

Supplementary Information

Transcriptome analysis provides a blueprint of coral egg and sperm functions

Julia Van Etten, Alexander K. Shumaker, Tali Mass, Hollie M. Putnam, Debashish Bhattacharya

Parameters:

```
Trim adapter list = New Trim Adapter List
Ambiguous trim = Yes
Ambiguous limit = 0
Quality trim = Yes
Quality limit = 0.05
Use colorspace = No
Create report = Yes
Also search on reversed sequence = No
Save discarded sequences = No
Remove 5' terminal nucleotides = No
Discard short reads = No
Remove 3' terminal nucleotides = No
Discard long reads = No
Save broken pairs = No
```

Figure S1. The parameters used to trim all sperm and egg RNA-seq runs in CLC Genomics Workbench.

DESEQ2 COMMANDS:

1. EGG vs. ADULT

```
> library(DESeq2)
> CountMatrix <- as.matrix(read.csv("PATH::TO::EGGCOUNTMATRIX", row.names = "Name"))
> ColData <- read.csv("PATH::TO::EGGCOLUMNDATA", row.names=1)
> dds_EggAnalysis <- DESeqDataSetFromMatrix(countData = CountMatrix, colData = ColData, design =
~ Treatment)
> DESeqDONEEvSA <- DESeq(dds_EggAnalysis)
> write.csv(results(DESeqDONEEvSA, contrast = c("Treatment", "Egg", "ATAC"), alpha = 0.05),
"NAME::OF::EGGOUTPUTFILE")
```

2. SPERM vs. ADULT

```
> library(DESeq2)
> CountMatrix <- (read.csv("PATH::TO::SPERMCOUNTMATRIX", row.names = "Name"))
> ColData<-(read.csv("PATH::TO::SPERMCOLUMNDATA", row.names=1))
> dds_SpermAnalysis <- DESeqDataSetFromMatrix(countData = CountMatrix, colData = ColData, design
= ~ Treatment)
> DESeqDONESvsA <- DESeq(dds_SpermAnalysis)
> write.csv(results(DESeqDONESvsA, contrast = c("Treatment", "unbleached", "ATAC"), alpha = 0.05),
"NAME::OF::SPERMOUTPUTFILE")
```

2. EGG vs. SPERM

```
> library(DESeq2)
> CountMatrix <- (read.csv("PATH::TO::EGGvsSPERMCOUNTMATRIX", row.names = "Name"))
> ColData<-(read.csv("PATH::TO::EGGvsSPERMCOLUMNDATA", row.names=1))
> dds_EggSpermAnalysis <- DESeqDataSetFromMatrix(countData = CountMatrix, colData = ColData,
design = ~ Treatment)
> DESeqDONEEvSS <- DESeq(dds_EggSpermAnalysis)
> write.csv(results(DESeqDONEEvSS, contrast = c("Treatment", "Egg", "unbleached"), alpha = 0.05),
"NAME::OF::EGGvsSPERMOUTPUTFILE")
```

Figure S2. The commands used for differential expression analysis in RStudio 1.2.1335 with the DESeq2 package.

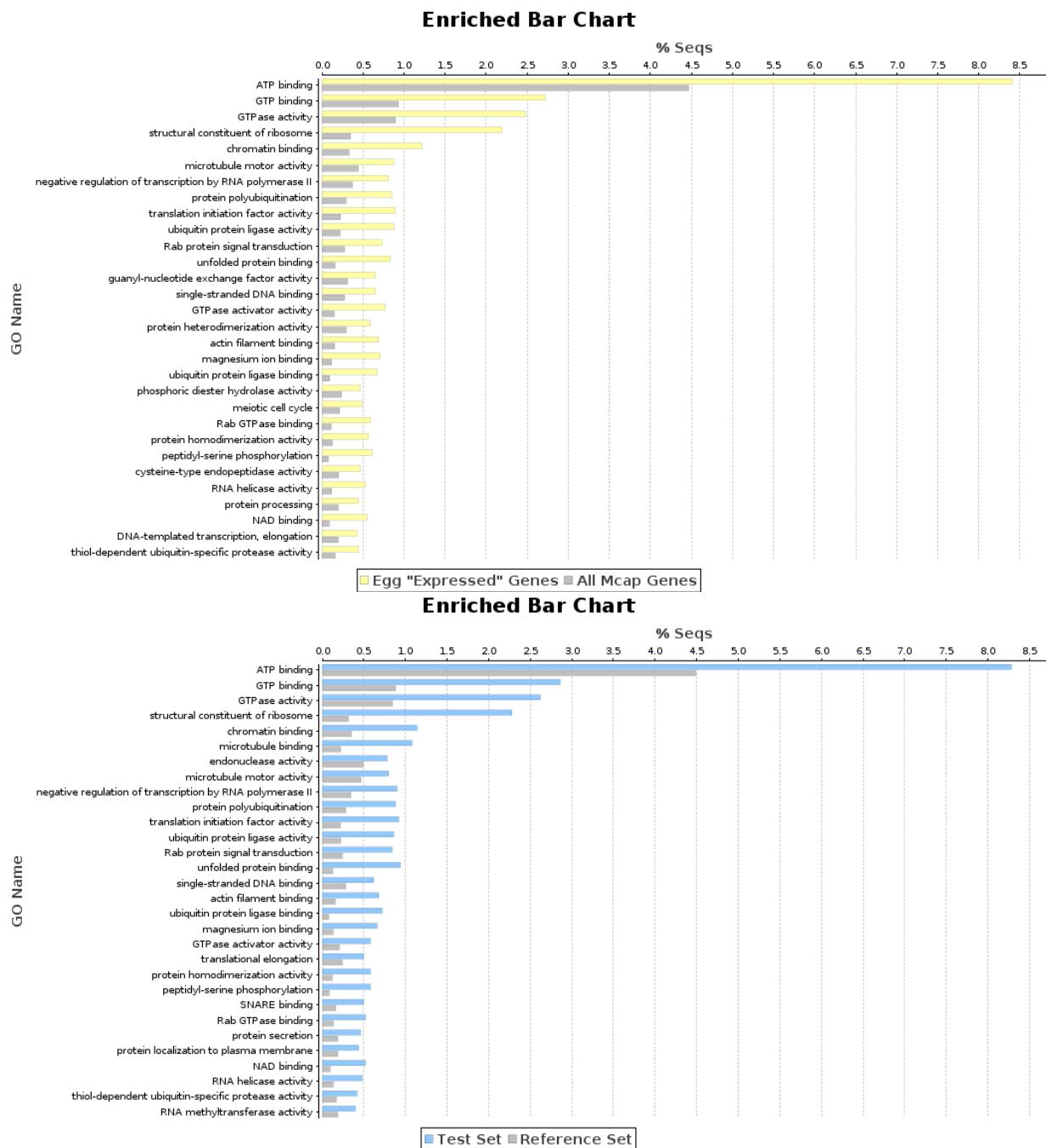


Figure S3. (above). (A) Fisher's Exact Test results showing enriched GOs for egg "expressed" (TPM>100) genes compared with all *M. capitata* genes. (B) Fisher's Exact Test results showing enriched GOs for sperm "expressed" (TPM>100) genes compared with all *M. capitata* genes. Both figures were made with Blast2GO OmicsBox 1.1.164.

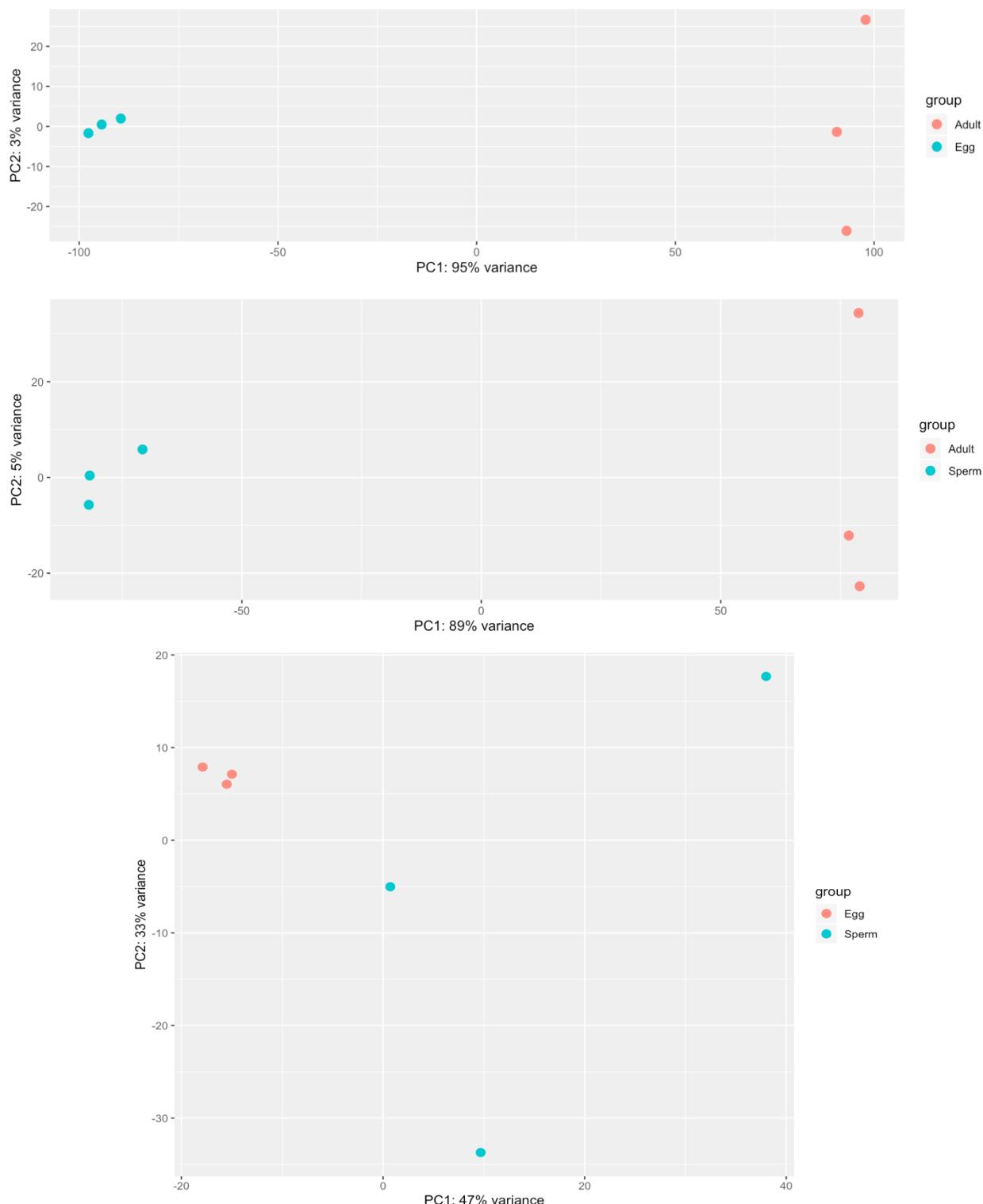


Figure S4. (A) Principal components analysis (PCA) plot of egg cDNA libraries against ATAC adult libraries. (B) Principal components analysis (PCA) plot of sperm cDNA libraries against ATAC adult libraries. (C) Principal components analysis (PCA) plot of egg cDNA libraries against sperm cDNA libraries. Both plots generated in RStudio 1.2.1335.

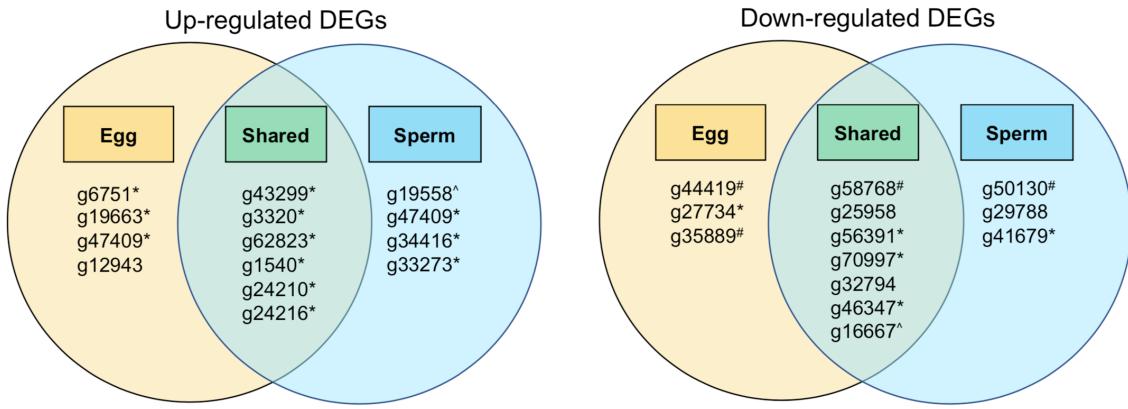


Figure S5. Differentially expressed genes in sperm and eggs.

(A) Venn diagram showing the top ten up-regulated DEGs in egg only (left), sperm only (right), and shared (center). (B) Venn diagram showing the top ten down-regulated DEGs in egg only (left), sperm only (right), and shared (center). * ^ # indicate the top BLAST annotation is to a “predicted”, “hypothetical”, “uncharacterized” protein, respectively.

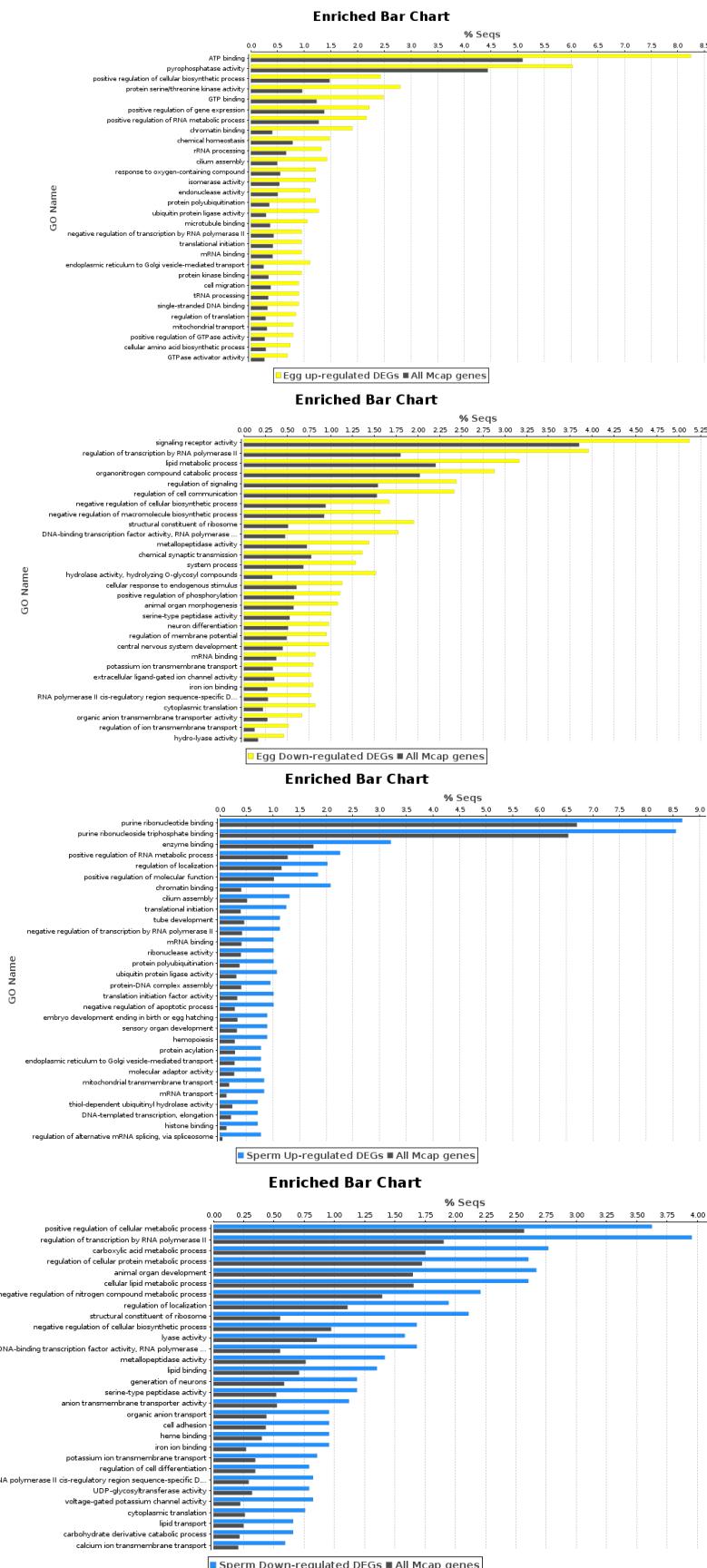


Figure S6. Results of GO-term enrichment analysis (Fisher's Exact Test) using gamete and adult RNA-seq data (A) DEGs up-regulated in egg; (B) DEGs down-regulated in egg; (C) DEGs up-regulated in sperm; (D) DEGs down-regulated in sperm. Test sets were all made from DESeq2 output data with adjusted p-values less than 0.05 and | Δ fc| > 1.5 in up-regulated DEGs and < -1.5 in down-regulated DEGs.

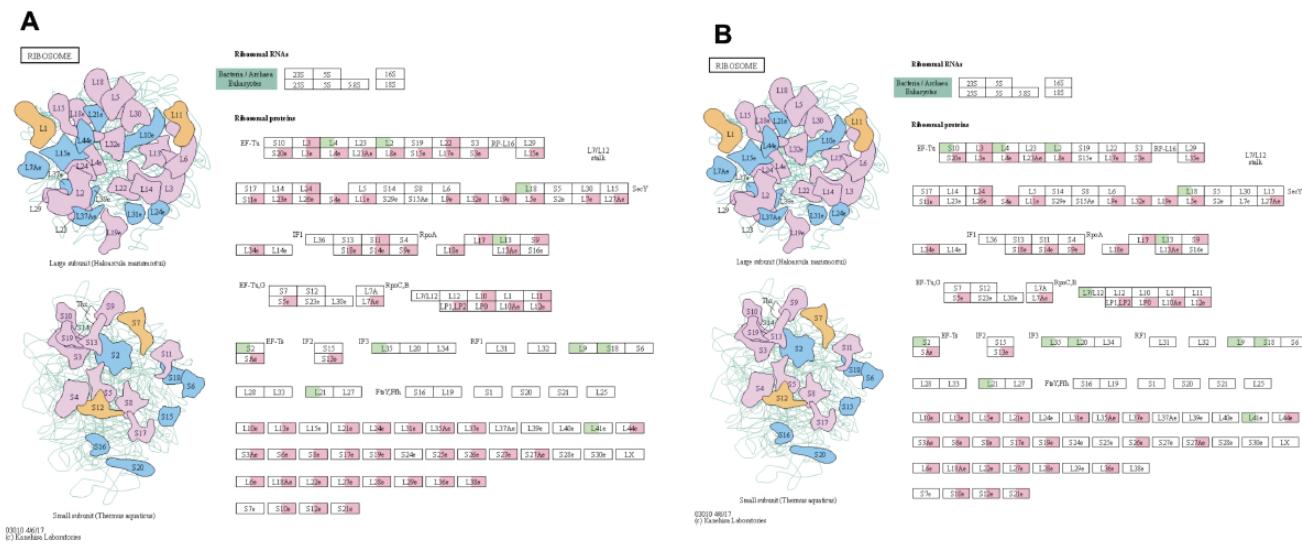


Figure S7. (A) KEGG diagram of a ribosome with egg DEGs superimposed; red squares correspond to down-regulated genes, green squares correspond to up-regulated genes. (B) KEGG diagram of a ribosome with egg DEGs superimposed; red squares correspond to down-regulated genes, green squares correspond to up-regulated genes. Images produced with permission from KEGG (Kanehisa 2017).

Name		Sperm TPM	Sperm DE 2fc	p-adj	Egg TPM	Egg DE 2fc	p-adj	Blastx Accession/Description
g63277	g63277.t1	532.1779	5.14502938	8E-13	445.15	5.00993	2E-16	>XP_015777656.1 PREDICTED: uncharacterized protein LOC107355583 isoform X1 [Acropora digitifera]
PID	E-value	Bitscore	KEGG	Annotation GO ID	Annotation GO Term			
88	1E-298	1034.6	N/A	GO:0003824;GO:0005509;GO:0016021;GO:0031514	catalytic activity;calcium ion binding:integral component of membrane:motile cilium			
88	9E-297	1028.5	N/A	GO:0003824;GO:0005509;GO:0016021;GO:0031514	catalytic activity;calcium ion binding:integral component of membrane:motile cilium			

Figure S8. Genes involved in the fertilization process that are interest for future investigation.

Table S1. Mapping results to the *M. capitata* reference genome using the five sperm cDNA libraries.

Run	Raw read count	Trimmed reads	Total mapped (% of trimmed)	# Unique Exon reads (% of total mapped)
Ub2	5,022,893	5,019,322	4,009,485 (79.88%)	347,116 (9.87%)
Ub3	7,444,969	7,423,204	6,127,430 (82.54%)	848,896 (15.75%)
Ub4	5,463,815	5,450,170	5,057,328 (92.79%)	595,703 (13.47%)

Table S2. Mapping results to the *M. capitata* reference genome using the three egg cDNA libraries.

Run	Raw read count (paired-end)	Trimmed reads	Total mapped (% of trimmed)	# Unique Exon reads (% of total mapped)
119 (E1)	42,861,620	41,749,425	8,074,398 (38.73%)	2,564,616 (31.76%)
120 (E2)	45,175,428	44,812,008	5,553,370 (24.79%)	1,620,591 (29.18%)
121 (E3)	38,684,570	38,492,780	6,018,664 (31.28%)	1,970,340 (32.74%)

Table S3. Annotation of the top ten DEGs with blast hits in the comparison of egg and adult ATAC RNA-seq libraries. PID is protein identity.

Up-regulated

Gene Name	Log2 fold change	p-adjusted	BLASTx Annotation	PID
g43299	17.2574431	1.9183E-44	>XP_015762538.1 PREDICTED: GTP-binding protein Di-Ras2-like [<i>Acropora digitifera</i>]	74.4
g3320	16.5927276	1.2382E-38	>XP_015753988.1 PREDICTED: histone H1-delta-like [<i>Acropora digitifera</i>]	76.5
g6751	12.9375991	1.0698E-24	>XP_015747403.1 PREDICTED: type I iodothyronine deiodinase-like [<i>Acropora digitifera</i>]	82.5
g62823	12.913316	7.9303E-24	>XP_015780785.1 PREDICTED: lysosomal acid lipase/cholesteryl ester hydrolase-like [<i>Acropora digitifera</i>]	70.5
g1540	12.4274331	1.5848E-22	>XP_015758457.1 PREDICTED: uncharacterized protein LOC107337742 [<i>Acropora digitifera</i>]	39.0
g24210	12.3950835	5.2798E-23	>XP_015767499.1 PREDICTED: uncharacterized protein LOC107346242 [<i>Acropora digitifera</i>]	69.9
g19663	12.1687266	6.9504E-42	>XP_015758457.1 PREDICTED: uncharacterized protein LOC107337742 [<i>Acropora digitifera</i>]	39.0
g47409	11.8694204	4.1555E-27	>XP_015776190.1 PREDICTED: cyclin-dependent kinases regulatory subunit-like [<i>Acropora digitifera</i>]	94.9
g24216	11.6985132	2.315E-20	>XP_015767499.1 PREDICTED: uncharacterized protein LOC107346242 [<i>Acropora digitifera</i>]	63.8
g12943	11.6662902	5.5762E-20	>XP_022810028.1 histone-arginine methyltransferase CARMER-like [<i>Stylophora pistillata</i>]	85.6

Down-regulated

Gene Name	Log2 fold change	p-adjusted	BLASTx Annotation	PID
g58768	-12.699101	4.7416E-07	>XP_022787132.1 uncharacterized protein LOC111327261 [<i>Stylophora pistillata</i>]	86.6
g25958	-12.867411	1.1268E-25	>sp D9IQ16.1 GXN_ACRM1 RecName: Full=Galaxin; Flags: Precursor	59.2
g56391	-12.928326	1.1717E-24	>XP_015754552.1 PREDICTED: zinc transporter ZIP14-like isoform X1 [<i>Acropora digitifera</i>]	62.8
g44419	-13.067867	1.6268E-25	>XP_020622435.1 uncharacterized protein LOC110060033 isoform X1 [<i>Orbicella faveolata</i>]	56.3
g27734	-13.285605	1.2925E-25	>XP_015759469.1 PREDICTED: dysferlin-like [<i>Acropora digitifera</i>]	72.8

g70997	-13.349726	1.4527E-12	>XP_015768671.1 PREDICTED: uncharacterized protein LOC107347285 [<i>Acropora digitifera</i>]	59.1
g35889	-13.392157	9.223E-11	>XP_020621079.1 uncharacterized protein LOC110058761 [<i>Orbicella faveolata</i>]	58.6
g32794	-13.478762	2.1422E-26	>XP_020616594.1 G1/S-specific cyclin-D2-like [<i>Orbicella faveolata</i>]	88.3
g46347	-14.723807	4.4928E-30	>XP_015773363.1 PREDICTED: choloylglycine hydrolase-like [<i>Acropora digitifera</i>]	84.7
g16667	-15.294337	8.2905E-35	>AET09742.1 hypothetical protein A014-G2 [<i>Acropora millepora</i>]	53.8

Table S4. Annotation of the top ten DEGs in the comparison of sperm and adult ATAC libraries.

Upregulated

Gene Name	Log2 fold change	p-adjusted	BLASTx Annotation	PID
g19558	22.1450262	2.5257E-07	>PFX33780.1 hypothetical protein AWC38_SpisGene1427 [<i>Stylophora pistillata</i>]	60.0
g43299	17.521645	2.457E-32	>XP_015762538.1 PREDICTED: GTP-binding protein Di-Ras2-like [<i>Acropora digitifera</i>]	74.4
g3320	16.8839084	2.0961E-34	>XP_015753988.1 PREDICTED: histone H1-delta-like [<i>Acropora digitifera</i>]	76.5
g1540	14.111518	1.2444E-19	>XP_015758457.1 PREDICTED: uncharacterized protein LOC107337742 [<i>Acropora digitifera</i>]	39.0
g24210	12.8454596	2.0921E-20	>XP_015767499.1 PREDICTED: uncharacterized protein LOC107346242 [<i>Acropora digitifera</i>]	69.9
g62823	12.6536856	8.8274E-21	>XP_015780785.1 PREDICTED: lysosomal acid lipase/cholesteryl ester hydrolase-like [<i>Acropora digitifera</i>]	70.5
g47409	12.2995242	2.1163E-19	>XP_015776190.1 PREDICTED: cyclin-dependent kinases regulatory subunit-like [<i>Acropora digitifera</i>]	94.9
g34416	12.1425323	7.9883E-15	>XP_015747104.1 PREDICTED: uncharacterized protein LOC107326866 [<i>Acropora digitifera</i>]	66.8
g33273	12.0821698	1.0349E-18	>XP_015758244.1 PREDICTED: uncharacterized protein LOC107337557 [<i>Acropora digitifera</i>]	81.9
g24216	12.0437682	2.7248E-18	>XP_015767499.1 PREDICTED: uncharacterized protein LOC107346242 [<i>Acropora digitifera</i>]	63.8

Downregulated

Gene Name	Log2 fold change	p-adjusted	BLASTx Annotation	PID

g70997	-11.174478	2.2222E-13	>XP_015768671.1 PREDICTED: uncharacterized protein LOC107347285 [<i>Acropora digitifera</i>]	59.1
g32794	-11.291512	1.1493E-21	>XP_020616594.1 G1/S-specific cyclin-D2-like [<i>Orbicella faveolata</i>]	88.3
g58768	-11.310221	3.2655E-09	>XP_022787132.1 uncharacterized protein LOC111327261 [<i>Stylophora pistillata</i>]	86.6
g25958	-11.487415	3.1229E-17	>spID9IQ16.1GXN_ACRM1 RecName: Full=Galaxin; Flags: Precursor	59.2
g56391	-11.548706	1.3949E-17	>XP_015754552.1 PREDICTED: zinc transporter ZIP14-like isoform X1 [<i>Acropora digitifera</i>]	62.8
g46347	-11.644295	1.6658E-20	>XP_015773363.1 PREDICTED: choloylglycine hydrolase-like [<i>Acropora digitifera</i>]	84.7
g50130	-11.957233	1.6898E-17	>XP_020612828.1 uncharacterized protein LOC110051165 [<i>Orbicella faveolata</i>]	61.3
g29788	-12.336762	4.2369E-12	>ABV24969.1 cnidCAP isoform 2, partial [<i>Montipora capitata</i>]	96.6
g41679	-12.875745	2.9519E-19	>XP_015777077.1 PREDICTED: collagen alpha-1(IX) chain-like isoform X1 [<i>Acropora digitifera</i>]	84.0
g16667	-12.961374	1.452E-23	>AET09742.1 hypothetical protein A014-G2 [<i>Acropora millepora</i>]	53.8