RUTGERS
NEW BRUNSWICK
Department of Ecology, Evolution and Natural Resources; Department of Marine and Coastal Sciences, Rutgers University, New Brunswick, NJ 08901

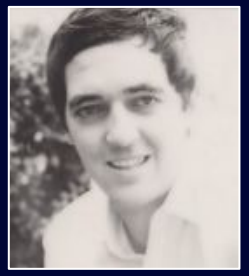
Analysis of coral genomes to elucidate the basis of biomineralization


Paul


Debashish


Jeana and Tali


Huan


Stan


Udi and Nicole

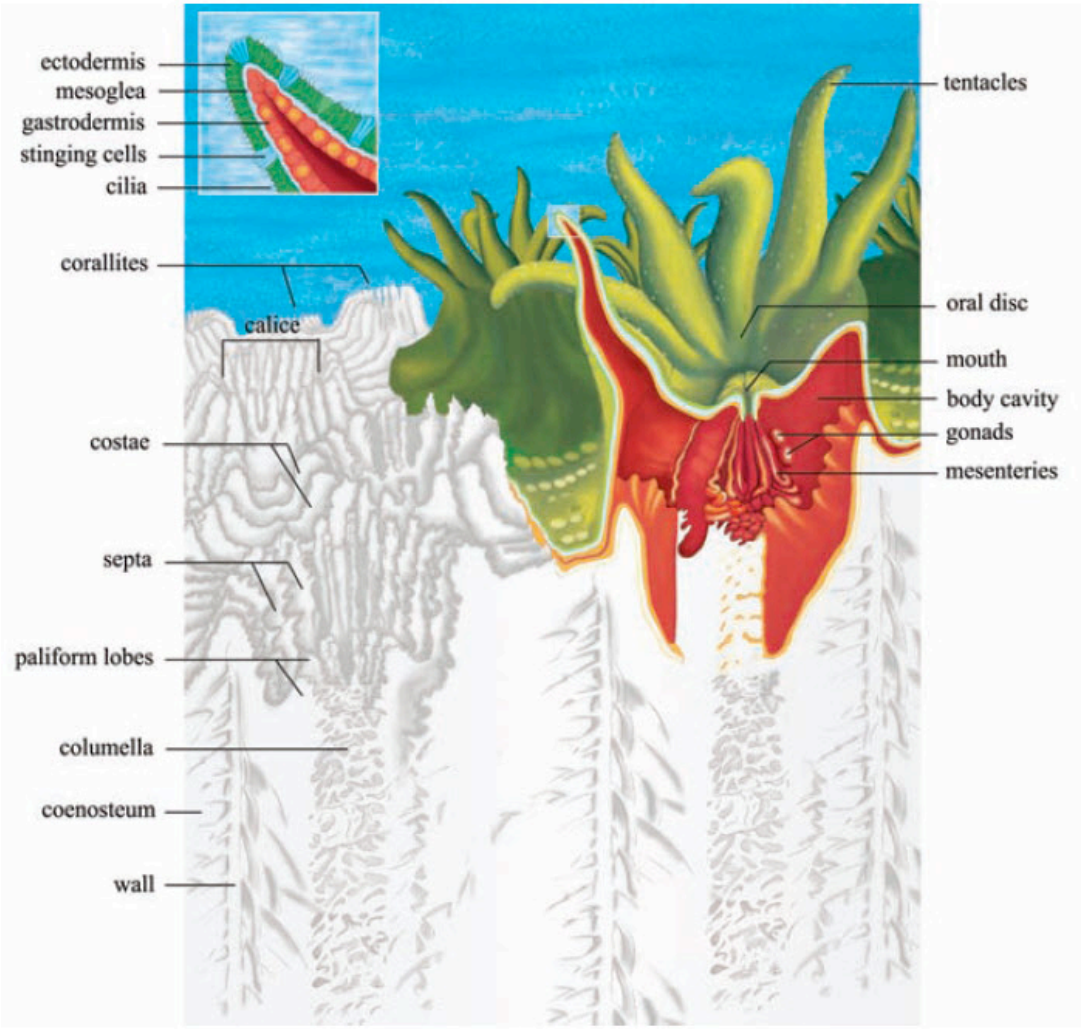
Debashish Bhattacharya (Co-PI)
Paul Falkowski (PI)
Tali Mass (Co-PI; Univ. Haifa, Israel)
Huan Qiu (post-doc)
Stan Von Euw (post-doc)
Jeana Drake (PhD student)
Rutgers Genome Cooperative
EF-1416785: 09/15/2014 – 08/31/2017
Ocean Acidification: Mechanisms of Coral Biomineralization

Goals of Talk

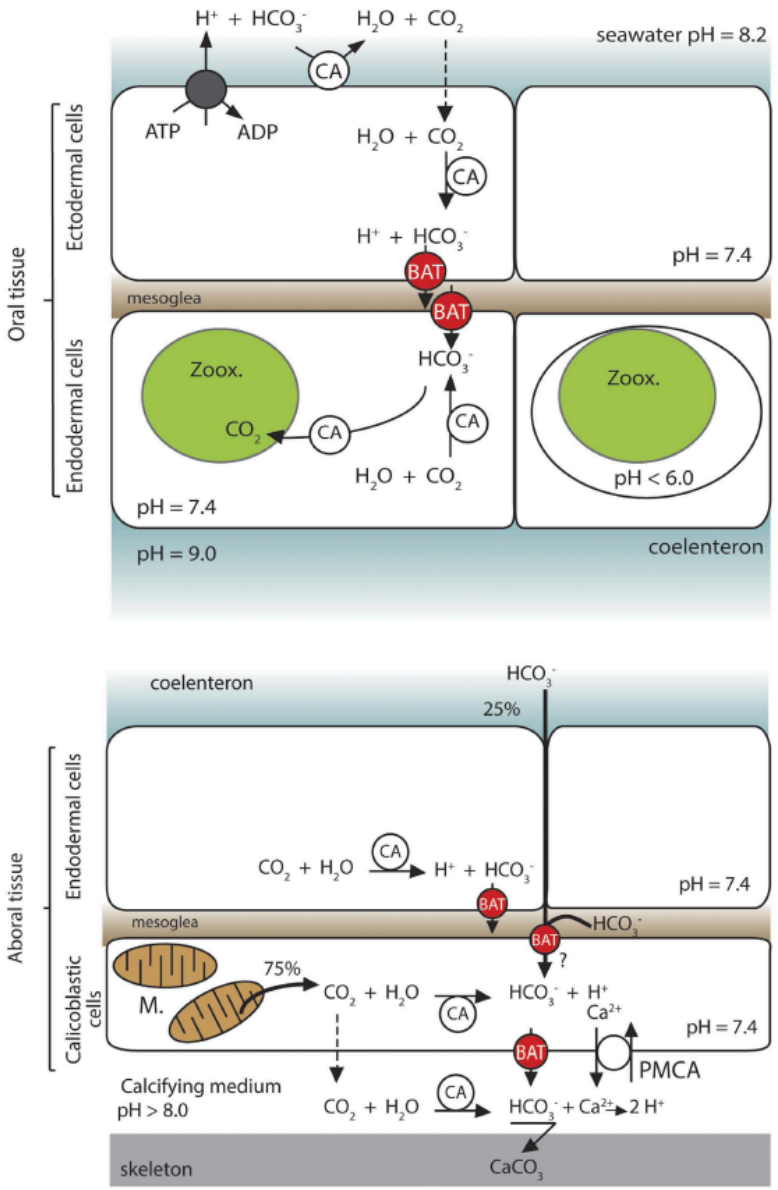
1. Brief summary of previous (relevant) coral work done in our lab by Paul's group.
2. The types of genomic analyses undertaken with some specific results.
3. Future directions.
4. *How do corals make rocks and what is the impact of ocean acidification?*
5. *What is the evolutionary history of coral biomineralization?*

Three dimensional structure $CaCO_3$ precipitation

Genomics, real-time PCR, immunolocalization

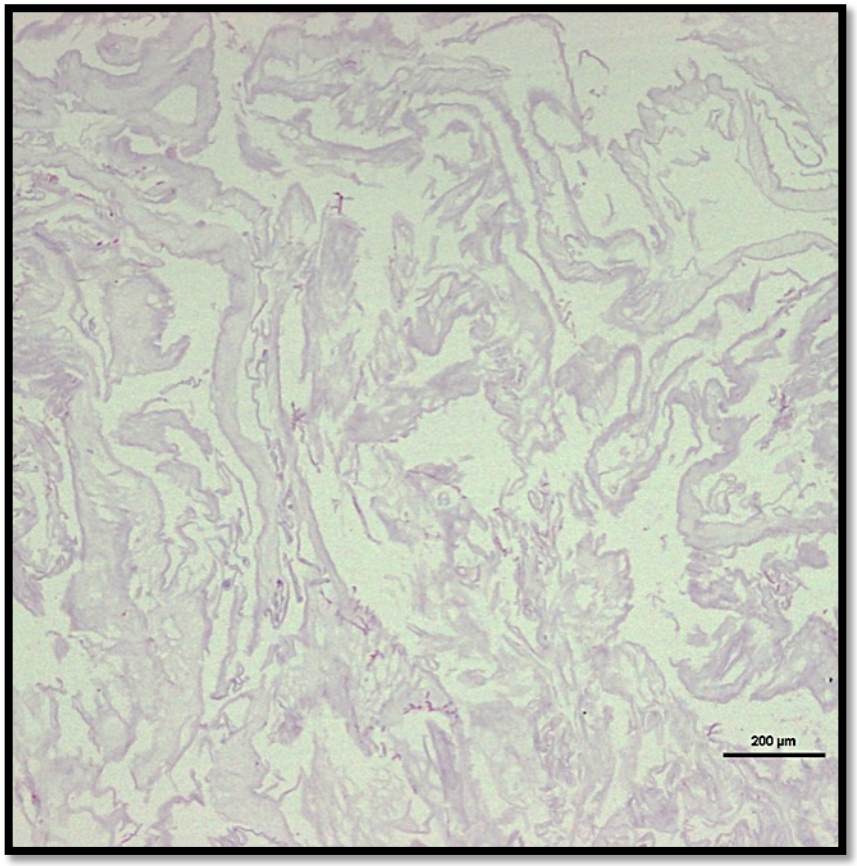


Based on Veron (1986)
http://coral.biota.biodiv.tw/book/export/html/2



Zoccola et al. 2015, Sci Rep

The skeletal organic organic matrix:
Key to understanding the biomineralization reaction



- Contains a suite of proteins, lipids and polysaccharides
- Composed of two fractions: the soluble (SOM) and insoluble organic matrix (IOM)
- Proteins of the IOM form a framework for crystal growth
- Proteins of the SOM play a role in nucleation and crystal growth

***Pocillopora verrucosa* SOM and IOM after decalcification of coral skeleton (photo*:* I. Brikner)**

Previous work in Falkowski lab:

- Establish a culture system that functions similar to corals (study biomineralization under controlled conditions)
- Characterize skeletal organic matrix proteins (36) that play a role in calcification
- Understand their role in nucleation and crystal growth
- Study the impact of seawater pH on these proteins

Proteomic analysis of skeletal organic matrix from the stony coral *Stylophora pistillata*

Jeana L. Drake[a], Tali Mass[a], Liti Haramaty[a], Ehud Zelzion[b], Debashish Bhattacharya[a,b], and Paul G. Falkowski[a,c,1]

[a]Environmental Biophysics and Molecular Ecology Program, Institute of Marine and Coastal Sciences, [b]Department of Ecology, Evolution, and Natural Resources, and [c]Department of Earth and Planetary Sciences, Rutgers University, New Brunswick, NJ 08901

3788–3793 | PNAS | March 5, 2013 | vol. 110 | no. 10

Cloning and Characterization of Four Novel Coral Acid-Rich Proteins that Precipitate Carbonates In Vitro

Tali Mass,[1] Jeana L. Drake,[1] Liti Haramaty,[1]
J. Dongun Kim,[1,3] Ehud Zelzion,[2] Debashish Bhattacharya,[2]
and Paul G. Falkowski[1,3,4,*]

Current Biology *23*, 1126–1131, June 17, 2013

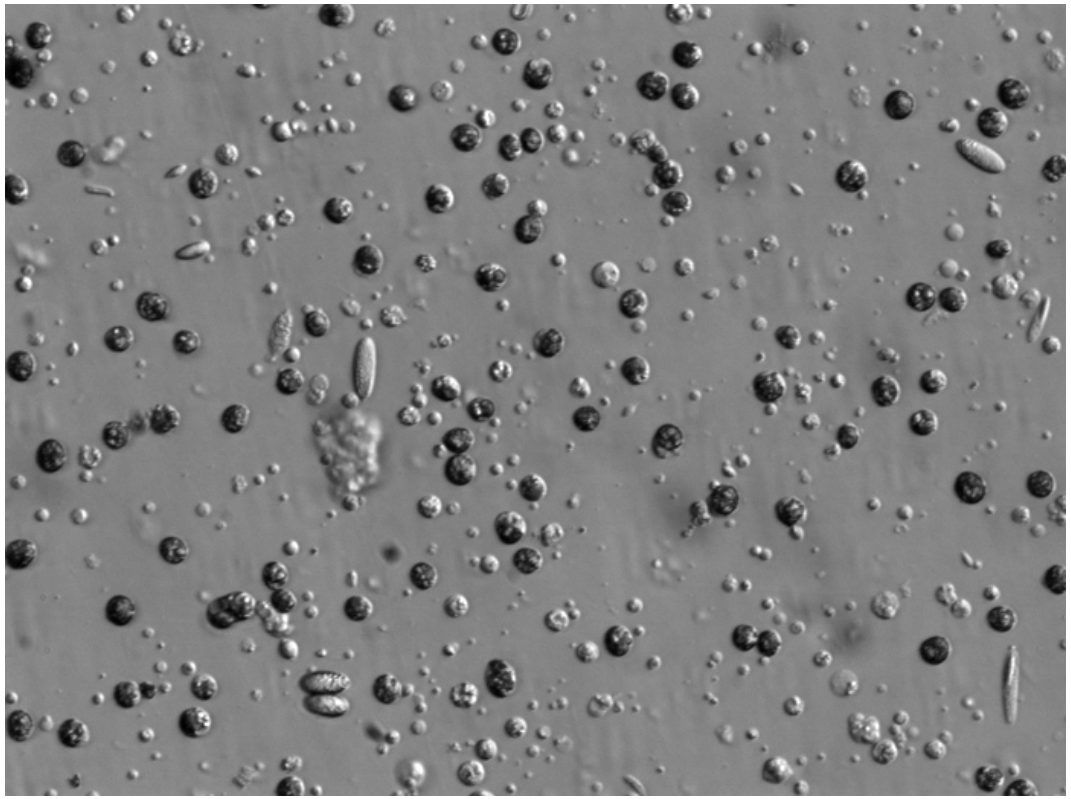
Immunolocalization of skeletal matrix proteins in tissue and mineral of the coral *Stylophora pistillata*

Tali Mass[a], Jeana L. Drake[a], Esther C. Peters[b], Wenge Jiang[c], and Paul G. Falkowski[a,d,1]

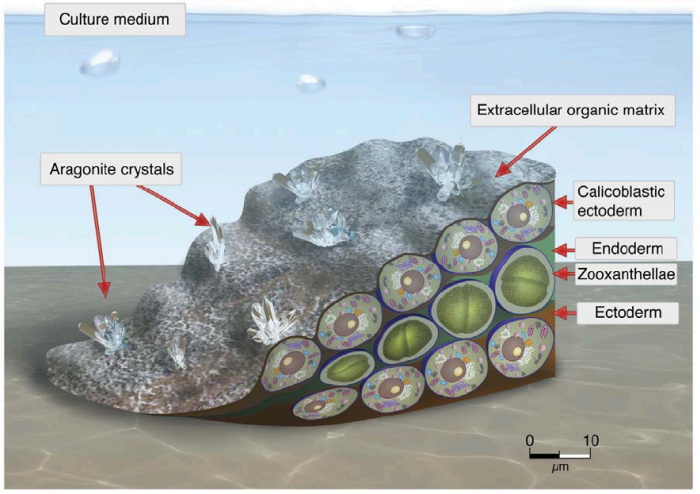
12728–12733 | PNAS | September 2, 2014 | vol. 111 | no. 35

Cell cultures “proto-polyp” formation: Tali Mass

Seriatopora sp.: T_0



- Coral “zoo”
  - Proteobacteria
  - Bacteria
  - Fungi
  - Archaea
  - Protozoa-ciliates



Dissociated tissue consists of a mixture of cell types including free zooxanthellae. and individual endoderm and ectoderm cells.

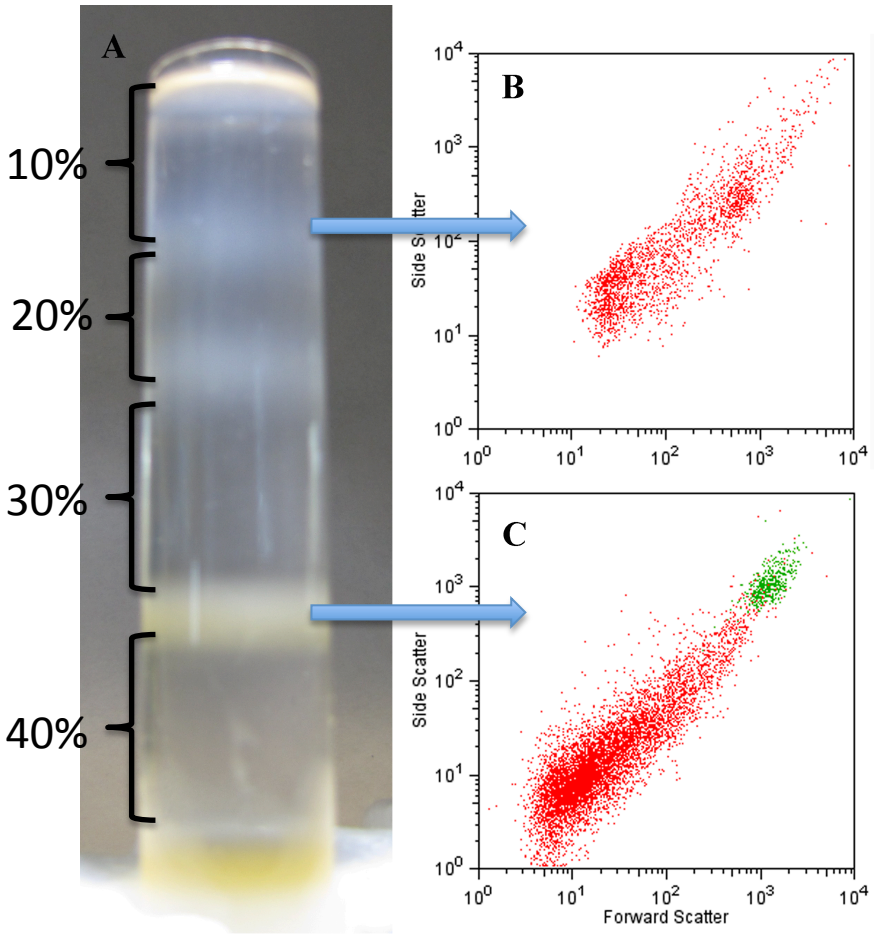
These cultures are a challenge to keep viable over a period longer then a month.

Proteomic Analyses of Stony Coral SOM

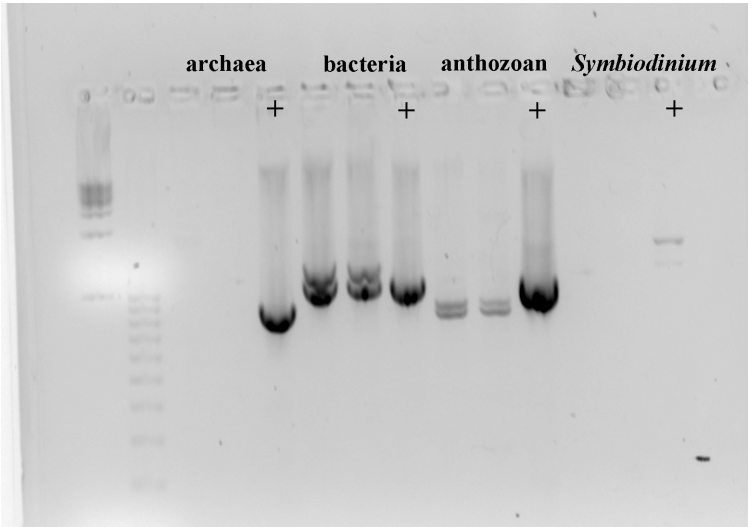
Protein	Coral Specific	Other Biomineralizers	Non-biomineralizers
Cadherins	-	+	+
Carbonic anhydrase	-	+	-
CARP4	+	-	-
Collagens	+	-	-
Galaxin	+	-	-
Laminin G domains	-	-	+
LDL-receptor domains	-	+	+
Other highly acidic proteins	+	-	-
Zona pellucida	-	-	+

Drake et al. 2013 PNAS,
Ramos-Silva et al. 2013 Mar. Biol. Evol.

Zooxanthellae-free Coral DNA



- Opti-Prep gradient followed by cell culture of the zooxanthellae-free fraction
- DNA extraction and PCR verification



Mass et al. Current Biology, 2013

Approach - Genomic

- Genomic DNA: 131,845,772 high quality reads (17.1 Gbp) used for genome assembly.
- ~530,000 contigs = 450 Mbp; average coverage of 40x
- The mRNA-Seq was 44,880,704 (3.8 Gbp) high quality reads, assembled into 44,219 contigs with a N50 = 871 bp and average coverage >5x.
- A total of 21,683 gene models were predicted using AUGUSTUS;
- assembled transcripts and Illumina RNA-seq data were incorporated as hints to aid in building these models.
- 39% GC content
  - ❖ Similar to the genome of *Acropora digitifera* (420 Mbp, 23,700 gene models)
- Where do we start the search for putative biomineralization proteins?

Mass et al. Current Biology, 2013

Acidic proteins in biomineralization:

>30% aspartic and glutamic acids

Pinctada fucata

Aspein, Pif: A shell-matrix protein (Tsukamoto et al. 2002, Suzuki et al. 2009)

Patinopecten yessoensis

MSP-1: A shell-matrix protein (Sarashina and Endo 1998)

Atrina rigida

Asprich family: A shell-matrix protein (Gotliv et al. 2004)

Panaeus japonicus

Crustocalcin: Ca^{2+} binding protein (Endo et al. 2000)

Danio rerio

Starmaker: Otolith biomineralization (Soellner et al. 2003)

Rattus norvegicus

Dentin sialophosphoprotein (Ritchie et al. 1994)

Homo sapiens

Dentin sialophosphoprotein (Gu et al. 2000)
Bone sialoprotein (Fisher et al. 1990)

Genome-wide search for acidic macromolecules

- Protein is longer than 100 amino acids
- contains >35% aspartic or glutamic acids
- exhibits an N-terminal secretion signal peptide

Stylophora pistillata (*Seriatopora* sp.)

Coral Acid-Rich Protein (CARPs)

- **CARP1**

MFHSWWMTLLILGSTVSFVFTEGDHLKPGHSEDEHDEDEHDEEMADHADEQNPADEEETEDEE
KDDDKMEDDSDDDEEDESQGDDEGEDENDQSHLEHDAFLDKDGKVSWEEFKKGHFSDDGKDE
DAKEQMKEDEEKFKFADEDGDGKLDLEEYMAFYHPGDNPRMTEFTIEDSLKKHDKDKDGQVSK
KEFLATFSDVNDDAKEEMEKDFNNNFDKDKNGRLNKEEMKSWLFPDDDFSTEEPKTLIKEADED
KDGKLTMDEIMKNYKVFIEDEPEDSSHDEL

- **CARP2**

MVLVLIQATHLLCSVLILVSSAPVENEIRIRGPKLEDEEEGNFPPIMPAQLELKEREFPKKEEERKEAK
EDENMLREELKHFRDEESLKNVITRLERELAFEKTEREENRETEDLSNEELVERELPEEVDEIPEEKG
ARELKEENGLEMFYRNLQRKLKEKQERDMPVKEMEYESPEDQEEEMQERELDEEFKEKSKRELEE
EDLEETGAEEREDKRELAEEVSSREELEENEEELALKRKRGEENMATEWEIPESVEHYDENKRSKH
PPKHMREREAERERERFDDHGHKEREREEFRERQRELALSNGGKLHERELEGRKQRQEIGLHGVR
REESERFRFRVRGE

- **CARP3**

MRNFLIALALIAIFAAVQSMPADTHEDKARNYVPESANATDPAVAEPSEAENDPAQSETEPAAEE
ASTDAASDTKEDDSAAADDSSDDDLDDDSVDENDEDDEDDEDDEDDEDDEDDEDDEDDS
DDSDDGDDGDDENDGDDEDDGDDEDDGDDE

- **CARP4**

AKMLSKSGKIMIVRDNDDDDDDDDDDDDDDDDDDDDDDDFSDDNEEMLSFEVDEVEEKDVN
GNDVDKRQRHSVDSFDDVDFTFTKVDTQAKYDGLPVTNVNLSATLLGFSSLEIMVYLFRQAGKV
AFGNETFRVEKGTVKFNIR

Mass et al. Current Biology, 2013

Coral Acid-Rich Protein

- **CARP1- isoelectric point=4.23**

MFHSWWMTLLILGSTVSFVFTEGDHLKPGHSEDEHDEDEHDEEMADHADEQNPADEEETE
DEEKDDDKMEDDSDDDEEDESQGDDEGEDENDQSHLEHDAFLDKDGKVSWEEFKKGHFS
DDGKDEDAKEQMKEDEEKFKFADEDGDGKLDLEEYMAFYHPGDNPRMTEFTIEDSLKKHD
KDKDGQVSKKEFLATFSDVNDDAKEEMEKDFNNNFDKDKNGRLNKEEMKSWLFPDDDFST
EEPKTLIKEADEDKDGKLTMDEIMKNYKVFIEDEPEDSSHDEL

- **CARP2- isoelectric point=4.78**

MVLVLIQATHLLCSVLILVSSAPVENEIRIRGPKLEDEEEGNFPPIMPAQLELKEREFPKKEEERK
EAKEDENMLREELKHFRDEESLKNVITRLERELAFEKTEREENRETEDLSNEELVERELPEEVDE
IPEEKGARELKEENGLEMFYRNLQRKLKEKQERDMPVKEMEYESPEDQEEEMQERELDEEFK
EKSKRELEEEDLEETGAEEREDKRELAEEVSSREELEENEEELALKRKRGEENMATEWEIPESV
EHYDENKRSKHPPKHMRERERERERFDDHGHKERERREEFRERQRELALSNGGKLHERELE
GRKQRQEIGLHGVRREESERFRFRVRGE

- **CARP3- isoelectric point=3.04**

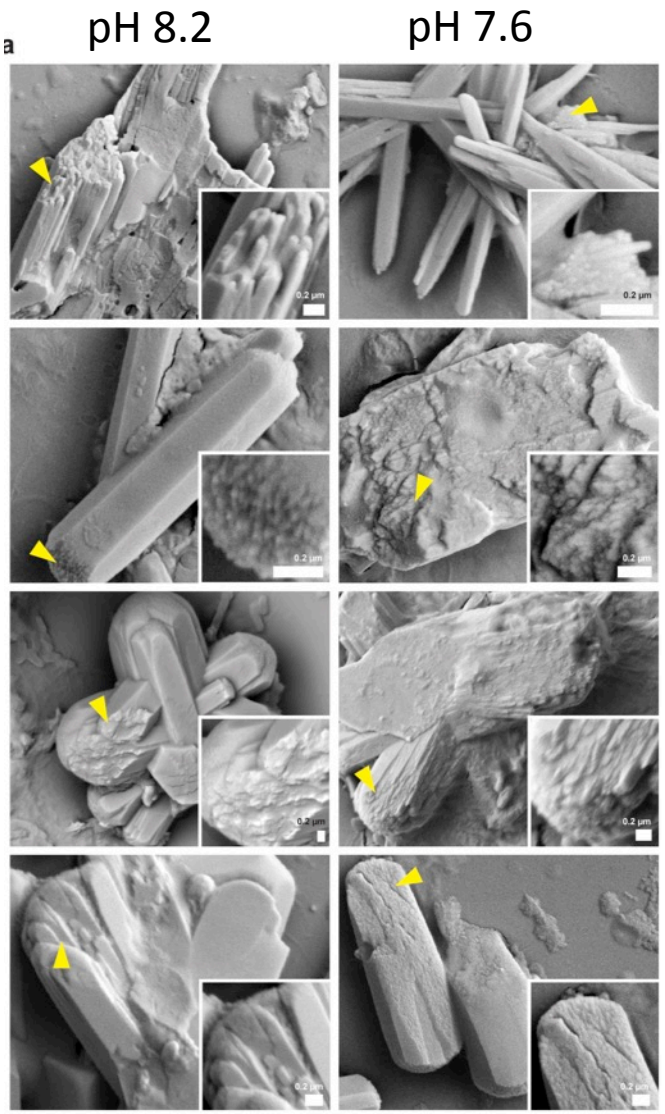
MRNFLIALALIAIFAAVQSMPADTHEDKARNYVPESANATDPAVAEPSEAENDPAQSETEPA
AEEASTDAASDTKEDDSAAADDSSDDDLDDDSVDENDEDDEDDEDDEDDEDDEDDEDDE
DDEDDSDDSDDGDDGDDENDGDDEDDGDDEDDGDDE

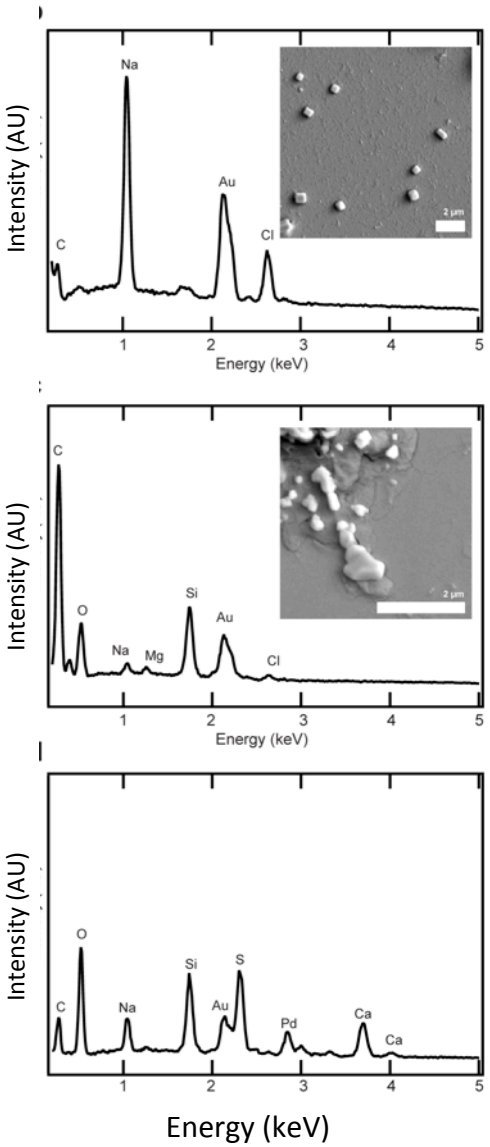
- **CARP4- isoelectric point=3.99**

AKMLSKSGKIMIVRDNDDDDDDDDDDDDDDDDDDDDDDDFSDDNEEMLSFEVDEVEEK
DVNGNDVDKRQRHSVDSFDDVDFTFTKVDTQAKYDGLPVTNVNLSATLLGFSSLEIMVYLFR
QAGKVAFGNETFRVEKGTVKFNIR

Mass et al. Current Biology, 2013

In-vitro $CaCO_3$ precipitation at pH 8.2 and 7.6 in the presence of CARPs





Left (SEM)
Images of $CaCO_3$ crystals grown in artificial seawater containing 0.1 μM CARP 1–4 at pH 8.2 and 7.6.

Right (elemental composition)
Top – protein-free artificial seawater;

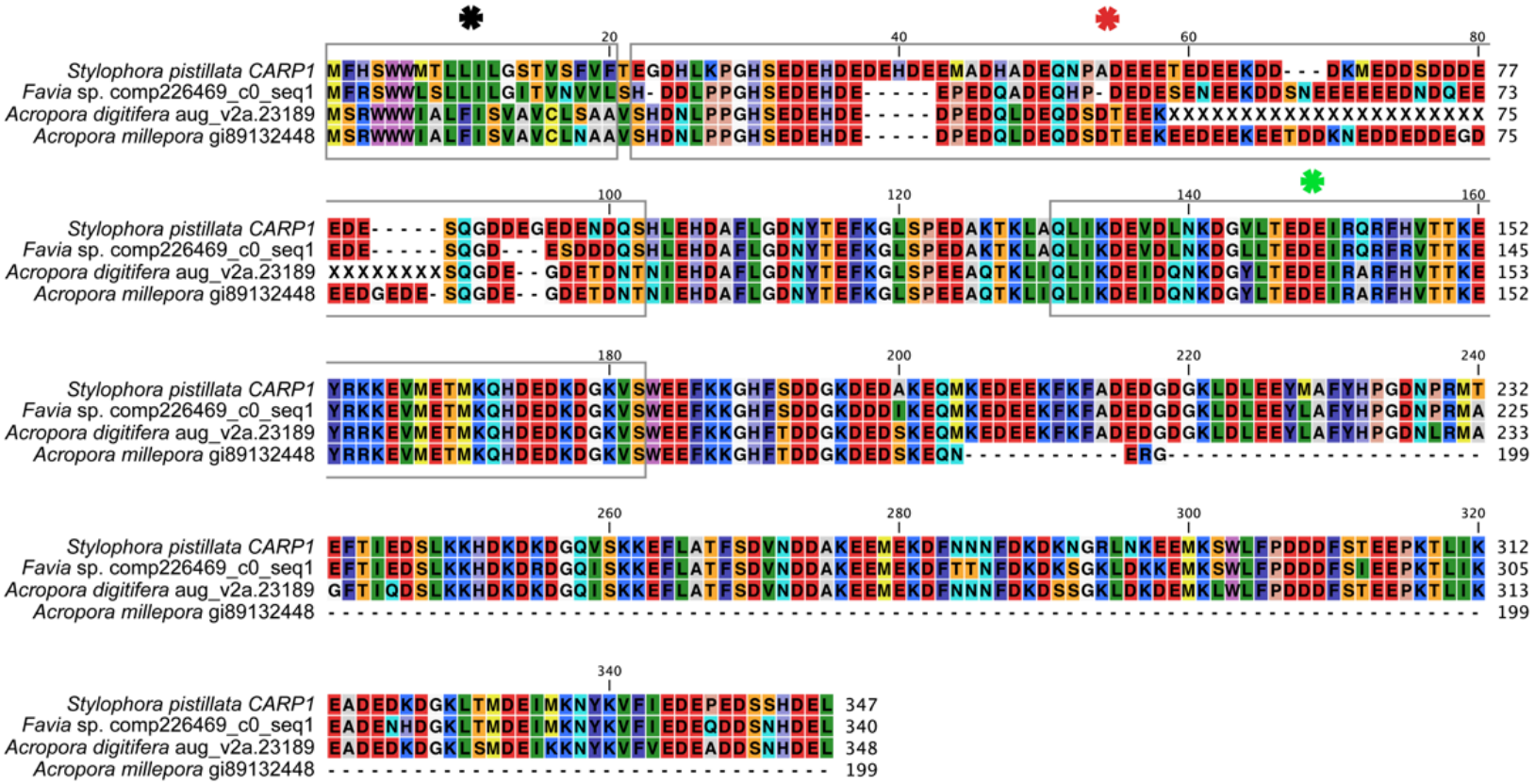
Middle - artificial seawater containing BSA;

Bottom - artificial seawater containing CARPs 1–4.
The Au and Pd peaks derive from the gold coating and the Si peak derives from the silica wafer base.

Mass et al. Current Biology, 2013

CARP Evolution

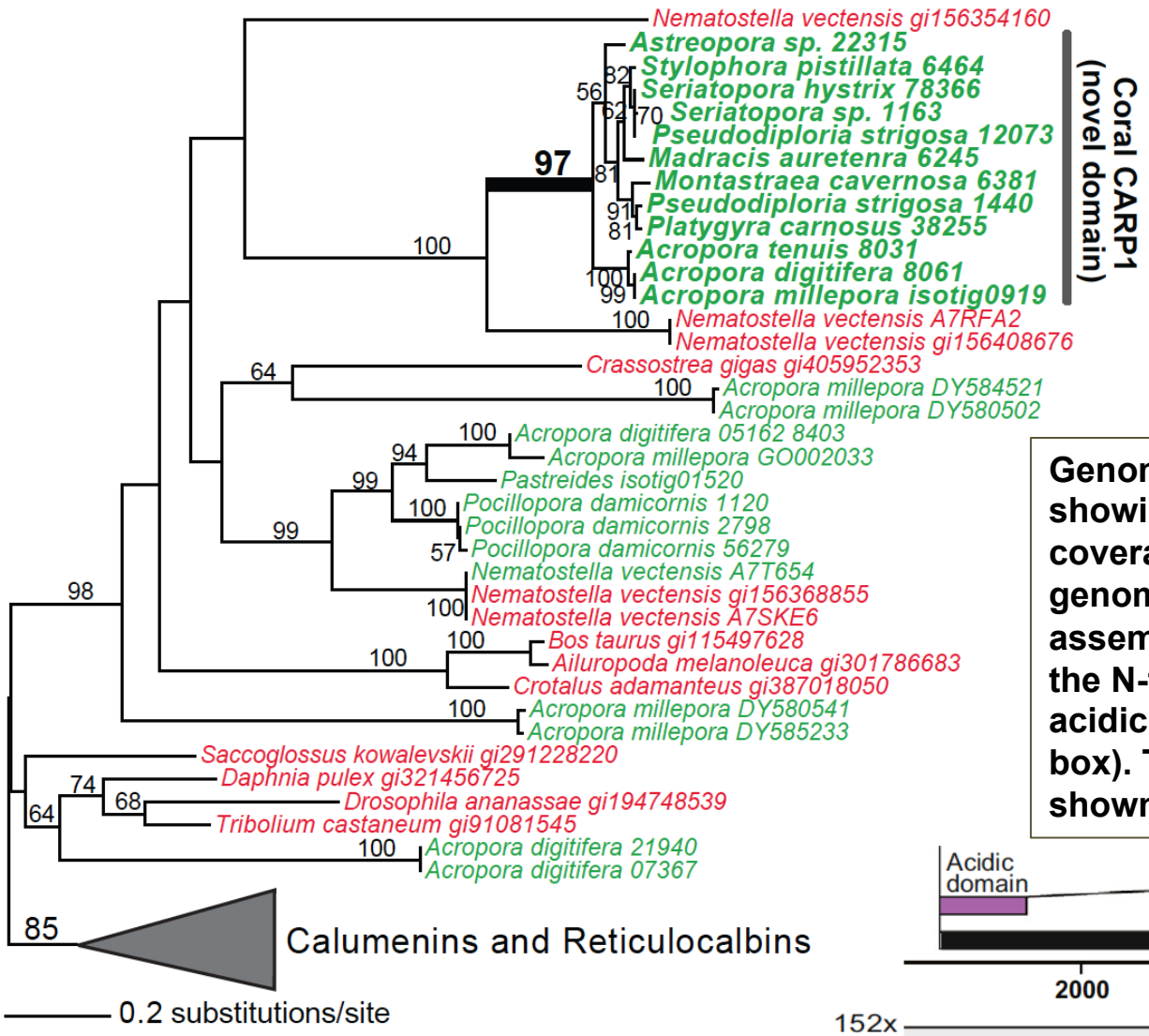
CARP1 gene-fusion



Favia sp *S. pistillata* *A. digitifera* *A. millepora*

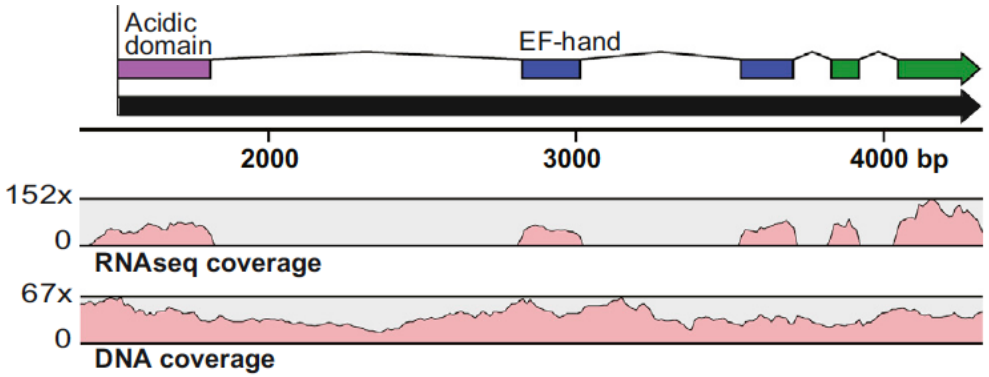
Mass et al. Current Biology, 2013

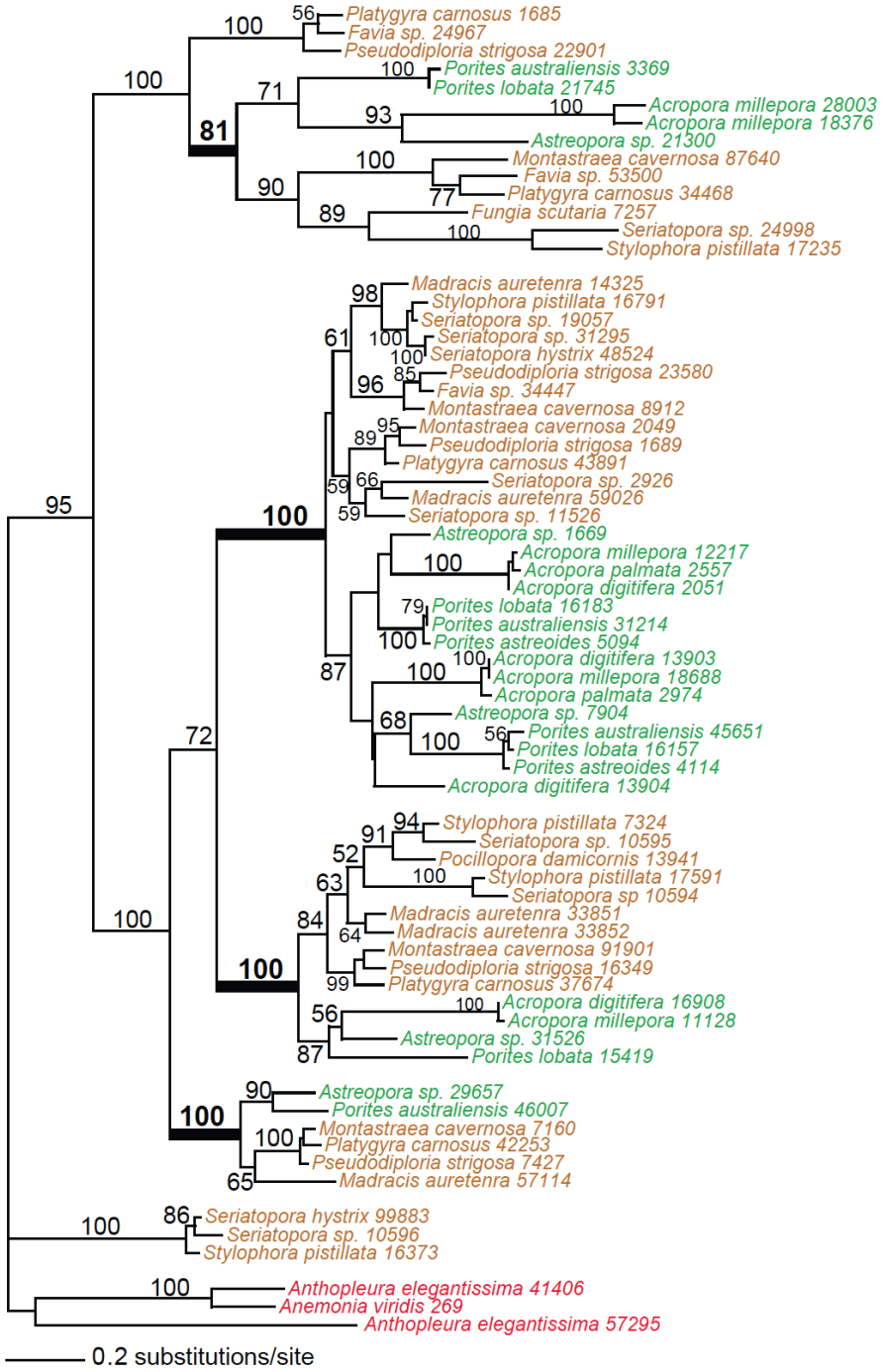
Evolution of CARP1: a Calcium binding acidic protein



Maximum likelihood (RAxML) tree showing the phylogenetic position of CARP1 among other calumenin-related homologs in corals (green text) and other taxa (red text). RAxML bootstrap values, when greater than 50%, are shown at the nodes.

Genome region that encodes *CARP1*, showing intron/exon structure, coverage using mRNA-seq, and genome coverage in the draft assembly. Note the strong support for the N-terminal exon that encodes the acidic domain (shown with the purple box). The EF hand-encoding exons are shown with the blue boxes.



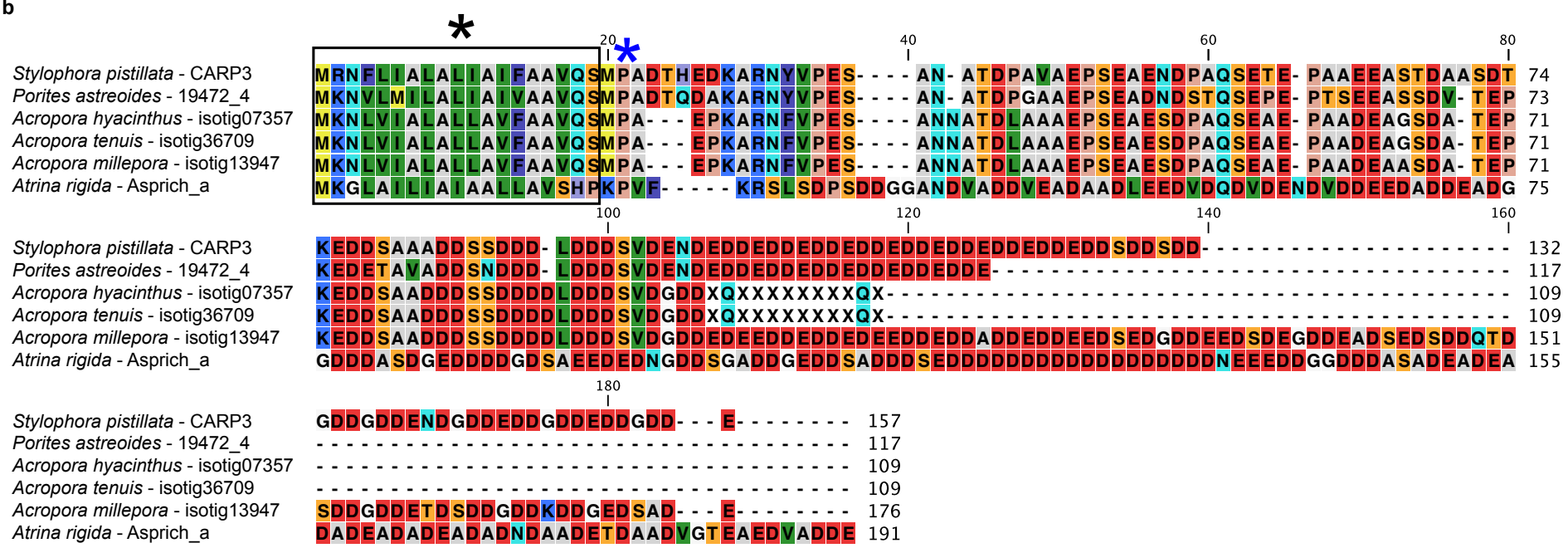


CARP 5

RAxML tree shows extensive history of CARP5 gene duplications that predates the split of robust (brown text) and complex (green text) corals.

CARP3

b



Pinctada rigida

P. astreoides

S. pistillata

A. tenuis

A. millepora

A. hyacinthus

Mass et al. Current Biology

Conclusions

- The *Seriatopora* sp. *draft* genome provides a tool to investigate biological processes in corals.
- CARPs bind Ca^{+2} and precipitate calcium carbonate *in vitro* in seawater at pH 8.2 and 7.6.
- The evolutionary history of CARPs indicates a coral-specific toolkit that has its roots deeper in metazoan evolution.

Comp
NSF S
Rutge
Febru

Manu
Mahd
Doug
Sylva
Ruth
Micha
Oren
Monic
Hollie
Chuya
Eiichi
Sylvie
Dan T
Christ
Andre
Didier
Bishoy Hanna
Debashish Bhattacharya
Jeana Drake
Paul Falkowski
Tali Mass
Ehud (Udi) ZelZion

Major areas addressed in collaborative paper:

1) Coral biomineralization;

2) Environmental and stress response systems;

3) Impacts of the symbiotic lifestyle in corals;

4) Horizontal gene transfer (HGT) and positive selection in coral genomes.

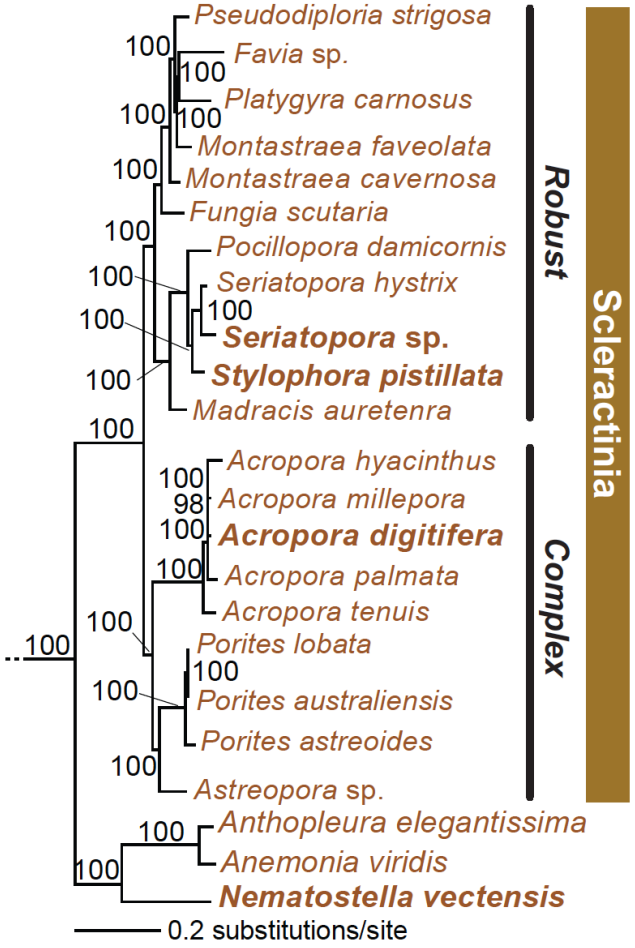
جامعة الـ
للـ
University of
hnology
ne City
niversity
ew York

Bar-Ilan University
Tradition of Excellence
NSF
IBSS
HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF
OSU
Oregon State
UNIVERSITY
Smithsonian
Institution
OIST
University of Haifa

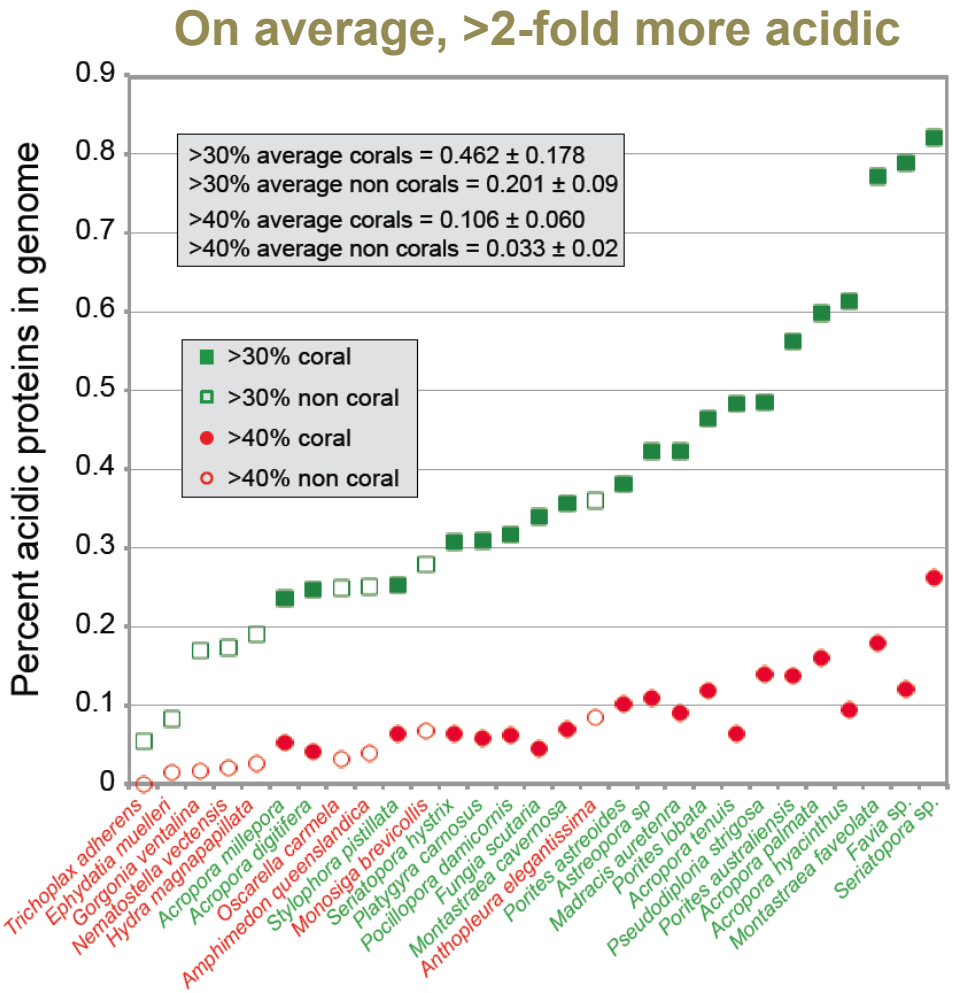
Coral genomic data used in analysis

Species	Number of sequences
Acropora digitifera	16977
Acropora hyacinthus	11589
Acropora millepora	28463
Acropora palmata	7522
Acropora tenuis	18419
Astreopora sp.	23921
Favia sp.	26627
Fungia scutaria	28265
Madracis auretenra	42119
Montastraea cavernosa	39938
Montastraea faveolata	5565
Platygyra carnosus	66449
Pocillopora damicornis	20509
Porites astreoides	15755
Porites australiensis	19567
Porites lobata	21062
Pseudodiploria strigosa	24345
Seriatopora hystrix	27680
Seriatopora sp.	35409
Stylophora pistillata	21810

http://comparative.reefgenomics.org/



Genome-based maximum likelihood tree inferred from analysis of 391 orthologs (63,901 amino acids) that was used as the reference tree to study coral evolution.



Comparison of coral (green text) and non-coral (red text) genomes with respect to percent of encoded proteins that contain either >30% or >40% negatively charged amino acid residues (i.e., aspartic acid [D] and glutamic acid [E]). The average composition and standard deviation of D + E is shown for the two cut-offs of these estimates. **On average, corals contain >2-fold more acidic residues than non-corals. This acidification of the coral proteome is postulated to result from the origin of biomineralization in this lineage.**

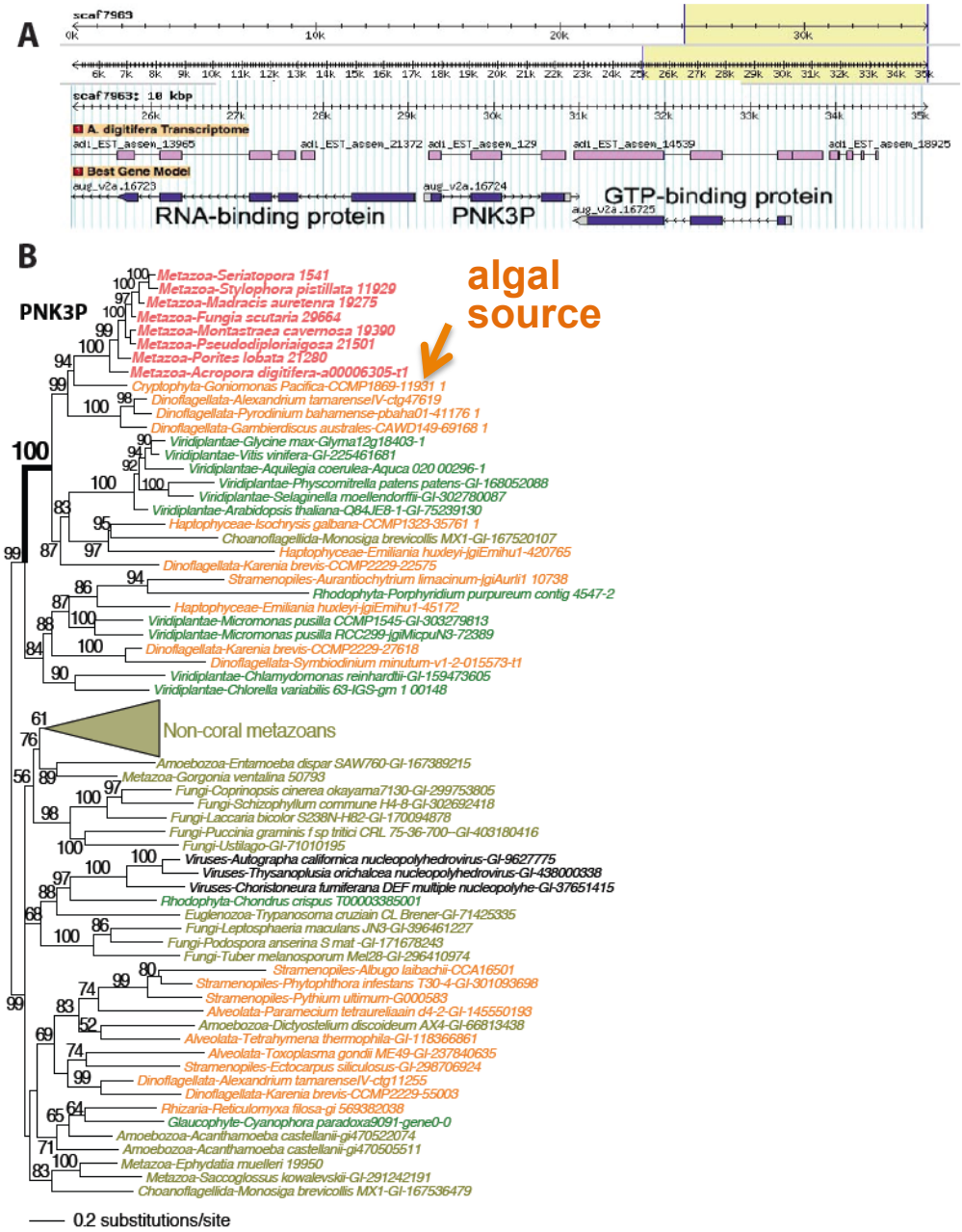
Rare but interesting examples of HGT in the coral animal…

A. digitifera and *Seriatopora* sp. proteomes used as queries in automated pipeline with 90% bootstrap support as cut-off.

This resulted in 13,256 and 19,700 alignments of which 21 and 43, respectively, supported HGT (64/32,956 trees = 0.2%).

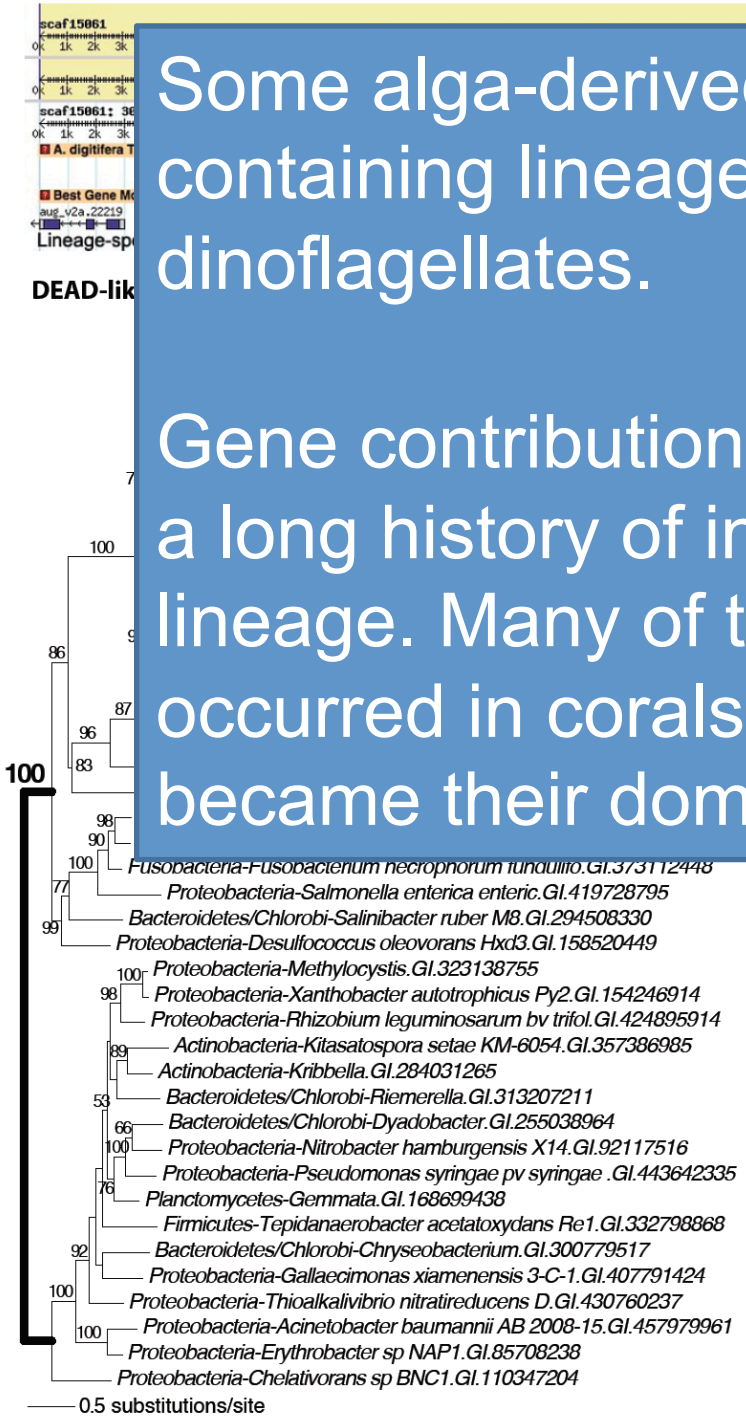
After accounting for gene duplicates and redundancy between the trees, we ended up with **42 unique instances of foreign gene acquisition from bacteria and algae; 14 specific to corals**.

Major targets of HGT were DNA repair (e.g., polynucleotide kinase 3 phosphatase [PNK3P] and uracil-DNA glycosylase) and reactive species detoxification (glyoxalase I [methylglyoxal from sugar metabolism]).



Some alga-derived HGTs are from chlorophyll *c*-containing lineages such as stramenopiles and dinoflagellates.

Gene contribution from these lineages suggests a long history of interaction with the anthozoan lineage. Many of these HGTs may have occurred in corals before *Symbiodinium* sp. became their dominant symbiotic partner.

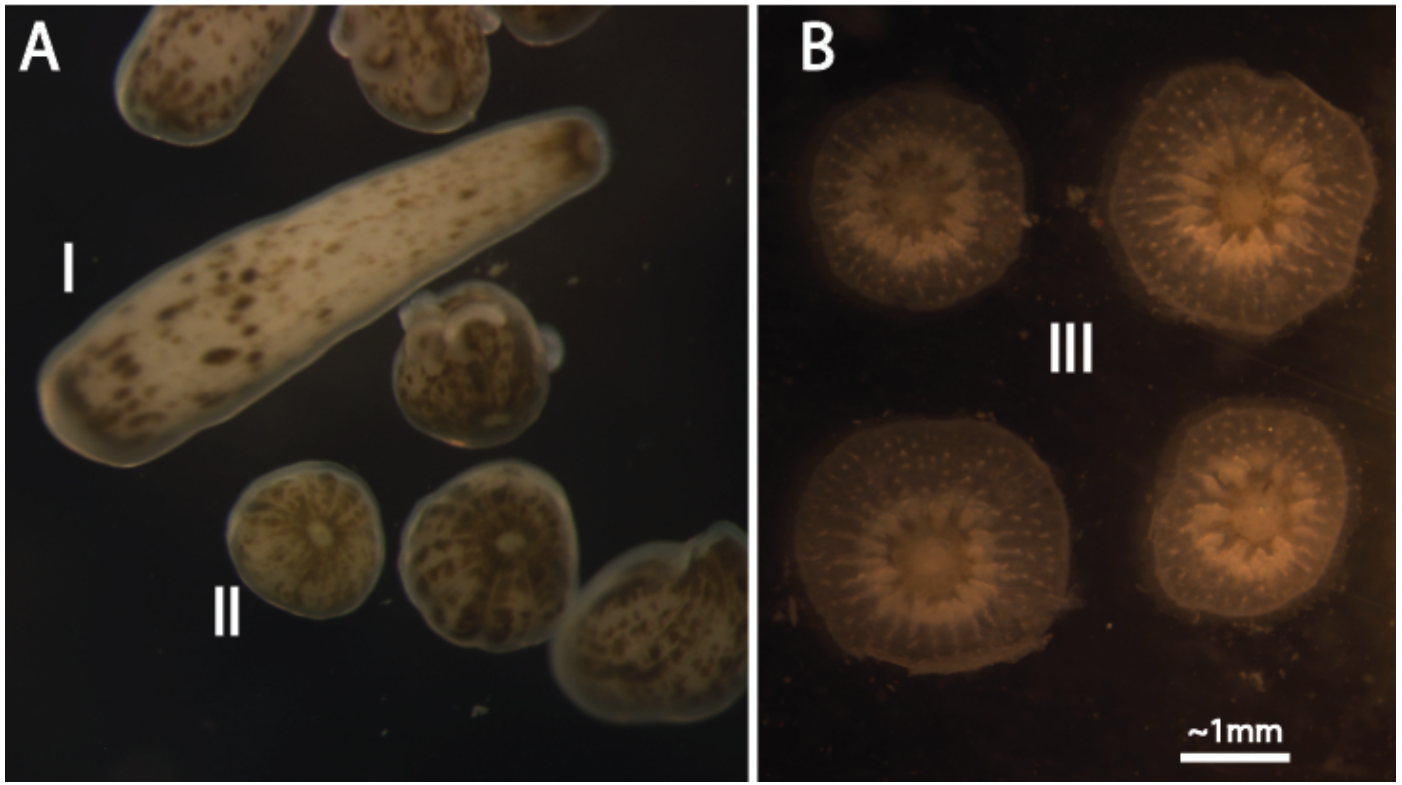


Ruth Gates
(HIMB)

CARP expression during development in *Pocillopora damicornis*

Hollie Putnam
(HIMB)

Motile larvae **Settled and calcifying spat**



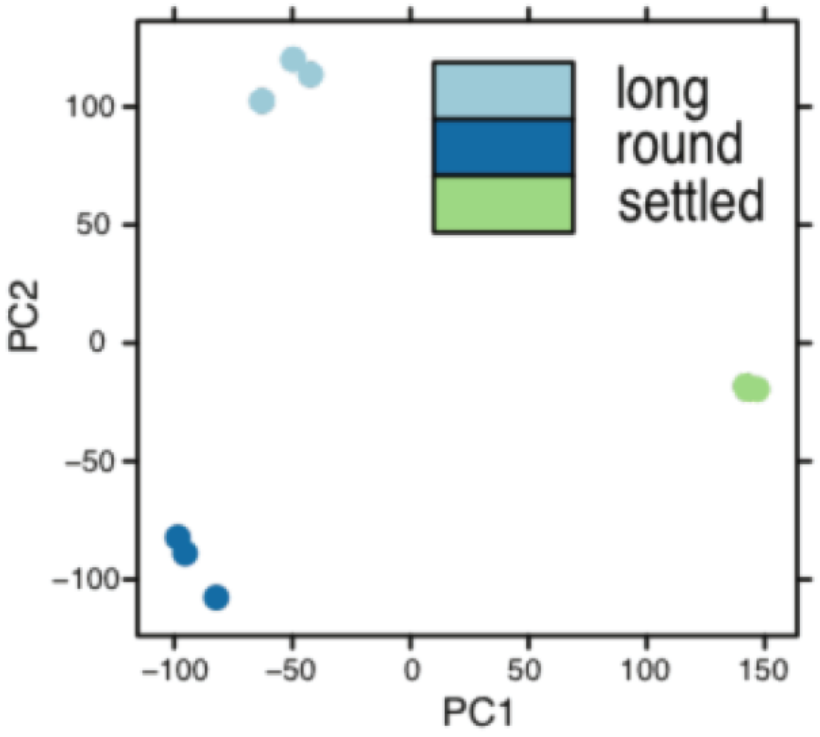
The different stages of brooded *P. damicornis* larvae that were used for our comparative analyses. Pictured are: A) stages I and II motile larvae and B) stage III settled and calcifying spat.

Mass et al. (In prep)

RNAseq data generated using triplicate samples of the 3 P. damicornis developmental stages; giving 104,402,226 high quality reads.

Reads used for de-novo transcriptome assembly, yielding 141,211 contigs with a 5x average coverage and N50 = 1,104bp.

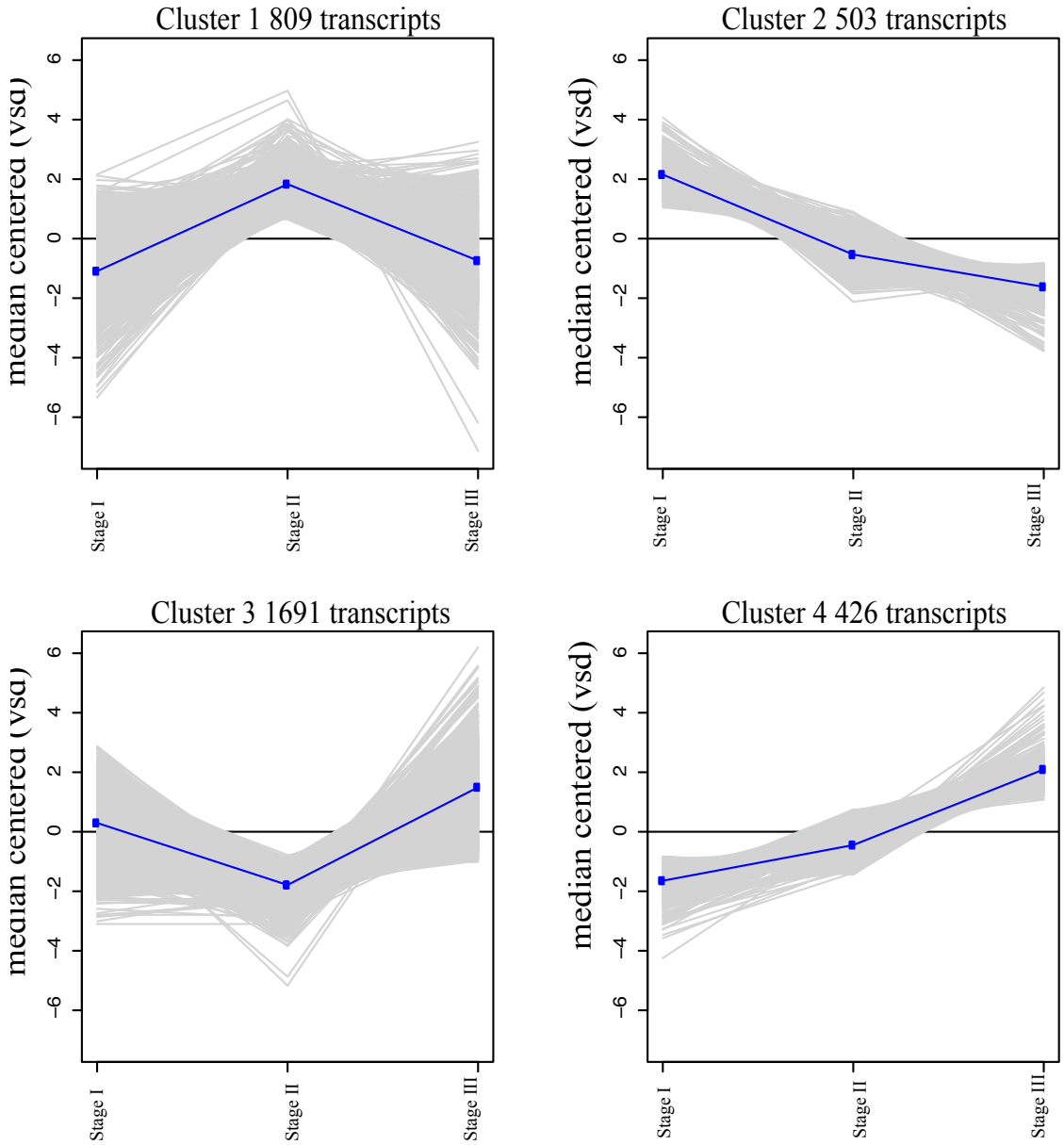
9 libraries (3 stages x 3 replicates) mapped to transcriptome assembly and unique reads used as input to DESeq2. Results show a strong correlation in gene expression within each developmental stage database, which supports the existence of stage-specific signal.



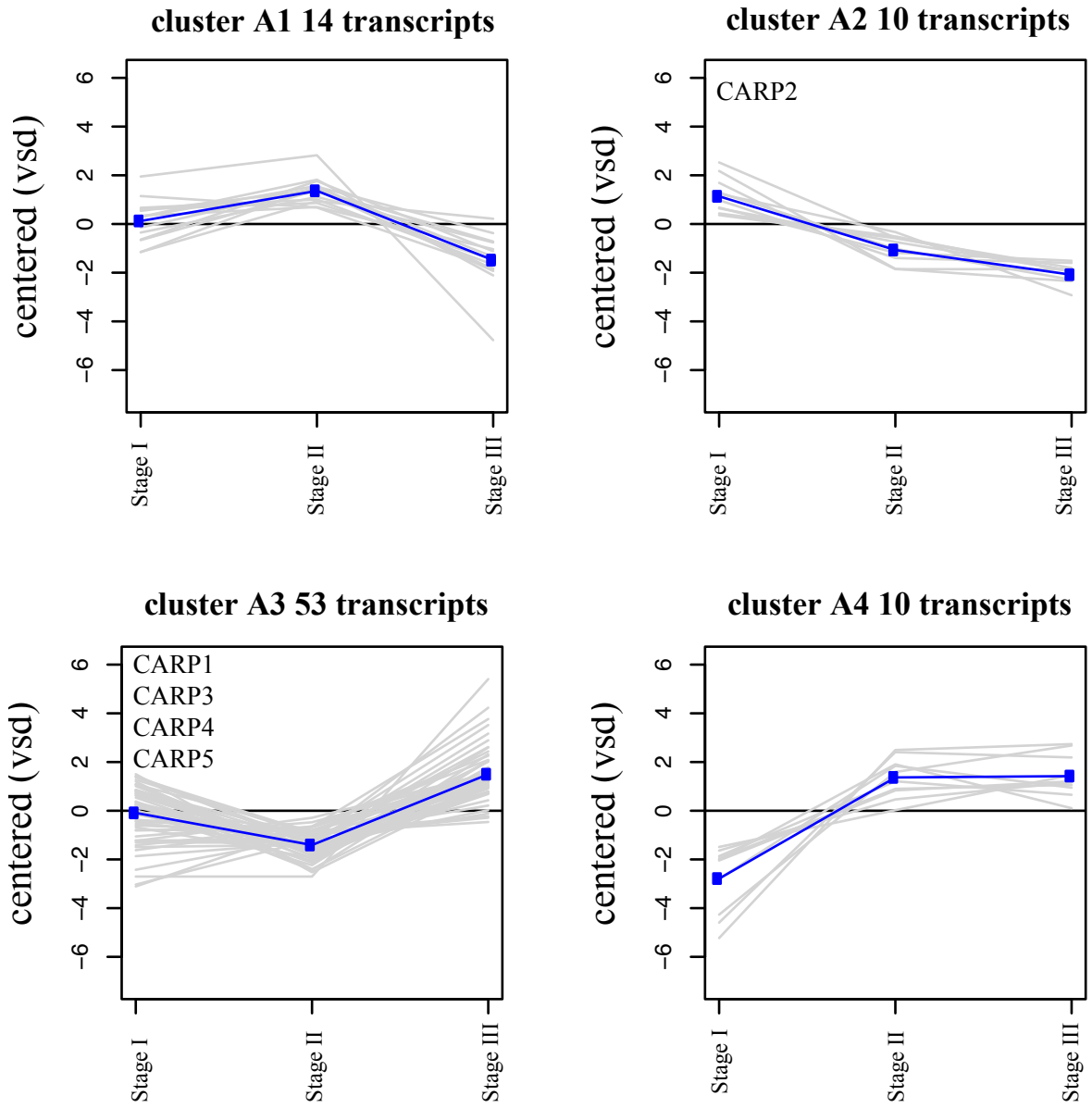
DE: stage III vs. stage II, stage III vs. stage I, and stage II vs. stage I. We collected all genes with DE in at least one of the comparisons.
Applied a fold-change (FC) cut-off of log2 (FC) > 3.5 (i.e., 11.31 fold difference) and log2 (1/FC) < -3.5 and a False Discovery Rate (FDR) < 0.001.

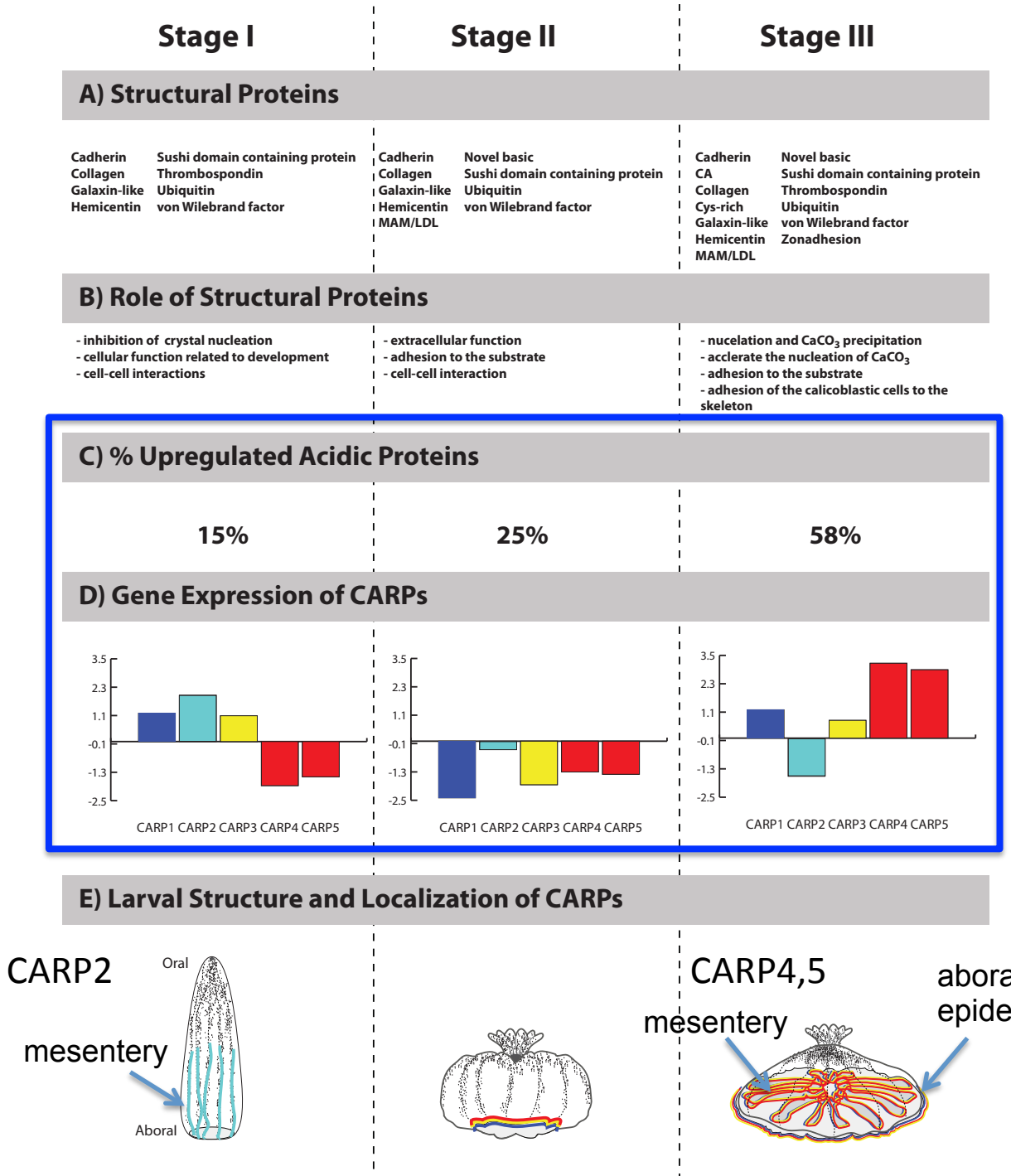
DE set was 9,986 genes, of which 3,429 had a BLASTP hit to metazoans, and 341 had hits with *e*-value <1e-20 to skeletal matrix proteins previously reported in *Stylophora pistillata* and *Acropora millepora*.

Gene expression clusters from RNAseq analysis of 3,429 DE genes in *P. damicornis*. These clusters are used to pose hypothesis about gene function.



Gene expression clusters from RNAseq analysis of CARPs and 82 other novel acidic proteins.





Model based on differential gene expression (DE) and protein immunolocalization during early development in *P. damicornis*.

A) The identity of DE structural proteins during the three stages of coral development;

B) the roles of these proteins in the animal;

C) the percent up-regulation of novel acidic proteins;

D) the relative expression levels of CARPs during development; and

E) immunolocalization of CARPs across the three life history stages. CARP 4 and 5 share high protein similarity and are immunolocalized by the same antibody.

Upcoming events in our group:

- *Pocillopora damicornis* development manuscript to be submitted.
- Follow-up study using RNA-seq data from adults is underway.
- Coral genome consortium paper to be submitted.
- Additional coral genomes underway: *Montipora capitata* and others at Genome Cooperative with Ruth Gates and Hollie Putnam (HIMB). Illumina 600 cycle and mate-pair approaches, as well as PacBio data for scaffolding.
- HHMI Janelia Farm Advanced Imaging Center opportunity to visualize coral crystal growth using live cell imaging (Structured Illumination Microscope, SIM).
- Work with Jess Adkins (CalTech) using NanoSIMS to localize CARP4 in newly precipitated CaCo3 to study the spatial relationship of the protein to the mineral.