGDCDDKKCCDKVQ EYCD SYACSD	67
118	TR4000 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLEDMATPRQESWDTGGLTFGQPTIGMSGAEVE DVWDDGAILKAFDDALSTHS RKPGAEARKWKKHAVQAAGG	96
119	TR4014 c0_g1_i1	DMMKEMMNSDMMEEVLGNPEKMEQARQAIANPTLKQAMSQLPGFVDMIESPEK WRENMEMALQMFKAPKA MMKQQGQAGGAEEQEEDFADVDDLDDD	98
120	TR406 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLFDRHQRKRIPGKQLSLIENLLDTSTSMYQGAKR PSENAQSRFSDSLYLHATI	78
121	TR420 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLGFLSFVLQAGLDQLSPSISVGRAFSFRQASLLP DLCRASILIFFG	71
122	TR4208 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLVRMMQTCGIISLLVYLSSVQVHFWGPSTRFG	56
123	TR4220 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLFPAAEHRGMSLCSVIDARRALNSSTLSRWVC VTILRIFFASRCCLSRNRRW	108

		AGVRVSSPSSPVLASPSNPSLFPLLSDSDA	
124	TR4246 c0_g2_i1	DMMKEMMNSDMMEEVLGNPEKMEQARQAIANPTLKQAMSQLPGFVDMIESPE KWRENMEMALQMFKAPK AMMKQQGQAGGAEEQEEDFADVDDLDD	98
125	TR4292 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLAANYPTPHPCSHRRDSSNGRYGHGQVRAP RDPTAHR	63
126	TR4312 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLGSSICTWIARGEMIFPGRCPEDRFSTKGWG RELDTAKLCLCRASQ	71
127	TR4314 c0_g2_i1	FFSSRRRHTRLQGDWSSDVCSSDLAQLLISVFSPLGLAASGFRSPRGAPAALVPP AWRWAGTA	63
128	TR4359 c0_g1_i1	MATAILEPSKKRSPNRMVVDEATSDDNSVVALSTAKMEELQLFRGDTVLIKGK KSRDTCIVLADDTVDDSSIRMNK VVRKNLRVRLGDLITVSACGDVPYGKRIHVLPVDDTIEGVTGNLFDVYLKPYFL EAYRPVKKGDLFLVRQAMHPVEF KVVECDPAPYCVAPDTVIHCEGEPIKREDE	185
129	TR4360 c0_g2_i1	LLVALKVRFKDRITILRGNHESRQITQVYGFYDECLRKYGNASVWNCFTELFDYL PMTALVEDKIFCLHGGLSPSIDTLD HARALDRIQEVPEGPMCDLVWSDPDDRCEIGRAHV	116
130	TR4367 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLHNNKNHTSTSVPLLIPLHSCPPSTRAADTIQ AAIVRCKRPVRPAFAGVTHHPW	78
131	TR4413 c0_g2_i1	MASMARKLAPLADRVLVRRLVAKAQTAGGVYLPESKAGKTNEAEVIAVGPGR LTDGGKTLPVSA	65
132	TR4428 c0_g2_i1	LHFFFFQAEDGIRDYKVTGVQTCALPISGGMLPGEVELIVSRALGRRGEVPRCLE CAPARKIFHHIFFSGKGSRWFLFRS DGVSGGLEWVQFTLAKNCLGFRRK	104
133	TR4439 c0_g2_i1	MATAILEPSKKRSPNRMVVDEATSDDNSVVALSTAKMEELQLFRGDTVLIKGKK SRDTCIVLADDTVDDSSIRMNKV VRKNLRVRLGDLITVSACGDVPYGKRIHVLPVDDTIEGVTGNLFDVYLKPYFLEA YRPVKKGDLFLVRQAMHPVEFKVV ECDPAPYCVAPDTVIHCEGEPIKREDE	185
134	TR4443 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLEVDAMAGSRAKCNMHEATTRILSILQKVE	118

		GFEKASKNTLICATNRKDDLDSALL SRFQLTIRFGLPDTSTRREVFAVYAKQLTRDELETLAS	
135	TR4451 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLSCYARGSYRFPGLMMPWPGMIHRRRAAGIY PFCFVFCIGYSMALDRLPFPWRS PIL	80
136	TR4458 c0_g2_i1	FFFFQAEDGIRDYKVTGVQTCALPICHVCPFTVSVSPRTMIARLFAFACLFLGAD AFFAGGPFRVPTRSASCRTSLSMA APADHPEIEEAEAAALDATKKFGATSEEARLAWDFYEEVAAADNSIATRAGLDE DCDVTHAAVCEEYQEQM	151
137	TR4461 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLELLAVSAPRRVELDDRELVFLQESVKNFCS QHDDLAVLGLYEGRAGSLLRGWR GLLGLHAVLDGLEVIDEV	96
138	TR4475 c0_g1_i1	FYQEMTKEYPDAKVILTTRDPDLWWESIKNTIHIHSPVVHTWVGIFFLQAFVPRFR KFWYMIRHIATLNMEETEAKAD FLRHSAQVKATIPAEKLLLEFSVK	101
139	TR4517 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVSLPISHFYFSARVFAFIFLCVIAWTPALRW	50
140	TR4569 c0_g2_i1	AYGFPGKQAKGVPRASLSDIATVLDNNSRTKPKVVCRSHHLMSPRPSYKQRCPSTS QRSEEHTSELQSPCNLVCRLLEKK	81
141	TR4595 c0_g2_i1	RFYPHLVLGVLRHVVEGSYVQPKLARFGELAEACTQRKERIAADARSLAHQSLA DIVDSVLMQSKTALFVGAVDEVLD VLAHVVGQFLEEHLFCFIVCERSHCRKSCFHHRFFCSKEIEAVAHSELQLGSTRS EEHTSELQSPCNLVCRLLEKKN	157
142	TR4599 c0_g2_i1	YNNSNFGEWGGGCRSSNGTLEFKAQVQFIGNVADYGGGLAVTGGDVTFHKGVL FEENGEDRKSTRLNSSHLVISY AVFCLKKK	83
143	TR4624 c0_g2_i1	LRAPIKALATSVNNALFDPANKRPILIDALRQATKDDSPFSAKIIRDILHLCDLIREC SHTPAGARWIADTLQSAAPDAA PPRYTPQSASSGSDSAAPGHRREDPASGRGNAKRTTFAGSDSDRSRRPTQDR RDYRRG	140
144	TR4669 c0_g2_i1	MAVAGLGADARQIVNRGQDEVDYVNTYSSHMPPSMLADRLAQYVHYFTLH GSLRPFASVLLAGYDPDVKTHELY	107

		VIETSGVMYRYFGAAVGKGRQGAKTELEKLLK	
145	TR4702 c0_g2_i1	FFFQAEDGIRYYKVTGVQTCALPISDQTAHASLLSFHILVYDVQLVVYPVLAAA PSAGFLRELISDFLMPKGNEALGLSK	80
146	TR4705 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLFPATKYFRRASAPWMSQVLLCCVAAVH LHT	55
147	TR4725 c0_g2_i1	TTCTARHHIMLSDRQKGFIEKVNADIDMPFISEGGEEKIIHKLLDMVDKVLETA LRVLIAEDYVGDRKSTRLNSSHLVISYA VFCLKKK	90
148	TR473 c0_g2_i1	PCRASRSPAAATNPSPFSLAPVSFAIKLSHQAIVWRFEAARWRLAQPPGSPLGLQ CLLERSWMNGTGMVYVFLTDTENL RLSPEELSRSEEHTSELQSPCNLVCRLLEKKK	111
149	TR4761 c0_g2_i1	PEEREIKAKVTAETRIPAILIQIQHDQLAKMGVSREEGREALNRYSRDRKSTRLN SSHLVISYAVFCLKKK	72
150	TR4780 c0_g2_i1	FCEVAMQKTADERNLSLFFFSSRRRHTRLQGDWSSDVCSSDLGAGGLSAVNAI GTQLRDPYISGLTRWRMAV	72
151	TR4812 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLRGRSQRNAHERLRRAFGPGYQAAGVLRVQ PHSLTRL	61
152	TR4843 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLHVLPKEYYDKLDEKDEEEDDMLDLAWG LTDT SRLGCQIKVTPELEGML	74
153	TR487 c0_g2_i1	RCKSVKKHKQRQGLHTAAGETVGVGRWGALGSENQASRSAPWERVSTGRER SEEHTSELQSPCNLVCRLLEKKK	75
154	TR4885 c0_g2_i1	FFFQAEDGIRDYKVTGVQTCALPISPDEIVIGNWRDDNFLKDRVFDTVLADYLI GAIDGFAPYFQDQVFERLKRHVAPG GMLYVVGMQPLP	91
155	TR4899 c0_g2_i1	MEEPPQPTCWAALQVLGTLGTMFGIASAAVLTPLKGEEGIIGIPAAPGVVIGSSL GVLAAQASGEIGI	69
156	TR4900 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLGCLWCSWTGFPPPCYCPLYLRLVYS	50
157	TR4925 c0_g2_i1	SDVCSSDLLKRTKHKTRARRMTDTKIIDRKTTVVQNTAFRRKTDGEEDLTF EECMKLARELKDFSRLPRSCRQFLDKK AVLHQACEGMTGD	92
158	TR4950 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLGLSTSSFGQQCGMLSRWLTNKNGAHV	62

		LLTAHKPPNIA	
159	TR4969 c0_g2_i1	MSYLLPHLRSGWAVDQAIINEEDRVVCLRFGHDHDMTCMQMDEVLAGIAE QVKNFCAIYVVDVTEVPDFNDMYE LYDPCTVMFFFRNKHIMIDLGTGNNNKISWAFNNKQAMIDIMEVVYRGARK GQIGRAHV	133
160	TR4976 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLVDSVWVAFKHTERNIAETGLEILYALL MNVGRTPSVAQAFYRQFLLS	72
161	TR4982 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLIGHQPYPVIDVVANPVRGLLDRKRSEEH LHSYNENKQTKHKRQKEK	71
162	TR4998 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLPTARIVRQMTGGNVQESGGEDDTPMGV QGVCLDEHDNKEKHPVSVCMIG TGEYTTGYVYGE	86
163	TR5005 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLADAVGVHRVHFDPFRRHQCLNDRQMSPC RCHDQRRESGGAHGPRIGPFRQQ	75
164	TR5018 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLVESVALLERAVGNGAYHGVLLLHLVRP VLQYLKVLVLAHV	67
165	TR5019 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLLFVCLFIWEGCVEFEVKRLVVQDTYCT TPGTCNTPL	61
166	TR5023 c0_g1_i1	FFFSSRRRHTRLQGDWSSDVCSSDLIAQVSPIECGVLQRRKLTHGSEDRNPFGR RSKRRAWKRCPCSGERNTDGRP HGFAGLEGHLRLGFLAGPVWRCIHFAL	104
167	TR5046 c0_g2_i1	LATGELALQNFVVLVGGDEIPRVHCCHSHLLVDGPQLVPETHFEAEQRLAVLE AQPWPPNLLFPQKTEEISKSHQFP RSEEHTSELQSPCNLVCRLLEKKK	102
168	TR5077 c0_g2_i1	MAFRMTGRVLGALQKERNDIQPALRKIWKTYEHTDGLVTRTISPYHLNPFGR SFAQSIPKCLKHRFIEHGVEVFMCGA IYGTYTAKKGFQDLHKDHRD	99
169	TR5130 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLES DIKRRLRLHSPSRGDWICADIAMKE PRIKGQVGTDRGGKLAHE	71
170	TR5176 c0_g2_i1	QNLTSIVESKAERARFGRFYRFPGLGESGLDVYNRATSFIA TMFRDFANQNI AREDLNVIIVTHGLTLRLLVMRWFQY	167

		SIADFEQTINPANGAFVVMERRTKKKTESRFFESTGERRDLGHWYELTPESR AHINFGEQRQFGSLWKLDRKSTRLN SSHLVISYAVFC	
171	TR5214 c0_g2_i1	FFSSRRRHTRLQGDWSSVVCSSDPKTSDPYLRGAVAVALAWRARAIHFV VILPTEGRRAATPPCAC	68
172	TR5234 c0_g2_i1	MDWKGLPITVKLTIINRQATVAVVPAASSLILKALNEPPDRKKVKHIKHD GNVTLDQVIDIAR	64
173	TR5267 c0_g2_i1	MDGSRSHSGRSARIPARGIAESPQQHADGPPSDDDNAAYLKNMQELLRKF WSEQMQEMQELEIGTEQDFKNHN DLPLARIKRIMKSDDEDVRMISAEAPVLFKACEMFILELTLRSWCYSEKNK RRTLQKEDIQAAIHKTEIFDFLVEAFSA GMS	155
174	TR5368 c0_g2_i1	MSDRATTETIQQDWKEREMIEIVHLNILKITKFLNEFEMSVRYKLSALAER VGTLEHLEFCETAVRTKEGQQGQNS	77
175	TR5378 c0_g2_i1	FFSSRRRHTRLQGDWSSDVCSSDLLDDGPQLFQVDPSGAYFGWKATAIGK NFVNAKNFLEKRFSEDAELDDAVHT ALLTMREGFEGEMTEKNIE	94
176	TR540 c0_g2_i1	MLDTQPILTKSLTSLVGFVAVGDILAQKFLEKKTEIDLPRMLKLASFGALIHG STGHFFYNFLDTKIPGTAALTVAKKVF IDQVIWNPIFGCFFF	94
177	TR5443 c0_g2_i1	EETIFFSSRRRHTRLQGDWSSDVCSSDLTVNSNLTSQTILSAIETTHIEHTV VVGDRHSGELSTT	66
178	TR5462 c0_g2_i1	FFSSRRRHTRLQGDWSSDVCSSDLLVLEELVPHETDNETRLPHGSIPKKHQ LKVTHSPRSHSPVNN	67
179	TR548 c0_g2_i1	FFFQAEDGIRDYKVTGVQTCALPISPPTFANDIVLTKAAAYVSDSVQPQL YKIPLDEDSGKLASGKDTT	69
180	TR5485 c0_g1_i1	FSSRRRQTRLQGDWSSDVCSSDLADPSPHRMLPLGFLRLIRHACDAILD GLDVSKVTAL	60
181	TR5486 c0_g2_i1	FFSSRRRHTRLQGDWSSDVCSSDLPNTHSLLPLTFPLCAYQSQISPVHN NPVHACPDSVTLTHVRFKNS	70

182	TR5496 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLSQNRRENSLQKRHAQFSGQKHEHM Q	50
183	TR5504 c0_g2_i1	FFFQAEDGIRDYKVTGVQTCALPIYEGDVFTCKRSSACYVKGRLQPTRSL GDAYLKYPEFNGQEDTHRSAGRFPMP PYTPPYITAEPEIFVHKVDHNLDDFVVLASDGLWDHVTNEEAV	119
184	TR5521 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLKGFDLSTAEMLYDSFLFFATAVIHPQ SINASA	57
185	TR5544 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLHPFRAKNRPVPSTATTYTPNPFRAFL TVSPLTADPASENTLFCPLCGRRQG GLSGGPPNYMVTTSIKTA	95
186	TR5589 c0_g2_i1	FLSSRRRHTRLQGDWSSDVCSSDLASFRPLRHFRAAVVDPPLPKRPR GFLKSFQISRCGETKNSSGIR	71
187	TR5594 c0_g2_i1	CCFFFSSRRRHTRLQGDWSSDVCSSDLTNAVFLHPSLPSFYRRHSNAYCS NTSSTA AVRIMLHETENSFTDSTSSI HQSAKADKLQTRASRSASCTSRFPACTTRCRKK	111
188	TR5599 c0_g2_i1	TCGGESLALVRFQAEEDGIRDYKVTGVQTCALPISEVKETTLPTEDTQP WIISTMGAEDVEMNEEVPQETTPA EVTEVKEISVLDALKEV	93
189	TR5700 c0_g2_i1	MADAPADRGGFSGGGDRGRGGFGRGRGRGRGRGRGRPRDDDKEW IPITKLGRLVKENKIKSLEDIYLFSMAIKEHQIVDFFLGTKLKDEVKIMPV QKVSAAGQITRFKAFVAVGDESGHVGLGVKCAKEVATAIRGAIIAKMSV VPVRRGYWGRRNGQPHTVPNKVTGKCGSVRRLIPAPRGTGLVAAPVPK KLLSMAGVEDCYTSSRGHTRTLGNFVKAVYFALAKTYGYL	238
190	TR5701 c0_g2_i1	FFFQAEDGIRDYKVTGVQTCALPISGTIIMGGDDWEKWEAVPLPAVRFSP WGDYAEIFDWYPWMRNFYTTW EIEYNT	78
191	TR5728 c0_g1_i1	FFFSSRRRHTRLQGDWSSDVCSSDLWYIISISTAPGTAACLIITPTMGILYT KTSIDVVARSIKYKDRNDHVLQY TSA YFK	81
192	TR5732 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLRDLRRRYGAASEGVACQSSRLLWG ADVFTAGNPPCLEYDQIMKVDQS	73

193	TR5752 c0_g2_i1	MTEVYIKMVLFFSSRRRHTRLQGDWSSDVCSSDLPVHASHAGERLTPLR LR	52
194	TR5814 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLGSKGGSKQPLKVGDECAKPKFD LLTTNGKWRLVFTTGDVKTQKKIGGKLTYPYIPIKAVQVFNPDYTITNGVYL GSFPLLEFAGTFTW	111
195	TR5821 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDIEKLVGRERLKSSTHKIPLFSRLRTSHT TPNRREETNPCVPYQRKTSVGVAF HCTKTILSHYFHNIHSRKQVSTLFPK	104
196	TR5822 c0_g2_i1	YPFVQLRLSGFMRVNVKVSNGFHVAVGEGVMREGRHIHLNYDDAMTFN TSHTIHDLSFGEPFSGMRDRKSTRL NSSHLVISYAVFCLKKK	90
197	TR5833 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLRLPGTTAAFALLSMPAVGRMLALGP CTAGGSTRGRSKLS	64
198	TR5859 c0_g2_i1	RRVLFRSAEVLFFQAEDGIRDYKVTGVQTCALPICGSRDFVFSSRRRH RLQGDWSSDVCSSDLRKS	69
199	TR5872 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLAAKSFRRKYNPRLCPLPKTALLTCL WVLSSSS	57
200	TR5914 c0_g2_i1	VVFFFSSRRRHTRLQGDWSSDVCSSDLGVPLTTSRTGNITRSIHTLLQMMV AIQSSTLERQYFYCFQLITFKVFTYMQVQENTVILVRVRWHNSNWHSTVT GTVDLSQKSYTAKSTMVTENTKAAKVSWNRFPTPPSHI	139
201	TR5919 c0_g2_i1	MVYPSMLCYNGRRSLRGPRLALGHPPEEEVSWCKLGLTLPKPKYTKLWME DEDAAILFLSGVRAPQVPVAGS	72
202	TR5920 c0_g1_i1	FFFSSRRRHTRLQGDWSSDVCSSDLGGIPKAINNFTTERLLVRAHWYVL ECGVAPTVPQSTNSQRYLINDGG	72
203	TR5945 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLKGGANLLVYVMSVMGDSSSHAP WFTSLSG	55
204	TR5963 c0_g2_i1	MRQATTFWRLAGMTYLEYASKATTVVRGALKEPARTQALARDVVSYN RAVFENGKPLPKIGRAHV	65
205	TR6002 c0_g2_i1	FFFFQAEDGIRDYKVTGVQTCALPIWGALGKASGSLRAGHFHSPSDWQG WQGIERVAESLRSGSEESLGRRYEQGPRLQTQYPECTMPLEDPCINTT AIGMVEAGYATYYDDTCAAQGYNGCFTQDIPECRTCYLSVEVYMEATNA	336

		TEANLPGWGDCPCCVPTTLASDHYDLEVQNFNYSIFGIAETCTDSPTQAPA GTPAPSNLLTPEPTMLSETPAPSEALSETPAPSIETPSPQLMSPTPTPSMTSTA TSGRGDVETPTPSSTSTASGTPESTASSSSQGVGTENIEESAGGGDDGLDTG AKIGIGVGVA VGVVVIGLASFGVLKASGRI	
206	TR6015 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLRVTLFLSNDFEGRSLRIEGTNRPC	50
207	TR6032 c0_g14_i	MMKTPSRAAALLGLAVCCCLALVSPSDARSPARHGRRMQTSGTAATGDLEE FYFFHSMVEGDENVVLLLEEG	70
208	TR6053 c0_g2_i1	MAGLKDVMVAGLVIATFIFTSAMAFISPSTPMTCRGVARASARSRNSER LLMASRPNAKKEKRQRNRDNMRKFTKKGTSRKKLMRVQRQERAAEAES DFMAQVFHTGMDDWEEDDIFAED	121
209	TR6105 c0_g2_i1	QGLGDHNVTVLGHWGCESTNGEFSTLNTPAEQTDGWEWEMKERMLRP KSVFFPILDVLTDPNASLHGGYSANILGGRSEEHTSELQSPCNLVCRLLEK	101
210	TR6151 c0_g2_i1	MFRTLVLTIATLLASANAFLPAIPAYSVPRNQRCGVQMTSPSVSTPSVTS PTEGTFKFSKGA YVDTLKKMEADCLKDLEKA FEETGKKAQAESDKIWAE	99
211	TR6199 c0_g1_i1	TESPSTNQGVRRPIRNRSQLRRKLEGAAQDLASVVFNFDQGVPAAEFMDG DVAAQNGHLELLKSGQAKEFSTQAMDLAATNGHLEVVKWLHENREEGC TTAAMDGAAKNGFLEVFKWLHENRQEGCTGGAMDTAARHGHLEVLKF MHNNRSEGCSTLAVDFAAHNGHVEVVKWLFENRADRKSTRLNSSHLVIS YAVFCLKKKK	204
212	TR620 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLEWALNAETGESLPHFPVRLMNRKGI STPITLVNLHELTAASMSDGERGS	75
213	TR6207 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLLSGGRTELLFFSALRLADLSSLILHC WRLALRARSAPRGAPRA	70
214	TR6210 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLPCMSLIYAERTDAISDDGPGLTHNRQ RSGDRRWKKE	61
215	TR6214 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLGAKAKAVVPCDALEGGGTAEAVRV QAMQDCVDLMTEKSKQGNEMV	71
216	TR6241 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLRVTFSTAAEGVFFSRRTFDITRPSGR KCRPTGDGKGAASAASGP	70
217	TR6250 c0_g2_i1	MMFLHCCGEHLAHRACRREHRPTVDGLHRTKKELDYRRHSFLRSEEHTSEL QSPCNLVCRLLEKKN	67

218	TR6272 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLPHKTFFFNSNQSSTPPRKMMADIVPA RLSYSYCPFE	61
219	TR6282 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLAHRPICAAHVQPLNLPQMCSSTT HPMNAKENTNTTVGPIRRPGASSV	74
220	TR6325 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLKDAAHGQSGQLYQVHHLPRPRL	49
221	TR6356 c0_g2_i1	FFFQAEDGIRDYKVTGVQTCALPISASLELAVTYYQYDVPVHGIFPSFMRITS IDEIQRKPII	63
222	TR6358 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLAIIFVRFLCVEAKRVTGEWWYHASAA RIAAHPMHGVRLVYASRALRLLFP GRGS	79
223	TR6428 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLAHSAMQCKAREQPRLNNPFTLTAFEA AEKSSRQKGL	61
224	TR644 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLVNTENTEFSAGLLGGIAGFVVGGPVIA ALLAVAANYGSKRDNEAGQAVRSISANAIEALNFITINSKYDVTGKAGASL DDAVSKLKESGEAGETVGKLEKT	127
225	TR6471 c0_g2_i1	IILHESCFFCSELARVVGFMVFFFSSRRRHTRLQGDWSSDVCSSDLLLPAHG FTVIIVIASVFLTVWAGVCVGGARK KFE	80
226	TR6492 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLFFESSTHSQASWVSARSRGSMNLQR RRCLILYSPLARGTSGSVVTSSPR WT	78
227	TR65 c0_g2_i1 m	MVDVFFFSSRRRHTRLQGDWSSDVCSSDLGMTVKASVETCIRAIEDSLG VAPEVGDTSPTSGRTGSQHASATELL DSVIHELEASLRTEGTGVKANGVGGVA	102
228	TR6503 c0_g2_i1	MEIATGPGKSLLMNCFMMYMSGGSINMISIMITFMGIMNPVKAIVNIPGTF RSVDDGKNSLLLPKLIYSALNFLG VCAALWKMAKMGLLPVTAADWTSYLMPKNVIEYSGVPL	113
229	TR6514 c0_g2_i1	MKRVDGDLAVSRALGDFQYKDEDLPPERCKVTAVPDTCIRRSPSDDEF IVACDGIWDVMTDEECARTIRQIFL DGESDLGTVCEEILDECLSQ	95
230	TR6534 c0_g2_i1	YDTNGDGVLRDEFAPKADLKLKLSLFSERAQEARLKEAEERVKEIEEH	74

		GDRKSTRLNSSHLVTSYAVFCLKK	
231	TR6547 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLRLGVFNFAVVYIATGAPLVTKLFLFF HGMVTSVVQVSTLFTFVTLAA	72
232	TR6551 c0_g1_i1	GGGFGGGGGPPAPRGFGGGRQTEFRTVVVTGLPMSASWQDLKDHMRKAG DVTYADVVDHKGGGVVHFGNRD DMEYALRKLDGSEFSNRYDSASIQVLPHRASVRYFIRI	107
233	TR659 c0_g2_i1	MKLNIAASPSTGSMFIDIDDEKKLRALYDKRISQEVEGEALGDEFAGYIFR TAVPLALS VWTSSGIDA	68
234	TR6599 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLPQRRGVRVIWGYNYGVCVESPSRAC	50
235	TR6619 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLWKRAFGGIIVWCGSSAWN VYGEMS PAVIDLDPAIMGKLGSGRCDLG PNGVIVRLCGEAVRAFELISQRKDQLGTNAWLLQR	107
236	TR6689 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLRETAGAITVLWFPRSR TMRKSIASAI GAAVLFVLPFVGGFT	66
237	TR672 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLLHLSNLVTVVGTGVVAPVATHVG SSGGVILDGGGVVSPVAASATIFLHLS APSRAG	81
238	TR6742 c0_g1_i2	MVRMSILRLAVSVAVVLSIVNPSTCEARPSPLEKLDEVDLADLLDAPLD ETFLQRKLEGGTTPAPTAPTEADREMMFSEVPTAAPTAVTREM FSEGPT AAPTAAPTMASTSGAMGMVQAAASLATAATV LVATAVIA	137
239	TR6824 c0_g1_i1	VRSAEAMVKKLQPATSGTVGRKKSATSSGNRKSALKKKGKAGTRSSKA DAGKKGKAVGALTSVFGGLYGGAKRAGALAWGSRNLMFFVSAVAVI HLRGDYLAV	102
240	TR6833 c0_g2_i1	GNMQGTDTALEWERNRKKTILEYLPKLVLVGLI WVVTILVYIYVRVQQ TGDPSYSTYDDL SRSEEHTSELQSPCNLVCRL LLEKKK	86
241	TR686 c0_g2_i1	LLSILGNL GILRQRLLDAGVDGDWQKAVLLPQLFCVSRVLRHLVFLGT VYRFSSKVGKCPWTRDTQLVPHGTQVPQIYLSALVSRQH LARSEEHTSE LQSPCNLVCRL LLEKKK	115
242	TR6860 c0_g2_i1	GGETGVTNGQPSAEQSLADSGIQSSDERRESKSSRNKGS DGTDARNNG IGSHGASGVMARSEEHTSELQSPCNLVCRL	79
243	TR692 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLQNIAVTRLSFTGT SAPVRRNMICLL	55

		CGNPP	
244	TR6958 c0_g2_i1	SSRRRHTRLQGDWSLDVCSSDLDAIGGSRFSEGTSADREIQRTLME LLNQ LDGFEDLGQVKVMATNRPDILDPALLRPGRLDRKIEIPLNETSRLDIM KIHSAKITKRGEIDFES	117
245	TR697 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLRELPSKMTSLLTCAFFETEGGELRLN EAQPRRRAAPTLGHGRVCPNRI AAYFGARVQG	83
246	TR6981 c0_g2_i1	IFFSSRRRHTRLQGDWSSDVCSSDLGRTDIFDTGPACEGKDGHKRGTYGY	51
247	TR7021 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLLNGKSSRACTSEKMAVNTAPVKWAQ RSDSIYLTLDLPDVKDEQLKLTKEMLSFSGTSGGKQYALDLKFLHVNSEK STWKVLPRSIQ	112
248	TR7040 c0_g2_i1	TRLQGDWSSDVCSSDLGRRRCGHPSRWNSQAVAPRGWKVRGRGLRDGLR GYTNGCGPCLVRGGLGSAPP SRSGCGGARHYVLGPPNDEDNHVGDASTPLHRKGLGLQSQ	110
249	TR7049 c0_g2_i1	RVHYKHMREVAHLIKGMKLSKAKQYLQDVLDYKRAVPFTMFTGGVGRH PQGKLP IAKRSEEHTSELQSPCN LVCRLLEKKNMPTRATPTKLHSMDCYVPV	102
250	TR7059 c0_g2_i1	FFFQAEDGIRDYKVTGVQTCALPISFFTHDLTRLSFITHSSGAYLSVPLNLR LVVELLATDKMVRLMFLVVGATLLGLCSNVNAWRGKKEGEGEPMPKG KDALDLARAGFQHAAGDVGMLQDLLKDLQDPETMKEVEKLMKDKDFT REANSYLSSDEWKKGSADAMKILSDPEAVAAMQKQANEYTDYLA KGPG KDGKRDAAMGLEAAARGMRDPGFLTDAMKMMNDPAVMKQVEETMKN PEFQKQREIMEAPAMKQVFTEAGANMEQLMKDPDLMATVQQQIAASM NEAAMA	295
251	TR7085 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLSTFCFATPRCFREFPRAKFAGSIRLLIV TCTYPKSFRGHQR	66
252	TR7109 c0_g2_i1	FFSSRRRHTRLQGDWSSDVCSSDLTSHHSSQKQALSSGKQHQT VVVVSN STPCSHAD	58
253	TR7110 c0_g1_i1	QHRYTFFFSSRRRHTRLQGDWSSDVCSSDLATRCGGTRSRPNVEEHPDK RLRRRAADRRDRVLETISTT STP	73

254	TR7116 c0_g2_i1	LADGEGDDCVTNTYQLIGKFLALRGPDKDGHEVDCVEHCDFWYWLQS KGINSNRSGIVNAIAEKVNT WIPGIYDADSYTA	81
255	TR7117 c0_g2_i1	FFSSRRRHTRLQGDWSSDVCSSDLGHRSNTPVGMSKLLKSLPASGPKLVL RKRHIDDGKRATAAVRKSP	68
256	TR7120 c0_g2_i1	MQIFVKTLTGKTITLDVEPSDTIDGVKQKIQDKEGIPPDQQLIFAGKQLED GRTLSDYNIQKES	65
257	TR7123 c0_g2_i1	FFSSRRRHTRLQGDWSSDVCSSDLPTRTPFLTPIIARDTYTLVPYTVYINIG MSH	56
258	TR7151 c0_g2_i1	FFFQAEDGIRDYKVTGVQTCALPIWAWQLCLEIVAIVGCLTNLVLAVFVCE HVSLYTPDSLSDHFGSMEAKF LLVVGVEHLLFVGKAVASGVIDKVPRVVR	101
259	TR716 c0_g2_i1	VAIQNPRQFKIPDWFLNRQRDWKTGKTTQISSQQIDQKLREDLERLKK IRSHRGIRHFWGVRVRGQHTKTT GRRGKIVAAGGRK	85
260	TR7194 c0_g2_i1	FFFQAEDGIRDYKVTGVQTCALPISPSESQYALVHGVAALDRSSDAAA TNERQRLHRLAVAQRLGRLPAAVH	73
261	TR7233 c0_g2_i1	TSYGIHMFDFLFFFQAEDGIRDYKVTGVQTCALPIFQVSSIGFASGFLSSTD TVYRSSNLGKSVLTCMPFRSDLEL LLAVARMISFSQYVTIRCNAMLHAK	101
262	TR726 c0_g2_i1	FFSSRRRHTRLQGDWSSDVCSSDLGDGDQESPKRARSDSENPSESKNTS LLED SGVVDEDSVTGMDVEDR SN	73
263	TR7293 c0_g2_i1	FFSSRRRHTRLQGDWSSDVCSSDLRCNPAAVAAACSSLTLCCWRSWN RRDTLPVSLMNFSAQTPMNLASL TSSRLLRKLLEQSSKHLKTRLLYIASESFL	101
264	TR734 c0_g2_i1	FELILAAANYMDIKPLLDLTCATVASMIGKTAEEIRRTFNIVNDFTPEEE AQVREENKWCEDA	63
265	TR739 c0_g2_i1	RREMSVGEANEQMMYK VFFFSSRRRHTRLQGDWSSDVCSSDLATSHS VNRPAATTALPLVRITTSHTA YSARTTRHWCDHRRQG	85

266	TR761 c0_g2_i1	FFFSSRRRHTRLQGDWSSEVCSSDLQTVVNRDPPLSPGPREEPEHLLVC PQLYTV	56
267	TR763 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLPKCRTIQTQRGGVCFLICSTHALFR QRGERVRVRLNSKSSRPVSSS	71
268	TR772 c0_g2_i1	VCYGVRNFAGDEAFFVVDTRAGNKTVGWHEVTEQSLAGCGANSTPSYS GAVRSEEHTSELQSPCNLVCR LLE	73
269	TR779 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLCYRGMRHFLFDRRAAPRRYPPCNSRA SQLQPLSRTPLMPQRQRQTL TRK	73
270	TR875 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLGIPSASLFSPKKLSRSCRALARRFLL SGIIM	57
271	TR876 c0_g2_i1	AGGEVADALVLEVLHPYRVAHLVDVLESVA AVTIHVAVPIGRTPIRK EDGNLVRALGAERSEEHTSELQS PCNLVCRLLEKKK	86
272	TR889 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLLPPPGALPLEEALRDGIVALAALPI	52
273	TR898 c0_g2_i1	FFFSSRRRHTRLQGDWSSDVCSSDLLRRLPASPITTALHPLMVLPHHSQDR	51
274	TR905 c0_g2_i1	RENVSDQPDVGGQSENKMTYGEWVGGFLAQGMVYTRLFDERLGVSQ TAKALAGKVQKFDETHKVSQTVTAAAMTAAEKARIVDNRLKVSETA KWAQKTKESTGRAIESNPRVA AGVKQAGDAFTKVVTDVSEFAQLSLRGGLEPPQQPQQGGDRKSTRLNS SHLVISYAVFCLKKK	178
275	TR92 c0_g2_i1 m	FFFQAEDGIRDYKVTGVQTCALPICYLRVVKYCHTKGRTPRHFQFDPSG NFLLSANQDTSITIF	66