


RESEARCH ARTICLE

Open Access



# Tracing key genes associated with the *Pinctada margaritifera* albino phenotype from juvenile to cultured pearl harvest stages using multiple whole transcriptome sequencing

Pauline Auffret<sup>1</sup>, Jérémy Le Luyer<sup>1</sup>, Manaarii Sham Koua<sup>1</sup>, Virgile Quillien<sup>1,2</sup> and Chin-Long Ky<sup>1,3\*</sup> 

## Abstract

**Background:** Albino mutations are commonly observed in the animal kingdom, including in bivalves. In the black-lipped pearl oyster *Pinctada margaritifera*, albino specimens are characterized by total or partial absence of colouration resulting in typical white shell phenotype expression. The relationship of shell colour with resulting cultured pearl colour is of great economic interest in *P. margaritifera*, on which a pearl industry is based. Hence, the albino phenotype provides a useful way to examine the molecular mechanisms underlying pigmentation.

**Results:** Whole transcriptome RNA-sequencing analysis comparing albino and black wild-type phenotypes at three stages over the culture cycle of *P. margaritifera* revealed a total of 1606, 798 and 187 differentially expressed genes in whole juvenile, adult mantle and pearl sac tissue, respectively. These genes were found to be involved in five main molecular pathways, tightly linked to known pigmentation pathways: melanogenesis, calcium signalling pathway, Notch signalling pathway, pigment transport and biomineralization. Additionally, significant phenotype-associated SNPs were selected ( $N = 159$ ), including two located in the *Pif* biomineralization gene, which codes for nacre formation. Interestingly, significantly different transcript splicing was detected between juvenile ( $N = 1366$ ) and adult mantle tissue ( $N = 313$ ) in, e.g., the tyrosinase *Tyr-1* gene, which showed more complex regulation in mantle, and the Notch1 encoding gene, which was upregulated in albino juveniles.

**Conclusion:** This multiple RNA-seq approach provided new knowledge about genes associated with the *P. margaritifera* albino phenotype, highlighting: 1) new molecular pathways, such as the Notch signalling pathway in pigmentation, 2) associated SNP markers with biomineralization gene of interest like *Pif* for marker-assisted selection and prevention of inbreeding, and 3) alternative gene splicing for melanin biosynthesis implicating tyrosinase.

**Keywords:** *Pinctada margaritifera*, Albinism, Pigmentation, RNA-seq, Notch signaling pathway, Tyrosinase, Biomineralization

\* Correspondence: [chinky@ifremer.fr](mailto:chinky@ifremer.fr)

<sup>1</sup>Ifremer, UMR EIO 241, Centre du Pacifique, BP 49, 98719 Taravao, Tahiti, Polynésie française, France

<sup>3</sup>IHPE, Université de Montpellier, CNRS, Ifremer, Université de Perpignan Via Domitia, F-34090 Montpellier, France

Full list of author information is available at the end of the article



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Background

Mutations affecting pigmentation were among the first in the animal kingdom to be studied for Mendelian inheritance. Some species became ideal models for studying the genetic mechanisms that determine colour phenotypes [1]. Colour traits in domestic animals have been widely used as unique phenotypes for morphological broodstock selection and identification. In the aquaculture industry, body pigmentation can be an economically important trait selected to directly enhance the commercial value of a given species, for example, the red Nile tilapia (*Oreochromis niloticus*) compared with the wild type [2], or the majority of freshwater/marine ornamental species such as clownfishes [3]. Distinctive colouration characteristics, such as albinism, are common in most fish and shellfish species. For instance, in albino rainbow trout *Oncorhynchus mykiss*, the associated phenotype is characterised by a white or yellow skin according to the species [4]. Albino yellow catfish *Pelteobagrus fulvidraco* have been observed in control breeding in China [5]. Albino individuals have also been observed among aquacultured Russian sturgeons, one of the most valuable freshwater fish species worldwide [6]. In bivalves, white specimens of the Akoya pearl oyster *Pinctada fucata* are already used in selective breeding programs [7, 8]. Albinism can also be very useful as a visual phenotypic marker for experimental tagging studies [9]. In genetic studies, it allows the evaluation of chromosome manipulation efficiency in gynogenic or triploid animals [10] as well as the optimisation of gene editing approaches [11, 12].

Albinism is a genetically inherited trait that is usually defined as the absence or deficiency of pigmentation due to dysfunction in melanin production. Many mutations associated with albinism have been identified and well described in mice and humans (<http://www.ifpcs.org/albinism/>). Dysfunction of pigmentation has been found to result from multiple and complex molecular processes occurring during melanin production, transport and transfer and, for bivalves, in shell biomineralization. Many genes are involved along the pathway and have been reported to be deregulated in white-shelled bivalves [13–15]. Among these, tyrosinases are key enzymes regulating the melanin biosynthetic pathway. Molluscan tyrosinases are secreted from the mantle and transported to the prismatic shell layer where they contribute to melanin biosynthesis and shell pigmentation [16]. The Notch signalling pathway is an essential cell interaction mechanism, which plays a fundamental role in metazoan development [17]. Recently, the Notch signalling pathway was found to play a crucial role in shell pigmentation in the clam *Meretrix meretrix*, following a gene-dosage dependent pattern [15]. The calcium signalling pathway may also be implicated through activation of the Notch pathway in this species. Other studies implicated the Wnt signalling pathway in the maintenance of

melanocyte and keratinocyte homeostasis and, therefore, its impact on pigmentation in the sea cucumber *Apostichopus japonicus* [18]. It is worth mentioning that the majority of the previous studies working on pigmentation in bivalves focused on mantle tissue samples.

Marine aquaculture of the black-lipped pearl oyster *Pinctada margaritifera* var. *Cumingii* for cultured pearl production is a major economic driver in French Polynesia for cultured pearl production (see [19] for current economic data). Cultured pearls form following a grafting operation, which consists in inserting a small piece of mantle tissue from a donor oyster (called a graft or *saibo*), into the gonad of a recipient oyster together with a spherical nucleus [20]. Nacre layers secreted by the mineralizing activity of the pearl sac (tissue which originated from the graft mantle) are progressively recovering the nucleus during culture time, forming the cultured pearl [21]. The whole process, from natural collection or hatchery production of spat, through juvenile growth to the adult stage at which the animals have reached a sufficient size to provide graft tissue or be grafted and finally the culture until pearl harvest, is a long process throughout which the biomineralization properties of specific tissues continue. *P. margaritifera* produces the largest range of cultured pearl colour in the genus [22, 23], including white pearls produced by the rare white albino shell morphotype, when these individuals are used as donors [24]. These albino individuals are characterised by total or partial absence of coloration resulting in a white shell (periostracum and calcitic layer) [14]. Studies on the *P. margaritifera* albino phenotype revealed Mendelian inheritance of shell colour under a three-allele model with no co-dominance, where the albino type is recessive to the black wild type and red phenotypes [24]. Suppressive subtractive hybridization (SSH) was previously used to characterise genes involved in shell colour by comparing black wild-type and albino phenotypes [14]. This earlier study revealed the implication of a small number of genes in albino phenotypes, including *ZINC*, a homolog of the *tyrosinase-related protein 1* gene, which codes for a protein with metallic zinc ion binding and a catalytic domain [14].

In the present study, we investigated the multiple albino phenotype expression in *P. margaritifera* through successive and complementary whole transcriptomic sequencing analyses on different tissues from oysters of different stages collected over the entire pearl culture cycle. Three different tissues were analysed: whole flesh of juveniles, mantle tissue of donor oysters and, at harvest, the pearl sacs of recipients that had been grafted using albino *saibo*. Three distinctive RNA-seq approaches were then used to compare gene expression between the white shell phenotype (albino) and the black shell phenotype (wild type) in the same cohort. The three RNA-seq datasets shared only a few deregulated

genes, underlining the importance of using multi-stage transcriptome analysis to provide a comprehensive overview of all processes potentially associated with the albino phenotype. We identified new candidate genes for white-shelled molluscs, together with genes already known to be expressed in the mantle tissue of *P. margaritifera*. Furthermore, we examined potential phenotype-associated SNPs and alternative gene splicing.

## Results

### Sequencing and mapping of libraries

The raw data of the three datasets were submitted to the National Center for Biotechnology (NCBI; BioProject ID: PRJNA642706). The sample identifiers were: 1) for the Juvenile (J) dataset: AN6 to AN10 (albino) and TB6 to TB10 (black wild-type); 2) for the Mantle (M) dataset: X11 to X15 (albino) and X6 to X9 (black wild-type); and 3) for the Pearl Sac (PS) dataset: B8, B11, B17 and B19 (albino) and N45, N47, N54 and N55 (black wild-type). A total of  $41.14 \cdot 10^9$  (J),  $83.10^8$  (M),  $5.61 \cdot 10^8$  (PS) clean reads filtered from  $1.17 \cdot 10^9$  (J),  $4.90 \cdot 10^8$  (M),  $5.41 \cdot 10^8$  (PS) paired-end 100 bp raw reads were obtained from the selected samples. Mapping yield against the multi-tissue and stages reference transcriptome of *P. margaritifera* [25] reached 75.61% (J), 80.81% (M) and 74.17% (PS). Finally, a total 90.33% (J), 90.89% (M) and 91.27% (PS) of the mapped reads were kept after filtering (Table 1).

### Differential expression analysis

After filtering for residual expression, we retained 37, 466; 28,286 and 32,313 transcripts in J, M and PS, respectively, for downstream analyses. By comparing albino with black wild-type results, a total of 1606 (J), 798 (M) and 187 (PS) differentially expressed genes (DEGs) were detected. Among these DEGs, 760 (J), 387 (M) and 121 (PS) were up-regulated in albino individuals, while 846 (J), 411 (M) and 66 (PS) were down-regulated (Additional file 1). There were notably fewer DEGs in the PS

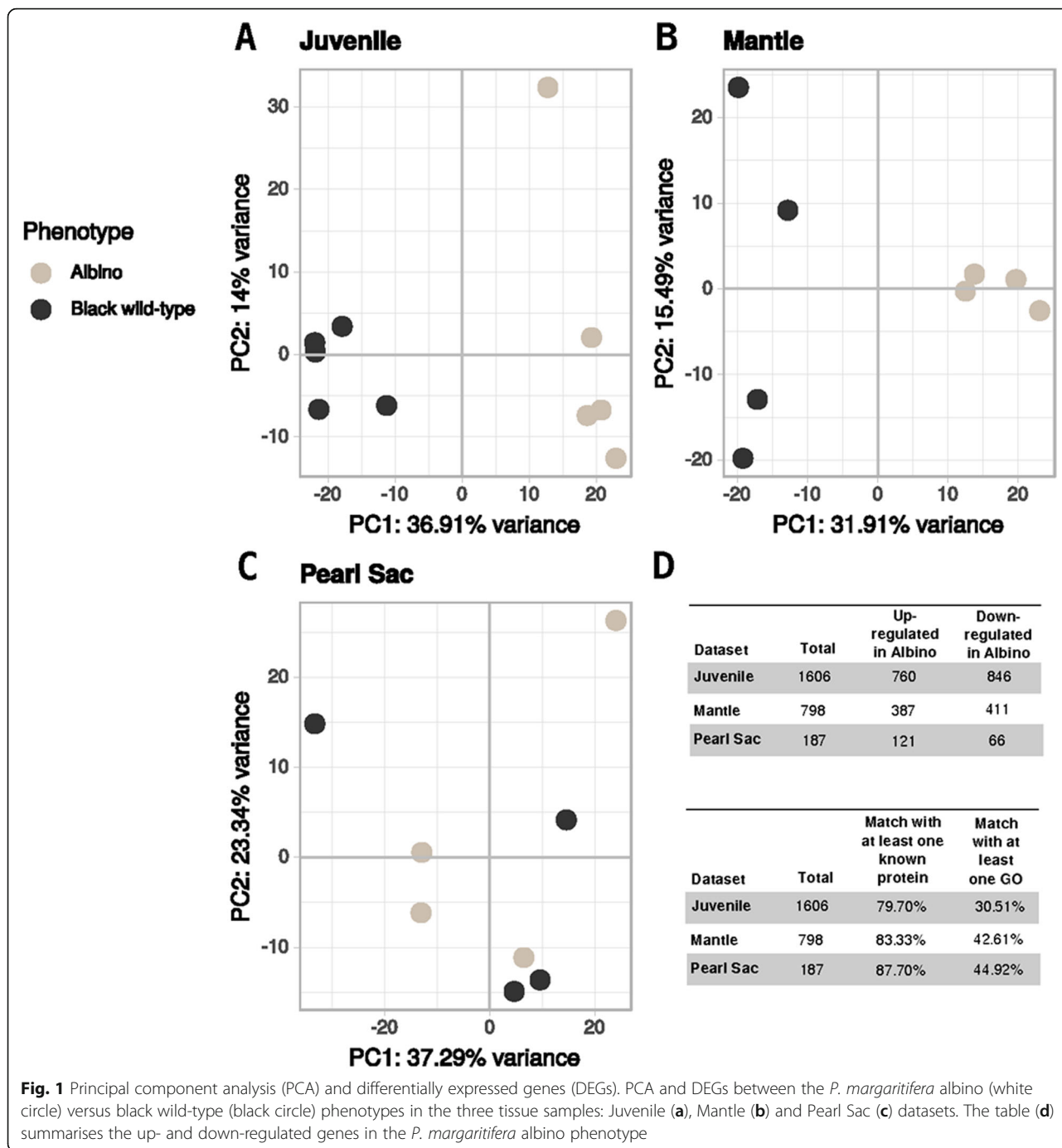
dataset. The principal component analysis (PCA) scatter-plots show how the samples clustered according to their transcriptomic profiles (Fig. 1). J and M datasets PCAs were able to discriminate the albino phenotype from black wild-type samples along the first principal component, accounting for 36.91% (J) and 31.91% (M) of the total variance. In the PS dataset PCA, no clustering of samples according to phenotype can be observed (Fig. 1). Out of the total number of DEGs, 79.70% (J), 83.33% (M) and 87.70% (PS) had at least one match with a known protein and 30.51% (J), 42.61% (M) and 44.92% (PS) had at least one associated GO (Additional file 1). The number of DEGs shared among the three datasets are displayed in an UpSet plot in Fig. 2. Interestingly, a high proportion of common DEGs in pairwise comparisons of the three datasets were not regulated in a consistent way: 62.82% (M & J), 10.26% (M & PS), 44.00% (J & PS) and 50.00% (M & J & PS). Only four DEGs overlapped the three datasets, of which two were consistently up-regulated in the albino phenotype and only one had a predicted function (endonuclease enzyme) in the available annotation (Additional file 2).

### DEG enrichment analysis

We performed gene ontology (GO) enrichment analysis on the identified DEGs to look into the main molecular mechanisms possibly involved in albinism in the different tissues (Additional file 3). When summarised separately into semantic-based clustering treemaps with the reduce visualize gene ontology tool (REVIGO), enriched GO in the J, M and PS datasets showed few or no overlapping Molecular Functions (MF). GO analysis based on down-regulated DEGs revealed some relevant MFs in M dataset such as ion transmembrane transporter activity, pigment binding, calcium ion binding (Additional file 4: Fig. S1). When considering common down-regulated DEGs between M and J, the GO analysis highlighted the Notch binding function (Additional file 4: Fig. S1). KEEG

**Table 1** General statistics of sequencing reads used in the three *P. margaritifera* RNA-seq datasets. Whole transcriptome sequencing libraries were generated on Illumina HiSeq4000 (paired-end (PE) sequencing  $2 \times 100$  bp) and mapped to the reference transcriptome of *P. margaritifera*. The three datasets came from three different tissue compartments. Dataset 'Juvenile': Juvenile stage dataset (albino phenotype  $N = 5$  and black wild-type phenotype  $N = 5$ ). Dataset 'Mantle': Adult stage, mantle tissue (albino phenotype  $N = 4$  and black wild-type phenotype  $N = 4$ ). Dataset 'Pearl Sac': Adult stage, pearl sac (albino phenotype  $N = 4$  and black wild-type phenotype  $N = 4$ )

Dataset	Juvenile	Mantle	Pearl Sac
PE raw reads ( $10^8$ )	11.66	4.90	5.61
PE trimmed reads ( $10^8$ )	11.37 (97.50%)	4.83 (98.58%)	5.41 (96.55%)
PE mapped reads ( $10^8$ )	8.59 (75.61%)	3.90 (80.81%)	4.02 (74.17%)
PE filtered mapped reads ( $10^8$ )	7.76 (90.33%)	3.55 (90.89%)	3.67 (91.27%)
Filtered counts ( $10^8$ )	3.86	1.76	1.82
Filtered SNPs	388,500	–	–
Splicing events	1366	313	–

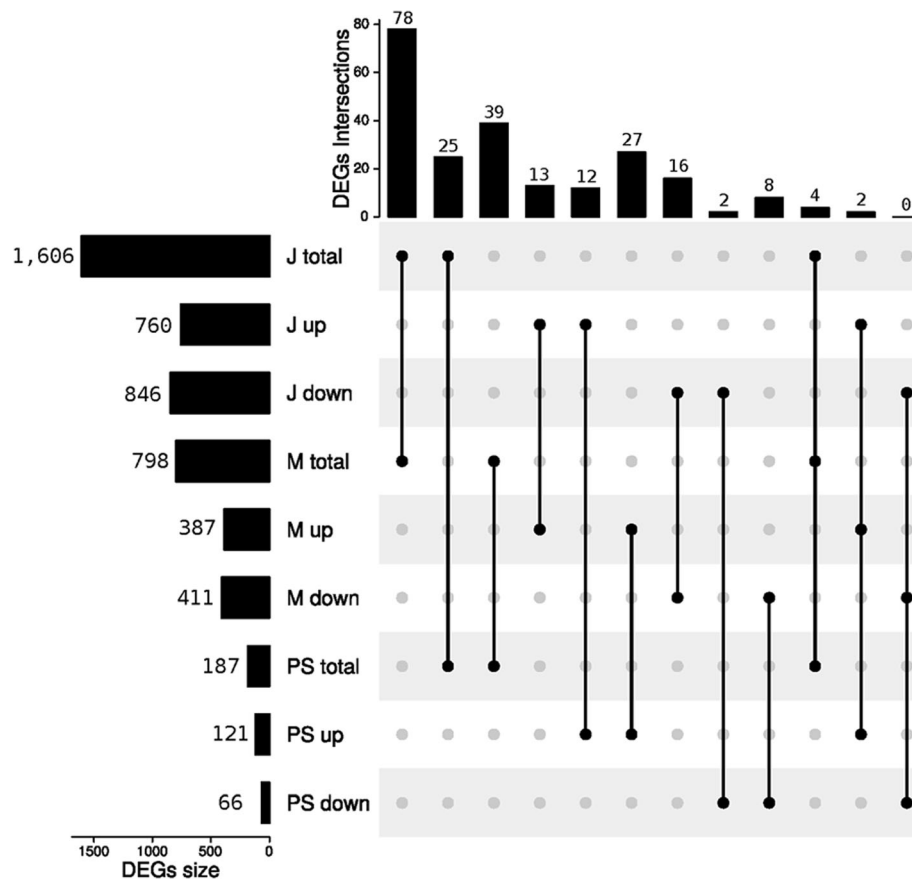


automatic annotation server (KAAS) results showed relevant Kyoto encyclopedia of genes and genomes (KEGG) pathways. Two DEGs in the J dataset were involved in ABC transporters, as was one DEG in the M dataset. We also noted some key genes encoding proteins involved in several important signalling pathways (sp), such as calmodulin (*k02183*), cell division control protein 42 (*cdc42*, *k04383*) and RTK (epidermal growth factor receptor, *k04361*), among others.

These are involved in RAS sp., Rap-1 sp., MAPK sp., calcium sp., melanogenesis (Additional file 4: Fig. S2). Notch sp. was also pinpointed in J and M (Additional file 4: Fig. S2).

**Identification of features of special interest**

From the annotated DEGs (Additional file 1), five molecular pathways involved in pigmentation were selected:



**Fig. 2** UpSet plot of differentially expressed genes (DEGs). The UpSet plot represents the intersections of DEGs between *P. margaritifera* albino versus black wild-type phenotype in three tissue compartments/datasets: juvenile (J), mantle (M) and pearl sac (PS). The number of total DEGs in the J, M and PS datasets are represented on the left (J total, M total, PS total) as are the number of up-regulated (J up, M up, PS up) and down-regulated genes (J down, M down, PS down) in the albino phenotype, as separate gene sets. The sizes of the intersections between these different sets are represented on the top barplot. The corresponding intersection is indicated by the connected dots

### Melanogenesis and Tyr genes encoding tyrosinase

Fifty-four tyrosinase-related transcripts were identified in *P. margaritifera* reference transcriptome, of which 31 were homologous to Tyrosinase-1 (*Tyr-1*), 13 were homologous to Tyrosinase-2 (*Tyr-2*) and 10 to Tyrosinase-3 (*Tyr-3*) (Additional file 5). No Tyrosinase-related genes were present in the PS DEGs dataset. In J samples, up-regulation of DEGs in the albino phenotype was observed for one *Tyr-1* ( $\log_2FC = 1.62$ ) and one *Tyr-2* ( $\log_2FC = 1.76$ ) homolog. In M tissue, up-regulation of DEGs in the albino phenotype was observed for one *Tyr-1* ( $\log_2FC = 3.32$ ) and one *Tyr-3* ( $\log_2FC = 2.00$ ) homolog, whereas down-regulated DEGs included four *Tyr-1* ( $\log_2FC$  ranging from  $-3.48$  to  $-2.26$ ) homologs, none of which corresponded to those DEGs found in the M dataset (Additional file 1). Moreover, in the M dataset, one P protein-related gene was down-regulated in the albino phenotype ( $\log_2FC = -2.13$ ).

### Calcium signalling pathway

Two calmodulin-like proteins were up-regulated in albino phenotype in the M dataset ( $\log_2FC = 3.22$  and  $2.49$ ). One was found down-regulated in the J dataset ( $\log_2FC = -1.95$ ).

### Notch signalling pathway

One Notch-related gene was found up-regulated in the albino phenotype in the J dataset ( $\log_2FC = 3.06$ ) and one Frizzled-related gene was down-regulated in the M dataset ( $\log_2FC = -4.59$ ).

### ABC transporter family

One *abcb* gene was found up-regulated in the albino phenotype in the M dataset ( $\log_2FC = -1.66$ ) and one down-regulated in the J dataset ( $\log_2FC = 2.15$ ). Two *abcc* genes were found down-regulated in the M dataset ( $\log_2FC = -1.55$  and  $\log_2FC = -3.01$ ) and one up-regulated in the J dataset ( $\log_2FC = -2.95$ ).



### Biom mineralization genes

One *Pif* gene was down-regulated in the M tissue ( $\log_2FC = -2.28$ ) and one shell protein was up-regulated in PS ( $\log_2FC = 2.70$  and  $2.54$ ). Two serine protease inhibitor genes were deregulated in both M ( $\log_2FC = -2.03$  and  $1.85$ ) and PS ( $\log_2FC = 3.85$ ). Three perlucin ( $\log_2FC$  from  $-2.92$  to  $3.62$ ) and four perlucin-like protein encoding genes ( $\log_2FC$  from  $-1253$  to  $2.97$ ) were deregulated in J; one perlucin ( $\log_2FC = 3.25$ ) and four perlucin-like protein encoding genes ( $\log_2FC$  from  $-4.16$  to  $-3.06$ ) in M; one perlucin-like protein ( $\log_2FC = 2.72$ ) in PS; one asparagin-rich protein (prism protein) in PS samples ( $\log_2FC = 2.70$ ); and one zinc metalloproteinase encoding gene in J tissue ( $\log_2FC = 2.86$ ) and two ( $\log_2FC = 5.97$  and  $-2.83$ ) in M samples.

### SNP outlier analysis

Population genetics analysis showed a strong family effect in the J and M datasets with maximal group assignment probability (Additional file 4: Fig. S3). When merged, variant calling files of J and M contained 388,500 SNPs located on 19,772 different transcripts. Of these, the lfm analysis returned 261 (0.07%) significant phenotype-associated SNPs located on 170 different transcripts. PCAdapt detected 50,083 (12.89%) outliers. There were 159 candidate SNPs, i.e., common SNPs among outliers detected with PCAdapt and phenotype-associated SNPs detected with lfm (Additional file 6). These were located on 115 different transcripts and 92 of them were annotated. Among these, two were non-synonymous substitutions affecting predicted proteins 'periostin' (Q29XZ0) and 'short-chain collagen C4' (P18503). Interestingly, two candidate SNPs were located on the *Pif* gene and were synonymous. There were two common transcripts between the concatenated M and J DEG set and the common outlier SNP set. Only one common transcript was annotated (probable RNA-directed DNA polymerase from transposon BS reverse transcriptase) in the reference transcriptome.

### Alternative gene splicing and exon usage

Significant differential splicing was observed for 1366 (J) and 313 (M) transcripts (adjusted  $p$  value  $< 0.01$ , Additional file 7). Of these, 1043 (J) and 251 (M) involved annotated transcripts. Some potentially interesting genes known to be involved in biom mineralization and/or pigmentation processes were identified, such as prism shell proteins, ABC transporter family members, calmodulin, collagen-alpha chain members, neurogenic locus notch homolog protein 1, PDZ, perlucin, *Pif*, putative Tyr-1, and Ras-related proteins in J; and ABCb transporter, one mantle protein encoding gene, *Pif*, Tyr-1 and zinc metalloproteinase in M. Differential splicing events were found to be deregulated for 21 (J) and 16 (M) transcripts, including Tyr-1 (Fig. 3)

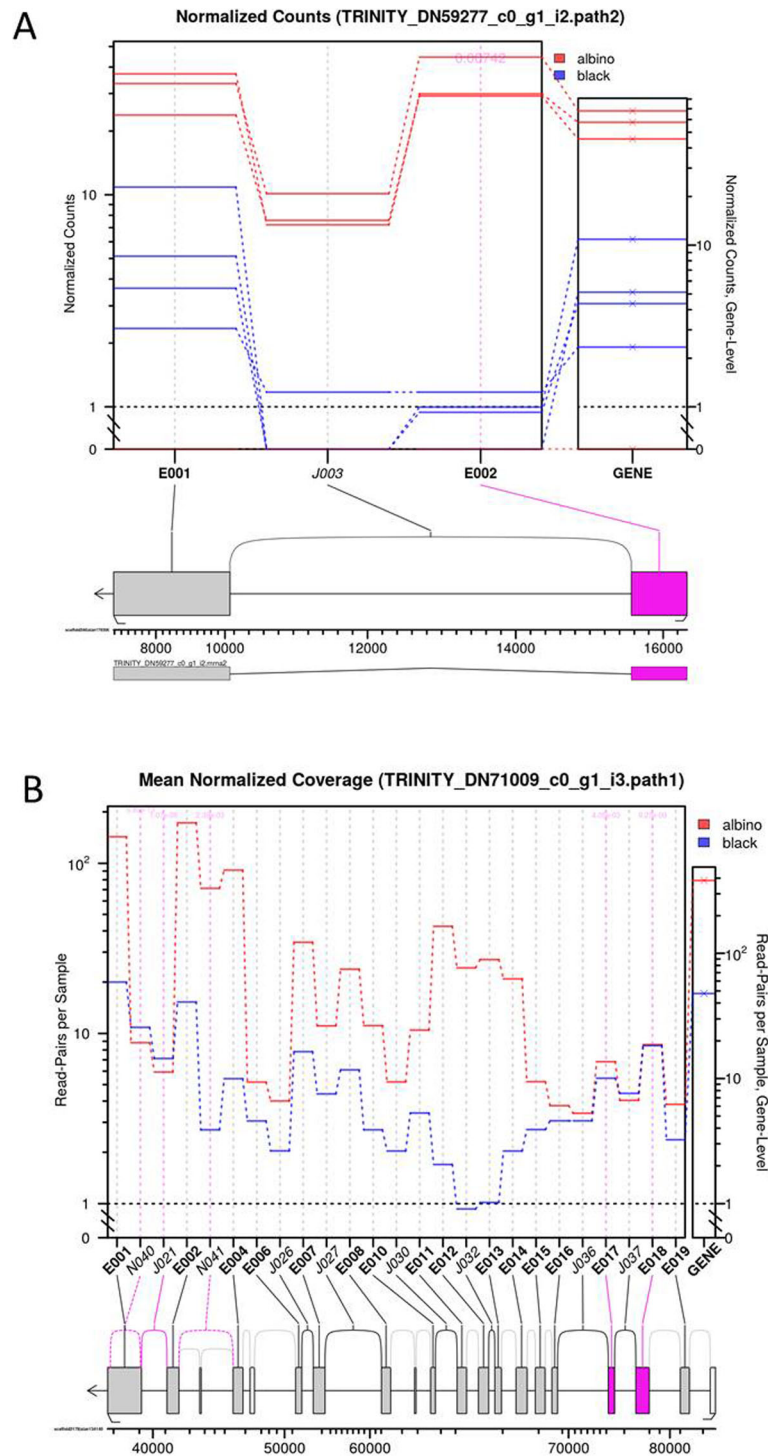
and Zinc metalloproteinase in M, and Notch in J (Fig. 3). There were 73 common transcripts between the J and M datasets showing significant differential splicing, among these, the *Pif* transcript carried two synonymous candidate SNPs.

### Discussion

Over the past decade, the advent of next-generation sequencing technologies has allowed significant progress to be made in understanding the genetic bases and molecular mechanisms involved in phenotypic trait expression in non-model species such as *P. margaritifera*. Previous studies used molecular markers or whole transcriptome to investigate pearl quality, growth, dynamics of nacre layer formation and biom mineralization [25–27]. In this study, we used RNA-seq to explore the molecular bases of the albino phenotype and to identify genes and molecular pathways associated with colour mutation across culture stages, from juvenile to adult, and tissue types (whole juvenile, mantle graft and pearl sac). We used independent multi-parental experiments to explore the genetic basis (SNPs), gene expression and alternative splicing underlying albinism in *P. margaritifera*. Our results show strong tissue and/or stage-specific responses, as suggested by the relatively low number of common DEGs across datasets. However, the related biological functions associated with the albino phenotype were redundant across J and M datasets and provide unprecedented information on the molecular effectors of albino pigmentation in *P. margaritifera*. Indeed, previous work investigating the genetic basis of rare flesh and shell colour variant polymorphism in *P. margaritifera* showed that shell colour is controlled by a three-allele model where albino is recessive to black wild type and not related to environmental effects [24]. Here, the use of three datasets from different tissues and representing different developmental stages adds value to the study, as the shell pigmentation of *P. margaritifera* is perceptible from the juvenile stage (three months old) [22] and can be related to pearl colour.

### The notch signalling pathway, involved in shell formation and pigmentation, was impacted in juveniles

The Notch signalling pathway is a highly conserved cell signalling system that regulates processes such as cell growth and proliferation, cell fate decisions, differentiation and stem cell maintenance [17]. The roles of the Notch signalling system have been well described in development [17] and mammalian hair pigmentation, where it is involved in the maintenance of melanoblasts [28]. Recently, it has been shown that the Notch signalling system is activated in a pearl oyster allograft (donor and recipient oysters from the same species) where it helps suppress the induced inflammation response [29].



**Fig. 3** Splicing event for two genes of *P. margaritifera*. Tyrosinase-1 in the mantle (a) and Notch homolog protein 1 in juvenile (b). Normalised counts are plotted for each gene section, either exon (E) or junction (J), and each individual (black wild-type in blue and albino in red). Values in the box plot represent *p*-values (Fisher’s test) for each gene section

In the clam *M. meretrix*, it was suggested that the Notch signalling pathway plays an important role in shell pigmentation following a gene dosage-dependent pattern and was also potentially involved in the shell colour

patterning [15]. The Notch signalling pathway has also been reported to play a role in the pigmentation process in Pacific oyster *C. gigas*, where it was proposed that white-shelled oysters used endocytosis to down-regulate

Notch level, thus causing melanoblast apoptosis, which prevents pigmentation [13]. Finally, in the sea cucumber *Apostichopus japonicus*, the Notch signalling pathway might be an upstream component of the pigmentation process by determining the location and boundaries of pigment occurrence [18]. Here we showed that the neurogenic locus Notch homolog protein 1 (Notch1), is up-regulated in albino Juvenile phenotype *P. margaritifera* individuals. These results strongly support the early role of Notch signalling regulation, which most likely results in the inhibition of shell pigmentation. Interestingly, in *Crassostrea gigas*, authors found Notch2, but not Notch1, to be down-regulated in white-shelled oysters [13]. Notch1 and Notch2 are the two closest related Notch paralogs with opposite biological functions to each other in mammals [30]. We further show differential exon usage of the same Notch1 gene, but no related SNPs in coding regions. Putative splice variants of Notch were also identified in the butterfly *Vanessa cardui* and suggested to be part of a conserved Notch signalling cassette. Certain of these genes were down-regulated during pupal stages, and their activation during the pre-mochrome stage might play a role in pattern positional information on the wing [31]. Admittedly, further studies will be needed to assess the precise role of alternative splicing in the Notch signalling pathway and any subsequent impact on shell pigmentation in *P. margaritifera*. In bivalves, shell construction starts early in larval development and differences in pigmentation can already be seen at the juvenile stage (three months old) [22]. Previous studies on shell pigmentation in *P. margaritifera* have, however, only focused on adult specimens. Here, the inclusion of juvenile specimens in our design allowed us to see the specific impact of the Notch signalling pathway in earlier shell formation and pigmentation.

#### **Adult mantle and juvenile transcriptome analysis highlights the association of key genes involved in melanogenesis with the *P. margaritifera* albino phenotype**

Tyrosinase is a key copper-containing enzyme known to be involved in shell structuration and pigmentation in bivalve molluscs through a large spectrum of biological processes, including pigment production, innate immunity, wound healing and exoskeleton building and hardening [32]. Tyrosinases are, for example, essential in the melanogenesis pathway as key regulators of melanin biosynthesis [33]. Tyrosinase deregulation has been reported in the bivalves *C. gigas*, *P. yessoensis* and *P. fucata martensii*, which show contrasting pigmentation phenotypes [13, 34, 35]. In bivalve molluscs, mantle tissue secretes proteins responsible for shell calcification and pigmentation, and pigments are probably formed in the secretory cells in the mantle and incorporated into

the shell along the growing edge [35]. It is thus not surprising that expression of Tyrosinase genes in *P. margaritifera* occurs predominantly in the mantle tissue [32]. In this study, no *Tyr-2* was found among the mantle DEGs, which is consistent with a previous study on *P. margaritifera* [14]. *Tyr-1*, however, showed isoform-specific regulation in mantle (one up-regulated and four down-regulated). We identified 54 tyrosinase or tyrosinase-like protein encoding genes in *P. margaritifera* reference transcriptome [25]. A complex evolutionary history has been described for tyrosinase genes in bivalves, in which tyrosinase duplicates may have been retained because of their functional diversification in the mantle [32]. Further studies will, therefore, be needed to examine more closely any isoform-specific function/s of the *Tyr* gene family in *Pinctada* sp. We also observed down-regulation of the P protein-encoding gene, a key gene involved in melanogenesis, in the albino phenotype mantle of *P. margaritifera*. P protein, known as the pink-eyed dilution protein encoded by the *OCA2* gene, could be involved in the transport of tyrosine, the precursor to melanin synthesis, within human melanocytes. Mutation in this protein may cause complete or partial albinism [36]. To our knowledge, this is the first report of such a gene in bivalves. Moreover, *frizzled-3* was also found to be down-regulated in the mantle of the *P. margaritifera* albino phenotype. Frizzled is an essential receptor associated with the Wnt signalling pathway, which is involved in the maintenance of melanocyte and keratinocyte homeostasis [18]. Our result is consistent with a previous study in zebrafish where inhibition of the Wnt signalling pathway decreased pigment cells [37]. We also observed deregulation of members of the ATP-binding cassette (ABC) transporter family, which is a large transporter family related to cellular detoxification, observed in all kingdoms of life [13, 38]. It has notably been detected in numerous marine species, including bivalves. In oysters, the involvement of *Abcb1* has been hypothesised in a molecular detoxification system through the transport of a variety of molecules and xenobiotics across the cell membrane [39]. Interestingly, in *C. gigas*, ATP-binding cassette members might be responsible for the transportation of pigment-related substrate that influences pigmentation [13]. Moreover, in human, calcium signalling plays a role in the co-regulation of retinal pigment epithelial cells [40]. In the present study, we showed that the calcium signalling pathway, especially *Calmodulin* gene expression, is associated with the albino phenotype in adult and juvenile *P. margaritifera*. Calmodulin and calmodulin-like proteins are involved in biomineralization and melanogenesis in bivalve molluscs [27]. Our results thus suggest that the albino white shell phenotype in *P. margaritifera* may result from dysfunction of melanin transfer from the mantle to the shell.



### Multiple transcriptome analysis confirms that shell pigmentation requires essential biomineralization genes

A previous study investigated the role of a subset of genes in shell biomineralization and pigmentation of *P. margaritifera* by comparing black wild-type and albino individuals [14]. Our present results are concordant for key biomineralization genes, including *Pif* in mantle and shell matrix proteins in pearl sac, but others are deregulated inconsistently, like serine protease inhibitor in mantle and pearl sac. We also report the deregulation of genes for perlucin and perlucin-like protein in pearl sac, both of which are involved in shell biomineralization [41]. The *Pif* gene codes for an important acidic matrix protein known to be involved in nacre formation and has a complex role in biomineralization [42]. Here, *Pif* shows significant exon usage in both J and M datasets, and also carries two significant synonymous phenotype-associated SNPs. These results suggest a tight association of *Pif* with the albino phenotype in M and J, but the elucidation of its precise role in *P. margaritifera* biomineralization, which has already been shown to be complex [25], would require further investigation. Finally, we observe limited albino specific signature of in PS tissue genes expression. These results possibly suggest that PS response is significantly affected by the PS environment, *i.e.* recipient oyster activity, supporting previous elegant work based on xenograft experiments and final pearl quality [43].

An understanding of the molecular and genetic basis of the *P. margaritifera* albino phenotype would be a valuable resource for its aquaculture in French Polynesia. Indeed, Polynesian pearl production is currently and mainly focused on dark-coloured pearls. Developing the production of lighter-coloured pearls could open new market opportunities for the Tahitian pearl industry. In fact, albino donor grafts into wild-type recipients produce light silver, white or light cream cultured pearls, with some interesting pinkish and blue reflected lustre. As albino pearl oysters are very rare in French Polynesia, setting up selection programs will require reliable genetic markers (not only limited to coding-regions) to monitor the level of inbreeding throughout production.

### Conclusions

This study was designed to identify key genes associated with the *P. margaritifera* albino phenotype, from juvenile to cultured pearl harvest stages, using comparative transcriptome analysis with black-shelled wild-type individuals. Results of the transcriptome analysis in three different datasets: juvenile whole flesh (J), adult mantle (M) and pearl sac (PS) revealed quite specific transcriptomic signatures with some related functions. We successfully identified a set of genes involved in pigmentation, including several genes not previously

described in pearl oyster *P. margaritifera*. The study also showed the putative association of significantly different exon usage and SNPs in the albino phenotype. For the first time in *P. margaritifera*, the involvement of the Notch signalling pathway in pigmentation was highlighted. *Notch1* was specifically deregulated and showed alternative splicing in the J dataset, suggesting an early role of this gene in shell formation and pigmentation. We confirmed the association of genes involved in melanin biosynthesis and transport, such as calmodulin and tyrosinase, with putative associated alternative splicing for the latter. We also discussed the possible involvement of P protein, hitherto only described in mammal albino phenotypes. This analysis also corroborates previous studies by calling into question the association of biomineralization genes like *Pif*, which carried two synonymous phenotype-associated SNPs. Future studies should further investigate the role of alternative splicing and SNPs using a complete annotated genome (presently under construction). The potential involvement of non-coding RNA in shell pigmentation should also be explored in *P. margaritifera*. Similarly, complementary approaches based on metabolites secretion would provide further evidence and validate regulatory network related to pigment synthesis in oyster.

### Methods

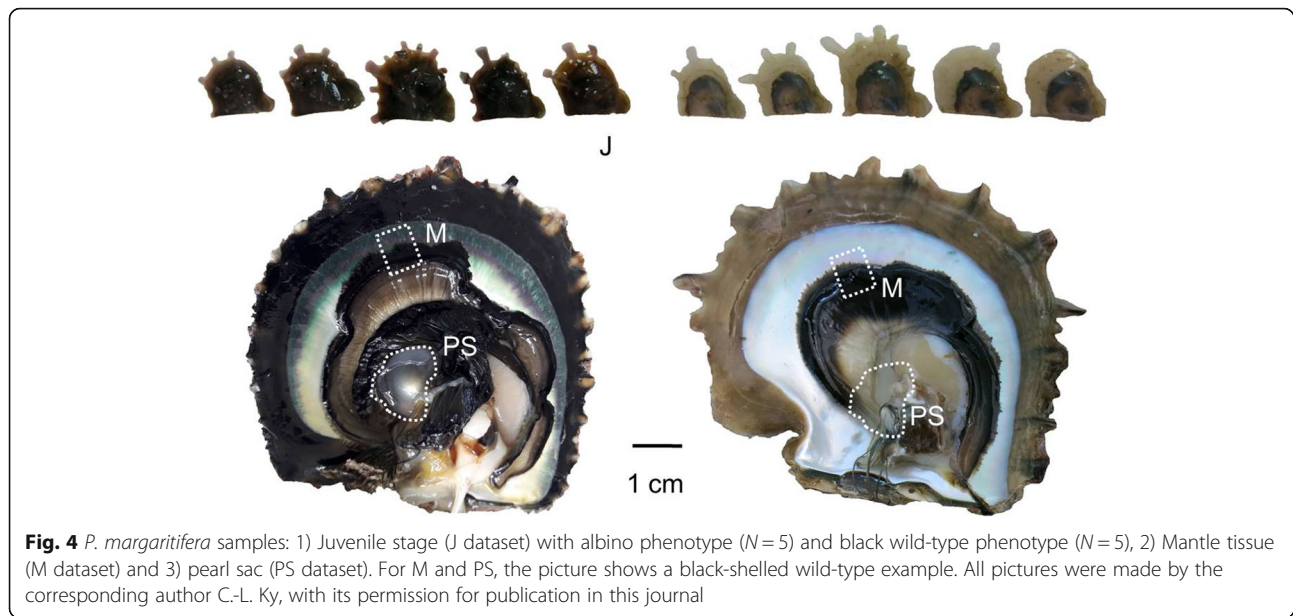
#### Animal tissue sampling

Three datasets were obtained from different tissue compartments, comparing two families with white albino and black wild-type phenotypes at three key stages of pearl production/ culture. These two families were both F1, produced simultaneously in a hatchery system at Ifremer facilities, following routine rearing procedures described previously [44, 45]. Juveniles and adults were reared according following standard procedures for *P. margaritifera* farming [46]. The F1 families were produced from multi-parental crosses made with wild broodstock originating from Takume atoll (Tuamotu archipelago, French Polynesia). After sampling, all tissues were preserved in ribonucleic acid (RNA) later (Qiagen) and kept at  $-80^{\circ}\text{C}$  until RNA extraction. The three samplings from the juvenile stage until pearl harvest were made over a four-year period encompassing the whole pearl culture cycle.

Dataset 1 Juvenile stage (J) (Fig. 4). *P. margaritifera* juveniles of 4.5 months were sampled (albino phenotype  $N=5$ ; black wild-type phenotype  $N=5$ ). Tissue samples consisted of whole soft tissues, *i.e.*, entire animals.

Dataset 2 Adult stage, mantle tissue (M) (Fig. 4). Tissue samples consisted of fragments of mantle (albino phenotype  $N=4$ ; black wild-type phenotype  $N=4$ ).

Dataset 3 Adult stage, pearl sac (PS) (Fig. 4). Eighty months before this sampling, recipient oysters had been



**Fig. 4** *P. margaritifera* samples: 1) Juvenile stage (J dataset) with albino phenotype ( $N=5$ ) and black wild-type phenotype ( $N=5$ ), 2) Mantle tissue (M dataset) and 3) pearl sac (PS dataset). For M and PS, the picture shows a black-shelled wild-type example. All pictures were made by the corresponding author C.-L. Ky, with its permission for publication in this journal

grafted using a set of white albino and black wild-type donors using a classical experimental graft designed as previously described [47]. At pearl harvest, pearl sacs were sampled to provide tissue for this dataset (albino phenotype  $N=4$  and black wild-type phenotype  $N=4$ ).

#### RNA extraction and sequencing

For all samples, 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 (NanoDrop Technologies Inc.) and a BioAnalyzer 2100 (Agilent Technologies), respectively. RNA was dried in RNA-stable solution (ThermoFisher Scientific) following manufacturer's recommendations. Samples were shipped at room temperature to McGill sequencing platform services (Montreal, Canada). TruSeq Sample Prep. (Illumina, San Diego, California CA, USA) RNA libraries were sequenced on a HiSeq 4000 100-bp paired-end (PE) sequencing device.

#### Differential expression analysis

Read quality was assessed with fastqc v0.11.5 [47] and multiQC [48]. Raw reads were filtered to remove sequencing adapters and for quality trimming ( $Q=28$ ) using: (i) Cutadapt v1.13 [49] for the M dataset and (ii) Trimmomatic v0.36 [50] for the PS and J datasets. For each dataset, only surviving paired-end reads were retained. Filtered reads were mapped on the multi-tissue reference transcriptome of *Pinctada margaritifera* (41,075 contigs) [25] using bwa v0.7.15 [51] with standard parameters. Reads with low mapping quality ( $Q \geq 5$ ), mispairing or multi-mapping were removed using

Samtools v1.4.1 [52]. A matrix of raw counts was built using HTSeq-count v0.6.1 [53]. To minimise the false-positive rate, count matrices were filtered for little expressed transcripts. For each dataset, all transcripts with less than 10 counts in at least two samples were discarded. We identified differentially expressed genes (DEGs) between albino and black wild-type individuals using DESeq2 v1.16.2 [54] with R v3.4.0 (<https://www.R-project.org/>) following the standard workflow. The DESeq2 method internally corrects for library size and uses negative binomial generalised linear models to test for differential expression. In this study, the statistical models were built using 'counts ~ Phenotype' design formula, where the Phenotype qualitative variable indicates oyster phenotype (albino/black wild-type). All features with absolute  $\log_2$  fold change greater than 1.5 and adjusted  $p$ -value smaller than 0.05 (Benjamini-Hochberg method) were reported as differentially expressed (DEGs). Overlapping DEGs between the three datasets were visualised using the UpSet function from ComplexHeatmap R package version 2.0.0 [55]. For this study, prior genes annotation with blastx (e-value  $10^{-5}$  [25];) against Swissprot-Uniprot, has been supplemented with searches against conserved protein domain databases. First, longest Open Reading Frame (ORF; min. Length 100 amino acid) for each gene was computed with TransDecoder v5.5.0 (<https://github.com/TransDecoder/TransDecoder>) tools. ORF where scan for protein conserved domains against CDD, TIGRFAM, PRINTS, SMART, SUPERFAMILY and Hamap databases using InterProScan v5.28–67.0 [56] and GO were retrieved using interPro2GO database (version 2020/04/18; <http://www.geneontology.org/external2go/interpro2go>). This search resulted in a total of 16,463

matches; out of which 9360 had an associated GO entry and 932 had no prior GO associated using Swissprot only. To test the overrepresentation of gene ontology (GO) terms in resulting DEGs, we used Goatools v0.8.4 [57] through the ‘go\_enrichment’ pipeline ([https://github.com/enormandeu/go\\_enrichment](https://github.com/enormandeu/go_enrichment)) with the go-basic.obo database (release 2018-05-07). The resulting lists of significant GO enriched terms were filtered for Fisher’s Test  $p$ -value  $< 0.05$  and used for semantic-based clustering using REVI GO with an allowed similarity of 0.5 [58]. DEG lists were submitted to the KAAS server ([59], last updated: April 3, 2015) to visualise related KEGG pathways.

### Population genetics analysis

In the J and M datasets, we investigated sites that were variable between the white albino and black wild-type populations. We did not include the PS dataset in this part of the study because of possible contamination with recipient RNA during pearl sac sampling. Single nucleotide polymorphisms (SNPs) were called from pre-processed aligned reads using Freebayes v1.1.0–3-g961e5f3 [60] with a required minimum mapping quality of 20. Pre-processing of aligned reads included marking and removing duplicates, correcting N cigar reads, sorting and indexing bam files using gatk v4.0.2.1 [61] and the Picard tools suite v1.119 [62]. Resulting variant calling files (VCF) were filtered for missing data and indels (none authorised), allele frequency ( $\geq 0.1$ ) and depth ( $\geq 20$ ) using vcftools v0.1.14 [63]. Filtered VCF for J and M were merged to minimise family effects and focus on phenotype-associated events. To investigate the structure of albino and black wild-type individuals within the two populations, we performed a population genetics analysis on the filtered VCF files using the following R packages: vcfr v1.8.0 [64], adegenet v2.1.1 [65] and genepop v1.0.5 [66] with R v3.4.0. We then looked for outlier SNPs (adjusted  $p$  value  $< 0.01$ , Benjamini-Hochberg procedure) with PCA-dapt v4.0.3 [67] PCAdapt tests for outliers using correlations between SNPs and first principal components (K) of PCA analysis. We used  $K = 2$  (J and M separately) and  $K = 3$  (M and J merged) and adjusted  $p$ -value  $< 0.01$ , according to the Benjamini-Hochberg procedure. We also performed phenotype-associated SNP analysis using lfmm v0.0 (<https://bcm-uga.github.io/lfmm/index.html>) to find potential genetic markers of albinism. The Lfmm program constructs latent factor mixed models (LFMMs), which are statistical regression models to test associations between a multidimensional set of response variables (here, genotypes) and a variable of interest (here, phenotype). LFMMs include unobserved variables, called latent factors, which correct the model for confounding effects due to population structure and other hidden causes. We selected phenotype-associated SNPs according to adjusted  $p$  values  $< 0.01$ . Phenotype-associated SNPs were annotated

using the ‘LongOrfs’ function implemented in Transdecoder v3.0.1 (<https://github.com/TransDecoder/TransDecoder/wiki>) and personal python script.

### Alternative gene splicing and exon usage

To detect differential splicing events in the three RNA-seq datasets, the filtered reads were mapped on the draft genome of *P. margaritifera* using GSNAP aligner v2017-03-17 [68], allowing five mismatches, splicing and using the ‘splitting-output’ function to retain only concordant and unique mapped paired-end reads, as described previously [25]. We used the QORTs [69] and JunctionSeq R packages [70] to detect significant differences in exon usage. Only exons and junctions with a minimal coverage of six were used for the analysis and only differences with FDR  $< 0.01$  were considered significant.

### Supplementary information

**Supplementary information** accompanies this paper at <https://doi.org/10.1186/s12864-020-07015-w>.

**Additional file 1.** Lists of DEGs between the *P. margaritifera* albino phenotype and black wild-type phenotype, associated statistics and annotation for the three datasets (Juvenile, Mantle and Pearl Sac).

**Additional file 2.** Lists of common DEGs between the *P. margaritifera* albino phenotype and black wild-type phenotype overlapping the three datasets (Juvenile, Mantle and Pearl Sac).

**Additional file 3.** Lists of enriched GO terms in the *P. margaritifera* albino phenotype compared with the black wild-type phenotype in the three datasets (Juvenile, Mantle and Pearl Sac).

**Additional file 4:** (docx) **Fig. S1.** Summarised REVIGO treemaps plot for gene ontology enrichment analysis between the *P. margaritifera* albino versus black wild-type phenotypes. **Fig. S2.** Signalling pathways (sp) potentially impacted by genes deregulated in albino *P. margaritifera* compared with the black wild-type. **Fig. S3.** Composition plots of albino and black wild-type populations of pearl oyster *P. margaritifera* based on filtered SNPs called with Freebayes in both Juvenile and Mantle datasets.

**Additional file 5.** List of tyrosinase-related transcripts in *P. margaritifera* reference transcriptome.

**Additional file 6.** List of 159 candidate SNPs, i.e., common SNPs among outliers detected with PCAdapt and phenotype-associated SNPs detected with lfmm, based on filtered SNPs called with Freebayes in both Juvenile and Mantle datasets.

**Additional file 7.** List of transcripts showing significant differential splicing (adjusted  $p$  value  $< 0.01$ ) in Juvenile and Mantle datasets.

### Abbreviations

DEGs: differentially expressed genes; DNA: Deoxyribonucleic acid; GO: gene ontology; KAAS: KEGG automatic annotation server; KEGG: Kyoto encyclopedia of genes and genomes; MF: Molecular Functions; NCBI: National Center for Biotechnology; PCA: principal component analysis; PE: Paired-end; REVIGO: reduce visualize gene ontology tool; RNA: Ribonucleic acid; RNA-seq: Ribonucleic acid sequencing; SNP: Single nucleotide polymorphism

### Acknowledgements

The authors would especially like to thank the host sites, Heimoana Poe Pearl Farm located in Raroia atoll (Tuamotu archipelago, French Polynesia) and Pahai Poe pearl farm located in Apataki atoll (Tuamotu archipelago, French Polynesia), for their generous support. The authors are also grateful to the Ifremer bioinformatics platform (Datarmor; SeBiMER) and *Station Biologique de Roscoff*



(ABiMS; <http://abims.sb-roscoff.fr>) for providing computing and storage resources, and to Léo Milhade for programming help and support.

#### Authors' contributions

CLK conceived the study. VQ conducted the molecular analysis. CLK and MSK performed the artificial crossbreeding and MSK reared the albino cohorts. PA and JLL analysed the RNAseq data. PA, JLL and CLK wrote the paper. All co-authors contributed to reviewing the manuscript and accepted the final version for publication. All authors read and approved the final manuscript.

#### Funding

This study was supported by grants from the *Direction des Ressources Marines*, through the ColoGEN-2 project (# 007110/VP/DRM). The role of the funders consisted on a financial support for all experiment realisations.

#### Availability of data and materials

The data for the genome of *P. margaritifera* is available following the direct web link: <https://sextant.ifremer.fr/eng/Donnees/Catalogue#/metadata/0d2dda11-db53-43f5-b0d4-1acb2ca0bdc4>. The full name of the data is: "Draft genome assembly of *Pinctada margaritifera*".

For the multi-tissue reference transcriptome of *Pinctada margaritifera*, the Transcriptome Shotgun Assembly project has been deposited at DDBJ/EMBL/GenBank under the accession GIUE00000000. The full names of the data are: 1) SUBID: SUB7928927, 2).

BioProject: PRJNA449941, 3) BioSample: SAMN02414627, SAMN02414631, SAMN10915869, SAMN10915871, SAMN10915877, SAMN15396325, SAMN15396326, SAMN15396328, 4).

Accession: GIUE00000000 and 5) Organism: *Pinctada margaritifera*.

All codes for RNA-seq analysis are available upon request.

#### Ethics approval and consent to participate

The study was approved by the ethics committee of the French institute of research and exploitation of the sea (Ifremer). The collection of *P. margaritifera* followed institutional and national guidelines.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup>Ifremer, UMR EIO 241, Centre du Pacifique, BP 49, 98719 Taravao, Tahiti, Polynésie française, France. <sup>2</sup>Ifremer, UMR LEMAR UBO CNRS Ifremer IRD 6539, ZI Pointe Diable CS 10070, F-29280 Plouzane, France. <sup>3</sup>IHPE, Université de Montpellier, CNRS, Ifremer, Université de Perpignan Via Domitia, F-34090 Montpellier, France.

Received: 28 April 2020 Accepted: 20 August 2020

Published online: 25 September 2020

#### References

- Hubbard JK, Uy JAC, Hauber ME, Hoekstra HE, Safran RJ. Vertebrate pigmentation: from underlying genes to adaptive function. *Trends Genet TIG*. 2010;26:231–9.
- Hillsdorf AWS, Penman DJ, Farias EC, McAndrew B. Melanophore appearance in wild and red tilapia embryos. *Pigment Cell Res*. 2002;15:57–61.
- Bagnara JT, Fernandez PJ, Fujii R. On the blue coloration of vertebrates. *Pigment Cell Res*. 2007;20:14–26.
- Dobosz S, Kohlmann K, Goryczko K, Kuzminski H. Growth and vitality in yellow forms of rainbow trout. *J Appl Ichthyol*. 2000;16:117–20.
- Zou M, Zhang X, Shi Z, Lin L, Ouyang G, Zhang G, et al. A comparative Transcriptome analysis between wild and albino yellow catfish (*Pelteobagrus fulvidraco*). *PLoS One*. 2015;10:e0131504.
- Gong Y, Hu M, Xu S, Wang B, Wang C, Mu X, et al. Comparative transcriptome analysis reveals expression signatures of albino Russian sturgeon, *Acipenseriformes gueldenstaedtii*. *Mar Genomics*. 2019;46:1–7.
- Wada KT, Komaru A. Color and weight of pearls produced by grafting the mantle tissue from a selected population for white shell color of the Japanese pearl oyster *Pinctada fucata martensii* (dunker). *Aquaculture*. 1996;142:25–32.
- Zou K, Zhang D, Guo H, Zhu C, Li M, Jiang S. A preliminary study for identification of candidate AFLP markers under artificial selection for shell color in pearl oyster *Pinctada fucata*. *Gene*. 2014;542:8–15.
- Jm B. H P. expression of family differences through within-lot competition in juvenile rainbow trout *Oncorhynchus mykiss*. *J World Aquac Soc*. 2003;34:425–32.
- Thorgaard GH, Spruell P, Wheeler PA, Scheerer PD, Peek AS, Valentine JJ, et al. Incidence of albinos as a monitor for induced triploidy in rainbow trout. *Aquaculture*. 1995;137:121–30.
- Boonanuntanasarn S, Yoshizaki G, Iwai K, Takeuchi T. Molecular cloning, gene expression in albino mutants and gene knockdown studies of tyrosinase mRNA in rainbow trout. *Pigment Cell Res*. 2004;17:413–21.
- Edvardsen RB, Leininger S, Kleppe L, Skaftnesmo KO, Wargelius A. Targeted mutagenesis in Atlantic salmon (*Salmo salar* L.) using the CRISPR/Cas9 system induces complete knockout individuals in the F0 generation. *PLoS One*. 2014;9:e108622.
- Feng D, Li Q, Yu H, Zhao X, Kong L. Comparative Transcriptome analysis of the Pacific oyster *Crassostrea gigas* characterized by Shell colors: identification of genetic bases potentially involved in pigmentation. *PLoS One*. 2015;10:e0145257.
- Lemer S, Saulnier D, Gueguen Y, Planes S. Identification of genes associated with shell color in the black-lipped pearl oyster, *Pinctada margaritifera* *BMC Genomics* 2015;16:568.
- Yue X, Nie Q, Xiao G, Liu B. Transcriptome analysis of shell color-related genes in the clam *Meretrix meretrix*. *Mar Biotechnol N Y N*. 2015;17:364–74.
- Nagai K, Yano M, Morimoto K, Miyamoto H. Tyrosinase localization in mollusc shells. *Comp Biochem Physiol B Biochem Mol Biol*. 2007;146:207–14.
- Artavanis-Tsakonas S, Rand MD, Lake RJ. Notch signaling: cell fate control and signal integration in development. *Science*. 1999;284:770–6.
- Xing L, Sun L, Liu S, Li X, Zhang L, Yang H. De novo assembly and comparative transcriptome analyses of purple and green morphs of *Apostichopus japonicus* during body wall pigmentation process. *Comp Biochem Physiol Part D Genomics Proteomics*. 2018;28:151–61.
- Ky C-L, Broustal F, Potin D, Lo C. The pearl oyster (*Pinctada margaritifera*) aquaculture in French Polynesia and the indirect impact of long-distance transfers and collection-culture site combinations on pearl quality traits. *Aquac Rep*. 2019;13:100182.
- Wada K. Formation and quality of pearls. *J Gemmol Soc Jpn*. 1999;20:47–62.
- 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.
- Ky C-L, Lo C, Planes S. Mono- and polychromatic inner shell phenotype diversity in *Pinctada margaritifera* donor pearl oysters and its relation with cultured pearl colour. *Aquaculture*. 2017;468 Part.1:199–205.
- Ky C-L, Koua M, Le Moullac G. Impact of spat shell colour selection in hatchery-produced *Pinctada margaritifera* on cultured pearl colour. *Aquac Rep*. 2018;9:62–7.
- Ky C-L, Nakasai S, Pommier S, Sham Koua M, Devaux D. The Mendelian inheritance of rare flesh and shell colour variants in the black-lipped pearl oyster (*Pinctada margaritifera*). *Anim Genet*. 2016;47:610–4.
- Le Luyer J, Auffret P, Quillien V, Leclerc N, Reisser C, Vidal-Dupui J, et al. Whole transcriptome sequencing and biomineralization gene architecture associated with cultured pearl quality traits in the pearl oyster, *Pinctada margaritifera*. *BMC Genomics*. 2019;20:111.
- Joubert C, Piquemal D, Marie B, Manchon L, Pierrat F, Zanella-Cléon I, et al. Transcriptome and proteome analysis of *Pinctada margaritifera* calcifying mantle and shell: focus on biomineralization. *BMC Genomics*. 2010;11:613.
- Blay C, Planes S, Ky C-L. Cultured pearl surface quality profiling by the Shell matrix protein gene expression in the Biomineralised pearl sac tissue of *Pinctada margaritifera*. *Mar Biotechnol N Y N*. 2018;20:490–501.
- Moriyama M, Osawa M, Mak S-S, Ohtsuka T, Yamamoto N, Han H, et al. Notch signaling via Hes1 transcription factor maintains survival of melanoblasts and melanocyte stem cells. *J Cell Biol*. 2006;173:333–9.
- Jiao Y, Yang S, Cao Y, Zheng Z, Deng Y, Wang Q, et al. Genome and transcriptome analyses providing insight into the immune response of pearl oysters after allograft and xenograft transplantations. *Fish Shellfish Immunol*. 2019;90:109–17.
- Liu Z, Brunskill E, Varnum-Finney B, Zhang C, Zhang A, Jay PY, et al. The intracellular domains of Notch1 and Notch2 are functionally equivalent during development and carcinogenesis. *Dev Camb Engl*. 2015;142:2452–63.

31. Connahs H, Rhen T, Simmons RB. Transcriptome analysis of the painted lady butterfly, *Vanessa cardui* during wing color pattern development. *BMC Genomics*. 2016;17:270.
32. Aguilera F, McDougall C, Degnan BM. Evolution of the tyrosinase gene family in bivalve molluscs: independent expansion of the mantle gene repertoire. *Acta Biomater*. 2014;10:3855–65.
33. D'Mello SAN, Finlay GJ, Baguley BC, Askarian-Amiri ME. Signaling pathways in *Melanogenesis*. *Int J Mol Sci*. 2016;17:1144.
34. Xu M, Huang J, Shi Y, Zhang H, He M. Comparative transcriptomic and proteomic analysis of yellow shell and black shell pearl oysters, *Pinctada fucata martensii*. *BMC Genomics*. 2019;20:469.
35. Sun X, Yang A, Wu B, Zhou L, Liu Z. Characterization of the mantle transcriptome of yesso scallop (*Patiniopecten yessoensis*): identification of genes potentially involved in biomineralization and pigmentation. *PLoS One*. 2015;10:e0122967.
36. Park S, Morya VK, Nguyen DH, Singh BK, Lee H-B, Kim E-K. Unrevealing the role of P-protein on melanosome biology and structure, using siRNA-mediated down regulation of OCA2. *Mol Cell Biochem*. 2015;403:61–71.
37. Dorsky RI, Moon RT, Raible DW. Control of neural crest cell fate by the Wnt signalling pathway. *Nature*. 1998;396:370–3.
38. Dermauw W, Van Leeuwen T. The ABC gene family in arthropods: comparative genomics and role in insecticide transport and resistance. *Insect Biochem Mol Biol*. 2014;45:89–110.
39. Marques-Santos LF, Hégaret H, Lima-Santos L, Queiroga FR, da Silva PM. ABCB1 and ABCG1-like transporters in immune system cells from sea urchins *Echinometra lucunter* and *Echinus esculentus* and oysters *Crassostrea gasar* and *Crassostrea gigas*. *Fish Shellfish Immunol*. 2017;70:195–203.
40. Smith-Thomas L, Haycock JW, Metcalfe R, Boulton M, Ellis S, Rennie IG, et al. Involvement of calcium in retinal pigment epithelial cell proliferation and pigmentation. *Curr Eye Res*. 1998;17:813–22.
41. Wang N, Lee Y-H, Lee J. Recombinant perlucin nucleates the growth of calcium carbonate crystals: molecular cloning and characterization of perlucin from disk abalone, *Haliotis discus discus*. *Comp Biochem Physiol B Biochem Mol Biol*. 2008;149:354–61.
42. Bahn SY, Jo BH, Choi YS, Cha HJ. Control of nacre biomineralization by Pif80 in pearl oyster. *Sci Adv*. 2017;3:e1700765.
43. McGinty EL, Evans BS, Taylor JUJ, Jerry DR. Xenografts and pearl production in two pearl oyster species, *P. maxima* and *P. margaritifera*: effect on pearl quality and a key to understanding genetic contribution. *Aquaculture*. 2010;302:175–81.
44. Ky C-L, Devaux D. Polynesian pearls. *Hatch Int* 2016;17. <https://archimer.ifremer.fr/doc/00341/45252/>. Accessed 13 Apr 2020.
45. Ky C-L, Nakasai S, Molinari N, Devaux D. Influence of grafter skill and season on cultured pearl shape, circles and rejects in *Pinctada margaritifera* aquaculture in Mangareva lagoon. *Aquaculture*. 2015;435:361–70.
46. 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.
47. Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>. Accessed 13 Apr 2020.
48. Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinforma Oxf Engl*. 2016;32:3047–8.
49. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal*. 2011;17:10–2.
50. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinforma Oxf Engl*. 2014;30:2114–20.
51. Li H, Durbin R. Fast and accurate long-read alignment with burrows-Wheeler transform. *Bioinforma Oxf Engl*. 2010;26:589–95.
52. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. *Bioinforma Oxf Engl*. 2009;25:2078–9.
53. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinforma Oxf Engl*. 2015;31:166–9.
54. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15:550.
55. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinforma Oxf Engl*. 2016;32:2847–9.
56. Jones P, Binns D, Chang HY, et al. InterProScan 5: genome-scale protein function classification. *Bioinforma Oxf Engl*. 2014;30(9):1236–40.
57. Klopfenstein DV, Zhang L, Pedersen BS, Ramirez F, Warwick Vesztrocy A, Naldi A, et al. GOATOOLS: a Python library for gene ontology analyses. *Sci Rep*. 2018;8:10872.
58. Supek F, Bošnjak M, Škunca N, Šmuc T. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One*. 2011;6:e21800.
59. Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M. KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res*. 2007;35(Web Server issue):W182–5.
60. Garrison E, Marth G. Haplotype-based variant detection from short-read sequencing. *ArXiv12073907 Q-Bio*. 2012. <http://arxiv.org/abs/1207.3907>. Accessed 20 July 2012.
61. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20:1297–303.
62. Picard Tools - By Broad Institute. <https://broadinstitute.github.io/picard/>. Accessed 13 Apr 2020.
63. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call format and VCFtools. *Bioinforma Oxf Engl*. 2011;27:2156–8.
64. Knaus BJ, Grünwald NJ. Vcfr: a package to manipulate and visualize variant call format data in R. *Mol Ecol Resour*. 2017;17:44–53.
65. Jombart T, Ahmed I. ADEGENET 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinforma Oxf Engl*. 2011;27:3070–1.
66. Rousset F. genepop'007: a complete re-implementation of the genepop software for windows and Linux. *Mol Ecol Resour*. 2008;8:103–6.
67. Luu K, Bazin E, Blum MGB. pcadapt: an R package to perform genome scans for selection based on principal component analysis. *Mol Ecol Resour*. 2017;17:67–77.
68. Wu TD, Reeder J, Lawrence M, Becker G, Brauer MJ. GMAP and GSNAP for genomic sequence alignment: enhancements to speed, accuracy, and functionality. *Methods Mol Biol*. 2016;1418:283–334.
69. Hartley SW, Mullikin JC. QoRTs: a comprehensive toolset for quality control and data processing of RNA-Seq experiments. *BMC Bioinformatics*. 2015;16:224.
70. Hartley SW, Mullikin JC. Detection and visualization of differential splicing in RNA-Seq data with JunctionSeq. *Nucleic Acids Res*. 2016;44:e127.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

