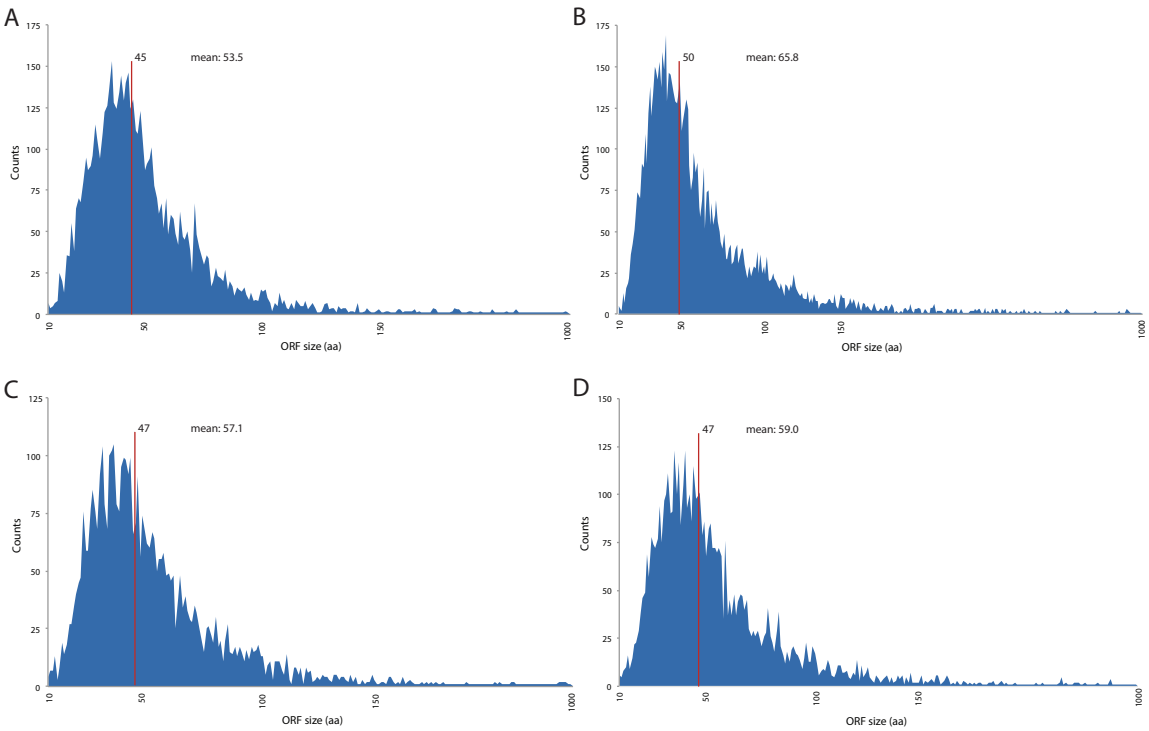
Supplementary Material for:

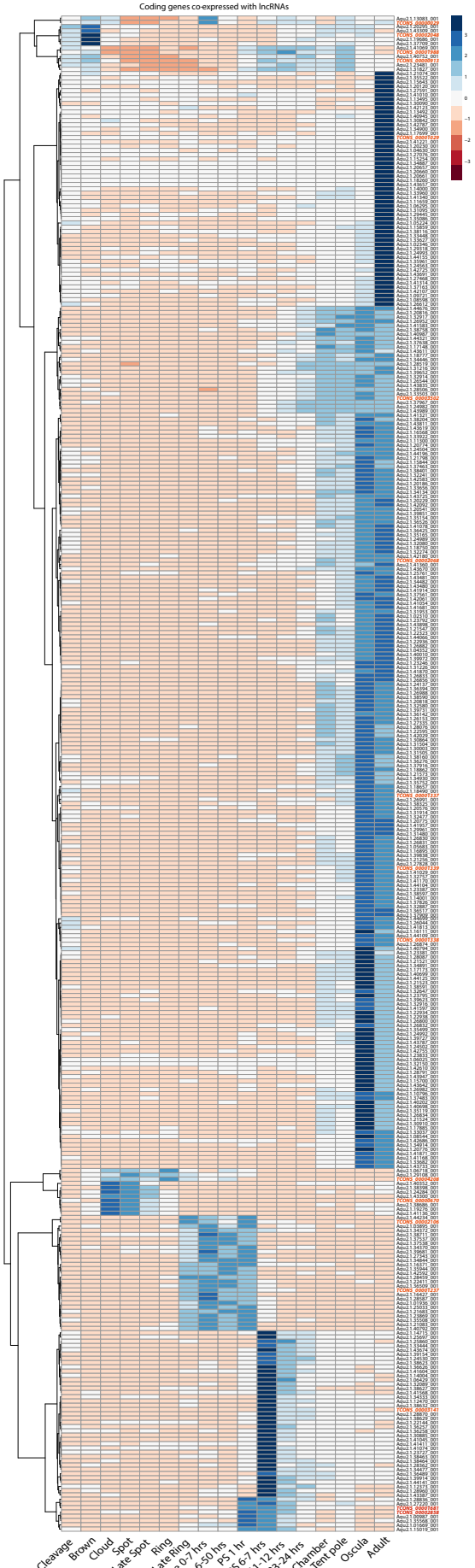
"Dynamic and widespread lncRNA expression in the sponge and the origin of animal complexity"

Authors: Federico Gaiti, Selene L. Fernandez-Valverde, Nagayasu Nakanishi, Andrew D. Calcino, Itai Yanai, Milos Tanurdzic and Bernard M. Degnan

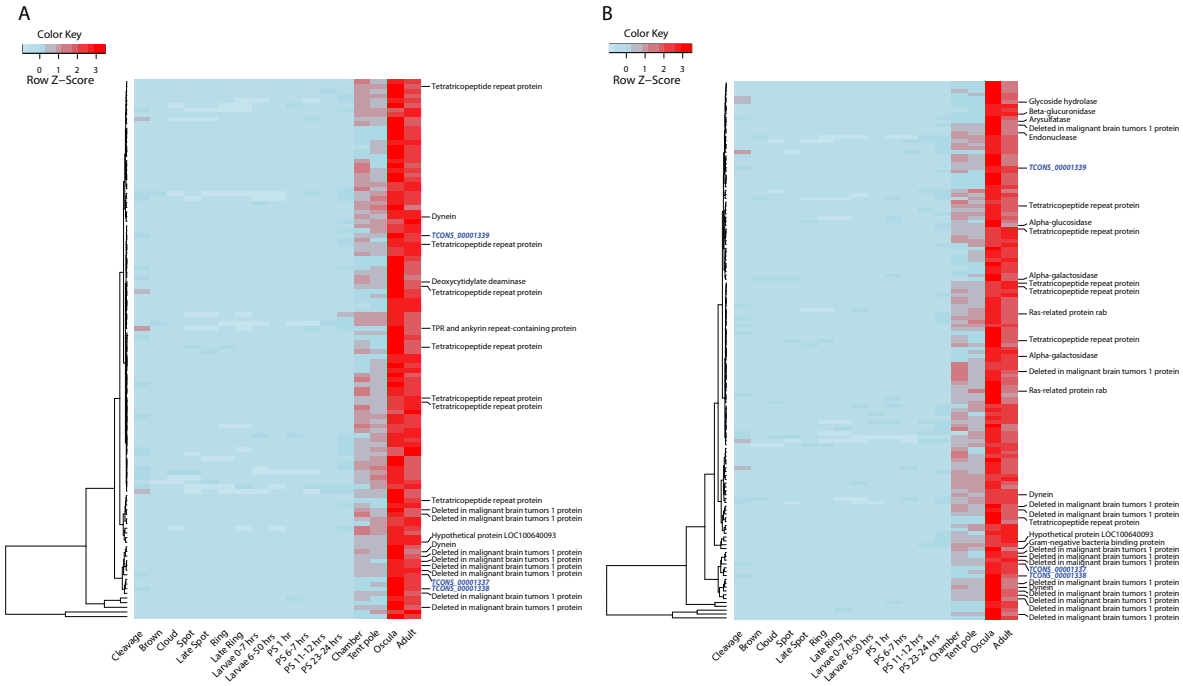
This PDF includes:



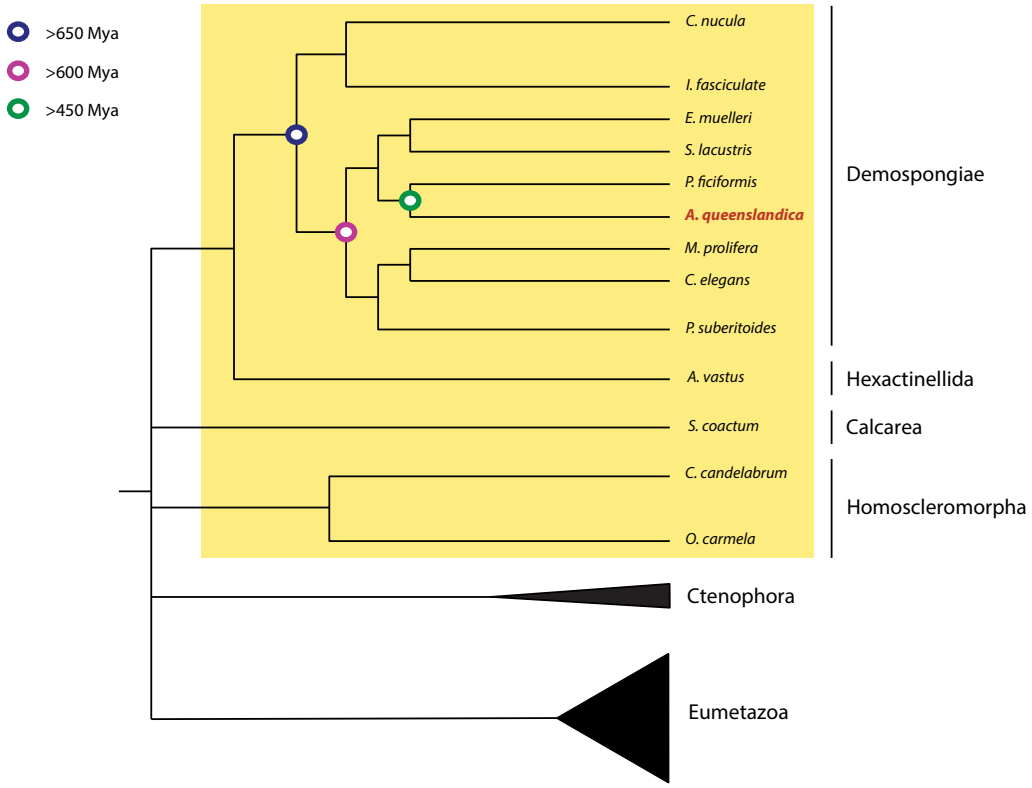
Supplemental Figure 1 – Validation of ORF size cut-off. The distribution of open reading frame size (x-axis) for the set of putative lncRNAs in (A) the precompetent larva, (B) the competent larva, (C) the juvenile and (D) the adult, when no ORF size cut-off was imposed (>300 amino acids), is shown here. The median ORF size is shown as a vertical red line.


Coding genes co-expressed with lncRNAs
3
2
1
0
−1
−2
−3
Cleavage
Brown
Cloud
Spot
Late Spot
Ring
Late Ring
Larvae 0-7 hrs
Larvae 6-30 hrs
PS 1 hr
PS 6-7 hrs
PS 11-12 hrs
PS 23-24 hrs
Chamber
Tent pole
Oscula
Adult

Supplemental Figure 2 – Heatmap showing expression abundance of all protein-coding genes co-expressed with the differentially expressed lncRNAs (rows) across the whole *Amphimedon* developmental time course, from early cleavage to adult (columns). Expression levels (normalized counts) were measured by CEL-Seq and rescaled by row. Genes were clustered by hierarchical clustering. To reduce noise, the differentially expressed lncRNAs and protein-coding genes with an overall expression of less than 100 CEL-Seq normalized counts throughout the whole developmental time course were discarded. Pearson's correlation and a Fisher's exact test were then used to correlate the expression level of each differentially expressed lncRNA with the protein-coding genes, using R (R Development Core Team 2010). Only genes that showed more than 0.95 correlation (positive and negative) and a *p-value* <0.05 were considered to be co-expressed. Blue shows high expression, red shows low expression. lncRNAs are shown in red. The codes used for this gene co-expression analysis are available for download at: https://bitbucket.org/selene_fernandez/amphimedon-lncrnas.



Supplemental Figure 3 – Developmental expression profiles of two of the six differentially expressed lncRNAs co-expressed with protein-coding genes involved in key metazoan developmental processes. Expression levels were measured by CEL-Seq and rescaled by row. Red indicates high expression level, light blue low expression. Rows corresponding to protein-coding genes with an enriched GO term (Fisher's exact test, *p-adj* <0.05) are shown on the *right.* For a complete list of enriched GO terms and relative protein-coding genes, see Supplemental Table 11. lncRNAs are shown in blue. Both *TCONS_00001337* (A) *and TCONS_00001339* (B), as also observed for *TCONS_00001338* (Fig. 5B), are expressed late in development and are co-expressed with protein-coding genes enriched for scavenger receptor activity, carbohydrate metabolic processes and hydrolase activity.



Supplemental Figure 4 – Schematic representation of the evolutionary relationship among the 13 sponge species used in this study. Yellow box highlights the sponge clade. *Amphimedon queenslandica* is shown in bold red. For detailed phylogenetic relationship analyses of sponges please refer to (Hill et al. 2013) and (Worheide et al. 2012).

References

Hill MS, Hill AL, Lopez J, et al. 2013. Reconstruction of family-level phylogenetic relationships within Demospongiae (Porifera) using nuclear encoded housekeeping genes. PloS one 8: e50437.

R Development Core Team. 2010. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Worheide G, Dohrmann M, Erpenbeck D, Larroux C, Maldonado M, Voigt O, Borchiellini C and Lavrov DV. 2012. Deep phylogeny and evolution of sponges (phylum Porifera). Advances in marine biology 61: 1-78.