**Oyster biomineralisation under ocean acidification: genes to shell**

Kanmani Chandra Rajan1, Meng Yuan2, Ziniu Yu3, Vengatesen Thiyagarajan1

**Supplementary information**

1The Swire Institute of Marine Science and School of Biological Sciences, The University of Hong Kong, Pokfulam, Hong Kong SAR, China.

2State Key Laboratory of Respiratory Disease, The Third Affiliated Hospital of Guangzhou Medical University, Guangzhou, China.

3South China Sea Institute of Oceanology, Guangzhou, China.

Correspondence: [u3005329@hku.hk](mailto:u3005329@hku.hk) +, [rajan@hku.hk](mailto:rajan@hku.hk) \*

**1. pH and salinity of the experimental region (Zhanjiang, Guangzhou, China)**

The experimental location (oyster hatchery) and the Donghai bridge, around where the pH measurements were performed are marked in the map in Figure S1. The average pH of the region calculated from the report is 7.74 ± 0.21. Based on this local adaptation and the projected IPCC scenarios, the experimental pH for the treatments were fixed as 7.7 and 7.4 as the oysters used in the experiment were obtained from this region. The average pH of this region might reach as low as 7.4 by the end of this century.

To estimate the average pH of the experimental region, pH data from the report titled, “Evaluation of the Marine Environment Tracking and Monitoring Results of the Donghai Bridge during the Construction of Yuzhan Expressway” prepared by Marine Resources and Environmental Monitoring Centre of Guangdong Ocean University, China, was used (report dated 03 March 2018, translated from Chinese).

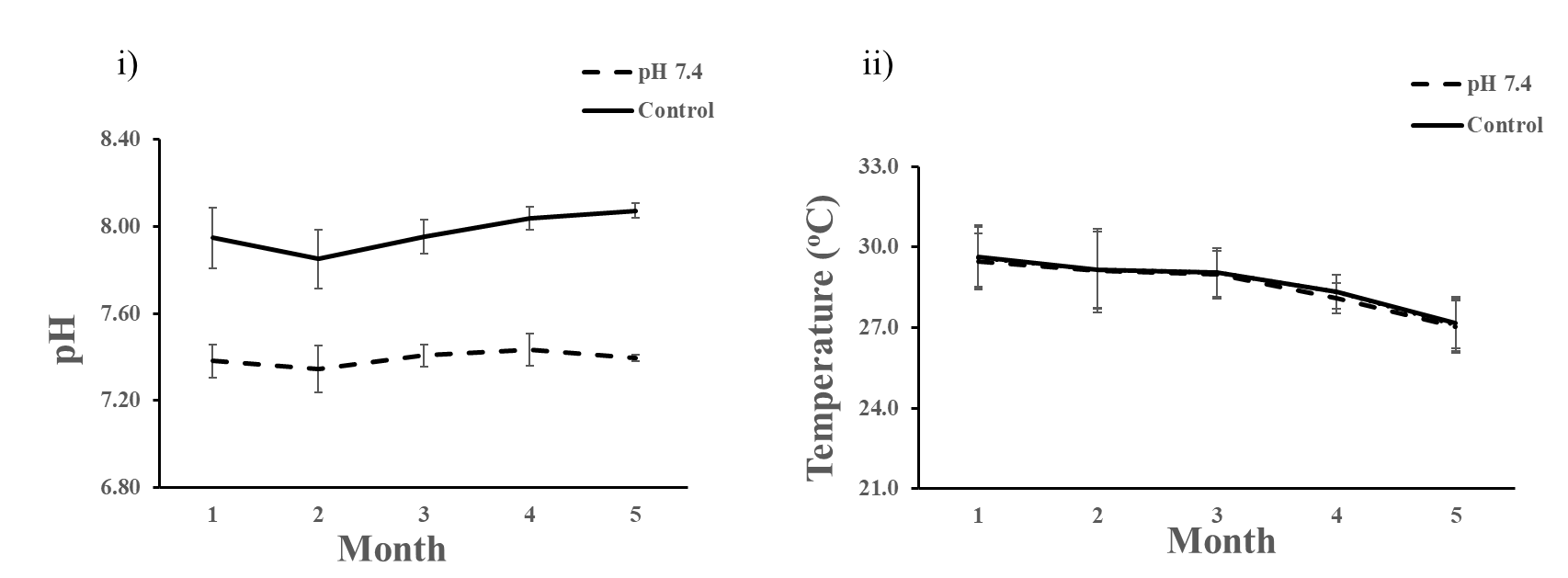


**Figure S1**: Experimental location and the pH sampling region (used for deciding the pH of the experiment) are marked in the map as (stars) 1 and 2, respectively. The environmental parameter data for the region was obtained from Environmental monitoring centre of the Guangdong Ocean university. [More details are provided in table S1 below]

**Table S1:** Water parameters from the experimental region from October 2017.

|  |  |  |  |
| --- | --- | --- | --- |
| **S.No.** | **Temperature** | **pH** | **Salinity (% 0)** |
| 1 | 28 | 7.92 | 21.0273 |
| 2 | 28.4 | 7.98 | 21.6053 |
| 3 | 28.4 | 7.56 | 14.6191 |
| 4 | 28.6 | 7.86 | 14.5473 |
| 5 | 27.9 | 7.54 | 14.8632 |
| 6 | 28 | 7.52 | 15.1747 |
| 7 | 27.9 | 7.51 | 14.9917 |
| 8 | 27.7 | 7.52 | 15.165 |
| 9 | 21.9 | 7.8 | 24.3651 |
| 10 | 23 | 8.2 | 25.0665 |
| 11 | 19.8 | 7.66 | 16.3415 |
| 12 | 20.8 | 8.04 | 16.08999 |
| 13 | 19.6 | 7.7 | 16.3193 |
| 14 | 19.8 | 7.67 | 16.6513 |
| 15 | 19.2 | 7.68 | 16.7883 |
| 16 | 19.4 | 7.63 | 16.9409 |
| **Average** | 24.28 ± 4.08 | 7.74 ± 0.21 | 17.54 ±3.47 |

**2. Experimental sea water physiochemical parameters**



**Figure S2:** Experimental pH and temperature – The graph (i) shows the mean pH maintained in OA and control conditions throughout the experimental duration where the average pH of treatments and control conditions were 7.39 ± 0.01 and 7.97 ± 0.03 respectively. Graph (ii) shows the mean temperature maintained in treatments and control conditions throughout the experimental duration where the average temperature of pH 7.4 (OA) and control conditions were 28.54 ± 0.98 and 28.67 ± 0.94, respectively. The mean was calculated for the months as follows: Month 1 – 25 May to 25 June, Month 2 – 26 June to 25 July, Month 3 – 26 July to 25 August, Month 4 – 26 August to 25 September, Month 5 – 26 September to 7 October. The pH values are averaged for both month and the replicate tanks. [Error bar: Standard deviation]

**Table S2**: Sea water physiochemical parameters of the experimental setup

|  |  |  |
| --- | --- | --- |
| **Parameters** | **pH 7.4** | **Control** |
| pH | 7.39 ± 0.01 | 7.97 ± 0.03 |
| Temperature | 28.54 ± 0.98 | 28.67 ± 0.94 |
| Salinity (%0) | 16.61 ± 1.82 | 16.63 ± 1.66 |
| TA (µmol L-1) | 1308.5 ± 193.6 | 1176.9 ± 221.7 |
| pCO2 (μatm) \* | 2384.5 ± 220.3 | 528.7± 78.9 |
| [CO3 2-] (μmol kg-1) \* | 15.10 ± 3.92 | 50.85 ± 27.23 |
| ΩCa \* | 0.43 ± 0.11 | 1.45 ± 0.77 |
| ΩAr \* | 0.27 ± 0.07 | 0.89 ± 0.48 |

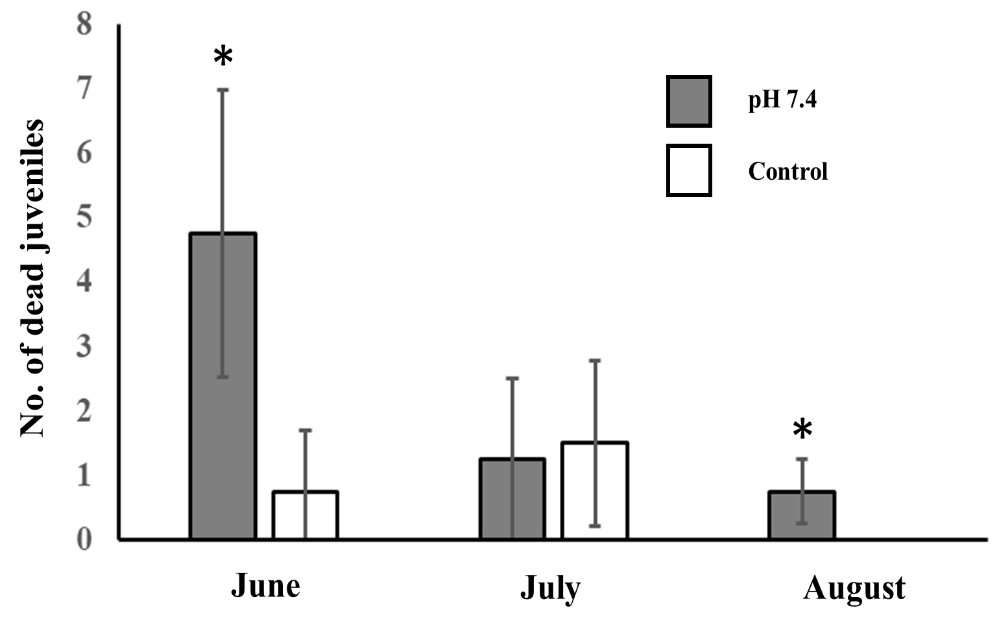
All the parameters in table S1 are given as Mean ± Standard deviation. The parameters were analysed on eleven different days (n=11) throughout the experimental duration. Mean of each day was calculated first by averaging the replicates and later the mean of all the eleven days was calculated. TA – Total alkalinity, pCO2 – carbon dioxide partial pressure; [CO3 2-] – concentration of carbonate ions; ΩCa – calcite saturation state, ΩAr – Aragonite saturation state.

\* Parameters were calculated using CO2SYS program (Pierrot et al. 2011) with equilibrium constants K1, K2 and KSO4 (Mehrbach et al. 1973; Dickson and Millero 1987)

**Table S3**: Number of individuals used in the experiment and mortality

|  |  |
| --- | --- |
| **Month** | **Total number of individuals in one replicate tank\*** |
| May | ~350 |
| June | 220 |
| July | 215 |
| August | 190 |
| September | 160 |
| October | 160 |

\* All the replicate tanks were maintained with similar biomass and number of individuals based on size



**Figure S3:** Mortality during the experiment for the months June, July, and August. No mortality observed during September and October. \* denotes p value <0.05

**3. RNA-seq analysis**

**3.1. Statistical results of reads and reference genome comparison rate**

HISAT (Kim, Langmead, and Salzberg 2015) was used to compare the sequence of clean reads with the specified reference genome to obtain the position information on the reference genome and the sequence characteristic information unique to the sequencing sample.

**Table S4: Statistical results of reads and reference genome comparison rate**

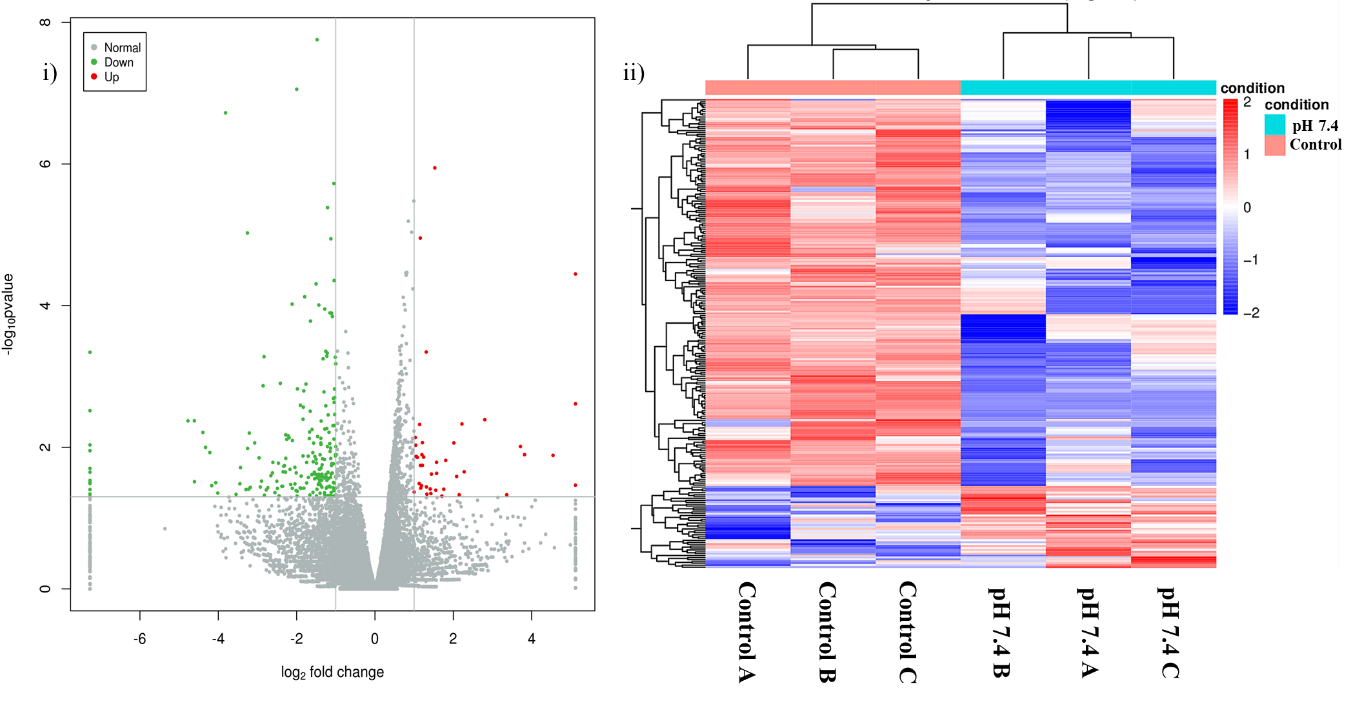
| **Sample** | **pH 7.4 A** | **pH 7.4 B** | **pH 7.4 C** | **Control A** | **Control B** | **Control C** |
| --- | --- | --- | --- | --- | --- | --- |
| **Total reads** | 46995352 | 46121688 | 46162608 | 46924946 | 46743480 | 46815156 |
| **Total mapped reads** | 12187368 (25.93%) | 11056428 (23.97%) | 11494879 (24.90%) | 12301228 (26.21%) | 11934047 (25.53%) | 11692542 (24.98%) |
| **Multiple mapped** | 1205918 (2.57%) | 1060968 (2.30%) | 1079270 (2.34%) | 1336268 (2.85%) | 1160422 (2.48%) | 1122769 (2.40%) |
| **Uniquely mapped** | 10981450 (23.37%) | 9995460 (21.67%) | 10415609 (22.56%) | 10964960 (23.37%) | 10773625 (23.05%) | 10569773 (22.58%) |
| **Read-1** | 5422623 (11.54%) | 4915164 (10.66%) | 5143382 (11.14%) | 5444893 (11.60%) | 5345865 (11.44%) | 5237083 (11.19%) |
| **Read-2** | 5558827 (11.83%) | 5080296 (11.01%) | 5272227 (11.42%) | 5520067 (11.76%) | 5427760 (11.61%) | 5332690 (11.39%) |
| **Reads map to '+'** | 5564061 (11.84%) | 5056848 (10.96%) | 5275663 (11.43%) | 5555953 (11.84%) | 5456599 (11.67%) | 5350326 (11.43%) |
| **Reads map to '-'** | 5417389 (11.53%) | 4938612 (10.71%) | 5139946 (11.13%) | 5409007 (11.53%) | 5317026 (11.37%) | 5219447 (11.15%) |
| **Non-splice reads** | 4410969 (9.39%) | 4085234 (8.86%) | 4254340 (9.22%) | 4554878 (9.71%) | 4439149 (9.50%) | 4259358 (9.10%) |
| **Splice reads** | 6570481 (13.98%) | 5910226 (12.81%) | 6161269 (13.35%) | 6410082 (13.66%) | 6334476 (13.55%) | 6310415 (13.48%) |
| **Reads mapped in proper pairs** | 7950626 (16.92%) | 6896958 (14.95%) | 7511808 (16.27%) | 8642870 (18.42%) | 8491634 (18.17%) | 8462352 (18.08%) |

Table S4 description: Total reads: the number of sequenced sequences filtered by sequencing data (Clean reads); Total mapped: the number of sequenced sequences that can be located on the genome; under normal circumstances, if it does not exist In case of contamination and the selection of the reference genome is appropriate, the percentage of this part of the data is greater than 70%; Multiple mapped: the number of sequencing sequences with multiple alignment positions on the reference sequence; Uniquely mapped: in the reference Statistics on the number of sequencing sequences with unique alignment positions on the sequence; Read-1, Read-2: the number statistics of the left reads and right reads aligned to the reference genome; Reads map to '+ ', Reads map to'-': Sequencing sequence is compared to the positive and negative strands of the genome; Splice reads: Totally sequenced reads, the sequence of the sequence is compared to the two exons ( (Also called Junction reads), Non-splice reads is the statistics of the entire sequence of the sequence alignment to the exon, and the percentage of Splice reads depends on the length of the sequencing fragments; Reads mapped in proper pairs: Statistics on the number of sequencing sequences on the double-ended alignment.

**3.2. Volcano plot and Cluster analysis**

We can understand the overall distribution of differentially expressed genes by a volcano plot.

The distance between samples were calculated and a distance matrix was formed. Two categories with the closest distance were combined into a new category. The distance between the new category and the current category was calculated, and then merged and calculated until there is only one category. The expression of different genes is used to calculate the direct correlation of samples. Generally speaking, samples of the same type appear in the same cluster by clustering. Genes clustered in the same cluster may have similar biological functions.



**Figure S4:** Up and down regulated genes i) Volcano plot showing upregulated genes in red, downregulated genes in green and normal genes in grey. The grey lines in the graph denoted the cut off for considering a gene as up or down regulated which is p value < 0.05 & |log2 FC| > 1 ii) Cluster map showing upregulated genes in red and down regulated genes in blue. The treatments are clustered together showing the robustness of the experimental replication and the similarity between the replicates.

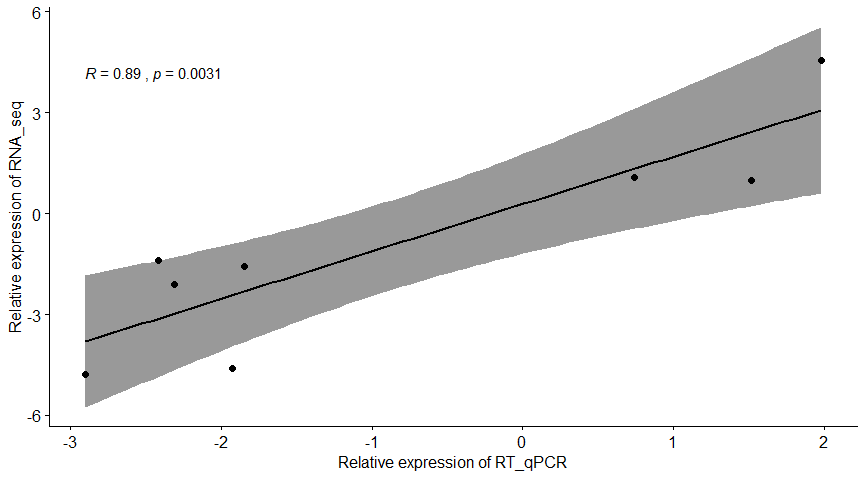
**3.3 RT-qPCR method**

To validate the RNA-seq data, eight differentially expressed genes were selected randomly, and RT-qPCR was performed. mRNA extraction was performed as mentioned in the methods section. Extracted mRNA was reverse transcribed into cDNA with the PrimerScript™ first strand cDNA synthesis kit (TAKARA, RR047A) and qPCR was performed using 2× RealStar Green Power mixture (GenStar, A311) in LightCycler® 480 II (Roche, Switzerland), according to the manufacturer’s instructions. Melting curve analysis was performed to confirm the formation of single fluorescent product. The reference housekeeping gene GAPDH was used as the internal control and the 2−ΔΔCT method was used to analyse the relative fold change in gene expression. The sequence of primers used are given in the table S2 and the result of RT-qPCR validation is given in the figure S5.

**Table S5**: Primer sequences used for RT qPCR and log2FC comparison between RNA seq and RT-qPCR for randomly selected genes.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Gene ID** | **Gene Description** | **Primers** | **Log2 FC** | |
|  |  |  | **RNA seq** | **RT-qPCR** |
| LOC105338822 | VDCC – Voltage-dependent calcium channel subunit alpha-2/delta-1 | F: TCTACAGCAGGTATCGTGCG  R: ACTGTCTTGGCAACCCCTTT | 1.00 | 1.52 |
| LOC105344773 | CACC - Calcium-activated chloride channel regulator 1-like | F: TTTCCTGGGTATTGCAGCGT  R: TGGTCCGGAAACGGTAAAGG | -1.38 | -2.42 |
| LOC105347882 | pEGF - pro-epidermal growth factor | F: TACGGATCCTCTGGCACATC  R: AGCGCTGCAGATTTGCTTCT | 4.55 | 1.98 |
| LOC105340761 | SerThrPRA - Serine/threonine-protein phosphatase 6 regulatory ankyrin repeat subunit B | F: GGCAACGGCGCTTCATAAAT  R: ACAGAGGTGAATTGGGCGAG | 1.08 | 0.74 |
| LOC105320559 | TRE homo - transcriptional  regulator ERG homolog | F: TAGGTGCAAAGAAGGGGAGGC  R: AGAGCAGGTTTGCACTGGAT | - 4.77 | -2.90 |
| LOC105347758 | Ankyrep - ankyrin repeat  and MYND domain-containing  protein 2 | F: CTAACGGAGCCGATGTGAAC  R: CACCCGCTTCTAGCACTTGT | - 4.61 | -1.93 |
| LOC105319934 | CollA1 - collagen alpha-  1(III) chain | F: GGGATGTTGGACCTCAAGGG  R: GATCCCCTTTGGAACCTGGG | -1.57 | -1.85 |
| LOC105318940 | AnnexA7 - annexin A7-like | F: GCCTATCGCATCATTGTTTGT  R: GTACCCTTGTTCGGGCTGAT | -2.12 | -2.31 |

In table S2, Gene ID refers to the locus ID of the gene. Gene description has a short form of the gene name first followed by the complete gene name. F and R refers to forward and reverse primers, respectively. Log2 FC (Fold Change) is given as mean.



**Figure S5:** Pearson correlation analysis between RNA-seq and RT-qPCR results. R in the graph refers to Pearson correlation coefficient showing a significant (p=0.0031) strong positive correlation between RNA-seq and RT-qPCR validating the RNA-seq results presented in this study. The genes used for this analysis with the primers and expression values is provided in the Table S2.

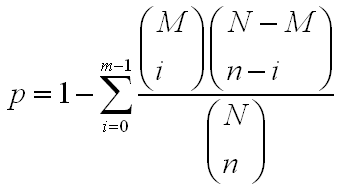
**Table S6:** Top DEGs in mantle tissue under OA (FDR adjusted p-value <0.05, |log2FC| >1)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **GENE ID** | **Description** | **FC** | **Log2FC** | **p-value** | **FDR p-value** |
| LOC105318316 | uncharacterized LOC105318316 | 0.07 | -3.81 | 1.90E-07 | 0.001 |
| LOC105334193 | uncharacterized LOC105334193 | 0.10 | -3.26 | 9.41E-06 | 0.01 |
| LOC105341416 | cystatin-A2 | 0.25 | -1.99 | 8.80E-08 | 0.0007 |
| LOC105336247 | ubiquitin-fold modifier 1 | 0.35 | -1.50 | 4.92E-05 | 0.045 |
| LOC105338114 | coactosin-like protein | 0.36 | -1.48 | 1.76E-08 | 0.0002 |
| LOC105332687 | Profilin | 0.43 | -1.21 | 4.12E-06 | 0.009 |
| LOC105347351 | E3 ubiquitin-protein ligase  rnf213-alpha-like | 0.46 | -1.13 | 1.14E-05 | 0.015 |
| LOC105340112 | Cofilin | 0.48 | -1.05 | 1.88E-06 | 0.006 |
| LOC105335623 | gelsolin-like protein 2 | 0.49 | -1.04 | 4.43E-05 | 0.04 |
| LOC105317600 | uncharacterized LOC105317600 | 2.23 | 1.16 | 1.11E-05 | 0.016 |
| LOC105346998 | spermidine synthase | 2.88 | 1.53 | 1.13E-06 | 0.004 |
| LOC105330900 | Uncharacterized LOC105330900 | 34.73 | 5.12 | 3.58E-05 | 0.038 |

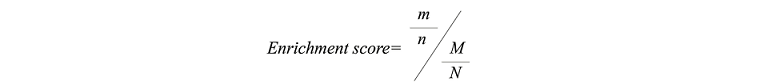
**3.4 Enrichment analysis**

**3.4.1 GO enrichment**

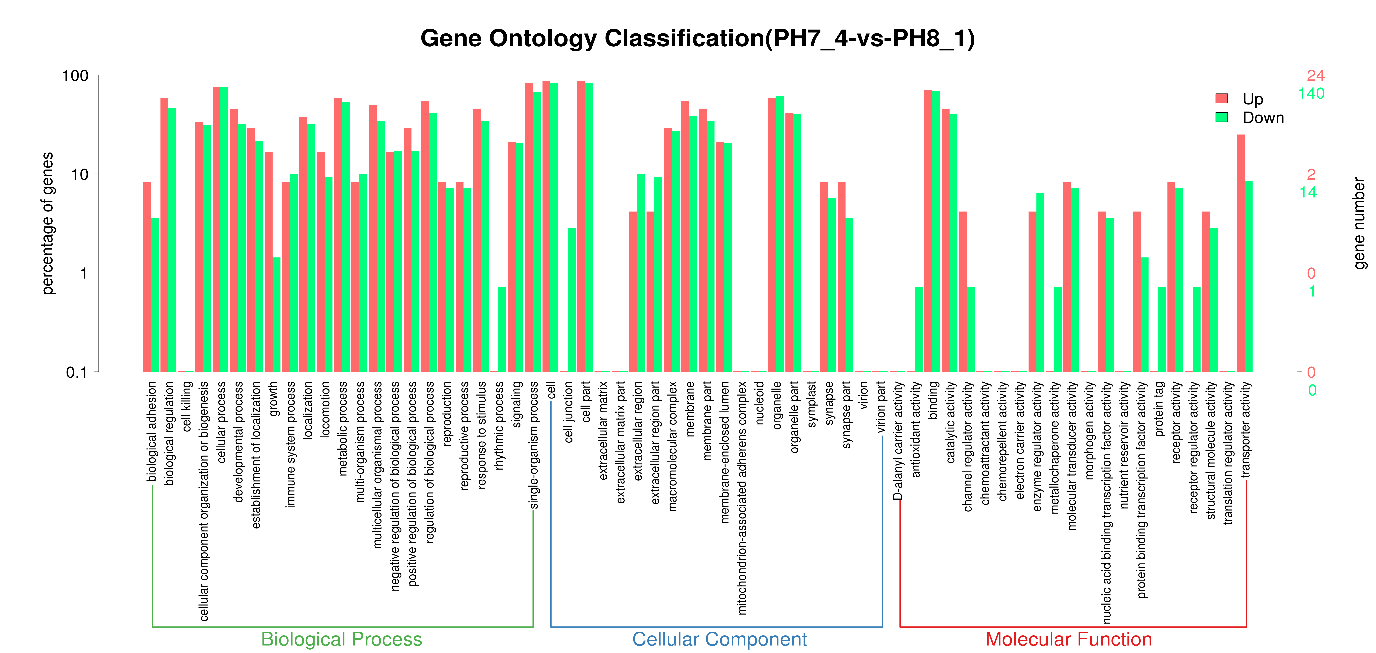
After obtaining differentially expressed genes, we performed GO enrichment analysis on the differentially expressed genes to describe their functions (combined with GO annotation results). GO function enrichment analysis method: all protein coding genes / transcripts are used as background lists, and differential protein coding genes / transcripts lists as candidate lists screened from the background list, and the hypergeometric distribution test is used to calculate the representative GO function set. The p value of the differential protein encoding gene / transcript list is significantly enriched, and then the p value is corrected by Benjamini & Hochberg multiple tests to obtain FDR. The formula for calculating the p-value by the hypergeometric distribution test is as follows:



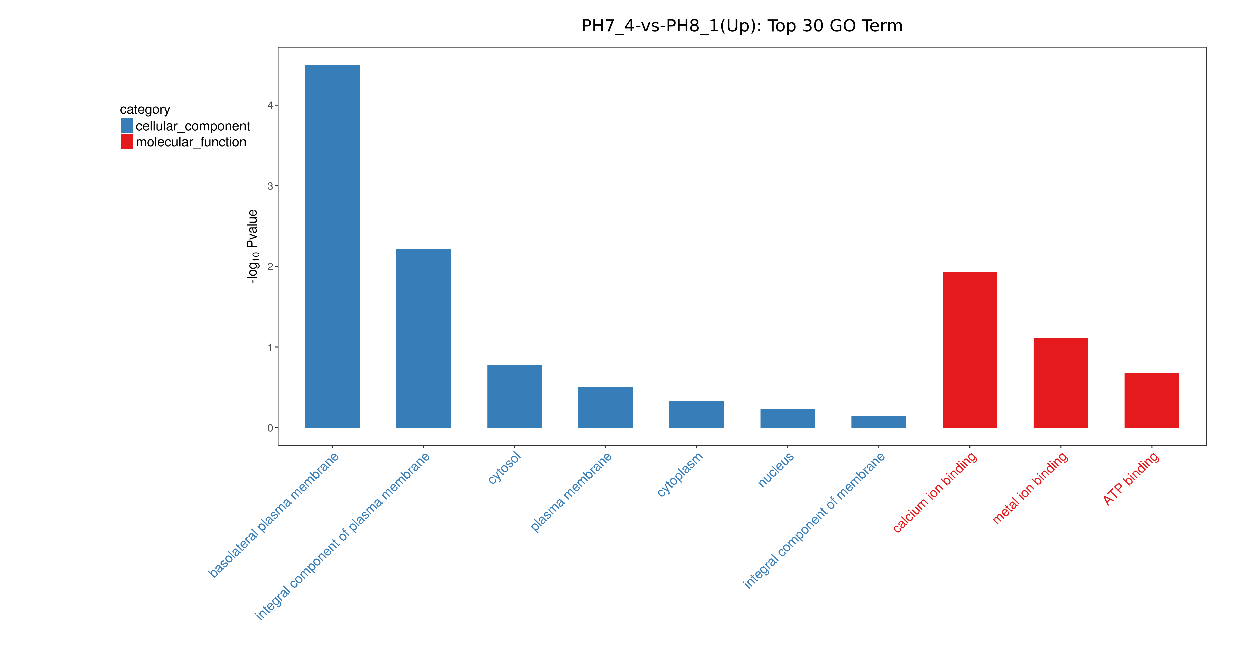
The calculation formula of Enrichment score is:



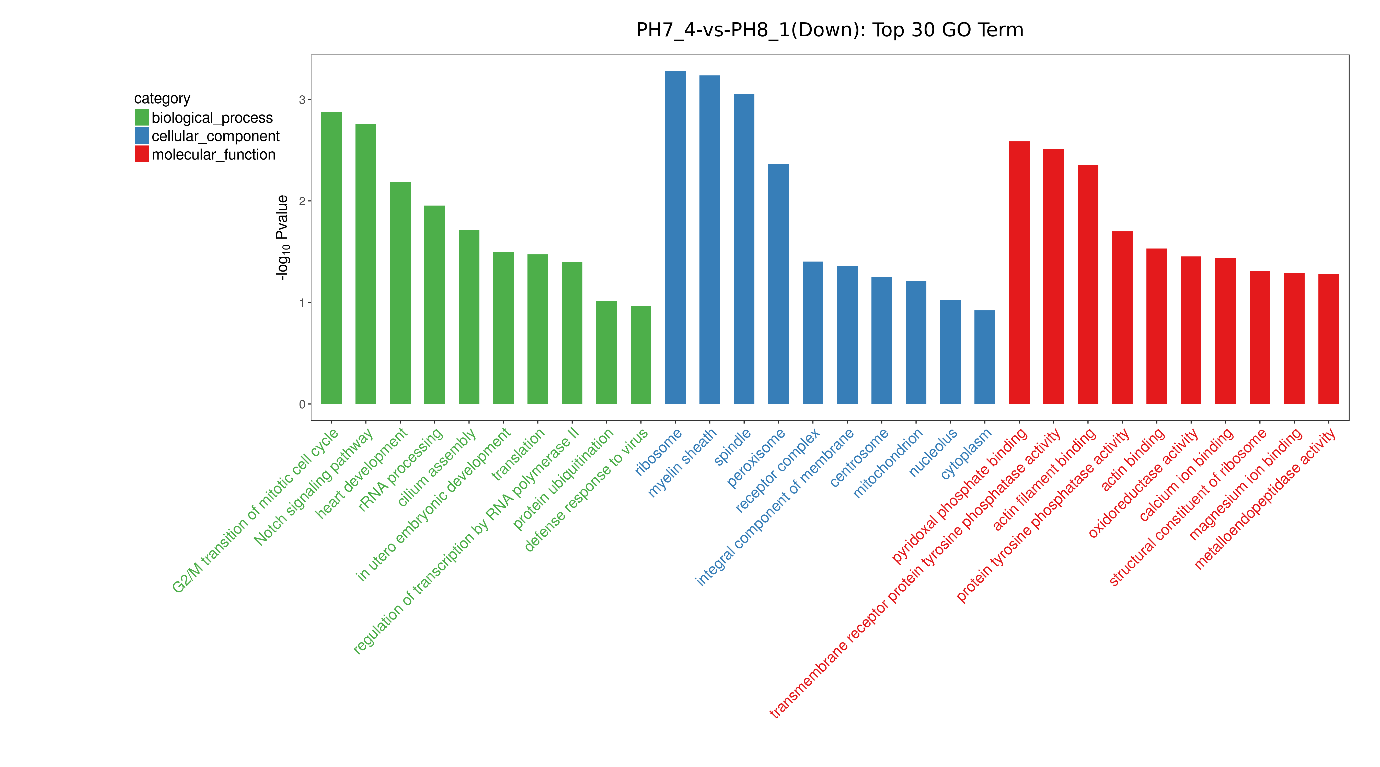
Where N is the number of genes with GO annotations in all genes; n is the number of genes with GO annotations in differentially expressed genes in N; M is the number of genes annotated with a specific GO term among all genes; m is the annotation with a specific GO term. The number of differentially expressed genes in GO term. Level 2 enrichment analysis involves biological process, cellular component, and molecular function.



**Figure S6: Level 2 Gene Ontology (GO) enrichment analysis** showing top upregulated (orange) and downregulated (green) gene categories under three major categories: biological process, cellular component, and molecular function. The top three upregulated molecular functional categories are “binding”, “catalytic activity” and “transporter activity”.



**Figure S7: Top GO categories of upregulated DEGs:** Cellular component in blue, molecular function in red.



**Figure S8: Top GO categories of downregulated DEGs:** biological process in green, cellular component in blue and molecular function in red.

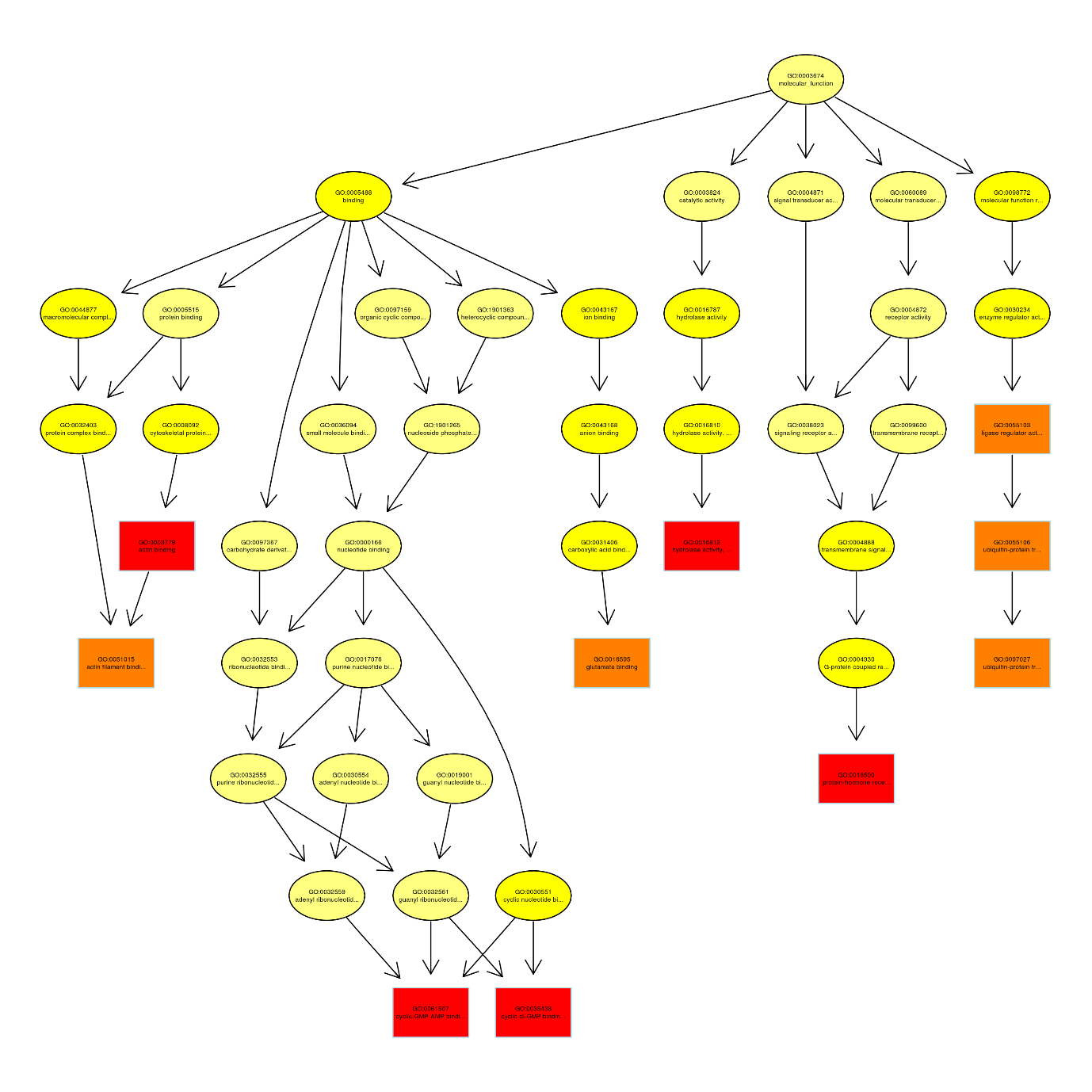
**Table S7: DEGs relevant for biomineralisation**

|  |  |  |  |
| --- | --- | --- | --- |
| **Functional category** | **Gene ID** | **Gene description** | **Log2FC** |
| **Ion channels or transporters** | LOC105342347 | TMC2 - Transmembrane channel-like protein 2 | 3.36 |
| LOC105324538 | Asph - aspartyl/asparaginyl beta-hydroxylase | -1.05 |
| LOC105343467 | VGPC - Voltage gated proton channel activity | -1.23 |
| LOC105332322 | Grik2 - Glutamate receptor ionotropic, kainate 2 | -1.54 |
| LOC105318785 | TRPC - Transient receptor potential channel | -1.40 |
| LOC105322233 | SLC9A1 - Na (+)/ H (+) exchanger beta | -1.06 |
| LOC105334654 | CTP - Copper transport protein ATX1-like | -1.43 |
| LOC105332811 | Orct - Organic cation transporter protein-like | X |
| **Ca2+ ion related**  **Ca2+ ion related**    **Ca2+ ion related** | LOC105333893 | CCR – Glutamate receptor (calcium channel regulator) | -2.25 |
| LOC105344773 | CACCR - Calcium-activated chloride channel regulator 1-like | -1.38 |
| LOC105338822 | VDCC - Voltage-dependent calcium channel subunit alpha-2/delta-1 | 1.00 |
| LOC105340783 | CPLA2 - Cytosolic phospholipase A2 (Ca2+ binding) | -1.17 |
| LOC105342224 | SPEC2A – Caltractin (Ca2+ binding) | -1.01 |
| LOC105317376 | CMR - C-type mannose receptor 2 (Ca2+ binding) | -1.28 |
| LOC105323183 | PC - protocadherin Fat 4  - Extracellular/secreted/plasma membrane (Ca2+ binding) | -1.37 |
| LOC105323822 | MLP Mucin-like protein (Ca2+ binding) | -2.11 |
| LOC105324714 | TPO - thyroid peroxidase-like (Ca2+ binding) | -1.67 |
| LOC105327999 | CALP - Calmodulin-A-like protein (Ca2+ binding) | -1.38 |
| LOC105338472 | MELC - myosin, essential light chain (Ca2+ binding) | -1.48 |
| LOC105345159 | GPD2 - glycerol-3-phosphate dehydrogenase, mitochondrial (Ca2+ binding) | -4.39 |
| LOC105348695 | Uncharacterised – Ca2+ binding | -1.82 |
| Testican-1 | Spock1 - Testican-1 (Ca2+ binding) | -1.59 |
| LOC105346195 | UNLP - Uncharacterised Notch like protein (response to Ca2+ & Ca2+ binding) | 1.169 |
| LOC105330515 | PH - protein hunchback - IKZF1 (positive regulation of cytosolic calcium ion concentration) | 1.14 |
| LOC105324638 | TKT - tyrosine-protein kinase receptor torso | 1.05 |
| LOC105342567 | UDD - uncharacterized discoidin domain | 1.81 |
| LOC105347882 | LDLR - pro-epidermal growth factor (Low-density lipoprotein-receptor YWTD domain) | 4.55 |
| Chitin  related | LOC105341736 | CBL1 - Chitin binding lectin 1 | -1.50 |

**Table S8: Top molecular function categories from GO enrichment**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **GO-ID** | **Molecular function** | **List hits** | **p value** | **ES** |
| **Down** | GO:0030170 | pyridoxal phosphate binding | 3 | 0.002 | 5.22 |
| GO:0005001 | transmembrane receptor protein  tyrosine phosphatase activity | 3 | 0.003 | 4.97 |
| GO:0051015 | actin filament binding | 4 | 0.004 | 3.72 |
| GO:0004725 | protein tyrosine phosphatase activity | 4 | 0.019 | 2.57 |
| GO:0003779 | actin binding | 4 | 0.029 | 2.32 |
| GO:0016491 | oxidoreductase activity | 3 | 0.035 | 2.43 |
| GO:0005509 | calcium ion binding | 13 | 0.036 | 1.57 |
| GO:0003735 | structural constituent of ribosome | 3 | 0.049 | 2.18 |
| **Up** | GO:0005509 | calcium ion binding | 4 | 0.011 | 2.82 |

The fisher algorithm was used to enrich molecular function for differentially expressed genes between pH 7.4 and control. TopGO was used to draw directed acyclic graphs of the enriched term. The topGO directed acyclic graph can directly display the GO nodes (Term) of the differential expression gene enrichment and their hierarchical relationship. It is a graphical display of the results of the GO enrichment analysis of differential expression genes.



**Figure S9**: **Differential gene topGO directed acyclic graph display**: the most significant 10 nodes are represented by rectangles. The colour of the rectangle represents the saliency of enrichment, from yellow to red the saliency becomes higher and higher. The basic information of each node is displayed in the corresponding graph as GO ID and GO term. Top molecular function or GO terms from the figure marked in red are: 1) GO:0003779 – actin binding 2) GO:0016812 – hydrolase activity 3) GO:0061507 – cyclic GMP-AMP binding 4) GO:0035438 – cyclic di GMP binding 5) GO:0016500 – protein hormone receptor binding. The most salient nodes from this figure are provided in Table S8.

**Table S9: Top ten molecular functional categories identified from all the DEGs by GO enrichment (acyclic graph display)**

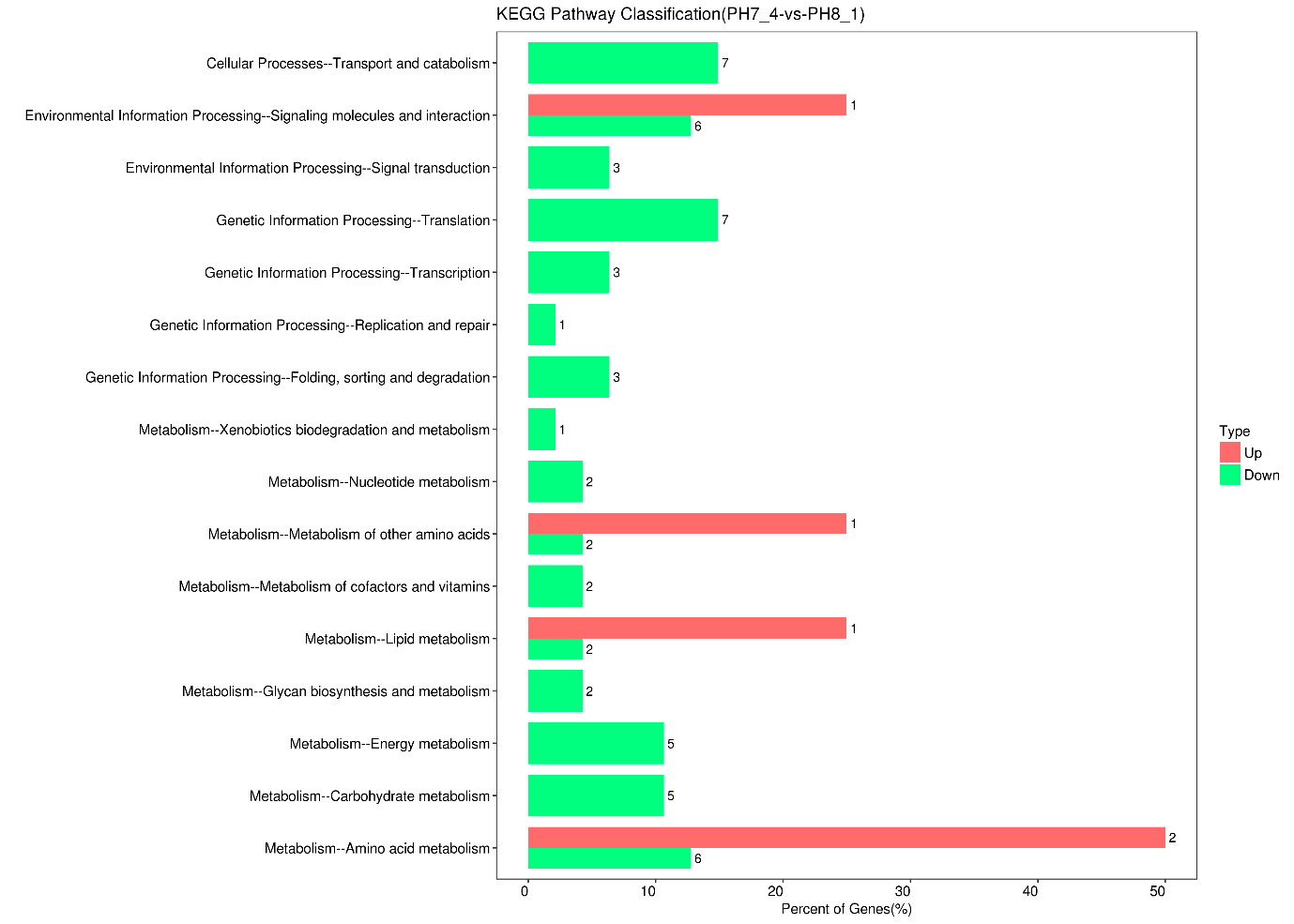
|  |  |  |
| --- | --- | --- |
| Code | GO-ID | Molecular function |
| Red | GO:0003779 | Actin binding |
| GO:0016812 | Hydrolase activity |
| GO:0061507 | Cyclic GMP-AMP binding |
| GO: 0035438 | Cyclic di-GMP binding |
| GO:0016500 | Protein-hormone receptor activity |
| Orange | GO: 0051015 | Actin filament binding |
| GO: 0016595 | Glutamate binding |
| GO: 0055106 | Ubiquitin protein transferase |
| GO: 0097027 | Ubiquitin protein transferase |
| GO: 0055103 | Ligase regulator activity |

**Table S10: GO terms related to calcium binding and ion homeostasis (repeated genes under the two categories are counted as one for plotting the graph in Figure 2iii).**

|  |  |  |  |
| --- | --- | --- | --- |
| **GO ID** | **GO term** | **List hits** | **Gene ID** |
| **Calcium ion binding** | | | |
| **Higher expression under OA** | | | |
| GO:0005509 | calcium ion binding | 4 | LOC105324638; LOC105342567;  LOC105346195; LOC105347882 |
| **Lower expression under OA** | | | |
| GO:0005509 | calcium ion binding | 13 | LOC105317376; LOC105323183; LOC105323822; LOC105324538; LOC105324714; LOC105327999; LOC105333893; LOC105338472; LOC105340783; LOC105342224; LOC105345159; LOC105348695; Testican-1 |
| **Ion transport or Channels (Ion homeostasis)** | | | |
| **Higher expression under OA** | | | |
| GO:0008381 | mechanosensitive ion channel activity | 1 | LOC105342347 |
| GO:0005246 | calcium channel regulator activity | 1 | LOC105338822 |
| GO:0005244 | voltage-gated ion channel activity | 1 | LOC105338822 |
| GO:0005216 | ion channel activity | 1 | LOC105342347 |
| GO:0005245 | voltage-gated calcium channel activity | 1 | LOC105342347 |
| GO:0005262 | calcium channel activity | 1 | LOC105338822 |
| **Lower expression under OA** | | | |
| GO:0031585 | regulation of inositol 1,4,5-trisphosphate-sensitive calcium-release channel activity | 1 | LOC105324538 |
| GO:0032237 | activation of store-operated calcium channel activity | 1 | LOC105324538 |
| GO:0030171 | voltage-gated proton channel activity | 1 | LOC105343467 |
| GO:0005234 | extracellularly glutamate-gated ion channel activity | 1 | LOC105332322 |
| GO:0005227 | calcium activated cation channel activity | 1 | LOC105318785 |
| GO:0005246 | calcium channel regulator activity | 1 | LOC105333893 |
| GO:0005229 | intracellular calcium activated chloride channel activity | 1 | LOC105344773 |
| GO:0005254 | chloride channel activity | 1 | LOC105344773 |
| GO:0005375 | copper ion transmembrane transporter activity | 1 | LOC105334654 |
| GO:0046933 | proton-transporting ATP synthase activity, rotational mechanism | 1 | LOC105338985 |

**3.4.2 KEGG enrichment**

The KEGG database is used for pathway analysis of differential protein coding genes (combined with KEGG annotation results), and the hypergeometric distribution test method is used to calculate the significance of differential gene enrichment in each pathway entry. The result of the calculation returns a p-value for the significance of enrichment. A small p-value indicates that the differential gene is enriched in the Pathway. For the corresponding calculation formula, see GO Enrichment Analysis. Through pathway analysis of differential genes, we can find pathway entries enriched in differential genes, and find out which differential protein coding genes in different samples may be related to changes in cellular pathways.



**Figure S10: Up and down regulated genes enriched by KEGG pathway analysis:** The horizontal axis is the ratio (%) of the up-regulated and down-regulated genes annotated to each Level2 metabolic pathway. The vertical axis represents the name of the Level2 pathway. The number on the right represents the number of differentially expressed genes annotated to the Level2 pathway.

**Table S11**: Gene categories based on KEGG functional enrichment (Figure 1 and Figure S11)

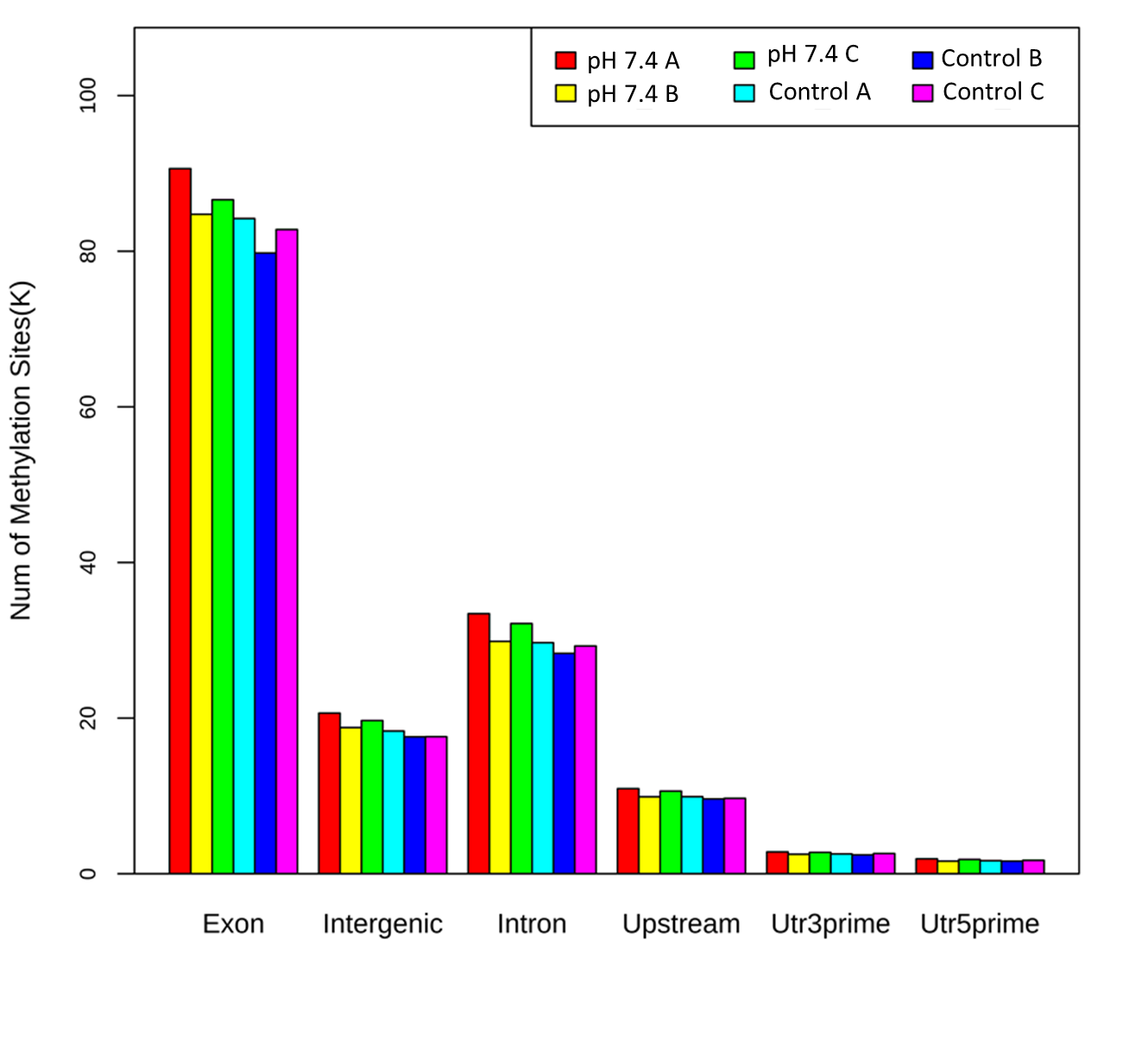
|  |  |  |  |
| --- | --- | --- | --- |
| **Functional category** | **Gene ID** | **Gene description** | **Log2FC** |
| **Amino acid metabolism** | LOC105339545 | arginase, hepatic (arginine biosynthesis & metal ion binding) | 1.03 |
| LOC105346998 | spermidine synthase (polyamine biosynthesis) | 1.53 |
| LOC105322265 | Glutamine synthetase (Glutamine biosynthesis) | -1.13 |
| LOC105324179 | DDO - D-aspartate oxidase (amino acid metabolism) | -1.49 |
| LOC105328238 | Oat - ornithine aminotransferase, mitochondrial | -3.07 |
| LOC105331916 | AGXT serine--pyruvate aminotransferase | -1.61 |
| LOC105335584 | ProDH2 - probable proline dehydrogenase 2 | -2.22 |
| LOC105339692 | CDO1 cysteine dioxygenase type | -1.67 |
| LOC105339426 | OPLAH - 5-oxoprolinase | -1.42 |
| **Other metabolisms** | LOC105319752 | OCRL - Inositol polyphosphate 5-phosphatase OCRL-1 | -2.29 |
| LOC105332291 | GPI - glucose-6-phosphate isomerase | -1.04 |
| LOC105345333 | PGM3 - phosphoacetylglucosamine mutase | -1.20 |
| LOC105329868 | ETHE1 - persulfide dioxygenase ETHE1, mitochondrial | -2.46 |
| LOC105332665 | UQCRC2 - cytochrome b-c1 complex subunit 2, mitochondrial | -1.30 |
| LOC105338985 | ATP syngamma - ATP synthase subunit gamma, mitochondrial | -1.12 |
| LOC105346116 | NDUFB5 - NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 5, mitochondrial-like | -2.61 |
| LOC105320595 | CHST15 carbohydrate sulfotransferase 15-like | -1.29 |
| LOC105340894 | MGAT4C- alpha-1,3-mannosyl-glycoprotein 4-beta-N-acetylglucosaminyltransferase C-like | -1.35 |
| LOC105346608 | smt1 - phosphoethanolamine N-methyltransferase 3 | 1.13 |
| LOC105340783 | PLA2G4A - cytosolic phospholipase A2 | -1.17 |
| LOC105320818 | Nmrk1 - nicotinamide riboside kinase 1-like | -1.03 |
| LOC105332807 | Thtpa - uncharacterized | -1.53 |
| LOC105327548 | TYMP - Thymidine phosphorylase | X |
| LOC105348214 | ALN - allantoinase, mitochondrial | -3.22 |
| **Genetic information processing**  **(Cell proliferation)** | LOC105319937 | FBXW7 beta-TrCP | -2.31 |
| LOC105330593 | EXOSC10 - exosome component 10 | -2.01 |
| LOC105336919 | Psma1 proteasome subunit alpha type 1 | -1.10 |
| LOC105343120 | GTF2H3 - general transcription factor IIH subunit 3-like | -1.16 |
| LOC105321191 | POLR3C - DNA-directed RNA polymerase III subunit RPC3 | -3.55 |
| LOC105326953 | Prpf40b pre-mRNA-processing factor 40 homolog B | -1.46 |
| LOC105318541 | RPL21 60S ribosomal protein L21 | -1.74 |
| LOC105319282 | RpL34 60S ribosomal protein L34-like | -1.04 |
| LOC105319541 | Eif2b4 translation initiation factor eIF-2B subunit delta | -1.00 |
| LOC105325583 | Rpp21 ribonuclease P protein subunit rpr2-like | -2.74 |
| LOC105327141 | WARS tryptophan--tRNA ligase, cytoplasmic | -1.76 |
| LOC105336731 | rpl7a 60S ribosomal protein L7a | -1.09 |
| LOC105339029 | mRpL24 39S ribosomal protein L24, mitochondrial | -1.39 |
| LOC105340312 | cysteine-rich protein 2-binding protein-like | -1.58 |
| **Environmental signal processing** | LOC105327454 | Gadd45a - uncharacterized | -1.05 |
| LOC105329517 | Tshr - follicle-stimulating hormone receptor | 1.19 |
| LOC105328932 | ITGB3 - integrin beta-3 | -1.23 |
| LOC105337786 | SIFaR - neuropeptide SIFamide receptor | -1.82 |
| LOC105338916 | collagen alpha-1 (I) chain | -2.90 |
| LOC105345324 | SPR - sex peptide receptor-like | -1.91 |
| LOC105348009 | 5-hydroxytryptamine receptor | -2.45 |
| **Transport and catabolism** | LOC105319768 | TUBA1A - tubulin alpha-1A chain | -1.03 |
| LOC105331919 | Pex2 peroxisome biogenesis factor 2-like | -2.51 |
| LOC105334653 | cathepsin L1-like | -1.16 |
| LOC105348613 | Gm2a - ganglioside GM2 activator-like | -4.22 |

**4. Methyl RAD analysis**

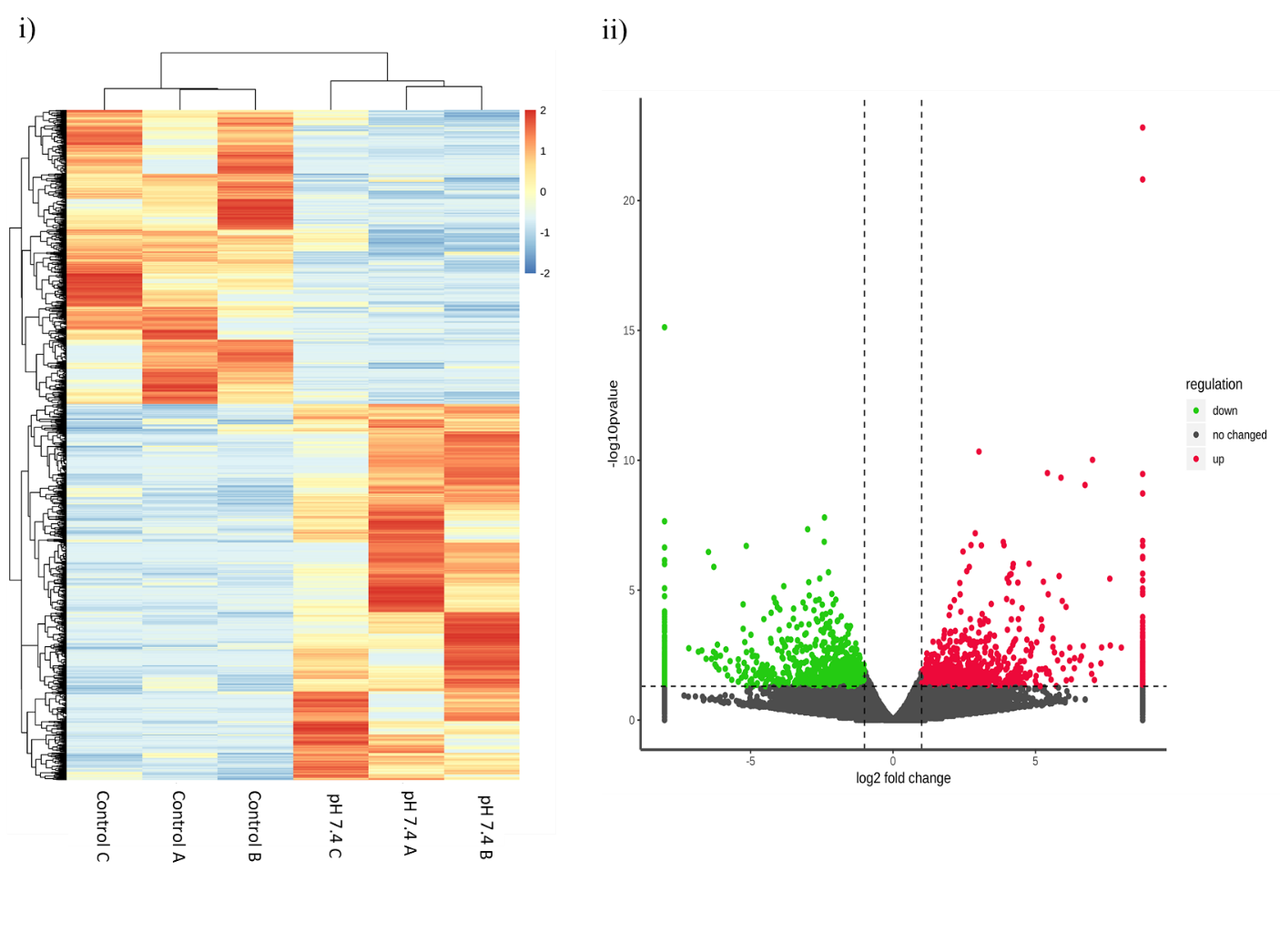
**Table S12: Sample sequencing data volume and comparison rate**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sample | Raw reads | Enzyme reads | Ratio | Mapping reads | Mapping ratio |
| pH 7.4 A | 16104023 | 11268040 | 69.97% | 2519744 | 22.36% |
| pH 7.4 B | 16122820 | 11280593 | 69.97% | 2481140 | 21.99% |
| pH 7.4 C | 16307152 | 11886792 | 72.89% | 2703034 | 22.74% |
| Control A | 16213180 | 11283479 | 69.59% | 2619817 | 23.22% |
| Control B | 16248891 | 10454206 | 64.34% | 2276948 | 21.78% |
| Control C | 16260295 | 10831211 | 66.61% | 2473788 | 22.84% |

Table S7 description: (1) Sample: sample name; (2) Raw\_Reads: offline sequencing data; (3) Enzyme Reads: raw reads After quality control, reads containing the expected cleavage sites (4) Ratio: ratio of enzyme reads and raw reads (%) (5) Mapping reads: enzyme reads with unique alignment positions on the reference sequence (6) Mapping ratio: ratio of mapping reads to enzyme reads.



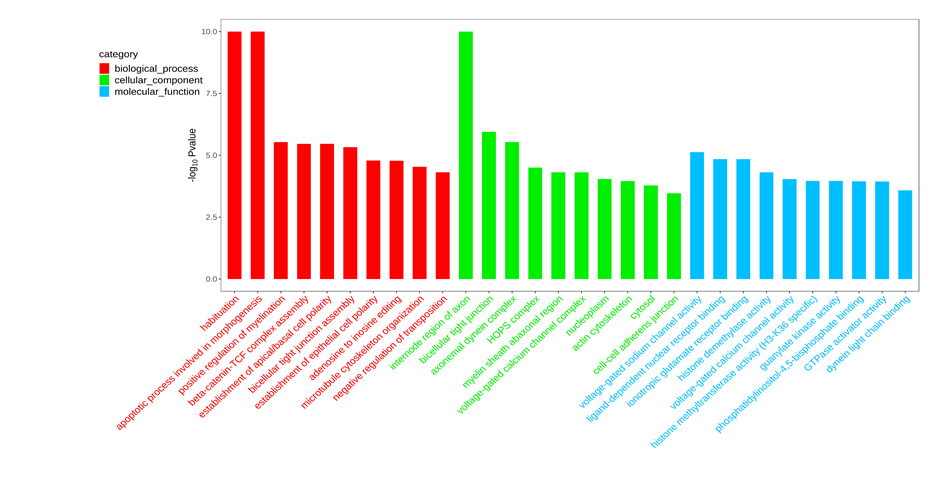
**Figure S11: Number of methylation sites found on various gene elements:** The number of methylation sites are uniformly distributed between the control and the treatment. The highest number of methylation sites are present in exon region followed by introns, intergenic and upstream regions.



**Figure S12: Clustered heat map and volcano plot:** i) The clustering of differentially methylated sites in the heat map shows clear grouping of pH 7.4 replicates and control. The methylation level is specified using the colour scale from red denoting hypermethylated level to blue denoting hypomethylation. ii) Volcano plot shows the significantly hypermethylated (up) and hypomethylated (down) above the cut off values (p value <0.05 and log2FC >1), each data point on the plot, represents a differentially methylated site. Green represents hypomethylated sites and red represents hypermethylated sites.

**Table S13: GO enrichment and top molecular function categories of DMSs**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **GO-ID** | **Molecular function category** | **p value** | **Enrichment score** |
| **Hypomethylated** | GO:0005248 | voltage-gated sodium channel activity | 7.47E-06 | 11.55 |
| GO:0016922 | ligand-dependent nuclear receptor binding | 1.42E-05 | 8.25 |
| GO:0035255 | ionotropic glutamate receptor binding | 1.42E-05 | 8.25 |
| GO:0032452 | histone demethylase activity | 4.84E-05 | 11.55 |
| GO:0005245 | voltage-gated calcium channel activity | 9.14E-05 | 7.70 |
| GO:0046975 | histone methyltransferase activity (H3-K36 specific) | 0.00010 | 9.90 |
| GO:0004385 | guanylate kinase activity | 0.00010 | 9.90 |
| GO:0005546 | phosphatidylinositol-4,5-bisphosphate binding | 0.00011 | 5.13 |
| GO:0005096 | GTPase activator activity | 0.00011 | 2.72 |
| GO:0045503 | dynein light chain binding | 0.00026 | 5.25 |
| **Hypermethylated** | GO:0008017 | microtubule binding | 1.60E-07 | 3.30 |
| GO:0003779 | actin binding | 2.70E-07 | 3.03 |
| GO:0004385 | guanylate kinase activity | 5.73E-06 | 11.50 |
| GO:0008026 | ATP-dependent helicase activity | 6.02E-06 | 15.10 |
| GO:0005085 | guanyl-nucleotide exchange factor activity | 2.13E-05 | 4.10 |
| GO:0003777 | microtubule motor activity | 2.13E-05 | 4.10 |
| GO:0005524 | ATP binding | 3.08E-05 | 1.49 |
| GO:0030676 | Rac guanyl-nucleotide exchange factor activity | 3.16E-05 | 8.94 |
| GO:0009374 | biotin binding | 8.335E-05 | 10.06 |
| GO:0051864 | histone demethylase activity (H3-K36 specific) | 0.00017 | 6.71 |

****

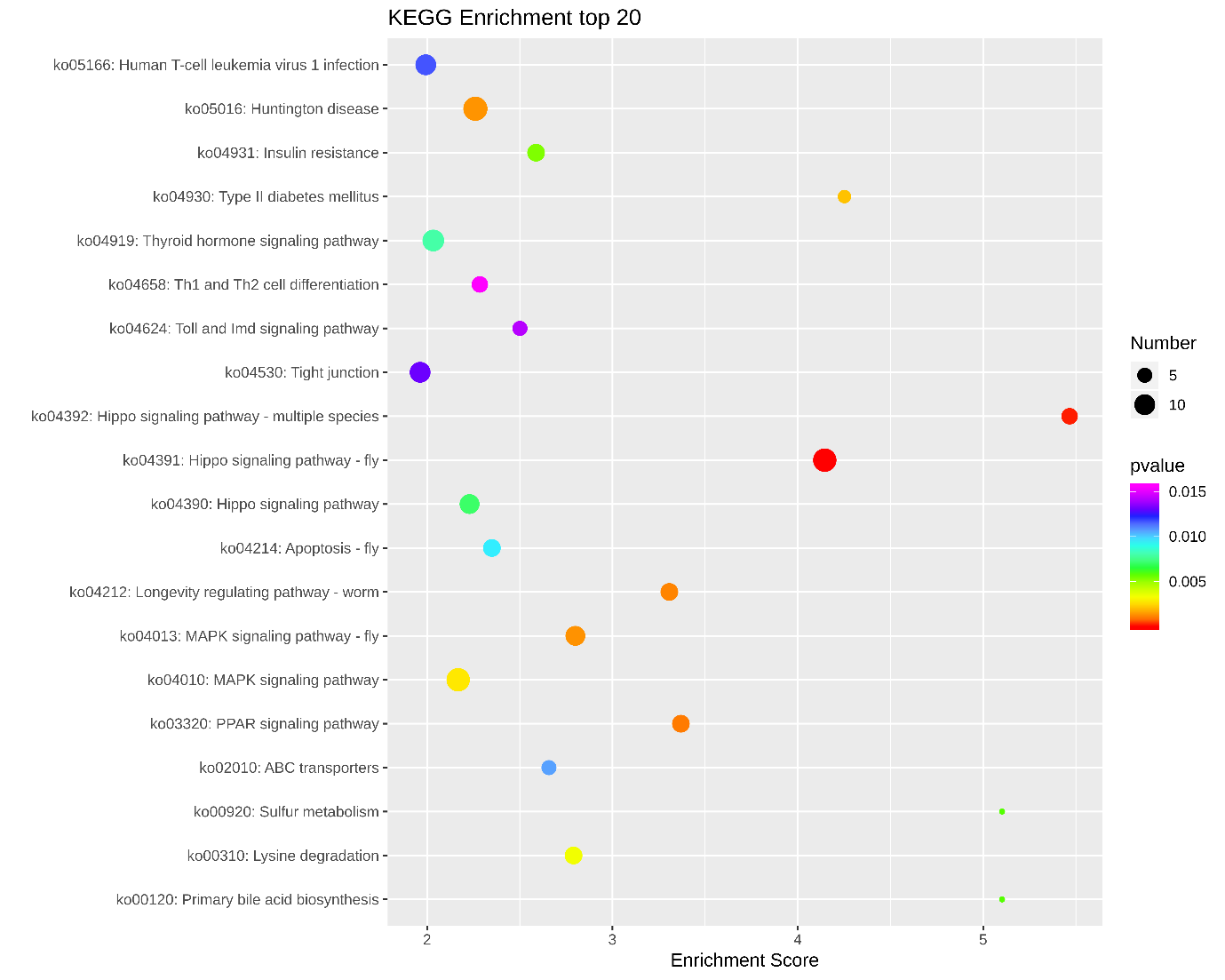
**Figure S13: Top hypomethylated GO functional categories of DMSs (differentially methylated sites)** Biological processes in red, cellular components in green and molecular function in blue.



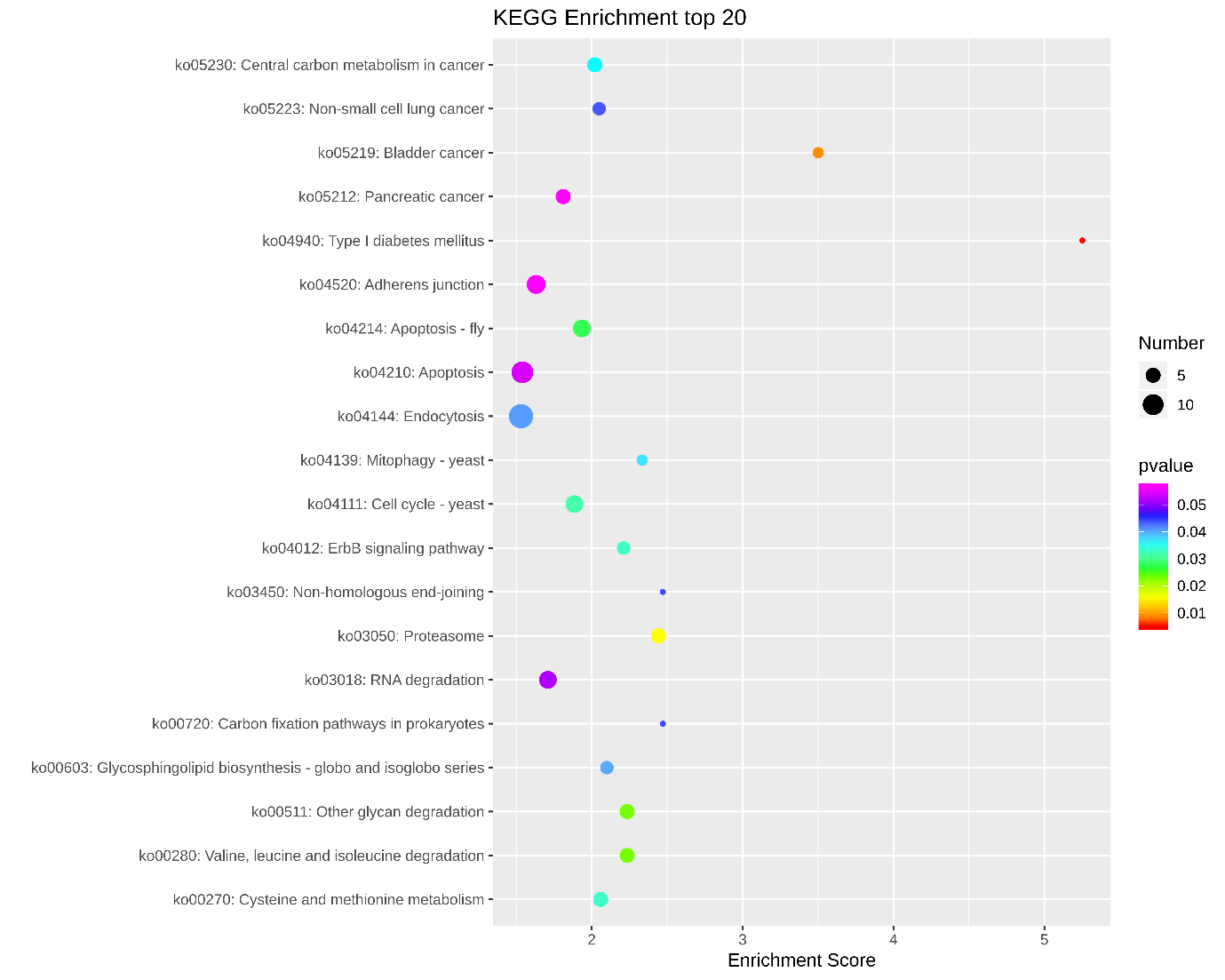
**Figure S14: Top hypermethylated GO functional categories of DMSs (differentially methylated sites)** Biological processes in red, cellular components in green and molecular function in blue.

**Table S14: GO categories of DMSs filtered using the term “calcium” relevant for biomineralisation**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Regulation** | **GO category** | | **Molecular function** | | **Gene ID** |
| Hyper  Hyper | GO:0005509 | | calcium ion binding | | LOC105342947; LOC105340209; LOC105327339; LOC105339146; LOC105330001; LOC105347050; LOC105342521; LOC105348209; LOC105320073; LOC105330338; LOC105325917; LOC105342123; LOC105339407; LOC105339725; LOC105330539; LOC105317836; LOC105344586; LOC105347404; LOC105344119; LOC105332676; LOC105330539; LOC105340285; LOC105347404; LOC105342567; LOC105339011; LOC105336164; LOC105339725 |
| GO:0005262 | | calcium channel activity | | LOC105342947; LOC105348061; LOC105338822 |
| GO:0008332 | | low voltage-gated calcium channel activity | | LOC105336481; LOC105336481 |
| GO:0005227 | | calcium activated cation channel activity | | LOC105338497; LOC105338497 |
| GO:0005245 | | voltage-gated calcium channel activity | | LOC105332264; LOC105336481 |
| GO:0005229 | | intracellular calcium activated chloride channel activity | | LOC105338497; LOC105338497 |
| GO:0086056 | | voltage-gated calcium channel activity involved in AV node cell action potential | | LOC105336481 |
| GO:0015369 | | calcium: proton antiporter activity | | LOC105342123 |
| GO:0072345 | | NAADP-sensitive calcium-release channel activity | | LOC105332264 |
| GO:0008273 | | calcium, potassium: sodium antiporter activity | | LOC105342947 |
| GO:0016286 | | small conductance calcium-activated potassium channel activity | | LOC105328597 |
| GO:0009931 | | calcium-dependent protein serine/threonine kinase activity | | LOC105335050 |
| GO:0005388 | | calcium-transporting ATPase activity | | LOC105340024 |
| GO:0004198 | | calcium-dependent cysteine-type endopeptidase activity | | LOC105318395 |
| GO:0048306 | | calcium-dependent protein binding | | LOC105347619 |
| GO:0005544 | | calcium-dependent phospholipid binding | | LOC105320073 |
| Hypo  Hypo | | GO:0008332 | low voltage-gated calcium channel activity | LOC105336481; LOC105336481 | |
| GO:0048763 | calcium-induced calcium release activity | LOC105347534 | |
| GO:0086056 | voltage-gated calcium channel activity involved in AV node cell action potential | LOC105336481 | |
| GO:0086059 | voltage-gated calcium channel activity involved SA node cell action potential | LOC105336481 | |
| GO:0005245 | Voltage-gated calcium channel activity | LOC105347534; LOC105341433; LOC105336481; LOC105345994 | |
| GO:0015369 | calcium: proton antiporter activity | LOC105342123 | |
| GO:0015278 | calcium-release channel activity | LOC105332384 | |
| GO:0099604 | ligand-gated calcium channel activity | LOC105332384 | |
| GO:0072345 | NAADP-sensitive calcium-release channel activity | LOC105341433 | |
| GO:0008273 | calcium, potassium: sodium antiporter activity | LOC105342947 | |
| GO:0005262 | calcium channel activity | LOC105342947; LOC105336504; LOC105329347; LOC105347534; LOC105320695 | |
| GO:0005246 | calcium channel regulator activity | LOC105335450; LOC105343140 | |
| GO:0005229 | intracellular calcium activated chloride channel activity | LOC105319711; LOC105319711 | |
| GO:0005509 | calcium ion binding | LOC105342947; LOC105347723; LOC105332384; LOC105344582; LOC105327957; LOC105317752; LOC105341103; LOC105342521; LOC105332900; LOC105330929; LOC105320073; LOC105327229; LOC105342123; LOC105347534; LOC105339725; LOC105321018; LOC105342681; LOC105321018; LOC105332962; LOC105326480; LOC105340285; LOC105340624; LOC105342567; LOC105325327; LOC105339011; LOC105333173; LOC105318088; LOC105339725; LOC105322203 | |
| GO:0015269 | calcium-activated potassium channel activity | LOC105326147 | |
| GO:0005388 | calcium-transporting ATPase activity | LOC105340024 | |
| GO:0005544 | calcium-dependent phospholipid binding | LOC105317752; LOC105320073; LOC105326480 | |
| GO:0015279 | store-operated calcium channel activity | LOC105333583 | |
| GO:0005227 | calcium activated cation channel activity | LOC105319711 | |
| GO:0048306 | calcium-dependent protein binding | LOC105330929 | |



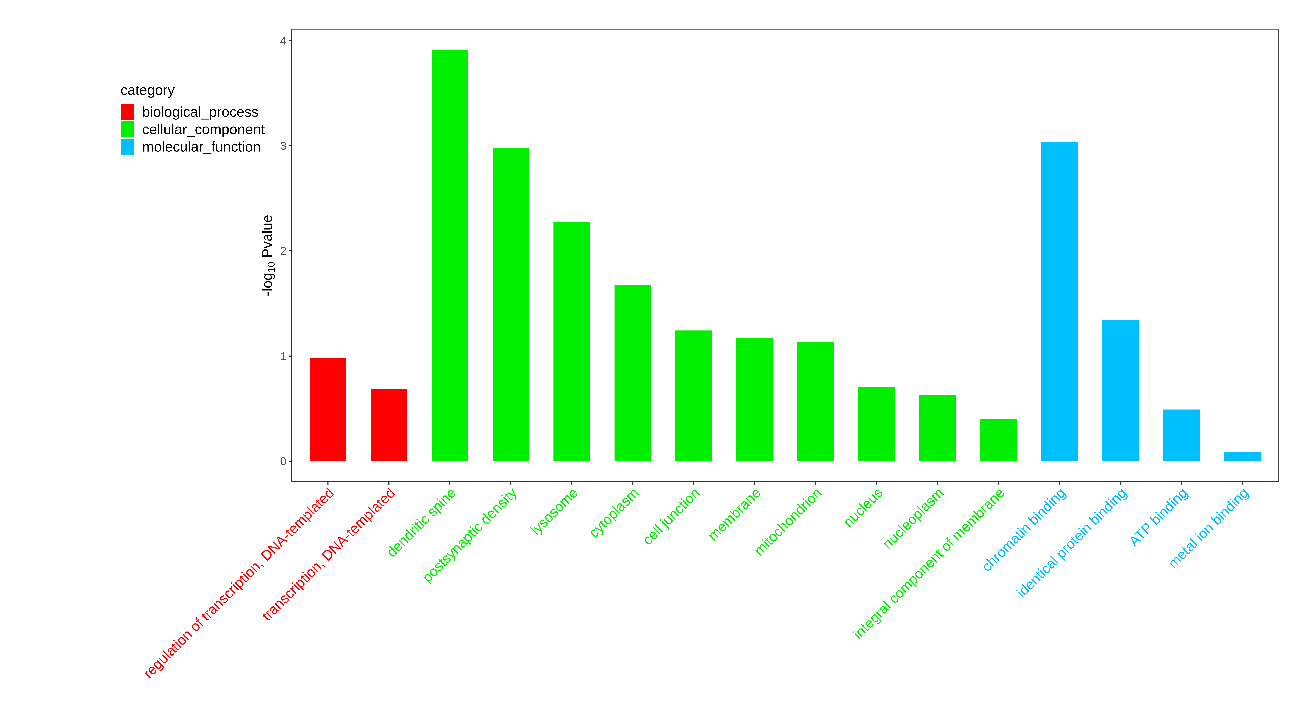
**Figure S15: Top hypomethylated DMSs by KEGG enrichment analysis:** the top pathway with many DMSs were ko04391 - Hippo signalling pathway (fly) and ko04390 Hippo signaling pathway. The bigger the dot, more the number of genes or sites under the pathway that are differentially methylated.



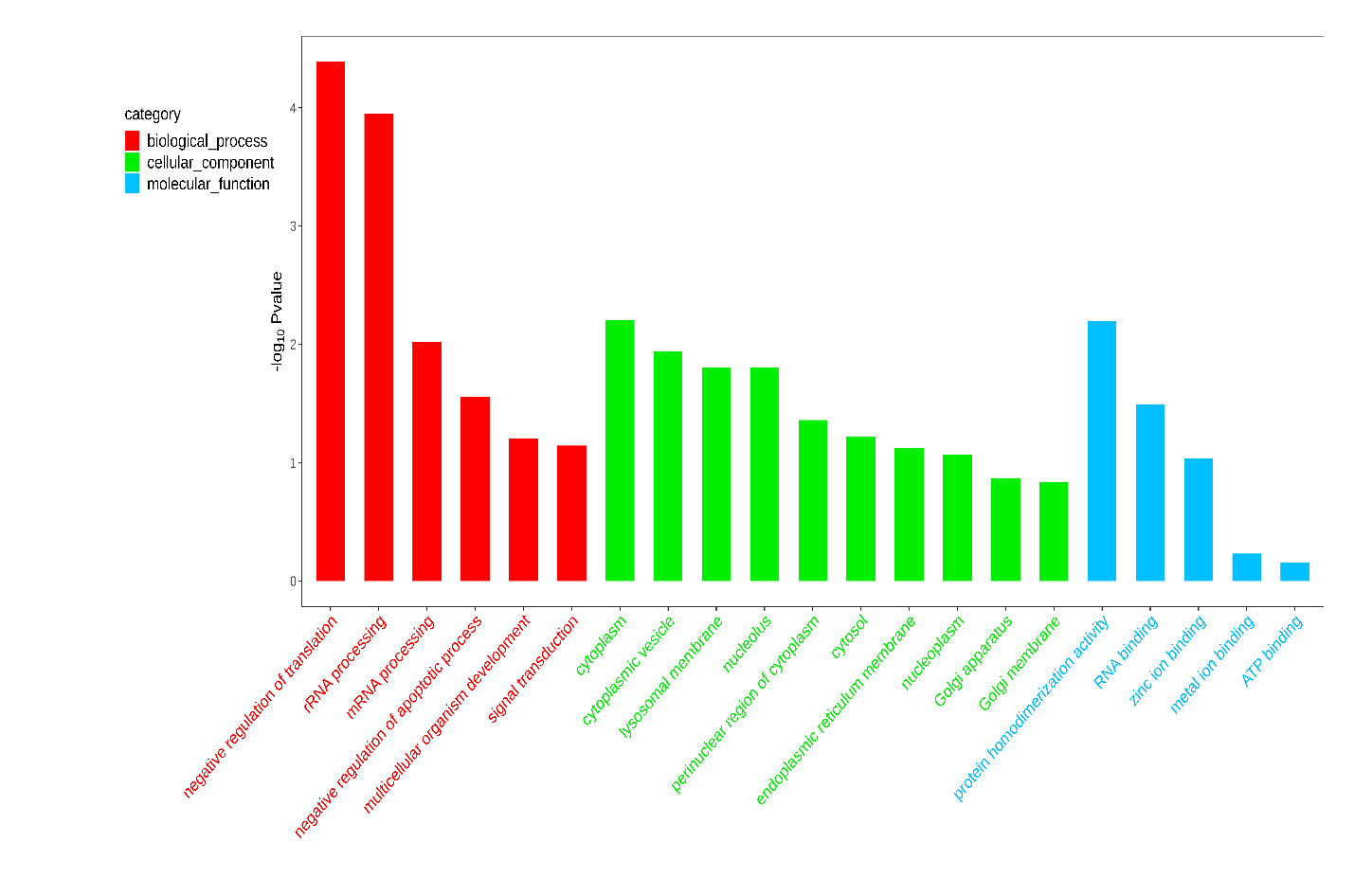
**Figure S16: Top hypermethylated** **DMSs by KEGG enrichment analysis:** The top hypermethylated pathway with many DMSs was endocytosis. The bigger the dot, more the number of genes or sites under the pathway that are differentially methylated.

**Table S15: GO enrichment and top molecular function categories of DMGs**

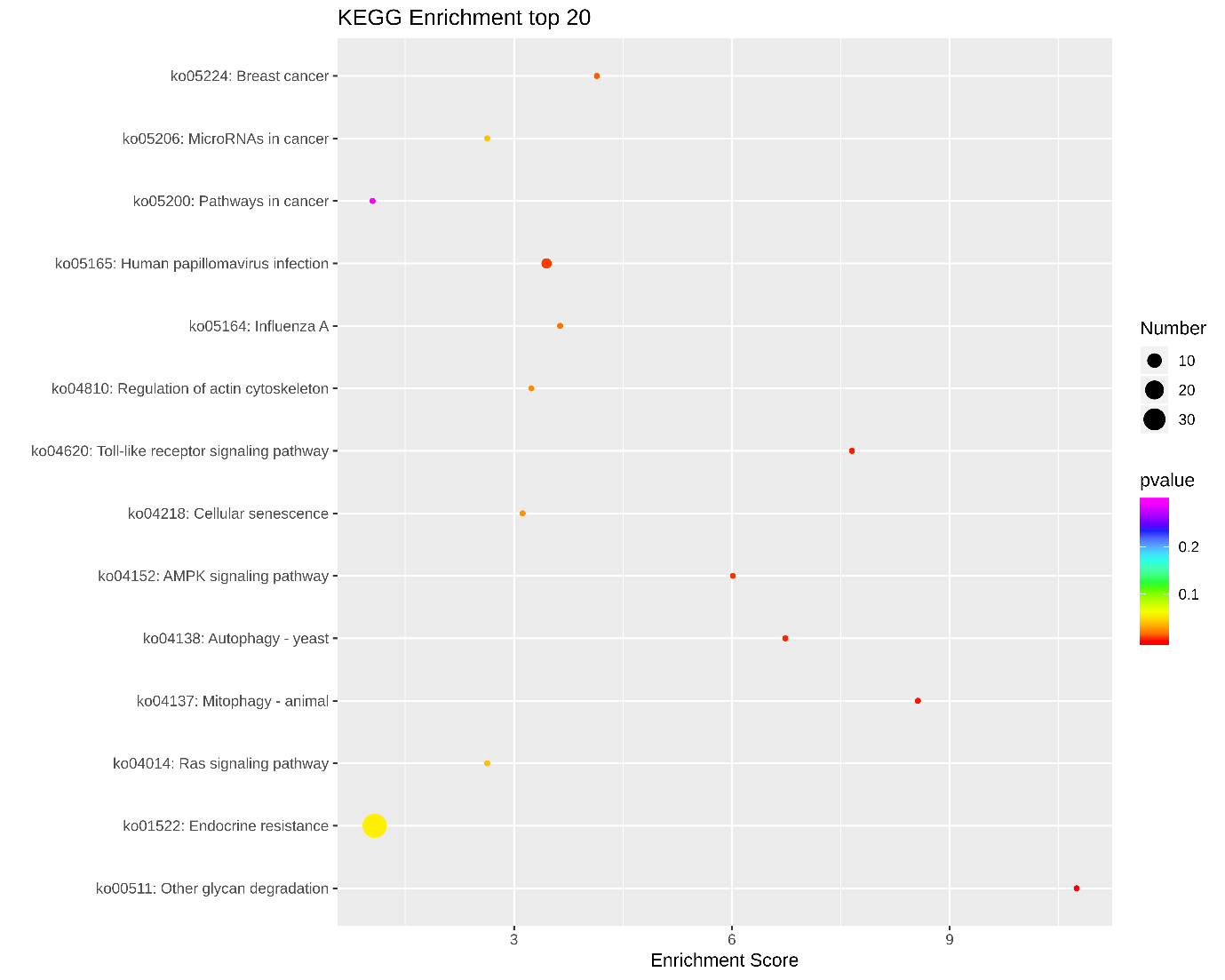
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **GO-ID** | **Molecular function category** | **p value** | **Enrichment score** |
| **Hypomethylated** | GO:0003682 | chromatin binding | 0.00092 | 5.25 |
| GO:0042802 | identical protein binding | 0.045 | 2.02 |
| GO:0005524 | ATP binding | 0.32 | 1.07 |
| GO:0046872 | metal ion binding | 0.82 | 0.54 |
| **Hypermethylated** | GO:0042803 | protein homodimerization activity | 0.0063 | 2.71 |
| GO:0003723 | RNA binding | 0.031 | 1.87 |
| GO:0008270 | zinc ion binding | 0.091 | 1.59 |
| GO:0046872 | metal ion binding | 0.58 | 0.85 |
| GO:0005524 | ATP binding | 0.70 | 0.72 |

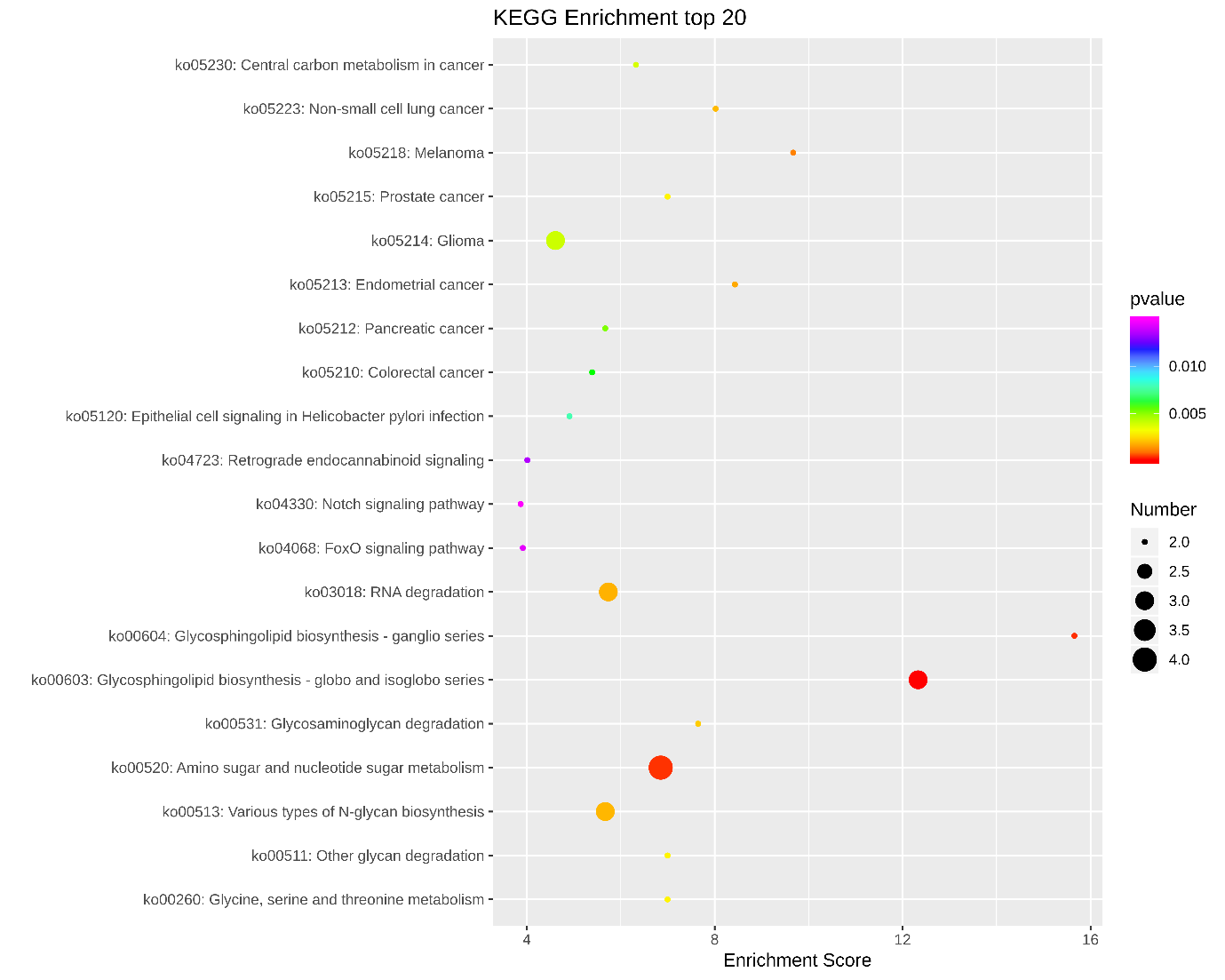


**Figure S17: Top GO categories of hypomethylated DMGs (differentially methylated genes):** Biological processes in red, cellular components in green and molecular function in blue.



**Figure S18: Top GO categories of hypermethylated DMGs (differentially methylated genes):** Biological processes in red, cellular components in green and molecular function in blue.

 **Figure S19: Top hypomethylated** **DMGs by KEGG enrichment analysis:** The top hypomethylated pathway with many DMSs was endocrine resistance – ko01522. The bigger the dot, more the number of genes or sites under the pathway that are differentially methylated.



**Figure S20: Top hypermethylated** **DMGs by KEGG enrichment analysis:** The top pathway with many hypermethylated DMSs was amino sugar and nucleotide sugar metabolism – ko01520. The bigger the dot, more the number of genes or sites under the pathway that are differentially methylated.

**Table S16: GO categories of DMGs filtered using the term “calcium” relevant for biomineralisation**

|  |  |  |  |
| --- | --- | --- | --- |
| **Regulation** | **GO category** | **Molecular function** | **Gene ID** |
| Hyper | GO:0009931 | calcium-dependent protein serine/threonine kinase activity | LOC105335050 |
| GO:0048306 | calcium-dependent protein binding | LOC105347619 |
| Hypo | GO:0005246 | calcium channel regulator activity | LOC105343140 |
| GO:0005509 | calcium ion binding | LOC105321018; |

**Table S17: DMGs filtered using the term “calcium” from GO enrichment.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Regulation** | **Gene ID** | **Molecular function** | **Log2 FC** | **p value** | **Gene Element** |
| Hyper | LOC105335050 | calcium/calmodulin-dependent protein kinase type IV | 5.35 | 0.01 | Intron |
| LOC105347619 | Battenin (Nucleoside transporter) | 1.25 | 0.04 | Exon |
| Hypo | LOC105343140 | Uncharacterised transmembrane receptor | -2.53 | 0.004 | Exon |
| LOC105321018 | eukaryotic elongation factor 2 kinase | -1.64 | 6.62E-05 | Exon |

**Table S18: RNA seq and DNA methylation – DEGs with DMSs**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Gene ID** | **Description**  **(or methylation nucleotide position in the gene)** | **Gene element** | **Regulation Log2 FC** | | **Correlation** |
| **MethylRAD** | **RNA seq** |
| LOC105338822 | voltage-dependent calcium channel subunit alpha-2/delta-1 | Intron | Hyper  (ON) | 1.00 | Positive |
| LOC105323822 | Mucin like protein | Intron | Hypo (OFF) | -2.11 | Positive |
| LOC105324638 | Tyrosine Kinase receptor torso | Intron | Hypo (OFF) | 1.05 | Negative |
| LOC105342567  Uncharacterised Ca2+ binding | 167107 | Exon | -1.89 | 1.81 | N/A |
| 183993 | Exon | -2.35 |
| 184436 | Exon | -4.50 |
| 184736 | Exon | -3.87 |
| 201788 | Exon | -3.81 |
| 209384 | Exon | Hyper (ON) |
| 209387 | Exon | Hyper (ON) |
| 229649 | Upstream | 2.22 |
| LOC105339426 | 5-oxoprolinase | Exon | -1.78 | -1.42 | Positive |
| LOC105332665 | cytochrome b-c1 complex subunit 2, mitochondrial | Intergenic | Hyper (ON) | -1.30 | Negative |
| LOC105319937 | beta-TrCP | Intergenic | 1.40 | -2.31 | Negative |
| LOC105318316  Uncharacterized  Transposase activity | 214178 | Utr3’ | Hyper (ON) | -3.81 | Negative |
| 216738 | Exon | Hyper (ON) |
| 217976 | Exon | Hyper (ON) |
| 218332 | Exon | Hyper (ON) |
| LOC105327883 | uncharacterised | Intergenic | Hyper (ON) | -3.21 | Negative |
| LOC105326758  glycine-rich cell wall structural protein 1.8-like | 28438 | Intergenic | 4.14 | -2.83 | N/A |
| 34295 | Intergenic | Hypo (OFF) |
| LOC105328651 | leucine-, glutamate- and lysine-rich protein 1 | Intron | 1.26 | -1.45 | Negative |
| LOC105324387 | ABC transporter F family member 4 | Exon | -3.22 | -1.28 | Positive |
| LOC105347351  E3 ubiquitin-protein ligase rnf213-alpha-like | 610407 | Intron | 3.51 | -1.13 | N/A |
| 610411 | Intron | 3.51 |
| 663481 | Exon | -1.31 |
| LOC105329990 | centrosomal protein of 162 kDa | Exon | 3.07 | -1.09 | Negative |
| LOC105317600 | uncharacterised | Exon | -1.67 | 1.157 | Negative |
| LOC105342036 | uncharacterised | Intergenic | Hypo (off) | 1.24 | Negative |
| LOC105326080 | uncharacterised | Exon | 1.17 | 1.31 | Positive |
| LOC105333678 | Protein memo1 | Intron | -2.38 | 1.57 | Negative |
| LOC105337216 | uncharacterised | UTR5’ | -3.24 | 2.80 | Negative |

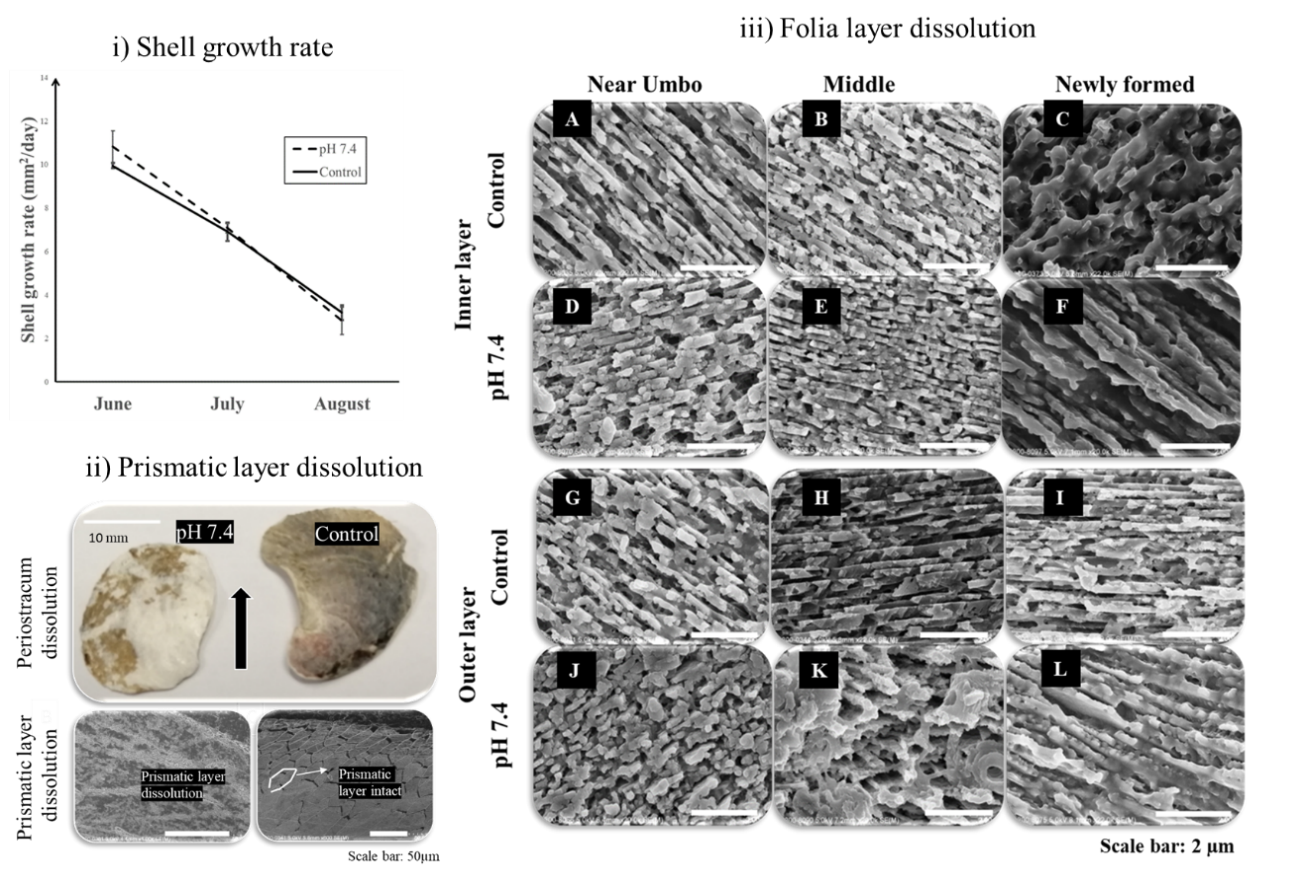
**Table S19: Correlation analysis between MethylRAD and RNA seq Log2 FC**

|  |  |  |  |
| --- | --- | --- | --- |
| Gene ID | Methyl RAD | RNA seq | Correlation |
| LOC105326080 | 1.17 | 1.31 | Positive |
| LOC105333678 | -2.38 | 1.57 | Negative |
| LOC105337216 | -3.24 | 2.8 | Negative |
| LOC105329990 | 3.07 | -1.09 | Negative |
| LOC105317600 | -1.67 | 1.157 | Negative |
| LOC105328651 | 1.26 | -1.45 | Negative |
| LOC105324387 | -3.22 | -1.28 | Positive |
| LOC105319937 | 1.4 | -2.31 | Negative |
| LOC105339426 | -1.78 | -1.42 | Positive |

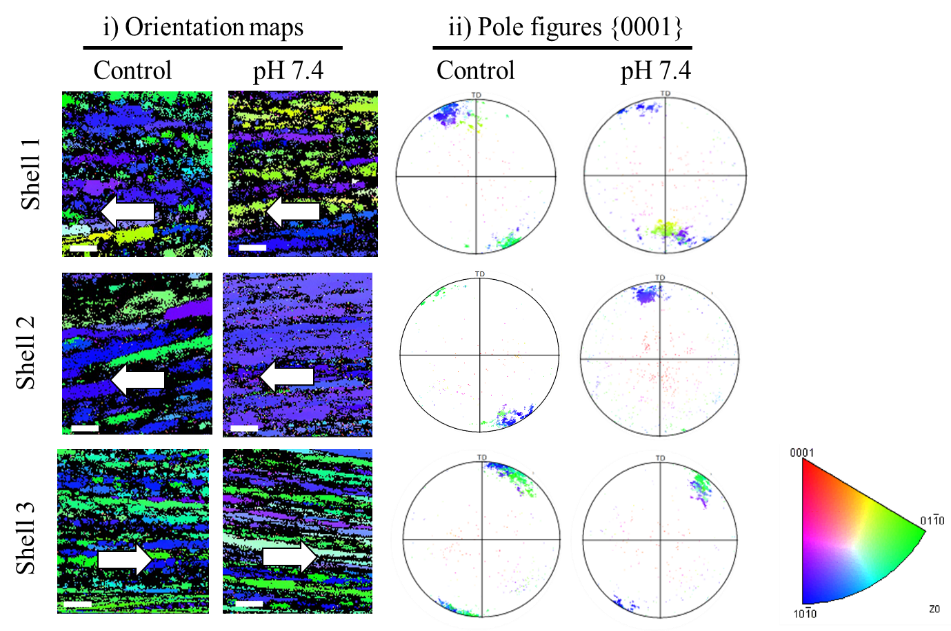
Pearson correlation was performed using the log2FC values of MethylRAD and RNA seq provided in the table and the r value was found to be – 0.46, indicating a weak negative correlation.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Gene ID | Description | Gene element | Methyl RAD | RNA seq |  |
| LOC105326080 | uncharacterised | Exon | 1.17 | 1.31 | Positive |
| LOC105333678 | Protein memo1 | Intron | -2.38 | 1.57 | Negative |
| LOC105337216 | uncharacterised | UTR5’ | -3.24 | 2.8 | Negative |
| LOC105329990 | centrosomal protein of 162 kDa | Exon | 3.07 | -1.09 | Negative |
| LOC105317600 | uncharacterised | Exon | -1.67 | 1.157 | Negative |
| LOC105328651 | leucine-, glutamate- and lysine-rich protein 1 | Intron | 1.26 | -1.45 | Negative |
| LOC105324387 | ABC transporter F family member 4 | Exon | -3.22 | -1.28 | Positive |
| LOC105319937 | beta-TrCP | Intergenic | 1.4 | -2.31 | Negative |
| LOC105339426 | 5-oxoprolinase | Exon | -1.78 | -1.42 | Positive |

**5. Shell properties**



**Figure S21: Shell properties:** i) Shell growth rate was maintained between pH 7.4 and control ii) Periostracum was almost completely dissolved in pH 7.4, further leading to loss of prismatic microstructure iii) The folia layer remains intact except slight dissolution towards the outer layer of the shell in near umbo and middle region. The newly formed folia remained intact in both pH 7.4 and control indicating the quality of shell formed is not affected under OA, however affected due to dissolution.

****

**Figure S22: Electron Back Scatter Diffraction of folia**: (i) Crystal orientation maps of the calcite crystals in the folia layer of the shell in reference to {0001} plane (Scale bar: 5 µm). Crystallographic planes of calcite are color-coded linked to the normal crystallographic direction using the colour code. (ii) Pole figures for calcite correspond to the crystallographic orientation maps and follows the same colour code (iii) Colour code for orientation maps and pole figure. The white arrow denotes the direction of elongation of the folia.

**References:**

Dickson, AG, and Frank J Millero. 1987. 'A comparison of the equilibrium constants for the dissociation of carbonic acid in seawater media', *Deep Sea Research Part A. Oceanographic Research Papers*, 34: 1733-43.

Kim, Daehwan, Ben Langmead, and Steven L Salzberg. 2015. 'HISAT: a fast spliced aligner with low memory requirements', *Nature methods*, 12: 357.

Mehrbach, Carl, CH Culberson, JE Hawley, and RM Pytkowicx. 1973. 'Measurement of the apparent dissociation constants of carbonic acid in seawater at atmospheric pressure 1', *Limnology and Oceanography*, 18: 897-907.

Pierrot, DE, DWR Wallace, E Lewis, D Pierrot, E Lewis, R Wallace, D Wallace, W Wallace, and DWR Wallace. 2011. 'MS Excel program developed for CO2 system calculations'.