Next generation sequencing of SSH libraries identifies spermatogenesis gene transcripts differentially expressed in inactive *vs.* maturing male gonad of the scallop *Nodipecten subnodosus*

Raúl A. Llera-Herrera

Alejandra García-Gasca

Arnaud Huvet

Ana María Ibarra

raul.llera@gmail.com aibarra@cibnor.mx

Nodipecten subnodosus
almeja mano de león : lion-paw scallop

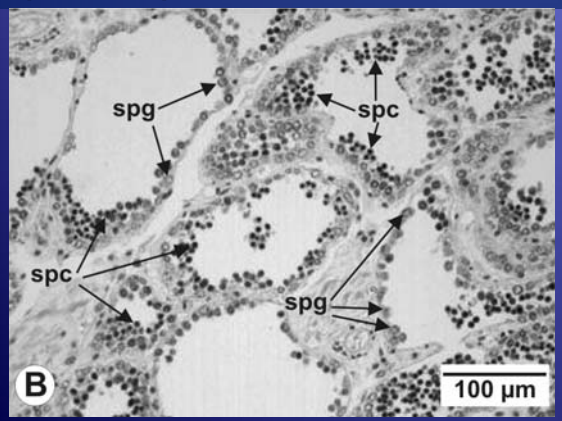
Petersen et al. 2010. Mar.Biol.

An important fisheries and aquaculture species in Northwest Mexico

Large adductor muscle size and high economic value

Maeda y Lodeiros (Eds) 2011. LIMUSA

Triploids present 95% sterility

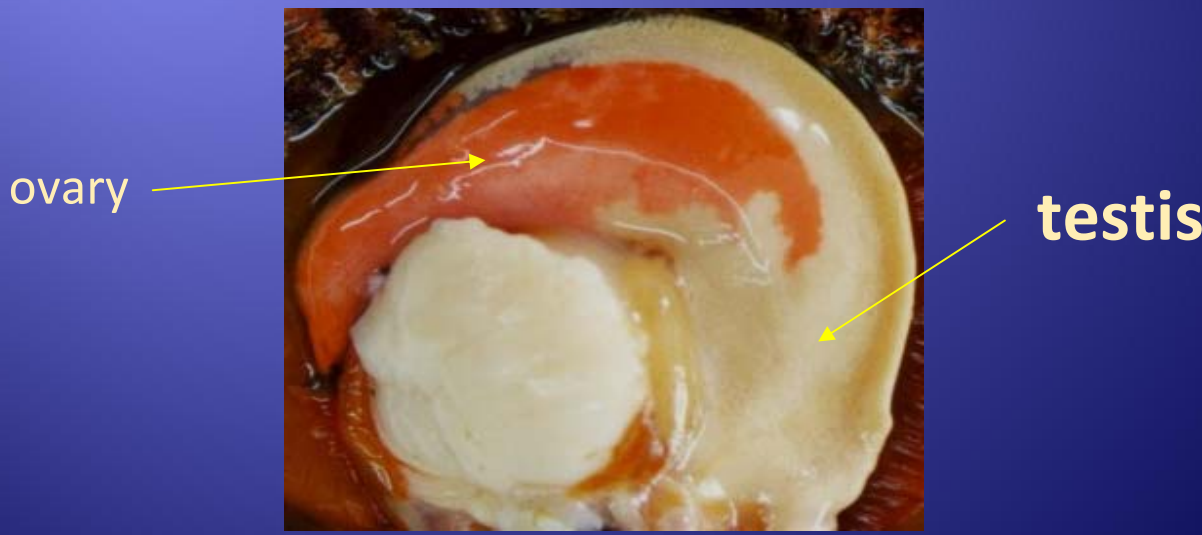


Maldonado-Amparo et al. 2004. Aquaculture

Understanding the molecular basis for such large sterility is important for future applications in controlling reproduction

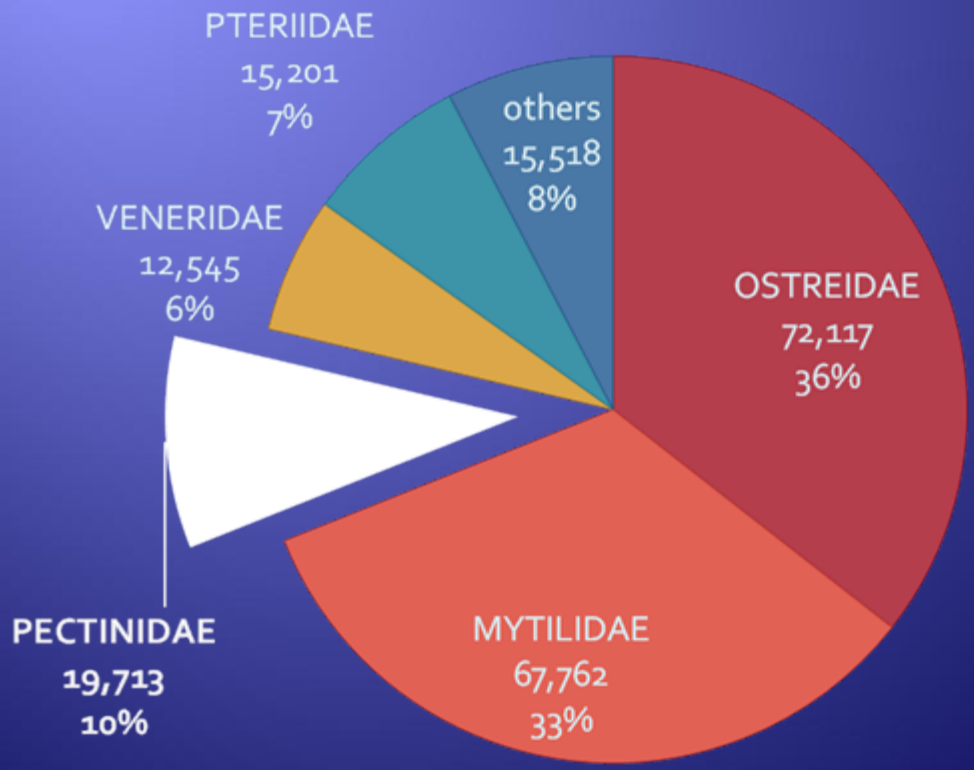
The lion- paw scallop is a good model for studying the molecular basis of gametogenesis

It's a functional hermaphrodite: Produces simultaneously both, sperm and oocytes, in the same ovotestis or gonad sac

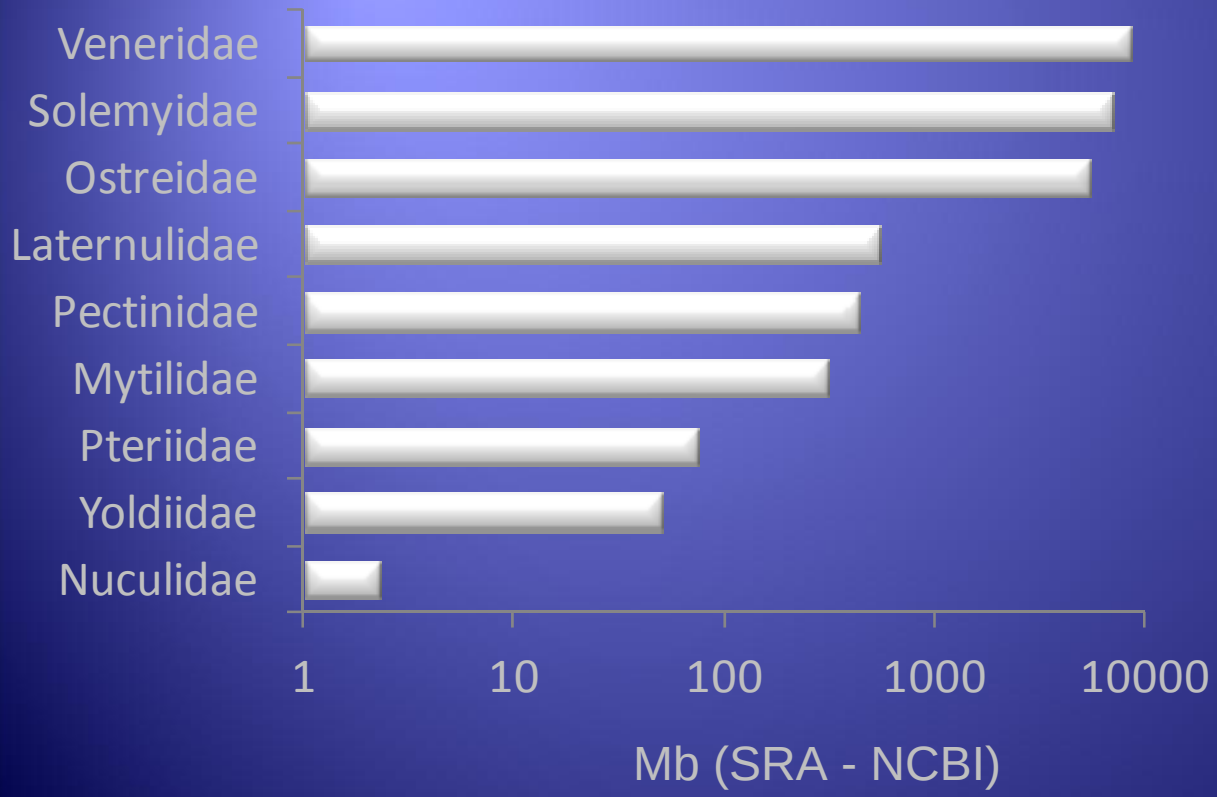


Transcriptomic information was not existent for the lion-paw scallop

October 2010: Transcriptomic information available for the BIVALVIA class (202,852 total EST's; NCBI dbEST)



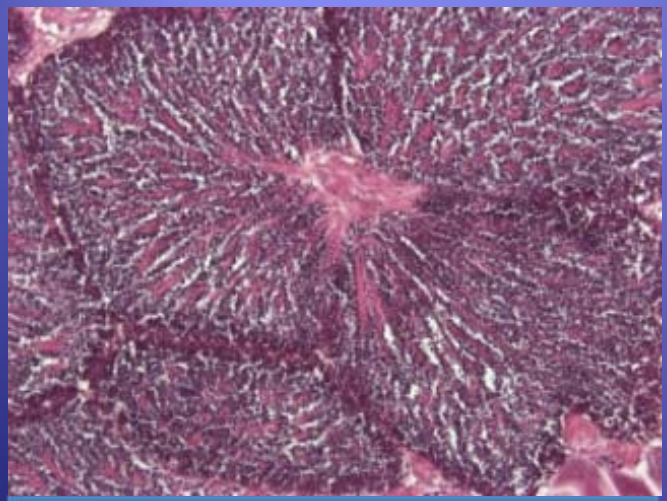
Today's effort in sequencing Mollusks transcriptomes - MegaBases



26	Illumina HiSeq2000 (1 exp), GAIIx
1	454-FLX: Illumina GAIIx
10	454-FLX; SOLiD; GA
3	454-FLX
1	454-FLX
7	454-FLX
1	454-FLX
1	454-FLX
1	454-FLX

In order to isolate genes up/down regulated during **spermatogenesis**
(meiosis, cell cycle control, and gamete development),

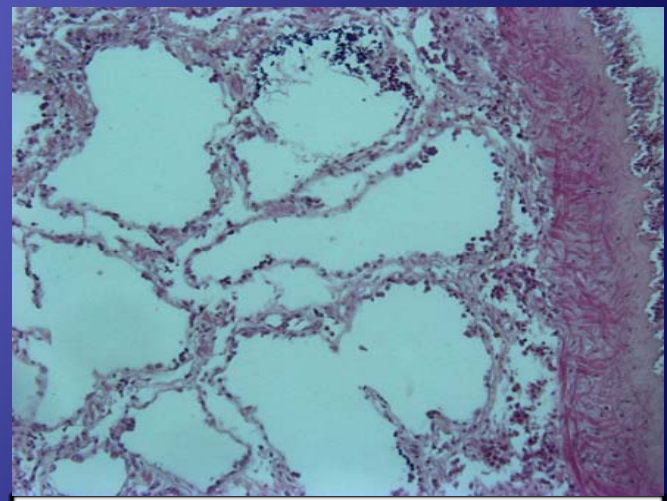
we obtained **two libraries** from the lion-paw scallop after reciprocally subtracting
maturing testis *vs*. **inactive gonad**



Male gonad in active spermatogenesis

A - I

I - A



Inactive male gonad (postspawn)

Reciprocal SSH libraries

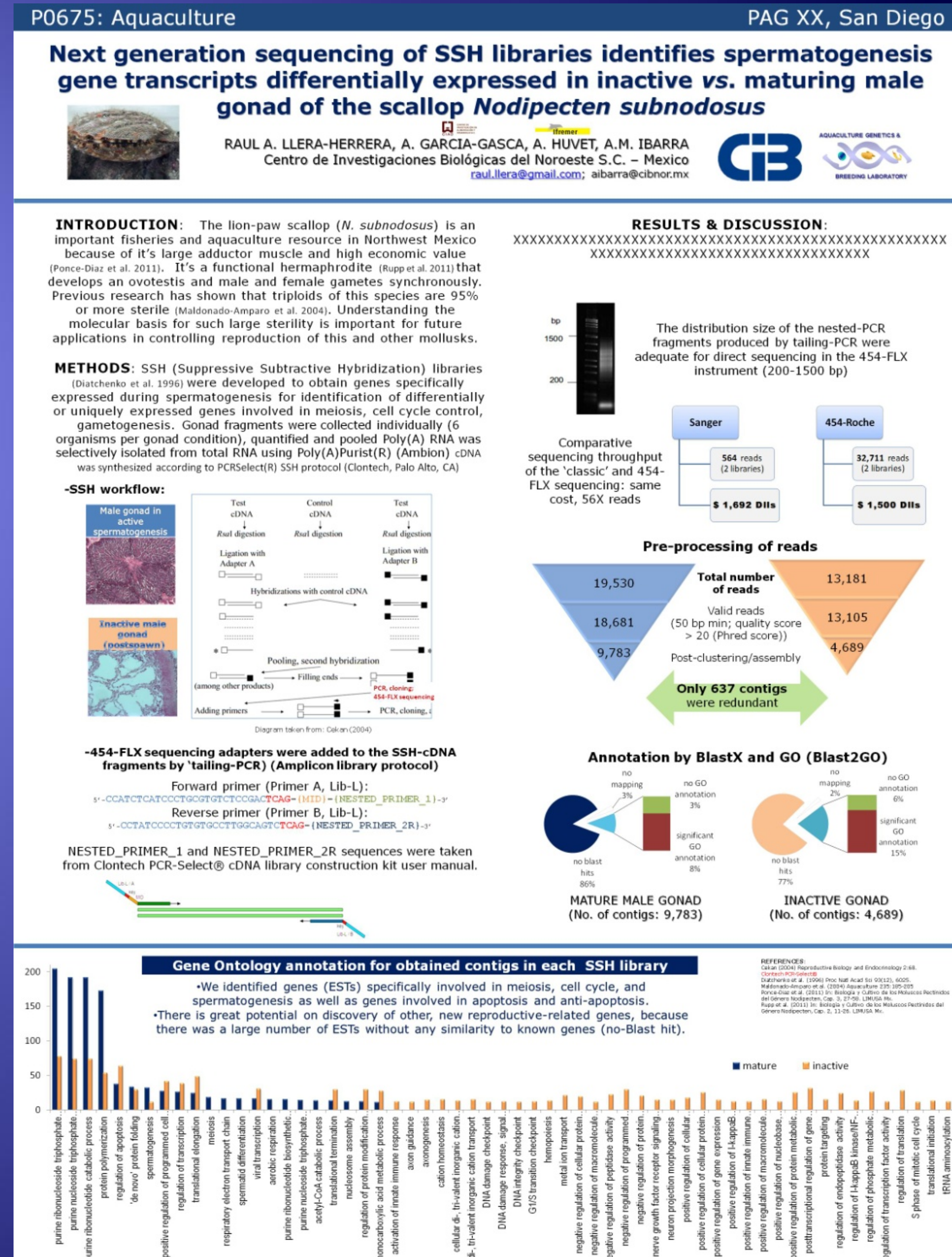
Why SSH libraries?

- Is a powerful method for generating subtracted cDNA
- Based primarily on a suppression polymerase chain reaction (PCR) technique
- Combines normalization and subtraction in a single procedure
- Significantly increases the probability of obtaining low-abundance differentially expressed cDNA's
- Applicable to many comparative and functional genetic studies for the identification of differentially or uniquely expressed genes

Lukyanov et al. 2007; Diatchenko et al. 1996

Methods

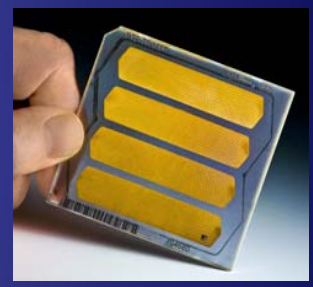
Please see poster
Po675 : Aquaculture
for details on SSH libraries preparation



Deep sequencing of the 'M' and 'I' SSH libraries by 454-FLX (Roche)

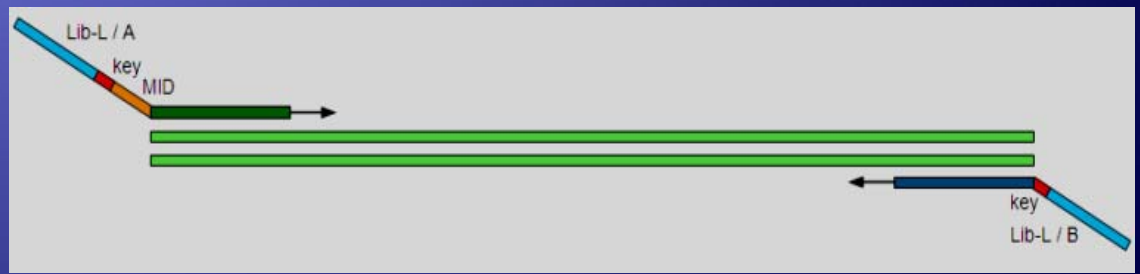
The 454-FLX instrument can sequence thousands of mid-size reads simultaneously - using an adapted protocol for amplicon library sequencing

Two ¼ lanes of a plate were used, mixing after tagging a total of 8 libraries; 2 of the 8 libraries were for the present study



454-FLX sequencing requires two adapters (Lib-L/A and Lib-L/B) in both sides of PCR-fragments; they were included in SSH amplicons by adapter-tailing-PCR

MID-tags (unique 10bp codes) were used to discriminate between fragments from each library (M or IM)



564 plasmid clones were sequenced also by Sanger

More details in the poster session

Bioinformatics Pipeline

Roche-454 FLX sequencing files (.sff)
(two ¼ lanes; 2 of 8 total libraries)

Libraries were split by using MID-adapters (tags library-specific)

↓

quality-trimming (quality score > 20), adapter removal (vector adapters, Roche and Clontech (SSH) adapters) - (*SeqClean, SnoWhite Perl scripts*)

↓

reads assembly (3 rounds of MIRA and one round of CAP3, chimera screening by MegaBlast; implemented in *iAssembler* script)

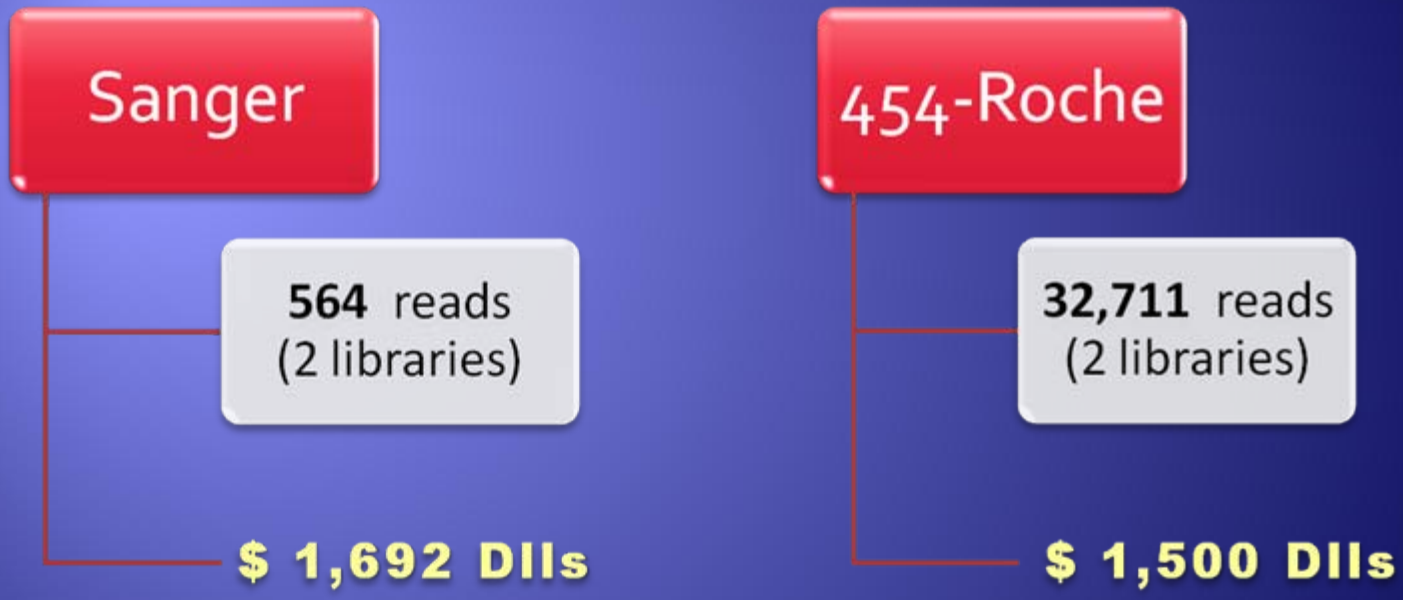
↓

BlastX and ontology annotation (*Blast2GO*)

Phred: Ewing & Green (1988). Genome Research 8: 186-194. *SnoWhite:* Barker et al (2010) Evolutionary Bioinformatics 6: 143-149.
iAssembler: Fei Lab (Boyce Thompson Institute for Plant Research). *Blast2GO*: Conesa et al (2003) Bioinformatics 21: 3674-3676.

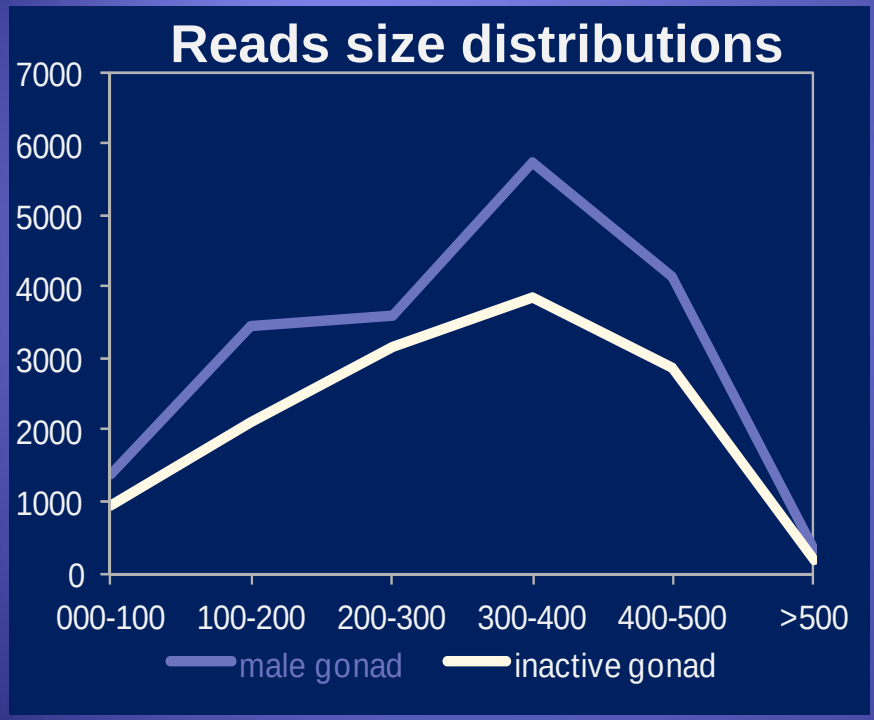
Results - Discussion

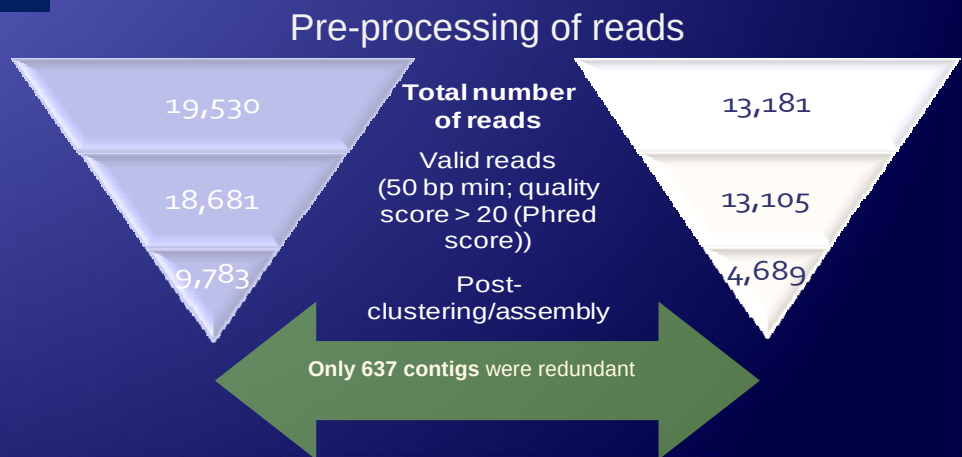
Transcriptomic information obtained for the lion's paw scallop (*Nodipecten subnodosus*) through two sequencing platforms:



Deep sequencing by 454FLX under a tailing PCR and multiplex strategy provides cost-effective, deeper coverage than traditional Sanger sequencing for the isolation of EST's in SSH libraries

454-Roche Sequencing Statistics



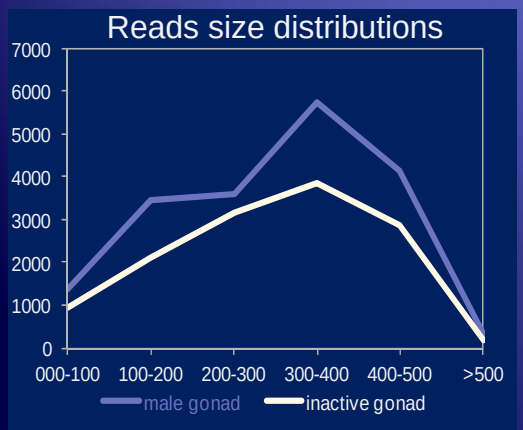


454-Roche Sequencing Statistics

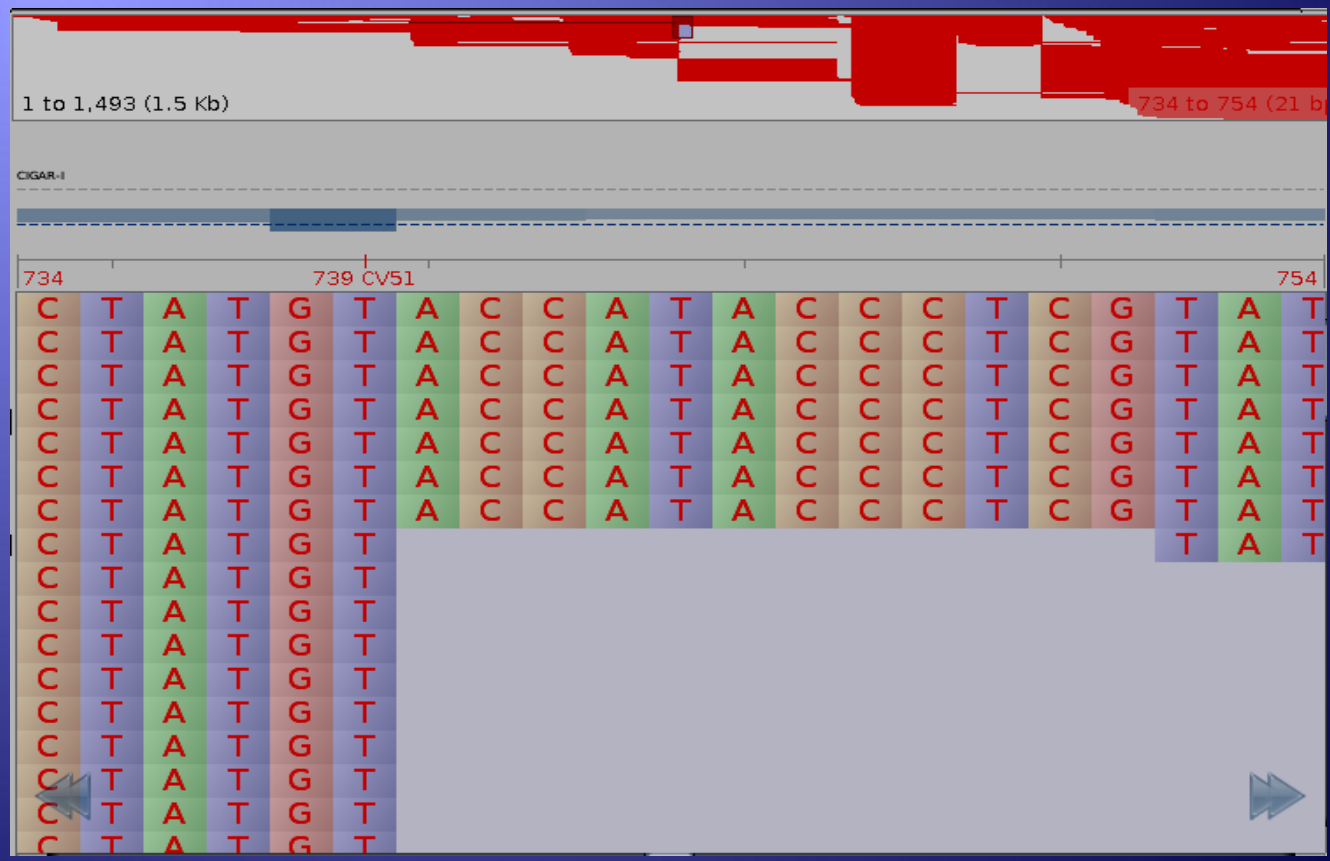
Pre-processing of reads

	Mature male	Inactive gonad
Total number of reads	19,530	13,181
Valid reads (50 bp min; quality score > 20 (Phred score))	18,681	13,105
Post-clustering/assembly	9,783	4,689

Only 637 contigs were redundant



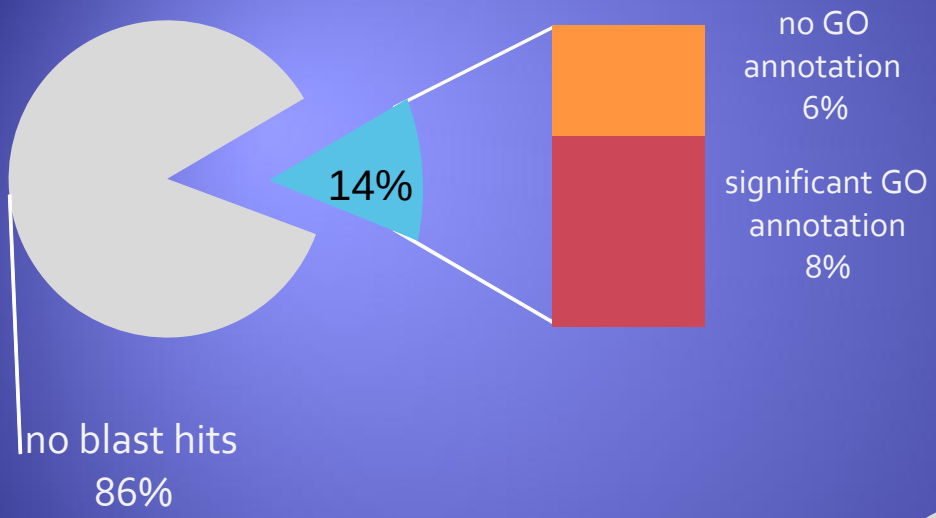
Even if SSH libraries include a restriction step of cDNA, it is still possible to obtain large contigs (i.e, ~1.5 kb) from assembled reads because of incomplete restriction occurs in some fragments as shown here, allowing assembly



***Rsa*I** restriction site: **5'..GT^AC..3'**

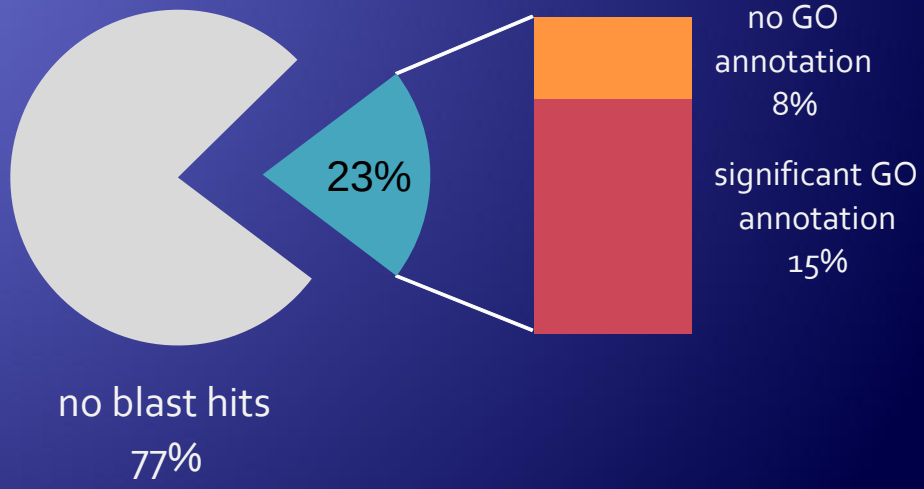
Annotation Statistics

Blast2go



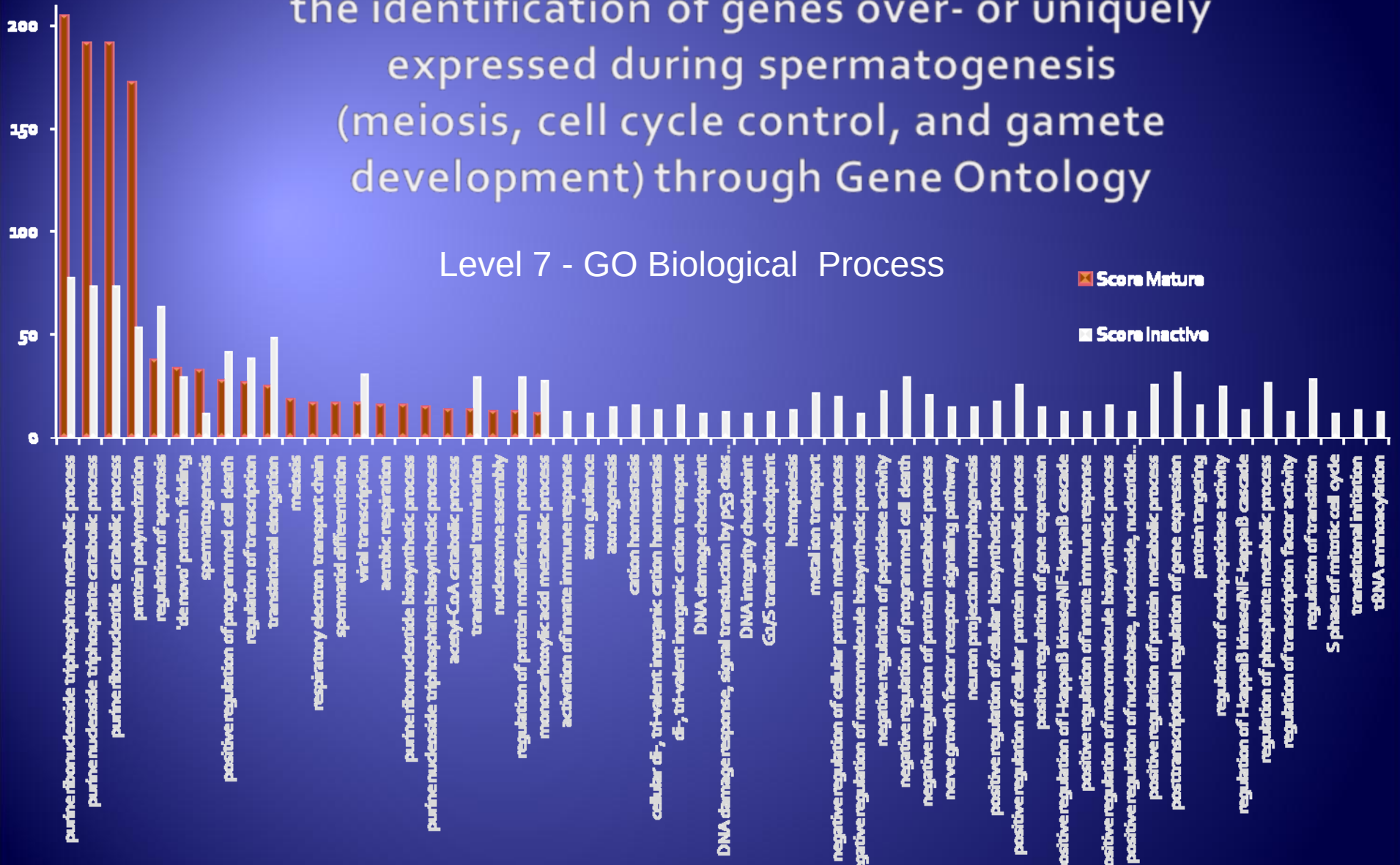
MATURE MALE GONAD
(No. of contigs: 9,783)

INACTIVE GONAD
(No. of contigs: 4,689)

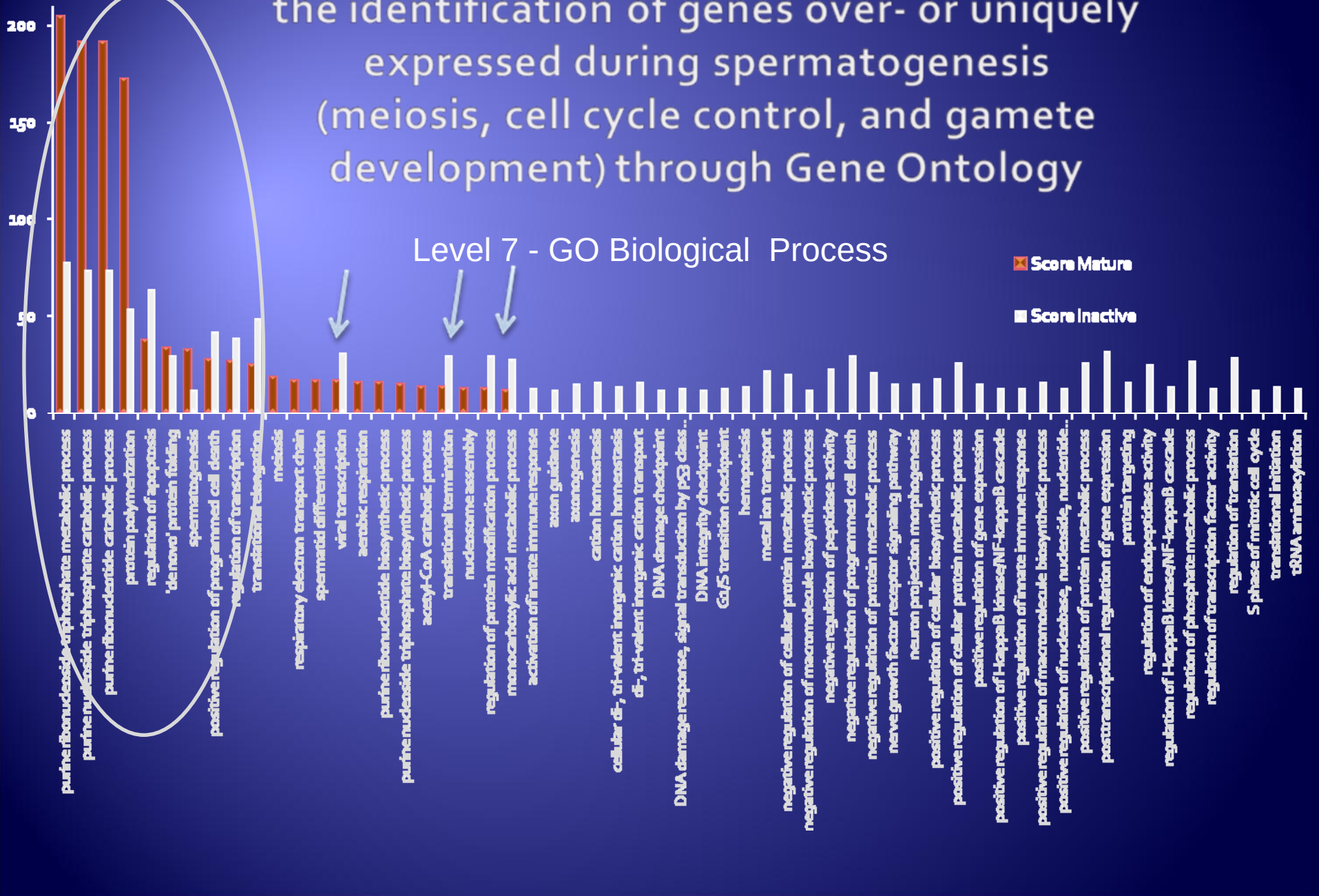


28% Blast hits for *Patinopecten yessoensis* (454-Roche full-plate). Hou et al. 2011

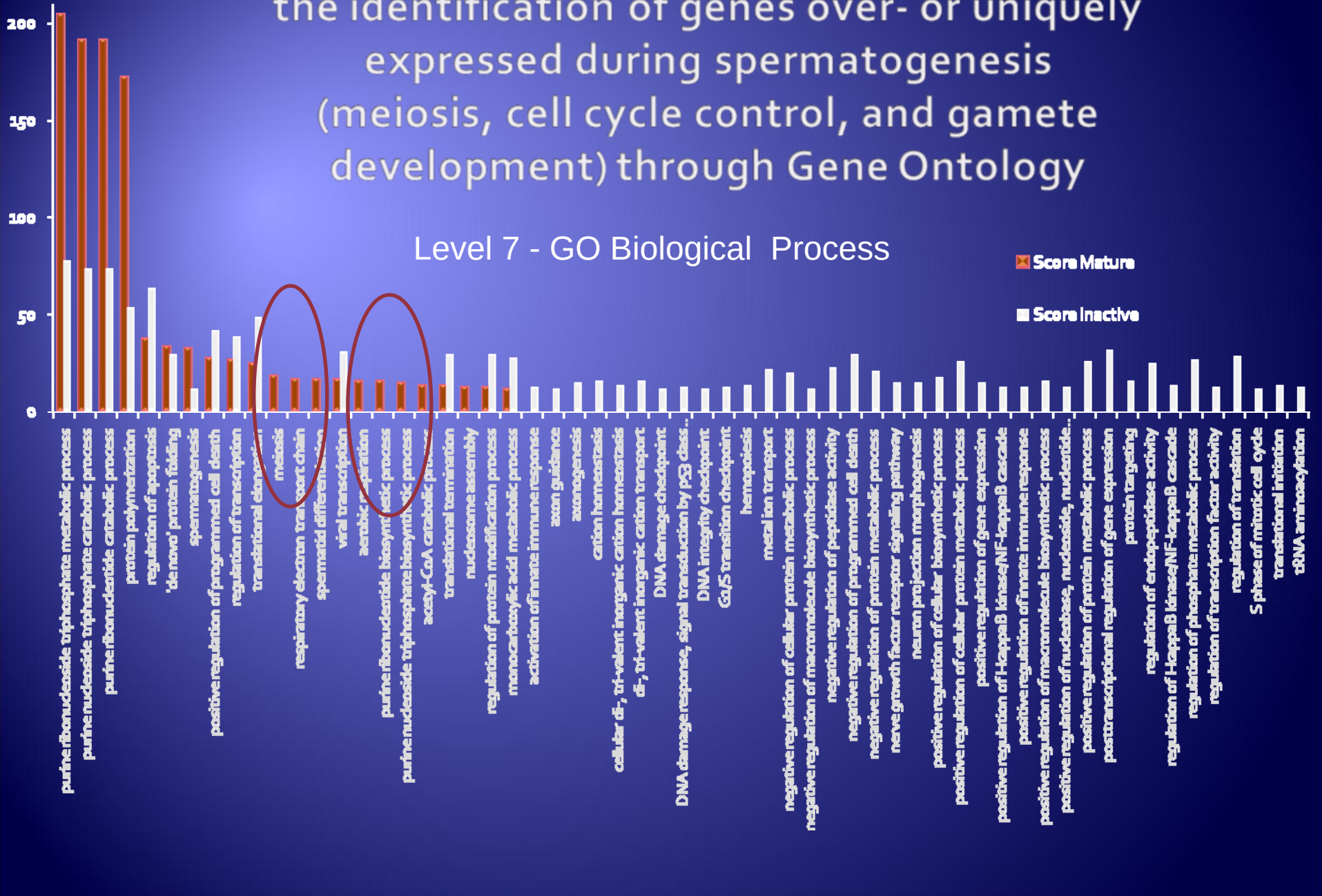
Suppressive subtractive hybridization allowed for the identification of genes over- or uniquely expressed during spermatogenesis (meiosis, cell cycle control, and gamete development) through Gene Ontology



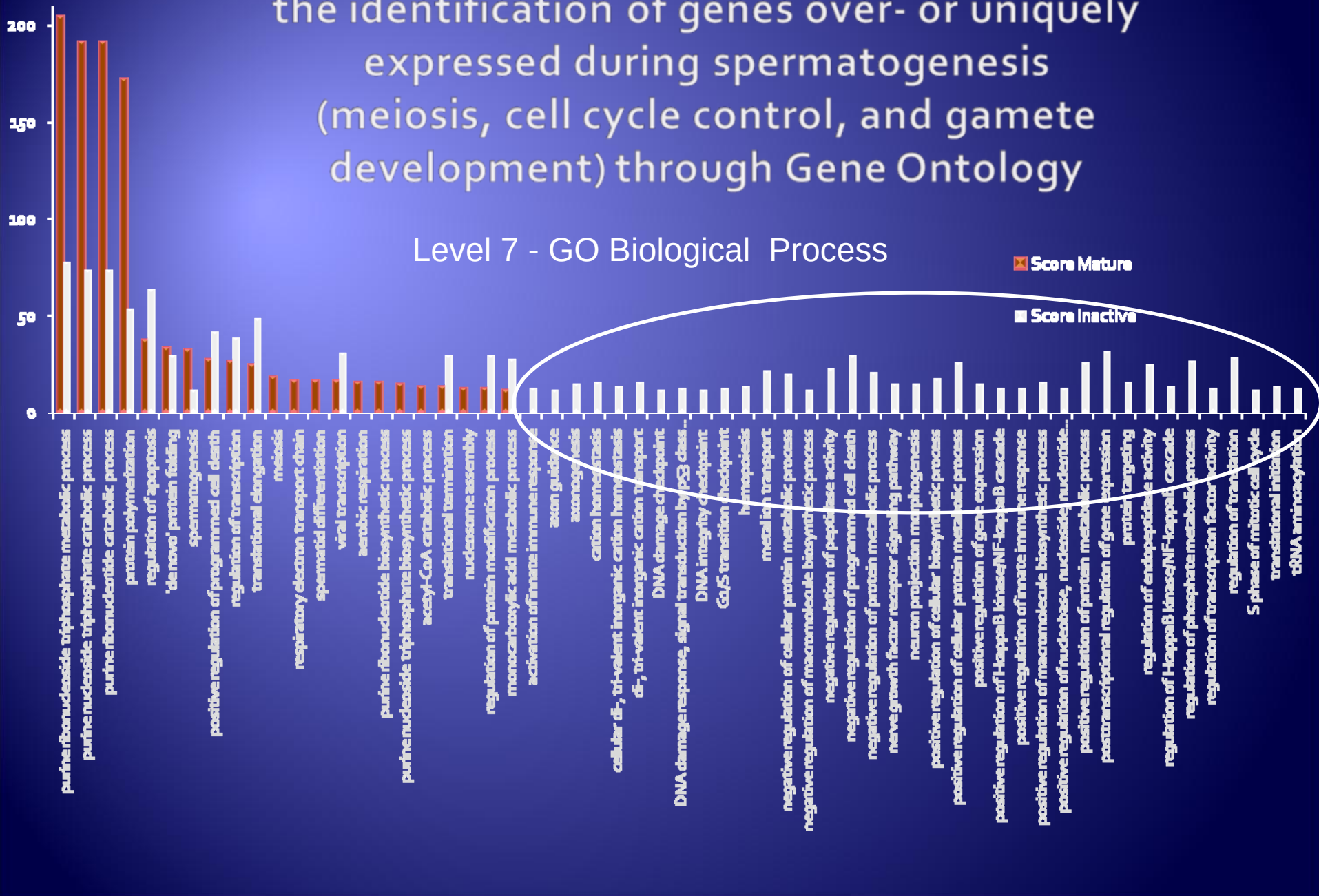
Suppressive subtractive hybridization allowed for the identification of genes over- or uniquely expressed during spermatogenesis (meiosis, cell cycle control, and gamete development) through Gene Ontology

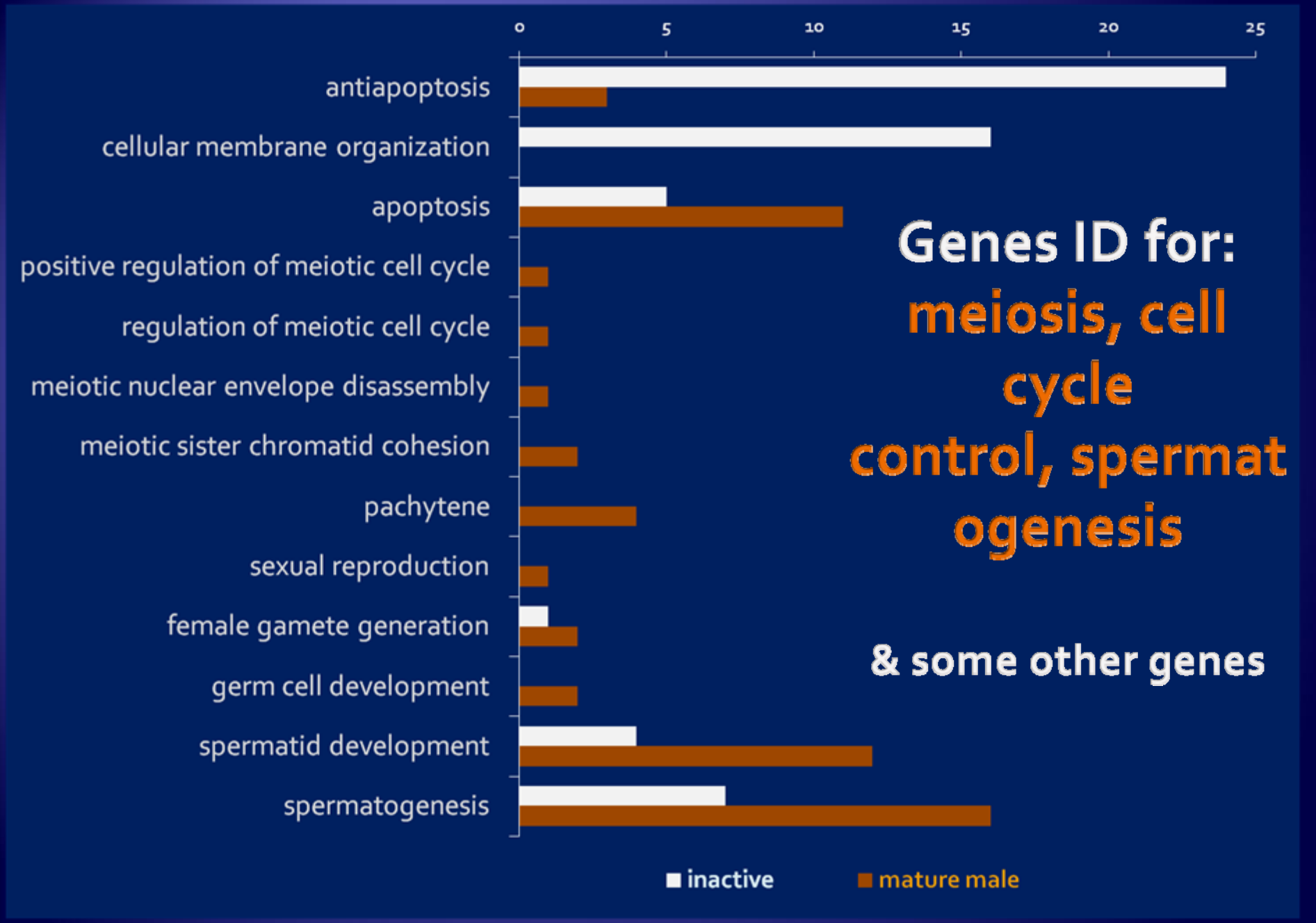


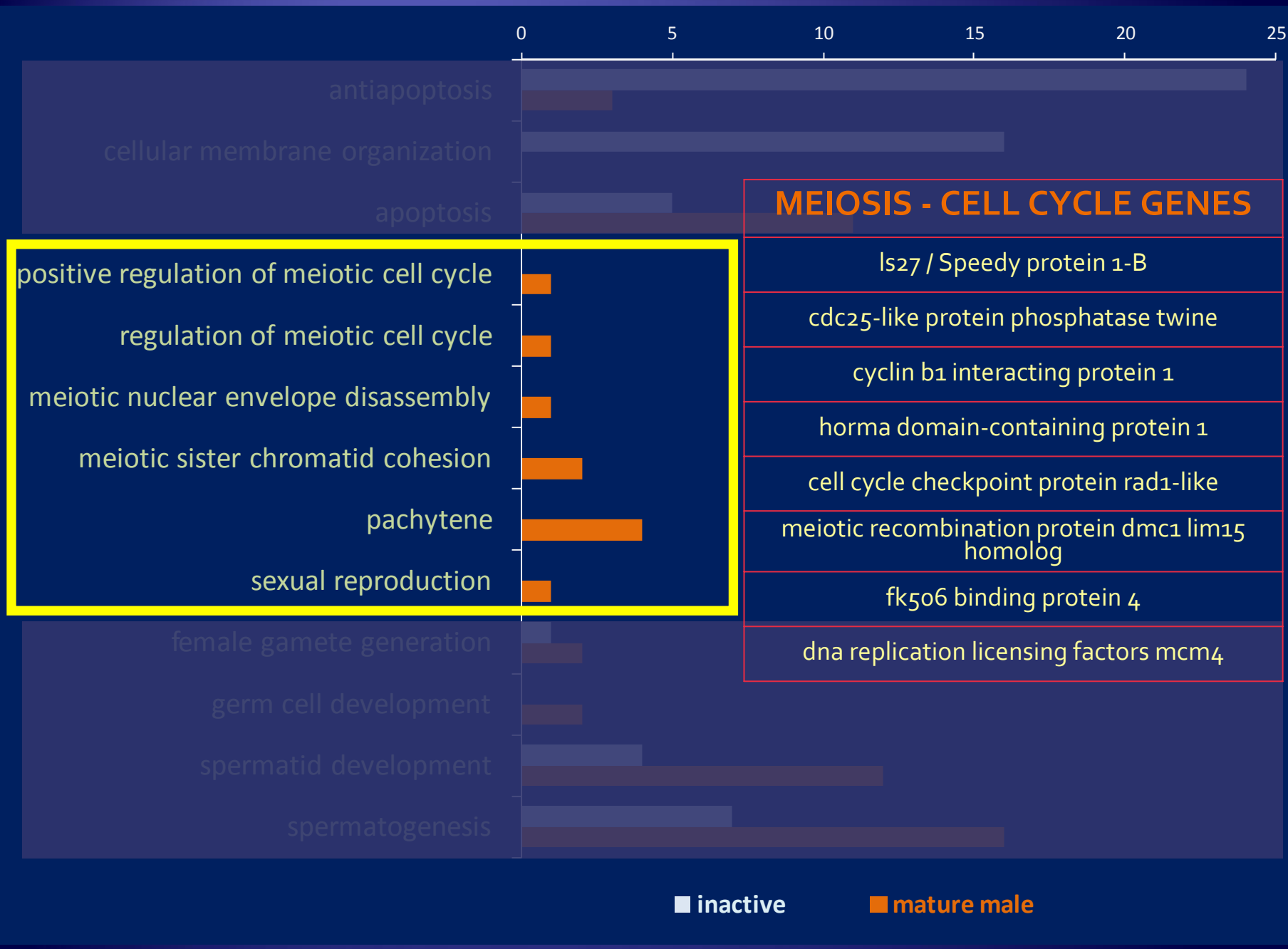
Suppressive subtractive hybridization allowed for the identification of genes over- or uniquely expressed during spermatogenesis (meiosis, cell cycle control, and gamete development) through Gene Ontology



Suppressive subtractive hybridization allowed for the identification of genes over- or uniquely expressed during spermatogenesis (meiosis, cell cycle control, and gamete development) through Gene Ontology

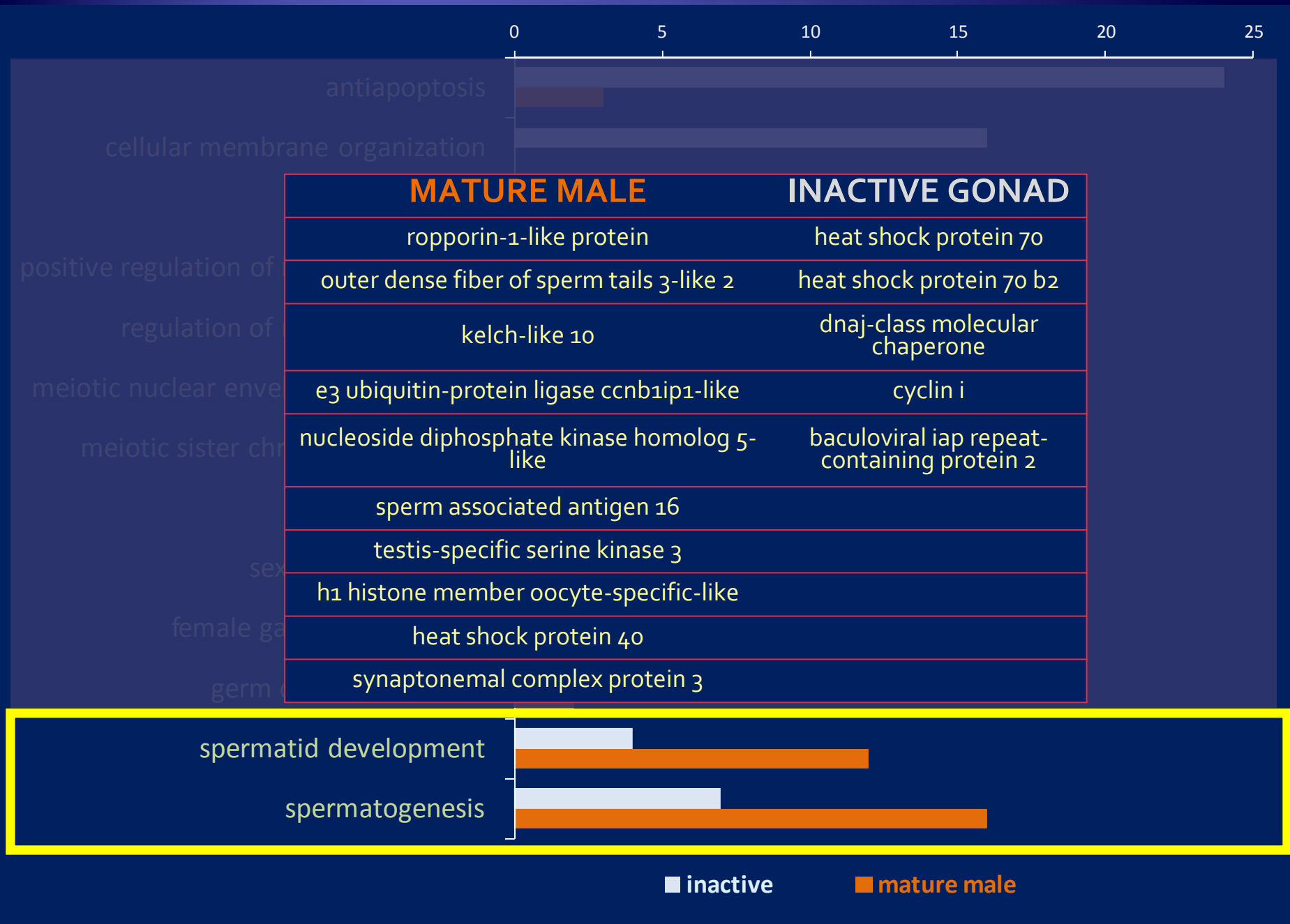



0
5
10
15
20
25
antiapoptosis
cellular membrane organization
apoptosis
positive regulation of meiotic cell cycle
regulation of meiotic cell cycle
meiotic nuclear envelope disassembly
meiotic sister chromatid cohesion
pachytene
sexual reproduction
female gamete generation
germ cell development
spermatid development
spermatogenesis
inactive
mature male
Genes ID for:
meiosis, cell cycle control, spermatogenesis
& some other genes

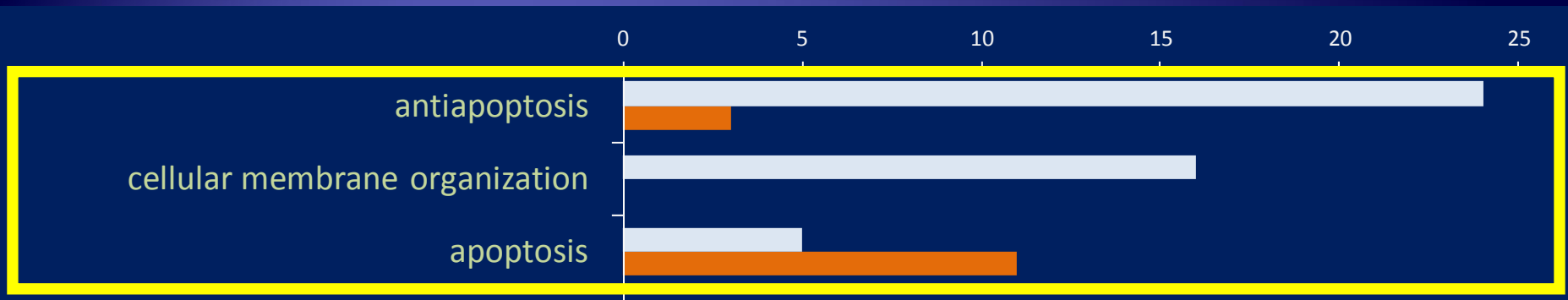


MEIOSIS - CELL CYCLE GENES

- ls27 / Speedy protein 1-B
- cdc25-like protein phosphatase twine
- cyclin b1 interacting protein 1
- horma domain-containing protein 1
- cell cycle checkpoint protein rad1-like
- meiotic recombination protein dmc1 lim15 homolog
- fk506 binding protein 4
- dna replication licensing factors mcm4



MATURE MALE	INACTIVE GONAD
ropporin-1-like protein	heat shock protein 70
outer dense fiber of sperm tails 3-like 2	heat shock protein 70 b2
kelch-like 10	dnaj-class molecular chaperone
e3 ubiquitin-protein ligase ccnb1ip1-like	cyclin i
nucleoside diphosphate kinase homolog 5-like	baculoviral iap repeat-containing protein 2
sperm associated antigen 16	
testis-specific serine kinase 3	
h1 histone member oocyte-specific-like	
heat shock protein 40	
synaptonemal complex protein 3	



MATURE MALE	INACTIVE GONAD	
son dna binding isoform cra_g	heat shock protein 70	ANTIAPOPTOSIS
nucleoside diphosphate kinase homolog 5-like	baculoviral iap repeat containing 2	
	tnf receptor-associated factor 6	
	ubiquitin c	
	bag family molecular chaperone regulator 3	
e3 ubiquitin-protein ligase ccnb1ip1-like	dnaj-class molecular chaperone	APOPTOSIS
trichoplein keratin filament-binding protein	histidine triad nucleotide-binding protein mitochondrial-like	
cyclin b1 interacting protein 1		
wd repeat-containing protein 92		
heat shock protein 40		
mitochondrial isoform d		
nadh dehydrogenase fe-s protein isoform cra_a		

In summary

- We identified genes (ESTs) specifically involved in meiosis, cell cycle, and spermatogenesis as well as genes involved in apoptosis and anti-apoptosis.
- There is great potential on discovery of other, new reproductive-related genes, because there was a large number of ESTs without any similarity to known genes (no-Blast hit).
- What is next: The expression analyses of this set of meiotic-relevant genes in the context of triploid-induced sterility.
- Deeper transcriptome sequencing with Illumina.

Acknowledgements

- CONACYT PhD studies fellowship num. 206375
- CONACYT Research Grants 48443 & 126427

- CIBNOR Graduate Division
  - for travel support
- CIBNOR Bioinformatics Group
  - Dr. E. Romero-Vivas & Dr. R. Vázquez-Juárez
- CIBNOR Aquaculture Genetics & Breeding Lab.
  - Susana Ávila & José L. Ramírez

Muchas gracias y hasta la vista!

Aquaculture Genetics

& Breeding Laboratory

Reproductive Biology and Endocrinology

BioMed Central

Review

Open Access

Methods to find out the expression of activated genes

Sten Z Cekan* 2004, 2:68