Whole transcriptome sequencing and biomineralization gene architecture associated with cultured pearl quality traits in the pearl oyster, *Pinctada margaritifera*

Le Luyer J.[1]*, Auffret P. [1], Quillien V. [1], Leclerc N. [1], Reisser C.[1], Vidal-Dupiol J.[2,1], Ky C.-L.[1,2]*

[1] IFREMER, UMR 241 Ecosystèmes Insulaires Océaniens, Centre Ifremer du Pacifique, BP 49, 98719 Tahiti, Polynésie française

[2] IFREMER, IHPE UMR 5244, Univ. Perpignan Via Domitia, CNRS, Univ. Montpellier, F-34095 Montpellier, France

***Corresponding authors:** jeremy.le.luyer@ifremer.fr; chinky@ifremer.fr

Keywords: *Pinctada margaritifera* – biomineralization– pearl quality – pearl sac – RNA-seq

Abstract

Background

Cultured pearls are unique gems produced by living organisms, mainly molluscs of the *Pinctada* genus, through the biomineralization properties of pearl sac tissue. Improvement of *P. margaritifera* pearl quality is one of the biggest challenges that Polynesian research has faced to date. To achieve this goal, a better understanding of the complex mechanisms related to nacre and pearl formation is essential and can now be approached through the use of massive parallel sequencing technologies. The aim of this study was to use RNA-seq to compare whole transcriptome expression of pearl sacs that had producing pearls with high and low quality. For this purpose, a comprehensive reference transcriptome of *P. margaritifera* was built based on multi-tissue sampling (mantle, gonad, whole animal), including different living stages (juvenile, adults) and phenotypes (colour morphotypes, sex).

Results

Strikingly, few genes were found to be up-regulated for high quality pearls (n = 16) compared to the up-regulated genes in low quality pearls (n = 246). Biomineralization genes up-regulated in low quality pearls were specific to prismatic and prism-nacre layers. Alternative splicing was further identified in several key biomineralization genes based on a recent *P. margaritifera* draft genome.

Conclusion

This study lifts the veil on the multi-level regulation of biomineralization genes associated with pearl quality determination.

Background

The mollusc *Pinctada margaritifera* var. *Cumingii* is a species of great economic importance in French Polynesia. The associated pearl industry represents the second most important source of income there, just after tourism. Cultured pearl production requires two distinct animals. A small piece of graft mantle tissue is dissected from a sacrificed donor oyster and inserted with a round bead of nacre (a nucleus, made of mussel shell) into the gonad of a recipient oyster [1, 2]. If the graft is not rejected and the recipient oyster survives the grafting operation, the implanted mantle tissue will grow to completely surround the nucleus and form a "pearl sac", capable of secreting biomaterial layers (calcite and aragonite) around the nucleus [3]. After 15 to 24 months of culture, pearls are harvested and usually sorted according to six main quality traits: size, shape, colour, surface complexion, lustre and grade [4]. It is estimated, however, that only 5% of the harvested pearls can be classified as grade A, which corresponds to the best quality according to the local regulatory quality standards [5]. Average export price of cultured pearls in French Polynesia has fallen dramatically over the past decade, mainly due to a combination of factors including overproduction; hence quantity has been favoured to the detriment of quality. The improvement of cultured pearl quality is an imperative aspect of a pearl farm's sustainability and as one of the biggest challenges that research is facing in French Polynesia.

Factors affecting pearl quality have diverse and non-exclusive origins including genetics, environment and/or genotype-by-environment interaction (GEI). This is further complexified in the *Pinctada* transplant model because of the phenotype transmission from the donor oyster to the recipient oyster, their interplay, and their interaction with the environment. One particularity of this animal model is the chimera system attributed to the pearl sac tissue, whose interaction between $Genome_{Donor}$ x $Genome_{Host}$ has a significant effect on pearl quality. A recent study in *P. margaritifera*, based on

64 controlled bi-parental crosses and the F1 generation, demonstrated heritability (h^2 from
65 0.21 to 0.37) for nacre weight and thickness, pigmentation darkness and colour, surface
66 defects and grade, signifying a donor oyster effect with a genetic basis, although there
67 were also important interaction components [6]. Previous studies reported that
68 location, temperature and food availability [7–9], pearl rotation [10], donor oyster
69 genotype [4, 11–14], age [15], position of graft mantle [16] and contamination during
70 the graft operation and/or graft operator skills [17] are all determinants of final pearl
71 quality and do not necessarily affect similar traits.

72 Various genomic approaches have been applied in pearl oyster with the
73 objectives of identifying key candidate markers related to pearl quality traits. For
74 instance, Lemer et al. [18] identified a set of genes differentially expressed between two
75 matle colour phenotypes (black *P. margaritifera* phenotypes and full albino individuals),
76 using a suppressive and subtractive hybridization (SSH) method. In silver-lipped oyster
77 (*Pinctada maxima*), genetic association analyses has permitted the identification of QTLs
78 linked to pearl surface complexion and colour as well as genetic associations of regions
79 and markers for pearl size, weight, colour and surface complexion [19, 20].
80 Transcriptome-wide and proteomic approaches have also been used to characterize the
81 pool of genes expressed during pearl formation and to discriminate markers
82 preferentially associated with nacreous and/or prismatic layers. Studies in the Japanese
83 pearl oyster (*Pinctada fucata*) showed that the genes *msi60* and *aspein* from the mantle
84 tissue were up-regulated in low quality pearls compared to high quality groups, while
85 the expression of the *msi30* gene from the pearl sac tissue was up-regulated in high
86 quality pearls [21, 22]. Finally, a recent study on *P. margaritifera* revealed that
87 *shematrin5 and 9*, *prismalin* and *aspein* encoding genes were up-regulated in the pearl
88 sacs of individuals producing low pearl surface quality [23]. Studies have been limited to

relatively few candidate genes, however, and an overall evaluation of the actors involved in pearl quality remains to be conducted.

The aim of our present study was to identify key genes involved in the regulation of pearl quality through a comparative RNA-seq analysis of pearl sacs producing high and low quality cultured pearls. We constructed a new comprehensive multi-tissue transcriptome assembly, covering several developmental stages, colour morphotypes and tissue origins, which will be useful for further transcriptomic studies in *P. margaritifera*. Furthermore, based on a recently assembled draft genome of *P. margaritifera*, we successfully explored the possibility that alternative gene splicing events are involved in the regulation of biomineralization processes.

Results

Transcriptome assembly. The raw transcriptome assembly contained 541,184 transcripts. After filtering for redundancy and functionality, we retained a total of 41,075 transcripts (assembly metrics given in Table 2). Transcriptome completeness evaluation indicated that 90.6 % of the highly conserved single-copy metazoan genes (n = 978) were present in our transcriptome (89.8 % are complete and in a single-copy). Similarly, mean mapping rate reached 66.31 ± 1.72 % with negligible differences in sample condition. Both the transcriptome completeness and satisfactory mapping rate suggest that the several steps of filtering applied did not have major impact on the overall transcriptome. Functional annotation identified a total of 33,532 transcripts (81.5%) with at least one match with deposited sequences (Table 2).

Differential expression in biomineralization-related genes. We found a total of 262 differentially expressed genes (DEGs), with 246 up- and 16 down-regulated genes in low quality pearls (Table S3). Out of the 262 DEGs, 216 (82.24%) had at least one match with

known protein sequences (Table S3). Finally, only 114 of the DEGs (43.5 %) had at least one associated GO term. GO analysis revealed enrichment for some relevant functions involved in pearl formation, including oxidoreductase activity (GO:0016491), peptidase inhibitor activity (GO:0030414), serine-type peptidase activity (GO:0008236), chitin binding (GO:0008061) and copper ion binding (GO:0005507). The GO enrichment analysis is summarized in Figure S1.

We identified several biomineralization genes discriminating high and low quality pearls (Figure 3; Table S3). Most of these genes are characteristic of prismatic and prism-nacre layers [24] and were found up-regulated in low quality pearls. The blue mussel shell protein-like (*BMSP-like*) coding gene is the single biomineralization gene up-regulated in high quality pearls. *BMSP-like* shares strong homology with the *P. fucata pif-177* gene and is notably involved in determining the polymorph of $CaCO_3$ [25]. We also identified Gypsy and Jockey-family transposable elements up-regulated in high quality pearls (Table S3). The qPCR analysis shows that the individual relative expression of the four biomineralization-related genes analysed is in accordance with results from pool RNA-seq data (Figure S2). We found the *MP10*, *shematrin-9* and *aspein* up-regulated in low quality pearls (p-value < 0.001) while no significant difference was observed for *pif-177* (p-value = 0.012).

Different alternative splicing in biomineralization-related genes. We found a total of 28 transcripts showing significant differential splicing (FDR < 0.001) (Figure 4). Several of these genes are known to be involved in biomineralization processes: *pif*, *aspein*, *pwap* and *wdf18*. For *aspein*, high quality pearls show lower exon 1 usage. This specific exon overlaps with the promoter domain of the transcript while the exon 2 codes for the D-domain rich in Asp amino acid that gives the aspein protein its crystal

binding affinity and its function in regulating crystal growth [21, 26–29]. The GO enrichment analysis identified the peptide biosynthetic process (GO:0043043), amide biosynthetic process (GO:0043604) and translation function (GO:0006412) as enriched for the genes showing different exon usage.

Discussion

Access to massive parallel sequencing technologies now greatly contributes to the understanding of molecular expression of the phenotype in non-model species. Here, we used a common transcriptomic approach (RNA-seq) to obtain an overview of differences in gene expression and alternative splicing between high and low cultured pearl quality. Recently, RNA-seq has been successfully used to explore genes related to pearl oyster growth and response to environmental stressors (*P. fucata*) [30, 31] and biomineralization (*P. martensii* and *P. penguin*) [32–34]. The use of RNA-seq in *P. margaritifera* for biomineralization related studies has been held back by the relatively limited coverage of previous transcriptome references (Roche 454; [35]) or their reduced tissue representation [36]. Recent work has however completed multi-tissue transcriptome assemblies of four species of pearl oysters including *P. margaritifera*. Yet, unfortunately, sampling did not include pearl sac tissues and to our knowledge assemblies were not made publicly available [37]. The present study provides a comprehensive reference transcriptome for *P. margaritifera*, encompassing whole-body tissue as well as phenotypic variation for a single common tissue (for gonads this was either female or male; for mantle tissue and whole individuals this covered a broad range of shell colour) as well as including individuals of different life stages (juveniles or adults). Even when applying stringent filtering, the satisfactory completeness as well as good individuals mapping rates and high gene content conservation with the closely

161 related species *P. fucata* (78% overlap) all suggested that this new reference
162 transcriptome should prove itself a useful genomic resource for a broad range of future
163 transcriptomic research for *P. margaritifera*.

164 **Complex responses of biomineralization genes are associated with pearl quality**.
165 The pearl sacs producing low quality pearls were characterized by a higher activity of
166 prismatic layer-specific genes. Among the differentially expressed genes, *aspein* and
167 *shematrin-9* were found to be up-regulated in low quality pearl RNA-seq pools, as well as
168 being validated by individual qPCR, which is consistent with previous candidate gene
169 expression work [23]. The present study also provides a novel set of biomarkers
170 involved in the biomineralization process, such as *perlwapin* and *BMSP-like* genes, and
171 supports previous findings showing deep conservation of biomineralization genes
172 within molluscs [38–40].

173 A considerable number of studies have focused on the identification of genes
174 involved in aragonite and/or calcite formation in pearl oyster species and other bivalves
175 capable of shell mineralisation [16,18,45–51]. Nonetheless, it has proven difficult to
176 extrapolate the role of key actors involved in determining cultured pearl quality across
177 studies. It is reasonable to hypothesize that contrasting results in the expression of
178 biomineralization genes might result from the pleiotropic effect of biomineralization
179 genes and/or major differences in the design of these different studies (genetic
180 background, geographical locations, pearl grades sampled, time of sampling post-graft,
181 environmental factors and tissue-analysed). For instance, differences in response of the
182 *pearlin* gene, a gene commonly used to monitor biomineralization, have been observed
183 in both graft and pearl sac tissues under similar environmental stress conditions
184 (temperature and food availability) [8, 47]. Similarly, expression of key

185 biomineralization genes such as *nacrein* or *pif* in pearl sacs was clearly dependent on the
186 time of sampling post-graft [8, 23]. *Pif*-related genes are particularly interesting because
187 they directly controlled crystal growth during an *in vitro* experiment [41], yet the clear
188 connexion between *pif*-related gene expression and pearl phenotypes remains unclear.
189 The present study showed that neither *pif* nor *pif-177* were differentially expressed
190 between high and low pearl quality despite these genes having already been correlated
191 with pearl weight (Rho = 0.259; *p*-value = 0.01) and pearl quality 3 months post-graft in
192 previous studies on *P. margaritifera* [23, 48]. Inversely, the *BMSP-like* gene, a gene
193 related to the *pif* family [49], was up-regulated in high quality pearls. Clearly, further
194 studies will be needed to unravel the complex role of *pif*-related genes during the
195 different stages of pearl formation. Nevertheless, this study supports previous findings
196 on the role of *aspein* and *shematrin 9* genes in controlling pearl quality, independently of
197 geographical location or time of post-graft sampling. Inversely, *pif* gene (*pif-177*, *pif* and
198 *BMPS-like*) expression was not consistent with previous studies and suggests that *pif*
199 expression variation is strongly time-dependent, which could be the basis of the
200 complex role of the *pif*-related gene family in pearl biomineralization.

201 **Different exon usage plays a role in shaping pearl quality.** Genes and related
202 proteins involved in biomineralization have complex structures and often require post-
203 translational modifications and proteolytic cleavages [50]. Alternative splicing in
204 biomineralization-related genes has already been suggested by the presence of
205 numerous isoforms for spicule matrix protein (SpSMs) coding gene in sea urchin [51]. In
206 the present study, four biomineralization-related genes, *pif*, *aspein*, *perlwapin* and *wfd18*,
207 were found to have significant differences in exon usage, and were all also differentially
208 expressed between low and high pearl quality. The present study could not, however,
209 assess whether the control of exon usage is under epigenetic and/or genetic control. In

an effort to reduce the variability inherent to the complex determination of pearl quality, our sampling design only included pooled samples from a single geographic location and mixed several pearl defect types. However, broadening our results on splicing events to the individual level and specific defect types could enable us to redefine the link between biomineralization gene architecture and pearl quality traits.

Transposable elements might be involved in the regulation of cultured pearl quality. Strikingly, this study only identified a few genes as being up-regulated in high quality pearls (n = 16) compared to the number of up-regulated genes in low quality pearls (n = 246). Among these up-regulated genes, there was a surprisingly high representation of transposable element- (TE) related genes (n = 3 out of 7 annotated genes) including long-transposable elements (LTR) of the Gypsy and Jockey families. Recent studies in humans and plants suggest that TEs and TEs insertion are effective regulators of gene expression and alternative splicing events [52–54]. Furthermore, it has been shown that both tandem repeats (TRs) and TEs are intimately linked with TRs derived from younger/more active TEs [55]. As an example, a study in maize supports the results of specific centromeric TRs originating from Ty3/gypsy retrotransposons [56]. From an evolutionary perspective, genes involved in biomineralization are structurally complex and often characterized by the presence of numerous TRs [57]. It is thus plausible that control of tandem repeat formations might result from TE insertion [44, 45]. However, by which mechanisms TEs (specifically Gypsy and Jockey family TEs) are involved in pearl quality remains to be elucidated. Further studies correlating TE insertion with the structure of transcripts (such as *aspein*), considering the specific genotypic background, should provide useful information that will help unravel the complex regulation of cultured pearl quality traits.

Conclusions

This study successfully made it possible to: 1) identify genes whose expression in pearl sacs was associated with cultured pearl quality in *P. margaritifera*, and 2) highlight other putative regulation levels for pearl quality determination through alternative splicing and TE regulation. Among the genes differentially expressed, new candidates were identified for pearl quality (*perlwapin*, *BMSP-like*), as were previously described biomarkers (*aspein* and *shematrin-9*). The present study also showed, however, that gene expression of some biomarkers previously associated with pearl quality or thickness (*pif-177*, *pif*, *msi30*, *pearlin* or *nacrein*), is not systematically correlated with pearl quality, suggesting that other factors might be involved. Further studies should focus on time-course experiments from the first stages of mineral deposition until harvest so as to unravel gene expression in the successive biomineralization events and the interplay between environmental and genetic factors in controlling specific quality traits.

Methods

Transcriptome reference assembly of *P. margaritifera*. The multi-tissue reference transcriptome was built with tissues obtained from a total of 10 *P. margaritifera* individuals: gonad tissues (n = 2, obtained from one male and one female; [36]), whole tissue of 4-month-old juveniles (whole individuals, n = 2) and mantle tissue (n = 6; Table 1). For all samples apart from the gonads, total RNA was extracted with TRIZOL Reagent (Life Technologies) at a ratio of 1 ml per 100 mg tissue, following manufacturer's recommendations. RNA quantity/integrity and purity were validated on a Nanodrop (NadoDrop Technologies Inc.) and on a BioAnalyzer 2100 (Agilent Technologies), respectively. RNA was dried in RNA-stable solution (ThermoFisher Scientific) following

manufacturer's recommendations and shipped at room temperature to McGill sequencing platform services (Montreal, Canada). TruSeq Sample Prep. (Illumina, San Diego, Ca, USA) RNA libraries were multiplexed (n = 10 pools by lane) and sequenced by HiSeq4000 100-bp paired-end (PE) sequencing technology. For the gonads, the samples were also sequenced by HiSeq2000 100-bp PE and were downloaded from the SRA database (Bioproject PRJNA229186; see [36] for more information on RNA preparation). Reads were filtered for adapter removal, minimum length (≥ 40-bp) and minimum quality (Q = 28) using Trimmomatic v0.36 [58]. The retained PE reads were assembled with Trinity v2.4.0 [59] using default parameters with a minimum transcript length of 200-bp. Read quality was assessed before and after read trimming with FastQC v0.11.5 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/).

Functional and mapping-rate filtering approaches were combined to reduce the redundancy present in the first version of the *P. margaritifera* transcriptome. Briefly, open-reading frames (ORFs) for each transcript were predicted using '*LongOrfs*' function implemented in Transdecoder v3.0.1 [59]. Only the transcripts containing an ORF of at least 100 amino acids were kept. Another filtering step included the removal of isoforms with residual expression; hence, only the most expressed isoforms for each gene with a mean mapping rate of 0.5 transcripts per million (TPM) were kept. We used a blastN approach against curated and non-redundant viral, bacterial, archeal, plasmid and fungal RefSeq transcripts databases (Download 19-09-17; ftp://ftp.ncbi.nlm.nih.gov/refseq/release/) to remove putative contamination. Transcripts matching a reference with an e-value < 10^{-10}, minimum identity of 75% and minimum query coverage of 70% were discarded. Finally, Illumina adapters were screened in the transcriptome using a blastN approach and adapter list (http://omicsoft.com/downloads/ngs/contamination_list/v1.txt). Assessment of the

283 final transcriptome completeness was conducted with BUSCO v1.1b [60] against the
284 conserved single-copy metazoan genes database (n = 978). Each filtering step was
285 validated with TransRate v1.0.3 [61]. Filtered reads were then mapped back on the
286 filtered reference transcriptome to evaluate individual mapping rate with GSNAP v2017-
287 03-17 [62].

288 For functional annotation, the transcripts were searched against Uniprot-
289 Swissprot [63], *Pinctada fucata*
290 (http://marinegenomics.oist.jp/pearl/viewer/download?project_id=20) and
291 *Crassostrea gigas* (ftp://ftp.ncbi.nih.gov/genomes/Crassostrea_gigas/RNA) protein
292 databases using a blastX approach (e-value < 10^{-4}) [64]. The best hit results are reported
293 in Table 2.

294 **Animal and tissue sampling.** An experimental graft was realised in order to obtain
295 cultured pearls and their corresponding pearl sacs. For this, a total of 20 pearl oyster
296 donors were used to perform 600 grafts (30 grafts per donor) over a 2-day period, using
297 2.4 BU nuclei (7.304 mm diameter, 0.59 g weight - Nucleus Bio, Hyakusyo Co., Japan) in
298 December 2015. All grafts were performed under standard production conditions by a
299 single expert at the Pahai Poe Pearl Farm (Apataki atoll, 15°34'S , 146°24' W, Tuamotu
300 archipelago, French Polynesia) so as to minimize the grafter effect on pearl quality traits
301 described in [65]. *Pinctada margaritifera* used in this part of the study had been
302 collected as wild spat in December 2013 using commercial collectors in the lagoon of
303 Takapoto atoll (14°32' S, 145°14' W, Tuamotu archipelago, French Polynesia), two years
304 before the experimental graft took place.

305 At the time of pearl harvest (18 months after the grafting operation, May 2017)
306 and in order to minimize the contamination by recipient tissues, the gonad was first cut

307 from the recipient oysters. The gonad tissue was then removed with a surgical blade
308 until only a thin (< 0.5 mm) layer of tissue surrounding the pearl remained. At this point,
309 only the pearl sac and the pearl remained [8]. Next, an incision was made in the pearl sac
310 in such a way as to remove the pearl, which was then placed in a numbered box for
311 traceability. The pearl sac was kept in a 2.0 ml tube with RNAlater until RNA extraction.
312 A total of 442 pearl sacs were sampled (73.6% of the total number of oysters grafted).
313 This pearl harvest rate represents the production yield, as nucleus rejection and
314 mortality were observed over the entire culture time [66].

315 Cultured pearl grade was evaluated as described in [67]. The pearls were cleaned
316 by ultrasonication in soapy water (hand washing) with a LEO 801 laboratory cleaner (2
317 L capacity, 80 W, 46 kHz); they were then rinsed in distilled water. To ensure
318 homogeneity in parameter assessment, all evaluations were made visually (no
319 magnification devices were used) by the same professional operator. Cultured pearl
320 grade is made up of two components: surface defects and lustre. When pearls are
321 graded, the appearance of their surface is evaluated. The degree of imperfection is
322 correlated with the number of defects (i.e. smooth surface *vs.* one or more spots). For the
323 grade classification, high pearl quality corresponds to grades of C and above, and low
324 quality to D1 and D2 grades (Figure 1). From the high and low quality pearls, twenty
325 pearl sacs from each group that had produced only green pearls were randomly
326 selected. Overall, high quality pearls were significantly heavier, with a thicker nacre
327 layer than the low quality pearls (t-test; p-value < 0.05). Our sampling included 10
328 donors equally allocated between several pools (either low or high pearl quality, as
329 illustrated in Figure 1) with the objectives of reducing putative donor effect [68–71].

RNA extraction and sequencing. The total RNA extraction procedure was identical to that for the samples used in the transcriptome assembly. For each condition (low or high pearl quality), an equimolar RNA quantity was pooled (n = 5 individuals/pool) to give a total of five pools by condition. RNA was dried in RNA-stable solution (ThermoFisher Scientific) following manufacturer's recommendations and shipped at room temperature to McGill sequencing platform services (Montreal, Canada). TruSeq Sample Prep. (Illumina, San Diego, Ca, USA) RNA libraries were multiplexed (N=10 pools by lane) and sequenced by HiSeq 4000 100-bp paired-end (PE) sequencing technology.

Differential expression analysis. Raw reads were trimmed using Trimmomatic v0.36 [58], using similar parameters as for the transcriptome assembly. Only paired-end reads were retained and mapped to the reference transcriptome, using GSNAP v2017-03-17 [62] with default parameters but allowing a minimum mismatch value of 5. Low mapping quality, multi-mapping (-q 5) and "non-properly paired reads" (-f 0x2) were removed using Samtools v1.4.1 [72]. A matrix of raw counts was built using HTSeq-count [73]. Low coverage transcripts (CPM < 1 in at least 10 individuals) were removed, resulting in a total of 40,633 transcripts and differential expression was assessed using the DESeq2 R package [74]. Transcripts were considered significant when FDR < 0.01 and |logFC| ≥ 1. Gene ontology (GO) enrichment was tested using GOAtools v0.6.5 [75] and the go-basic.obo database (release 2017-04-14) using Fisher's test. Our background list included the ensemble of genes used for differential expression after filtering for low coverage (n = 40,633 transcripts). Only GO terms with *p*-value < 0.05 and including at least three differentially expressed genes were considered. Significant GO enriched terms were used for semantic-based clustering in REVIGO (http://revigo.irb.hr/).

Individual gene expression validation. RT-qPCR was used to validate the expression patterns in the pearl sac, observed in RNA-seq, for key genes commonly used as markers

of pearl quality traits, namely *pif-177*, *aspein, shematrin-9* and the *mantle protein 10* (MP10). For each pool, total RNA was treated with DNAse I using a DNA-free Kit (Ambion). First, strand cDNA was synthesized from 500 ng total RNA using the Transcriptor First Strand cDNA Synthesis Kit (Roche) and a mix of poly (dT) and random hexamer primers. Real-Time PCR amplifications were carried out on a Roche Light Cycler 480. The amplification reaction contained 5 µL LC 480 SYBR Green I Mast (Roche), 4 µL cDNA templates, and 1 µL of primer (1µM), in a final volume of 10 µL. Each run included a positive cDNA and a blank control for each primer pair. The run protocol was as follows: initial denaturation at 95°C for 10 min followed by 40 cycles of denaturation at 95°C for 30 s, annealing at 60°C for 30 s and extension at 72°C for 60 s. Lastly, the amplicon melting temperature curve was analysed using a melting curve program: 45–95°C with a heating rate of 0.1°C s^{-1} and continuous fluorescence measurement.

All measurements were made in duplicate and all analyses were based on the Ct values of the PCR products. Relative gene expression levels were calculated using the delta–delta method, normalized with two reference genes (*elongation-factor 1* and *glyceraldehyde-3-phosphate dehydrogenase*), to compare the relative expression results as follows: Relative expression $_{(\text{target gene, sample x})} = 2^{\wedge -(\Delta Ct \text{ sample, sample x} - \Delta Ct \text{ calibrator, sample x})} = 2^{-\Delta\Delta Ct}$ [76]. Here, the ΔCt calibrator represents the mean of the ΔCt values obtained for the tested gene. The delta threshold cycle (ΔCt) is calculated as the difference in Ct for the target and reference genes. The relative stability of the GAPDH and EF-1 combination was confirmed using NormFinder [77]. PCR efficiency (E) was estimated for each primer pair by determining the slopes of standard curves obtained from serial dilution analysis of a cDNA to ensure that E ranged from 90 to 110%. The primers used for amplification are listed in Table S2. Wilcoxon non-parametric tests were used to compare relative

380 expression between conditions, differences were considered significant when p-value <
381 0.01. This complete list of primers is given in the supplementary file (Table S3).

382 **Alternative gene splicing and exon usage.** To detect putative differential splicing
383 variants, the reads were mapped on the scaffolds of the assembled draft genome,
384 available for *P. margaritifera*. To reduce non-biological redundancy inherent to the
385 current assembly state of the genome, only scaffolds with length >3,000 bp were used
386 for the mapping. From the 757,552 scaffolds initial set, only the 32,705 longest scaffolds
387 were retained for downstream analysis, on which 37,662 (92.7%) of the transcripts in
388 our set could be positioned using the GMAP v2017-03-17 aligner with the default
389 parameters for annotation [78]. Reads were mapped on the filtered reference genome
390 using GSNAP v2017-03-17 aligner allowing five mismatches, splicing and using the
391 '*splitting-output*' function to retain only concordant and unique mapped paired-end
392 reads [78]. We used the QORTs [79] and JunctionSeq R packages [80] to detect
393 significant differences in exon usage. Only exons and junctions with a minimal coverage
394 of six were used for the analysis and only differences with FDR < 0.01 were considered
395 significant.

396 ## List of abbreviations

397 BMSP-like: Blue mussel shell protein-like

398 Pwap: Perlwapin

399 wfd18: Wap four-disulfide core domain protein

400 ## Declarations

401 **Availability of data and material.** Raw transcriptome sequences data have been
402 deposited at NCBI (Bioproject: PRJNA449941). Genome and transcriptome assemblies

403 will be made publicaly available upon acceptance of the manuscript. Codes for RNA-seq
404 analysis are available upon request.

405 **Competing interests.** Authors declare no competing interest

406 **Funding.** This study was supported by grants from the "Direction des Ressources
407 Marines et Minières", through the AmeliGEN project (# 10065/MEI/DRMM).

408 **Authors' contributions.** CLK conceived the study. NL, VQ conducted the pearl sac
409 sampling and the RT-qPCR analysis, respectively. PA and JLL assembled the
410 transcriptome. CR and JVD assembled the reference genome of *P. margaritifera*. JLL
411 analysed the RNAseq data. JLL and CLK wrote the paper. All co-authors contributed to
412 reviewing the manuscript and accepted the final version for publication.

413 **Acknowledgements.** The authors would especially like thank the host site, Pahai Poe
414 Pearl Farm located in Apataki atoll (Tuamotu archipelago, French Polynesia) for their
415 generous support, and D. Potin for his assistance with the cultured pearl photographs.
416 The authors thank the Ifremer (Datarmor; RIC) and Station Biologique de Roscoff
417 (ABiMS; http://abims.sb-roscoff.fr) bioinformatic platforms for providing computational
418 resources and analysis support.

419 ## Figures and tables

420 Figure 1: High and low quality cultured pearl samples from *P. margaritifera*. Each row
421 within each condition (high or low quality), represents a specific pool (n = 4 cultured
422 pearls / pool / condition). Numbers represent pools by condition (high or low pearl
423 quality).

424 Figure 2: Genes differentially expressed between *P. margaritifera* pearl sacs having
425 produced high and low quality pearls: A) Principal component analysis (PCA); B)
426 Heatmap of differentially expressed genes. LogCPM (+2) were computed based on raw
427 counts normalized for library size reported to the mean gene expression over all
428 individuals; and C) Table showing the number of differentially expressed genes.

Figure 3: Bar plots of mean expression of biomineralization-related genes in pearl sac of *P. margaritifera*. Only genes with significant differential expression (FDR < 0.01 and |log2FC| > 1) are reported for clarity. Values are expressed as the mean (logCPM + 2) per condition ± standard deviation. Orange = high pearl quality; Blue = low pearl quality. Asterisks indicate genes for which multiple isoforms were reported in the transcriptome assembly. For each of these genes at least one of the isoforms was differentially expressed.

Figure 4: Splicing event visualization for the aspein gene in the pearl sac of *P. margaritifera*. Normalized counts are plotted for each gene section, either exon (E) or junction (J), and each individual; blue = high quality and orange = low quality. Values in the box plot represent p-values (Fisher's test) for each gene section.

Table 1: Transcriptome assembly, annotation statistics and differential expression results

Supplementary

Figure S1: Summarized REVIGO semantic plot for gene ontology enrichment analysis

Figure S2: Relative gene expression for biomineralization genes analysed by qPCR in the pearl sac of *P. margaritifera*. Values are expressed as means of relative expression ± standard deviation. Asterisks indicate significant differences (Wilcoxon test, *p*-value < 0.01).

Table S1: *P. margaritifera* individuals used for the transcriptome assembly. NA = Not identified.

Table S 2: Complete list and statistics on differentially expressed genes and their annotation.

Table S 3: Set of forward and reverse primers used for the biomineralization gene expression (real-time PCR) analysis in *Pinctada margaritifera*.

References

1. Taylor JJ, Strack E. Pearl production. In: Southgate PC, Lucas JS, editors. Pearl oyster. Amsterdam, The Netherlands: Elsevier. p. 273–302.

2. Gervis MH, Sims NA. The biology and culture of pearl oysters (Bivalvia pteriidae). London, England : Manila, Philippines: Overseas Development Administration of the United Kingdom ; International Center for Living Aquatic Resources Management; 1992.

3. Kishore P, Southgate PC. A detailed description of pearl-sac development in the black-lip pearl oyster, *Pinctada margaritifera* (Linnaeus 1758). Aquac Res. 2016;47:2215–26. https://onlinelibrary.wiley.com/doi/epdf/10.1111/are.12674. Accessed 27 Jun 2018.

4. Ky C-L, Blay C, Sham-Koua M, Vanaa V, Lo C, Cabral P. Family effect on cultured pearl quality in black-lipped pearl oyster *Pinctada margaritifera* and insights for genetic improvement. Aquat Living Resour. 2013;26:133–45. doi:10.1051/alr/2013055.

5. Ellis S, Haws M. Producing pearls using the black-lip pearl oyster (*Pinctada margaritifera*). Aquafarmer Inf. Sheet. 1999;:8.

6. Blay C, Planes S, Ky C-L. Crossing phenotype heritability and candidate gene expression in grafted black-lipped pearl oyster *Pinctada margaritifera*, an animal chimera. J Heridity. 2018;in press.

7. Ky C-L, Okura R, Nakasai S, Devaux D. Quality trait signature at archipelago scale of the cultured pearls produced by the black-lipped pearl oyster (*Pinctada margaritifera* Var. cumingi) in French Polynesia. J Shellfish Res. 2016;35:827–35.

8. Latchere O, Le Moullac G, Gaertner-Mazouni N, Fievet J, Magré K, Saulnier D. Influence

478 of preoperative food and temperature conditions on pearl biogenesis in *Pinctada*
479 *margaritifera*. Aquaculture. 2017;479 Supplement C:176–87.
480 doi:10.1016/j.aquaculture.2017.05.046.

481 9. Li S, Huang J, Liu C, Liu Y, Zheng G, Xie L, et al. Interactive effects of seawater
482 acidification and elevated temperature on the transcriptome and biomineralization in
483 the pearl oyster *Pinctada fucata*. Environ Sci Technol. 2016;50:1157–65.
484 doi:10.1021/acs.est.5b05107.

485 10. Gueguen Y, Czorlich Y, Mastail M, Le Tohic B, Defay D, Lyonnard P, et al. Yes, it turns:
486 experimental evidence of pearl rotation during its formation. R Soc Open Sci.
487 2015;2:150144. doi:10.1098/rsos.150144.

488 11. Tayalé A, Gueguen Y, Treguier C, Grand JL, Cochennec-Laureau N, Montagnani C, et al.
489 Evidence of donor effect on cultured pearl quality from a duplicated grafting experiment
490 on *Pinctada margaritifera* using wild donors. Aquat Living Resour. 2012;25:269–80.
491 doi:10.1051/alr/2012034.

492 12. Ky C-L, Blay C, Sham-Koua M, Lo C, Cabral P. Indirect improvement of pearl grade
493 and shape in farmed *Pinctada margaritifera* by donor "oyster" selection for green pearls.
494 Aquaculture. 2014;432 Supplement C:154–62. doi:10.1016/j.aquaculture.2014.05.002.

495 13. Blay C, Sham-Koua M, Vonau V, Tetumu R, Cabral P, Ky C-L. Influence of nacre
496 deposition rate on cultured pearl grade and colour in the black-lipped pearl oyster
497 *Pinctada margaritifera* using farmed donor families. Aquac Int. 2014;22:937–53.
498 doi:10.1007/s10499-013-9719-5.

499 14. Ky C-L, Blay C, Aiho V, Cabral P, Le Moullac G, Lo C. Macro-geographical differences
500 influenced by family-based expression on cultured pearl grade, shape and colour in the

501 black-lip 'pearl oyster' *Pinctada margaritifera*: a preliminary bi-local case study in
502 French Polynesia. Aquac Res. 2017;48:270–82. doi:10.1111/are.12880.

503 15. Blay C, Planes S, Ky C-L. Optimal age of the donor graft tissue in relation to cultured
504 pearl phenotypes in the mollusc, *Pinctada margaritifera*. PLOS ONE. 2018;13:e0198505.
505 doi:10.1371/journal.pone.0198505.

506 16. Ky C-L, Nakasai S, Parrad S, Broustal F, Devaux D, Louis P. Variation in cultured pearl
507 quality traits in relation to position of saibo cutting on the mantle of black-lipped pearl
508 oyster *Pinctada margaritifera*. Aquaculture. 2018;493:85–92.
509 doi:10.1016/j.aquaculture.2018.04.059.

510 17. Ky C-L, Molinari N, Moe E, Pommier S. Impact of season and grafter skill on nucleus
511 retention and pearl oyster mortality rate in *Pinctada margaritifera* aquaculture. Aquac
512 Int. 2014;22:1689–701. doi:10.1007/s10499-014-9774-6.

513 18. Lemer S, Saulnier D, Gueguen Y, Planes S. Identification of genes associated with shell
514 color in the black-lipped pearl oyster, *Pinctada margaritifera*. BMC Genomics. 2015;16.

515 19. Jones DB, Jerry DR, Khatkar MS, Raadsma HW, Zenger KR. A high-density SNP genetic
516 linkage map for the silver-lipped pearl oyster, *Pinctada maxima*: a valuable resource for
517 gene localisation and marker-assisted selection. BMC Genomics. 2013;14:810.
518 doi:10.1186/1471-2164-14-810.

519 20. Jones DB, Jerry DR, Khatkar MS, Moser G, Raadsma HW, Taylor JJ, et al. Quantitative
520 trait loci and genetic association analysis reveals insights into complex pearl quality
521 traits in donor silver-lipped pearl oysters. Aquaculture. 2014;434:476–85.
522 doi:10.1016/j.aquaculture.2014.08.038.

523 21. Inoue N, Ishibashi R, Ishikawa T, Atsumi T, Aoki H, Komaru A. Comparison of
524 expression patterns of shell matrix protein genes in the mantle tissues between high-
525 and low-quality pearl-producing recipients of the pearl oyster, *Pinctada fucata*. Zoolog
526 Sci. 2011;28:32–6. doi:10.2108/zsj.28.32.

527 22. Inoue N, Ishibashi R, Ishikawa T, Atsumi T, Aoki H, Komaru A. Can the quality of
528 pearls from the Japanese pearl oyster (*Pinctada fucata*) be explained by the gene
529 expression patterns of the major shell matrix proteins in the pearl sac? Mar Biotechnol N
530 Y N. 2011;13:48–55. doi:10.1007/s10126-010-9267-1.

531 23. Blay C, Planes S, Ky C-L. Cultured pearl surface quality profiling by the shell matrix
532 protein gene expression in the biomineralised pearl sac tissue of *Pinctada margaritifera*.
533 Mar Biotechnol. 2018;in press in press.

534 24. Marie B, Joubert C, Tayalé A, Zanella-Cléon I, Belliard C, Piquemal D, et al. Different
535 secretory repertoires control the biomineralization processes of prism and nacre
536 deposition of the pearl oyster shell. Proc Natl Acad Sci. 2012;109:20986–91.
537 doi:10.1073/pnas.1210552109.

538 25. Marie B, Jackson DJ, Ramos-Silva P, Zanella-Cléon I, Guichard N, Marin F. The shell-
539 forming proteome of *Lottia gigantea* reveals both deep conservations and lineage-
540 specific novelties. FEBS J. 280:214–32. doi:10.1111/febs.12062.

541 26. Isowa Y, Sarashina I, Setiamarga DHE, Endo K. A comparative study of the shell
542 matrix protein aspein in pterioid bivalves. J Mol Evol. 2012;75:11–8.
543 doi:10.1007/s00239-012-9514-3.

544 27. Takeuchi T, Sarashina I, Iijima M, Endo K. *In vitro* regulation of CaCO3 crystal
545 polymorphism by the highly acidic molluscan shell protein Aspein. FEBS Lett.

546 2008;582:591–6. doi:10.1016/j.febslet.2008.01.026.

547 28. Tsukamoto D, Sarashina I, Endo K. Structure and expression of an unusually acidic
548 matrix protein of pearl oyster shells. Biochem Biophys Res Commun. 2004;320:1175–
549 80. doi:10.1016/j.bbrc.2004.06.072.

550 29. Elhadj S, Yoreo JJD, Hoyer JR, Dove PM. Role of molecular charge and hydrophilicity
551 in regulating the kinetics of crystal growth. Proc Natl Acad Sci. 2006;103:19237–42.
552 doi:10.1073/pnas.0605748103.

553 30. Shi Y, He M. Differential gene expression identified by RNA-Seq and qPCR in two
554 sizes of pearl oyster (*Pinctada fucata*). Gene. 2014;538:313–22.
555 doi:10.1016/j.gene.2014.01.031.

556 31. Li S, Liu C, Huang J, Liu Y, Zhang S, Zheng G, et al. Transcriptome and
557 biomineralization responses of the pearl oyster *Pinctada fucata* to elevated CO_2 and
558 temperature. Sci Rep. 2016;6:18943. doi:10.1038/srep18943.

559 32. Zhao X, Wang Q, Jiao Y, Huang R, Deng Y, Wang H, et al. Identification of genes
560 potentially related to biomineralization and Immunity by transcriptome analysis of
561 pearl sac in pearl oyster *Pinctada martensii*. Mar Biotechnol. 2012;14:730–9.
562 doi:10.1007/s10126-012-9438-3.

563 33. Shi Y, Yu C, Gu Z, Zhan X, Wang Y, Wang A. Characterization of the pearl oyster
564 (*Pinctada martensii*) mantle transcriptome unravels biomineralization genes. Mar
565 Biotechnol. 2013;15:175–87. doi:10.1007/s10126-012-9476-x.

566 34. Li H, Liu B, Huang G, Fan S, Zhang B, Su J, et al. Characterization of transcriptome and
567 identification of biomineralization genes in winged pearl oyster (*Pteria penguin*) mantle

568 tissue. Comp Biochem Physiol Part D Genomics Proteomics. 2017;21:67–76.
569 doi:10.1016/j.cbd.2016.12.002.

570 35. Joubert C, Piquemal D, Marie B, Manchon L, Pierrat F, Zanella-Cléon I, et al.
571 Transcriptome and proteome analysis of *Pinctada margaritifera* calcifying mantle and
572 shell: focus on biomineralization. BMC Genomics. 2010;11:613. doi:10.1186/1471-2164-
573 11-613.

574 36. Teaniniuraitemoana V, Huvet A, Levy P, Klopp C, Lhuillier E, Gaertner-Mazouni N, et
575 al. Gonad transcriptome analysis of pearl oyster *Pinctada margaritifera*: identification of
576 potential sex differentiation and sex determining genes. BMC Genomics. 2014;15:491.
577 doi:10.1186/1471-2164-15-491.

578 37. Huang X-D, Zhao M, Liu W-G, Guan Y-Y, Shi Y, Wang Q, et al. Gigabase-scale
579 transcriptome analysis on four species of pearl oysters. Mar Biotechnol. 2013;15:253–
580 64.

581 38. Marin F, Luquet G, Marie B, Medakovic D. Molluscan shell proteins: primary
582 structure, origin, and evolution. Curr Top Dev Biol. 2008;80:209–76.
583 doi:10.1016/S0070-2153(07)80006-8.

584 39. Marie B, Arivalagan J, Mathéron L, Bolbach G, Berland S, Marie A, et al. Deep
585 conservation of bivalve nacre proteins highlighted by shell matrix proteomics of the
586 Unionoida *Elliptio complanata* and *Villosa lienosa*. J R Soc Interface. 2017;14:20160846.
587 doi:10.1098/rsif.2016.0846.

588 40. Kocot KM, Aguilera F, McDougall C, Jackson DJ, Degnan BM. Sea shell diversity and
589 rapidly evolving secretomes: insights into the evolution of biomineralization. Front Zool.
590 2016;13.

591 41. Bahn SY, Jo BH, Choi YS, Cha HJ. Control of nacre biomineralization by Pif80 in pearl
592 oyster. Sci Adv. 2017;3.

593 42. Gu Z, Yin X, Yu C, Zhan X, Shi Y, Wang A. Expression profiles of nine biomineralization
594 genes and their relationship with pearl nacre thickness in the pearl oyster, *Pinctada*
595 *fucata* martensii Dunker. Aquac Res. 2016;47:1874–84. doi:10.1111/are.12645.

596 43. Arivalagan J, Yarra T, Marie B, Sleight VA, Duvernois-Berthet E, Clark MS, et al.
597 Insights from the Shell Proteome: Biomineralization to Adaptation. Mol Biol Evol.
598 2017;34:66–77. doi:10.1093/molbev/msw219.

599 44. Jackson DJ, McDougall C, Woodcroft B, Moase P, Rose RA, Kube M, et al. Parallel
600 evolution of nacre building gene sets in molluscs. Mol Biol Evol. 2010;27:591–608.

601 45. Wang X, Liu Z, Wu W. Transcriptome analysis of the freshwater pearl mussel
602 (*Cristaria plicata*) mantle unravels genes involved in the formation of shell and pearl.
603 Mol Genet Genomics MGG. 2017;292:343–52.

604 46. Li S, Liu Y, Huang J, Zhan A, Xie L, Zhang R. The receptor genes PfBMPR1B and
605 PfBAMBI are involved in regulating shell biomineralization in the pearl oyster *Pinctada*
606 *fucata*. Sci Rep. 2017;7.

607 47. Joubert C, Linard C, Moullac GL, Soyez C, Saulnier D, Teaniniuraitemoana V, et al.
608 Temperature and food influence shell growth and mantle gene expression of shell
609 matrix proteins in the pearl oyster *Pinctada margaritifera*. PLOS ONE. 2014;9:e103944.
610 doi:10.1371/journal.pone.0103944.

611 48. Blay C, Parrad S, Cabral P, Aiho V, Ky C-L. Correlations between cultured pearl size
612 parameters and PIF-177 biomarker expression in *Pinctada margaritifera* families reared

613 in two contrasting environments. Estuar Coast Shelf Sci. 2016;182 Part B:254–60.
614 doi:10.1016/j.ecss.2016.05.020.

615 49. Feng D, Li Q, Yu H, Kong L, Du S. Identification of conserved proteins from diverse
616 shell matrix proteome in Crassostrea gigas: characterization of genetic bases regulating
617 shell formation. Sci Rep. 2017;7.

618 50. Suzuki M, Saruwatari K, Kogure T, Yamamoto Y, Nishimura T, Kato T, et al. An acidic
619 matrix protein, Pif, is a key macromolecule for nacre formation. Science.
620 2009;325:1388–1390. doi:10.1126/science.1173793.

621 51. Livingston BT, Killian CE, Wilt F, Cameron A, Landrum MJ, Ermolaeva O, et al. A
622 genome-wide analysis of biomineralization-related proteins in the sea urchin
623 *Strongylocentrotus purpuratus*. Dev Biol. 2006;300:335–48.
624 doi:10.1016/j.ydbio.2006.07.047.

625 52. Chuong EB, Elde NC, Feschotte C. Regulatory activities of transposable elements:
626 from conflicts to benefits. Nat Rev Genet. 2017;18:71–86. doi:10.1038/nrg.2016.139.

627 53. Hirsch CD, Springer NM. Transposable element influences on gene expression in
628 plants. Biochim Biophys Acta BBA - Gene Regul Mech. 2017;1860:157–65.
629 doi:10.1016/j.bbagrm.2016.05.010.

630 54. Trizzino M, Park Y, Holsbach-Beltrame M, Aracena K, Mika K, Caliskan M, et al.
631 Transposable elements are the primary source of novelty in primate gene regulation.
632 Genome Res. 2017;:gr.218149.116. doi:10.1101/gr.218149.116.

633 55. Ahmed M, Liang P. Transposable elements are a significant contributor to tandem
634 repeats in the Human genome. Comparative and Functional Genomics. 2012.

635 56. Sharma A, Wolfgruber TK, Presting GG. Tandem repeats derived from centromeric
636 retrotransposons. BMC Genomics. 2013;14:142. doi:10.1186/1471-2164-14-142.

637 57. Liu J, Yang D, Liu S, Li S, Xu G, Zheng G, et al. Microarray: a global analysis of
638 biomineralization-related gene expression profiles during larval development in the
639 pearl oyster, *Pinctada fucata*. BMC Genomics. 2015;16. doi:10.1186/s12864-015-1524-
640 2.

641 58. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina
642 sequence data. Bioinformatics. 2014;:btu170.

643 59. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, et al. De novo
644 transcript sequence reconstruction from RNA-seq using the Trinity platform for
645 reference generation and analysis. Nat Protoc. 2013;8:1494–1512.

646 60. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO:
647 assessing genome assembly and annotation completeness with single-copy orthologs.
648 Bioinformatics. 2015;31:3210–2.

649 61. Smith-Unna R, Boursnell C, Patro R, Hibberd JM, Kelly S. TransRate: reference-free
650 quality assessment of de novo transcriptome assemblies. Genome Res. 2016;26:1134–
651 44.

652 62. Wu TD, Reeder J, Lawrence M, Becker G, Brauer MJ. GMAP and GSNAP for genomic
653 sequence alignment: enhancements to speed, accuracy, and functionality. Statistical
654 Genomics: Methods and Protocols. 2016;:283–334.

655 63. Bairoch A, Apweiler R. The SWISS-PROT protein sequence database and its
656 supplement TrEMBL in 2000. Nucleic Acids Res. 2000;28:45–8.

http://www.ncbi.nlm.nih.gov/pmc/articles/PMC102476/. Accessed 20 Dec 2016.

64. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997;25:3389–402.

65. Ky C-L, Lau C, Koua MS, Lo C. Growth performance comparison of *Pinctada margaritifera* juveniles produced by thermal dhock or gonad scarification spawning procedures. J Shellfish Res. 2015;34:811–7.

66. Ky C-L, Molinari N, Moe E, Pommier S. Impact of season and grafter skill on nucleus retention and pearl oyster mortality rate in *Pinctada margaritifera* aquaculture. Aquac Int. 2014;22:1689–701.

67. Ky C-L, Okura R, Nakasai S, Devaux D. Quality trait signature at archipelago scale of the cultured pearls produced by the black-lipped pearl oyster (*Pinctada margaritifera* Var. cumingi) in French Polynesia. J Shellfish Res. 2016;35:827–35.

68. McGinty EL, Zenger KR, Jones DB, Jerry DR. Transcriptome analysis of biomineralisation-related genes within the pearl sac: Host and donor oyster contribution. Mar Genomics. 2012;5:27–33.

69. Tayalé A, Gueguen Y, Treguier C, Grand JL, Cochennec-Laureau N, Montagnani C, et al. Evidence of donor effect on cultured pearl quality from a duplicated grafting experiment on *Pinctada margaritifera* using wild donors. Aquat Living Resour. 2012;25:269–80.

70. Ky C-L, Blay C, Sham-Koua M, Vanaa V, Lo C, Cabral P. Family effect on cultured pearl quality in black-lipped pearl oyster *Pinctada margaritifera* and insights for genetic improvement. Aquat Living Resour. 2013;26:133–45.

679 71. Ky C-L, Blay C, Sham-Koua M, Lo C, Cabral P. Indirect improvement of pearl grade
680 and shape in farmed *Pinctada margaritifera* by donor "oyster" selection for green pearls.
681 Aquaculture. 2014;432 Supplement C:154–62.

682 72. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence
683 Alignment/Map format and SAMtools. Bioinformatics. 2009;25:2078–9.

684 73. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-
685 throughput sequencing data. Bioinformatics. 2015;31:166–9.

686 74. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for
687 RNA-seq data with DESeq2. Genome Biol. 2014;15:550.

688 75. Klopfenstein DV, Zhang L, Pedersen BS, Ramírez F, Warwick Vesztrocy A, Naldi A, et
689 al. GOATOOLS: A Python library for Gene Ontology analyses. Sci Rep. 2018;8:10872.
690 doi:10.1038/s41598-018-28948-z.

691 76. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time
692 quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods San Diego Calif.
693 2001;25:402–8. doi:10.1006/meth.2001.1262.

694 77. Andersen CL, Jensen JL, Ørntoft TF. Normalization of real-time quantitative reverse
695 transcription-PCR data: a model-based variance estimation approach to identify genes
696 suited for normalization, applied to bladder and colon cancer data sets. Cancer Res.
697 2004;64:5245–50.

698 78. Wu TD, Reeder J, Lawrence M, Becker G, Brauer MJ. GMAP and GSNAP for genomic
699 sequence alignment: enhancements to speed, accuracy, and functionality. Statistical
700 Genomics: Methods and Protocols. 2016;:283–334. http://dx.doi.org/10.1007/978-1-

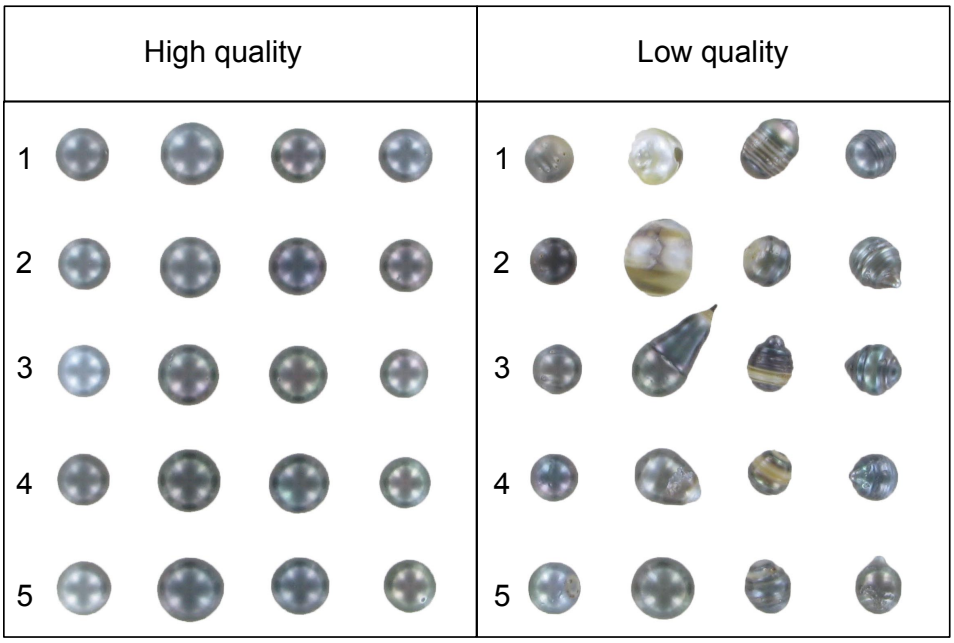
4939-3578-9_15.

79. Hartley SW, Mullikin JC. QoRTs: a comprehensive toolset for quality control and data processing of RNA-Seq experiments. BMC Bioinformatics. 2015;16.

80. Hartley SW, Mullikin JC. Detection and visualization of differential splicing in RNA-Seq data with JunctionSeq. Nucleic Acids Res. 2016;:gkw501.

708

709 Figure 1.



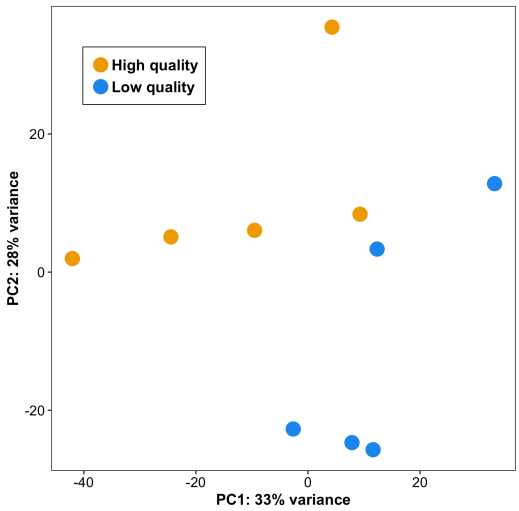
710

711

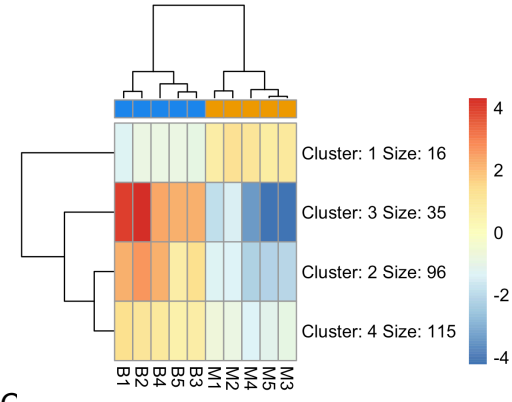
712

713 Figure 2.

A



B



C

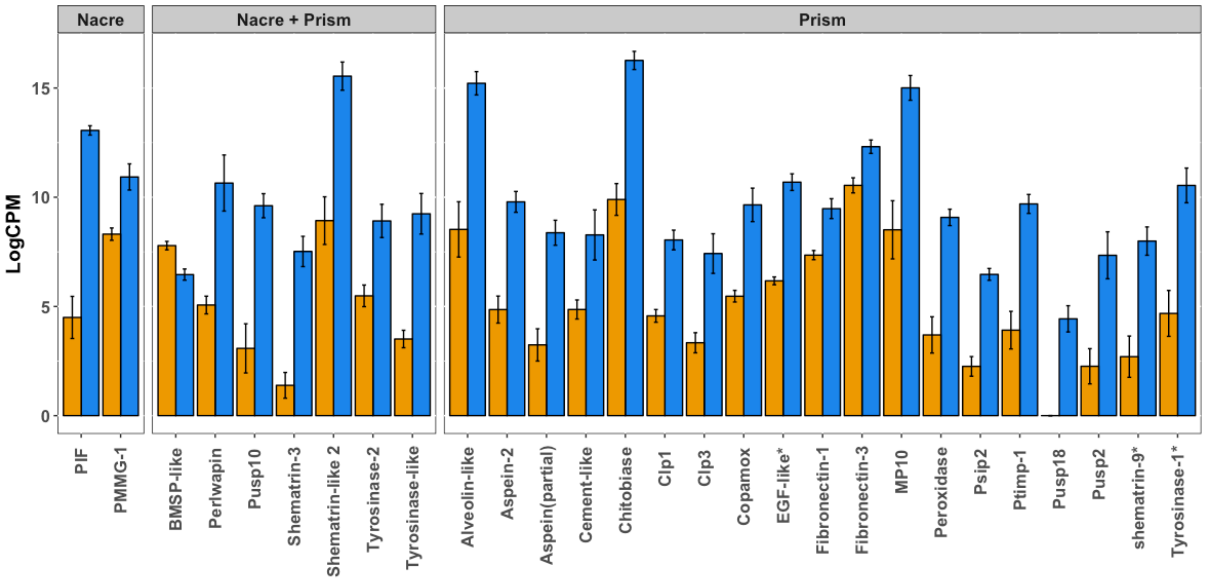
Condition	Up-regulated in Low	Down-regulated in Low	Total
High / Low	246	16	262

714

715

716

717 Figure 3.

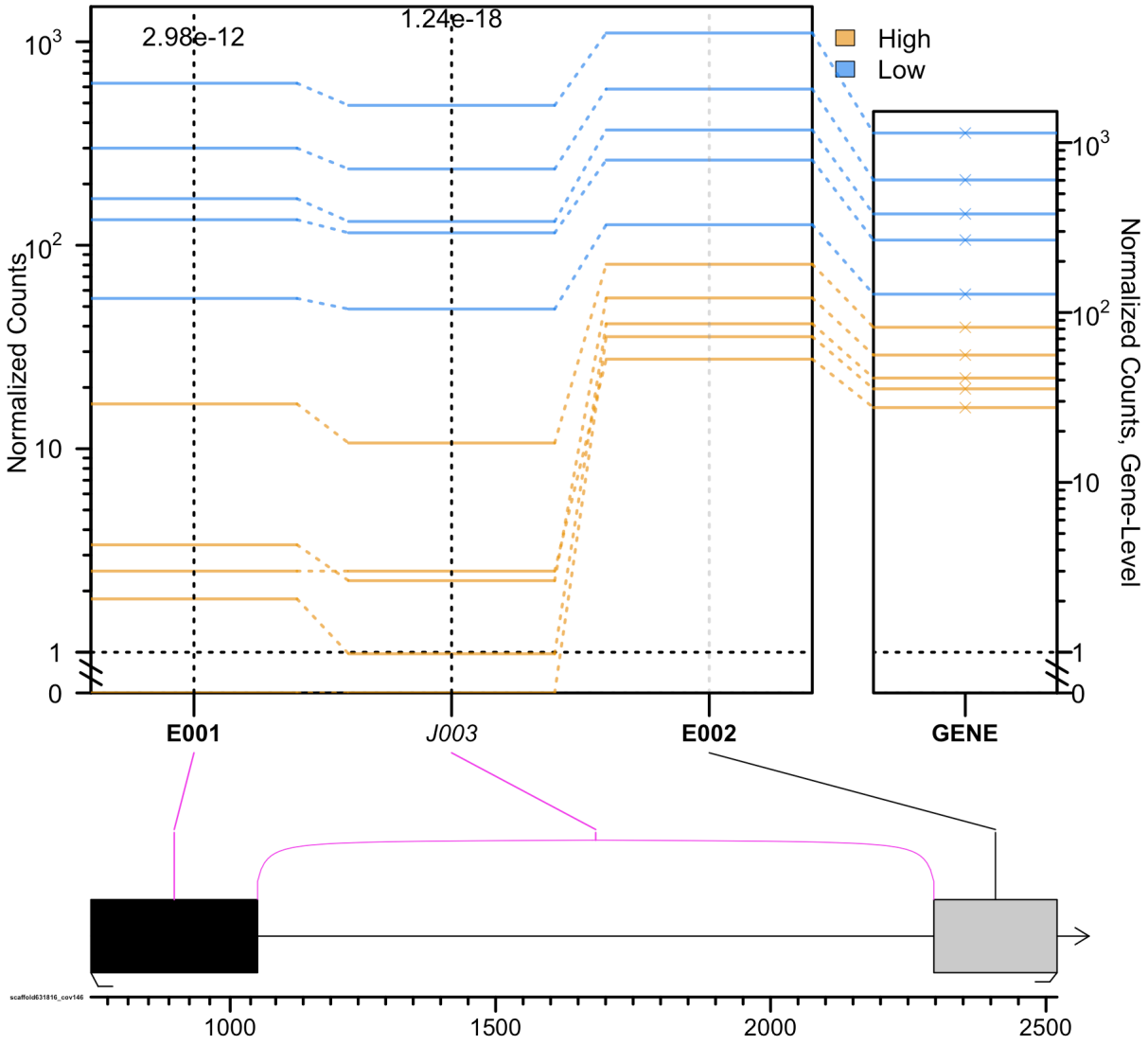


718

719

720

721 Figure 4.



722

723

724

725 Table 1.

Transcriptome statistics	
Raw number of contigs	541,184
Total number of contigs	41,075
Percent GC	39.1
Contigs N50 (bp)	2,087
Total assembled bases	57,671,860
Median contig length (bp)	958
Average contig length (bp)	1,404.06
Annotation	
Contig with *P. fucata* match (e-value 10^{-4})	32,489 (19,850)*
Contig with *C. Gigas* match (e-value 10^{-4})	27,568 (12,837)*
Contig with Uniprot-Swissprot match (e-value 10^{-4})	20,177 (845)*
Contig with GO identifier annotation**	18,163

726 * Best hit

727 ** Based on Uniprot-Swissprot blastX results

728