

Supporting Information

Table S4 - Differential expression of genes involved in *Acropora millepora* biomineralization according to a previous experiment on primary polyps (Moya *et al.* 2012): up-regulated genes (green), down-regulated genes (red), not available (-). Fold-changes (P value > 0.05) were obtained through the analysis of the count data available on the NCBI Gene Expression Omnibus database (GEO) under accession number GSE33016, using the edgeR package (Robinson *et al.* 2010). Transcript levels were originated from *Acropora millepora* primary polyps at 380 (control), 750 and 1000 ppm CO2 after 3 days exposure.

Protein Groups	Cluster	NCBI Ac. No.	Protein Names	Fold Change	
				CO2 750 ppm (vs. Control)	CO2 1000 ppm (vs. Control)
Acidic	Cluster009205	JT001945	SAARP 1	1.41	2.32
	Cluster008253	JR972076	Acidic SOMP	-	-1.79
	Cluster017392	JR991407	SAARP2	-	-2.29
	Cluster018838	JT006291	SAP1	2.01	3.23
	Cluster037255	JT018094		3.77	5.85
	Cluster022029	JR983041	SAP2	1.81	2.05
	Cluster014254	JR983175	Glu-rich protein	1.93	1.51
Extracellular matrix proteins	Cluster001173	JR987773	Mucin-like	-	2.08
	Cluster001025	JT016638	Coadhesin	1.90	1.95
	Cluster000033	JT011118	MAM and LDL-receptor domain- containing protein 1	3.81	3.94
	Cluster000006	JR994474	MAM and LDL-receptor domain- containing protein 2	2.29	2.51
	Cluster014354	JT013896	Thr-rich protein	1.87	1.80
	Cluster010848	JR978035	Ectin	-	3.24
	Cluster012957	JT013217	MAM and fibronectin- containing protein	6.70	8.95
	Cluster007429	JT016410	MAM and fibronectin containing protein (isoform)	-	1.77
	Cluster000133	JR991141	PKD1-related protein	1.98	2.15
	Cluster011245	JR973492 (JN631095)	Zona pellucida domain-containing protein	-1.42	-1.86
	Cluster001085	JR980881	EGF and laminin G domain-containing protein	-	1.44
	Cluster000035	JT011093	Protocadherin-like	1.62	-
	Cluster050343	JR991083	Collagen	8.07	11.68
	Cluster000565m	JR993827	Neuroglian-like	2.60	1.88
	Cluster015162	JR989025	CUB domain-containing protein	1.87	2.87
Enzymes	Cluster002345	JT019463	Hephaestin-like	2.23	3.52
	Cluster020494	JR998014	Carbonic anhydrase	-	-

	Cluster023283	JR970990	CUB and Ser protease domain-containing protein 1	7.24	9.54
	Cluster005989	JT008002	CUB and Ser protease domain-containing protein 2	-	-
Uncharacterized proteins	Cluster013356	JR993391 (HM163215)	Galaxin	1.69	2.81
	Cluster015317	JR976690	Galaxin 2	-	-
	Cluster013623	JT021412	USOMP-1	2.03	-
	Cluster008498	JR982706	USOMP-2	-3.14	-2.29
	Cluster017073	JR997000	USOMP-3	1.53	1.96
	Cluster026302	JT004498	USOMP-4	-	-
	Cluster020453	JR973117	USOMP-5	3.96	5.03
	Cluster012833	JR971508	USOMP-6	-	-
	Cluster001446p	JR998260	USOMP-7	4.16	6.67
Cluster006620	JT014391	USOMP-8	-	-	
Toxin	Cluster001924	JR986059	Protein similar to cephalotoxin	-	-

Table S5: **Results from the comparison of the domains from *Acropora millepora* SOMPs versus those identified in other skeletal proteomes from *Strongylocentrotus purpuratus* (tooth, spicules, test and spine) (Mann et al. 2008a, 2008b; Mann et al. 2010), *Gallus gallus* (eggshell) (Jonchère et al. 2010; Mann et al. 2006; Miksik et al. 2010), *Lottia gigantea* (shell) (Marie et al. 2012), *Pinctada margaritifera* and *P. maxima* (shell) (Marie et al. 2012), *Stylophora pistillata* (Drake et al. 2013) and *Crassostrea gigas* (shell) (Zhang et al. 2012).** + indicates domains from proteins that were identified through proteomics and are expressed in skeleton secreting-tissues, or have further experimental evidence of involvement in biomineralization, (+) indicates domains from proteins identified in the organic matrix only by proteomics but for which no other evidence related to biomineralization is currently available. * Domains corresponding to more than one InterPro entry (*i.e.* with parent/child relationship), ^a Domains identified only in corals and ^b Databases containing intracellular proteins.

<i>Acropora millepora</i>		Versus species:	<i>S. purpuratus</i> ^b	<i>G. gallus</i> ^b	<i>L. gigantea</i>	<i>P. margaritifera</i> <i>P. maxima</i>	<i>S. pistillata</i> ^b	<i>C. gigas</i> ^b
Key domains (as in Figure 4)	InterPro entries identified in the SOMPs	Structure:	Tooth, spicules, test and spine	Eggshell	Shell	Shell	Skeleton	Shell
		Interpro no:						
Thrombospondin	Thrombospondin, type 1 repeat	IPR000884	(+)	+	-	-	(+)	(+)
Nidogen	Nidogen, extracellular domain	IPR003886	(+)	(+)	-	-	-	(+)
AMOP	AMOP	IPR005533	(+)	-	-	-	(+)	(+)
von Willebrand factor, type D	von Willebrand factor, type D domain	IPR001846	(+)	(+)	-	-	(+)	(+)
von Willebrand factor, type A	von Willebrand factor, type A	IPR002035	(+)	(+)	+	+	(+)	+
Epidermal growth factor*	Epidermal growth factor-like domain	IPR000742	(+)	+	+	+	(+)	+
	EGF-like calcium-binding	IPR001881	(+)	+	-	-	(+)	(+)
Coagulation factor 5/8 CT type domain*	Coagulation factor 5/8 C-terminal type domain	IPR000421	+	(+)	-	-	(+)	(+)
	Galactose-binding domain-like	IPR008979	+	(+)	-	-	(+)	(+)
CAP	CAP domain	IPR014044	(+)	-	+	+	-	(+)
MAM domain*	MAM domain	IPR000998	(+)	(+)	-	-	(+)	(+)
	Concanavalin A-like lectin/glucanase	IPR008985	(+)	(+)	-	+	(+)	+
Ricin B lectin domain	Ricin B lectin domain	IPR000772	-	(+)	-	-	-	-
	Fibronectin, type III	IPR003961	+	+	-	+	-	+
Fibronectin type III*	Fibronectin type III C-terminal domain ^a	IPR026966	-	-	-	-	-	-

ZP sperm-binding	Zona pellucida sperm-binding protein	IPR001507	-	(+)	+	+	(+)	+
CUB	CUB	IPR000859	(+)	+	+	-	-	(+)
Laminin G*	Laminin G domain	IPR001791	(+)	(+)	-	-	(+)	(+)
	Concanavalin A-like lectin/glucanase, subgroup	IPR013320	(+)	+	-	+	(+)	+
Carbohydrate-binding WSC*	Carbohydrate-binding WSC	IPR002889	(+)	-	-	-	-	-
	Carbohydrate-binding WSC, subgroup ^a	IPR013994	-	-	-	-	-	-
PKD/Chitinase domain*	PKD domain	IPR000601	-	(+)	-	-	-	(+)
	PKD/Chitinase domain ^a	IPR022409	-	-	-	-	-	-
	PKD/REJ-like protein	IPR002859	-	-	-	-	-	(+)
PKD/REJ-like protein*	Egg jelly receptor, REJ-like ^a	IPR014010	-	-	-	-	-	-
GPS	GPS domain	IPR000203	(+)	-	-	-	-	-
Cadherin*	Cadherin	IPR002126	+	+	-	-	(+)	(+)
	Cadherin-like	IPR015919	(+)	+	-	-	(+)	(+)
P-type trefoil ^a	P-type trefoil ^a	IPR000519	-	-	-	-	(+)	-
Fibrillar collagen, CT	Fibrillar collagen, C-terminal	IPR000885	+	+	-	-	-	-
Collagen triple helix repeat	Collagen triple helix repeat	IPR008160	+	+	-	-	-	(+)
	Immunoglobulin subtype 2	IPR003598	+	(+)	-	-	-	+
	Immunoglobulin subtype	IPR003599	+	(+)	-	-	-	+
Immunoglobulin-like*	Immunoglobulin-like	IPR007110	+	+	-	-	-	+
	Immunoglobulin I-set	IPR013098	+	(+)	-	-	-	+
	Immunoglobulin-like fold	IPR013783	+	+	-	+	-	+
Low-density lipoprotein receptor	Low-density lipoprotein (LDL) receptor class A repeat	IPR002172	(+)	(+)	-	-	(+)	-
Lipoxygenase ^a	Lipoxygenase, LH2 ^a	IPR001024	-	-	-	-	-	-
	Lipase/lipoxygenase, PLAT/LH2 ^a	IPR008976	-	-	-	-	-	-
Cupredoxin*	Cupredoxin	IPR008972	(+)	-	-	-	-	+
	Multicopper oxidase, type 2	IPR011706	-	-	-	-	-	+
	Multicopper oxidase, type 3	IPR011707	(+)	-	-	-	-	+
Alpha carbonic anhydrase	Alpha carbonic anhydrase	IPR001148	+	+	+	+	+	+
Peptidase cysteine/serine, trypsin-like*	Peptidase S1/S6, chymotrypsin/Hap	IPR001254	(+)	+	-	-	-	+
	Peptidase cysteine/serine, trypsin-like	IPR009003	(+)	(+)	-	-	-	+
Polycystin cation channel, PKD1/PKD2	Polycystin cation channel, PKD1/PKD2	IPR013122	-	-	-	-	-	(+)
Neurexin/syndecan/glycophorin C	Neurexin/syndecan/glycophorin C	IPR003585	(+)	-	-	-	-	-
Cadherin, cytoplasmic domain	Cadherin, cytoplasmic domain	IPR000233	-	(+)	-	-	-	(+)

Table S6: **Comparison between *Acropora millepora* SOMPs** and the proteins identified in the skeletal organic matrix from *Stylophora pistillata* (Drake et al. 2013). Pairs of related proteins are indicated by x – for more than 35% of identity (min. 100 aa) and by X – for homologous pairs. Homology could not be determined for protein fragments (*).

S. pistillata																																										
A. millepora	Protocadherin fat-like (P1)	CARP4 (P2)	Thrombospondin (P3)*	Viral inclusion protein (P4)	Hemicentin (P5)*	Actin (P6)	Actin (P7)*	Major yolk protein (P8)*	Protocadherin fat-like (P9)*	Cadherin (P10)*	Actin (P11)*	Unknown protein (P12)	Sushi domain-containing (P13)*	Collagen-alpha (P14)*	CARP5 (P15)*	Unknown protein (P16)*	Glyceraldehyde 3-phosphatase dehydrogenase (P17)*	Collagen alpha (P18)*	Contactin-associated protein (P19)*	MAM domain anchor protein (P20)*	Zona pellucida (P21)*	Unknown protein (P22)	Protocadherin (P23)*	Vitellogenin (P24)*	Ubiquitin (P25)*	Vitellogenin (P26)*	Integrin-alpha (P27)*	Late embryogenesis protein (P28)*	Tubulin-beta (P29)*	Myosin regulatory light chain (P30)*	Neurexin (P31)*	Kielin/Chordin like (P32)*	Flagellar associated protein (P33)*	MAM/LDL receptor domain containing protein (P34)*	Carbonic anhydrase (STPCA2) (P35)*	Zonadhesion-like precursor (P36)*						
	SAARP 1		X													X													X													
Acidic SOMP		X													X													X														
SAARP2*		X													X													X														
SAP1*																																										
SAP2*																																										
Glu-rich protein																																										
Mucin-like*					X								X																													
Coadhesin*			X		X									X				X																								
MAM and LDL-receptor domain-containing protein 1*																				X																X				X		
MAM and LDL-receptor domain-containing protein 2*																				X																X				X		
Thr-rich protein*																																										
Ectin*					X																																					
MAM and fibronectin-containing protein*																																										
MAM and fibronectin containing protein (isoform)*																																										
PKD1-related protein*																																										
Zona pellucida domain-containing protein																						X																				
EGF and laminin G domain-containing protein																			X													X										
Protocadherin-like	X								X	X																																

- Marie B, Joubert C, Tayalé A, Zanella-Cléon I, Belliard C, Piquemal D, Cochenec-Laureau N, Marin F, Gueguen Y, Montagnani C. 2012. Different secretory repertoires control the biomineralization processes of prism and nacre deposition of the pearl oyster shell. *Proc Natl Acad Sci USA*. 109:20986-20991.
- Miksík I, Sedláková P, Lacinová K, Pataridis S, Eckhardt A. 2010. Determination of insoluble avian eggshell matrix proteins. *Anal Bioanal Chem*. 397:205-214.
- Moya A, Huisman L, Ball EE, Hayward DC, Grasso LC, Chua CM, Woo HN, Gattuso J.-P, Forêt S, Miller DJ. 2012. Whole Transcriptome Analysis of the Coral *Acropora millepora* Reveals Complex Responses to CO₂-driven Acidification during the Initiation of Calcification. *Mol Ecol*. 21:2440-2454.
- Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 26:139-140.
- Zhang G, Fang X, Guo X, et al. (85 co-authors). 2012. The oyster genome reveals stress adaptation and complexity of shell formation. *Nature*. 490:49-54.