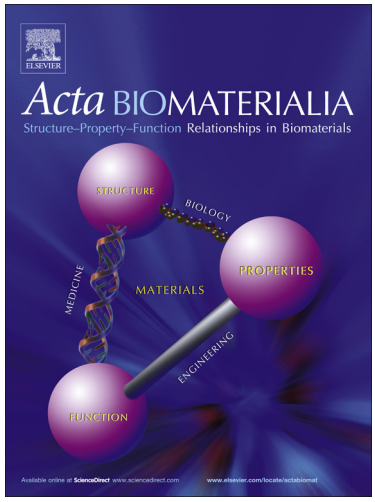


Evolution of the tyrosinase gene family in bivalve molluscs: independent expansion of the mantle gene repertoire

Felipe Aguilera, Carmel McDougall, Bernard M Degnan

PII: S1742-7061(14)00151-2
DOI: http://dx.doi.org/10.1016/j.actbio.2014.03.031
Reference: ACTBIO 3182

To appear in: *Acta Biomaterialia*

Please cite this article as: Aguilera, F., McDougall, C., Degnan, B.M., Evolution of the tyrosinase gene family in bivalve molluscs: independent expansion of the mantle gene repertoire, *Acta Biomaterialia* (2014), doi: http://dx.doi.org/10.1016/j.actbio.2014.03.031

Evolution of the tyrosinase gene family in bivalve molluscs: independent expansion of the mantle gene repertoire

Authors: Felipe Aguilera, Carmel McDougall and Bernard M Degnan*

Authors Affiliations: Centre for Marine Sciences, School of Biological Sciences, The University of Queensland, Brisbane, Australia, 4072.

Corresponding author: Bernard M Degnan, Centre for Marine Sciences, School of Biological Sciences, The University of Queensland, Brisbane, Australia, 4072. Ph: (+61) 7 3365 2467 Fax: (+61+ 7 3365 1199. Email. b.degnan@uq.edu.au

Abstract

Tyrosinase is a copper-containing enzyme that mediates the hydroxylation of monophenols and oxidation of *o*-diphenols to *o*-quinones. This enzyme is involved in a variety of biological processes, including pigment production, innate immunity, wound healing, and exoskeleton fabrication and hardening (e.g. arthropod skeleton and mollusc shell). Here we show that the tyrosinase gene family has undergone large expansions in pearl oysters (*Pinctada* spp.) and the Pacific oyster (*Crassostrea gigas*). Phylogenetic analysis reveals that pearl oysters possess at least four tyrosinase genes that are not present in the Pacific oyster. Likewise, *C. gigas* has multiple tyrosinase genes that are not orthologous to the *Pinctada* genes, indicating that this gene family has expanded independently in these bivalve lineages. Many of the tyrosinase genes in these bivalves are expressed at relatively high levels in the mantle, the organ responsible for shell fabrication. Detailed comparisons of tyrosinase gene expression in different regions of the mantle in two closely-related pearl oysters, *P. maxima* and *P. margaritifera*, reveals that recently-evolved orthologous tyrosinase genes can have markedly different expression profiles. The expansion of tyrosinase genes in these oysters and their co-option into the mantle's gene regulatory network is consistent with mollusc shell formation being underpinned by a rapidly evolving transcriptome.

Keywords: Tyrosinase; bivalve; gastropod; mantle; shell formation.

1. Introduction

Tyrosinases, tyrosinase-related proteins, hemocyanins and catechol oxidases are members of the type-3 copper protein superfamily. These enzymes possess a conserved pair of copper-binding domains, known as Cu(A) and Cu(B), each of which is coordinated by three conserved histidines [1, 2]. Members of this superfamily are present in both eukaryotes and prokaryotes, and are involved in a wide array of biological processes, including pigmentation, innate immunity, oxygen transport, sclerotisation, and wound healing [3-6]. The type-3 copper protein superfamily can be classified into three subclasses based on domain architecture and conserved residues in the copper-binding sites – secreted (α), cytosolic (β) and membrane-bound (γ) subclasses – and is typified by multiple and independent lineage-specific gene expansions and gene losses [7].

Tyrosinases (EC 1.14.18.1) catalyse both the initial hydroxylation of monophenols (e.g., tyrosine) and the further oxidation of *o*-diphenols (e.g., DOPA and DHI) to *o*-quinones [8], to produce melanin. In vertebrates, tyrosinase and its related proteins regulate pigment synthesis [3, 4]. In some invertebrates, melanin can physically encapsulate pathogens [5], and is therefore an important component of the immune system. Moreover, in insects other products of the melanin pathway participate in cuticle sclerotisation and wound healing [6]. In molluscs, tyrosinase is secreted (α-subclass) and appears to contribute to shell pigmentation and formation by the cross-linking of *o*-diphenols and quinone-tanning to form the non-calcified periostracal layer [9-12]. Tyrosinase gene expression and spatial localisation in the organ responsible for shell formation

and patterning in molluscs, the mantle, is consistent with a role in shell fabrication [13].

Here, we reveal through comparative genomics and transcriptomics that the tyrosinase gene family has undergone substantial expansions in at least two bivalve lineages, and that the resulting gene duplicates have been co-opted into the mantle gene regulatory network. Unique expression profiles of orthologous, lineage-restricted tyrosinase genes in the mantles of two closely-related pearl oysters, *Pinctada maxima* and *P. margartifera*, which are estimated to have diverged 8 million years ago [14], indicates that regulatory evolution further contributes to the neofunctionalisation of these new tyrosinase genes in shell formation.

2. Materials and methods

2.1 Genome- and transcriptome-wide surveys of tyrosinase genes

All potential tyrosinase genes were identified by HMMER searches using default parameters, an inclusive *E*-value of 0.05 and the tyrosinase domain (PF00264) as the profile HMM (www.hmmer.org). The analysed molluscan genomes included *Lottia gigantea* (http://genome.jgi-psf.org/Lotgi1/Lotgi1.home.html) [15], *Crassostrea gigas* (http://oysterdb.cn/) [16] and *Pinctada fucata* (http://marinegenomics.oist.jp/pinctada_fucata) [17]; the non-redundant protein database at the NCBI (National Centre for Biotechnology Information) was also analysed. Additionally, publicly available mantle transcriptome data from *P. margaritifera* (NCBI SRA: SRR057743, [18]), *P. fucata* (DDBJ SRA: DRS000687 and DRS000688, [19]), *C. gigas* (http://gigadb.org/dataset/view/id/100030, [16]), *Mytilus edulis*

(http://www.ebi.ac.uk/ena/data/view/PRJEB4516, [20], *Hyriopsis cumingii* (NCBI SRA: SRR530843, [21], *Laternula elliptica* (NCBI SRA: SRA011054, [22]), *L. gigantea* (NCBI EST: FC558616-FC635770), *Patella vulgata* [23], *Haliotis asinina* (NCBI EST: EZ420605-EZ421271, [24], *H. rufescens* (http://datadryad.org/resource/doi:10.5061/dryad.85p80, [25] were downloaded. *P. maxima* mantle transcriptome was obtained using 454 GS-FLX Plus sequencer (F. Aguilera et al. 2013, unpublished data).

For transcriptome datasets, low quality reads were removed and the remaining sequences *de novo* assembled using Trinity software [26] with default settings, followed by clustering of redundant contigs using CAP3 [27]. All transcripts from each species were translated into open reading frames and surveyed for tyrosinase sequence signatures using HMMER profiling. Tyrosinase sequences are available in supplementary data File S1. *P. maxima* tyrosinase sequences have been submitted to NCBI (accession numbers KJ533301-15). The derived protein sequences were BLASTP searched against the NCBI non-redundant protein database with an e-value of 1e-5 in order to corroborate tyrosinase as the best-hit matches.

2.2 Phylogenetic analyses

The retrieved protein sequences were aligned using the MAFFT algorithm [28] and then manually inspected to remove those hits fulfilling one of the following conditions: 1) not possessing all six conserved histidine residues in the copper-binding sites; 2) incomplete sequence with >99% sequence identity to a complete sequence from the same taxa and 3) sequences that showed extremely long branches in the preliminary maximum likelihood trees. The final alignment

was refined using the RASCAL webserver [29] and analysed with Gblocks 0.91b [30] to select conserved regions. Neighbour-Joining (NJ) reconstructions were performed using MEGA 5.2.2 [31] using the JTT substitution model [32] (4 gamma categories) and 1,000 bootstrap replicates. Maximum-likelihood (ML) trees were constructed using RAxMLGUI v1.3 [33] and the WAG substitution model [34], gamma distribution ("PROTGAMMA" implementation), four discrete rate categories, starting from a random tree and 1,000 bootstrap replicates. Bayesian inferences (BI) were performed using MrBayes v3.2 [35] and the WAG model [34] (4 gamma categories). The inference consisted of 1,500,000 generations with sampling every 100 generations, starting from a random starting tree and using 4 chains. Two runs were performed to confirm the convergence of the chains. Trees were visualised and edited using FigTree v.1.3.1 (http://tree.bio.ed.ac.uk/software/figtree/). All alignments are available upon request.

2.3 Gene architecture and synteny analysis

The draft assembly genomes of *L. gigantea, C. gigas* and *P. fucata* were downloaded from each genome portal mentioned above. In brief, the genomes were searched using the tyrosinase genes retrieved by HMM searches and the TBLASTN algorithm. Any identified scaffolds with similarity to tyrosinase genes were extracted for further analysis. Next, the intron-exon architectures of these genes were determined by alignment to the transcript. Each alignment was manually annotated with Geneious v6.0.5 (Biomatters Ltd.) and viewed using CLC Genomics Workbench v6.5.1 (CLC Bio).

To test whether the genes adjacent to the tyrosinase genes are shared across mollusc species (indicating syntenic conservation), scaffolds containing tyrosinase genes were analysed by Augustus v2.7 [36] to predict protein-coding sequences. All predicted sequences were BLASTX and BLASTP searched against NCBI non-redundant protein database, using an *e-value* cut off of 1e-5, and the best-hit match was recorded. In a pairwise approach, predicted amino acid sequences for gene models adjacent to *P. fucata, C. gigas* and *L. gigantea* tyrosinase genes were reciprocally BLASTP searched and the genomic location of five genes upstream and downstream of each tyrosinase genes was compared. Due to the limited length of *P. fucata* scaffolds, additional TBLASTN searches were performed between the genes adjacent to *C. gigas* and *L. gigantea* tyrosinases against the *P. fucata* genome to identify the scaffolds of these neighbours within this species and determine synteny conservation.

2.4 Tissue sampling, total RNA extraction and cDNA synthesis

P. margaritifera were collected from the reef flat at Heron Island Reef, the Great Barrier Reef, Queensland, Australia, and *P. maxima* were provided by Clipper Pearls/Autore Pearling, Broome, WA, Australia. Four individuals of each pearl oyster species were sampled. The gill, foot, adductor muscle, mouth, labial palp, mantle edge and mantle pallial were dissected from these individuals. Additionally, a section of mantle from the outer edge to the centre of four individuals of both pearl oyster species was divided into four equal sections in order to evaluate tyrosinase gene expression across the mantle.

Total RNA was extracted from the tissues and mantle sections with Tri reagent (Sigma-Aldrich) following a protocol modified from Gao et al. [37] to

remove inhibitory pigments. RNAs (500 ng) were treated with Amplification Grade DNase following the instructions of the manufacturer (Invitrogen). cDNA was synthesised using Superscript III reverse transcriptase (Invitrogen) according to the manufacturer's instructions.

2.5 Transcriptome profile analysis and real-time quantitative-reverse transcription PCR (qPCR)

Tyrosinase transcript abundances were assessed for five bivalve species (*P. maxima, P. margaritifera, P. fucata, C. gigas* and *L. elliptica*) using the single- and pair-end read sequences retrieved from each species. All mantle transcriptomes were sequenced from adult animals [16, 18, 19, 22], allowing for direct RNA-Seq comparisons.

Tyrosinase quantification from RNA-Seq data was conducted with RSEM v1.2.8 [38]. It allows for an assessment of transcript abundances based on the mapping of RNA-Seq reads to the assembled transcriptome. Gene level expression was multiplied by one million to obtain a measure given as transcripts per million (TPM) for each gene. Because gene length may vary between samples (isoforms) and species (orthologues), we prefer the use of TPM values over RPKM (read per kilobase per million) values. TPM is independent of the mean expressed transcript length and thus more comparable between different species and samples even if mRNA lengths differ [38, 39].

Nine genes encoding tyrosinase proteins (*P. maxima-TyrA2, -TyrB1.1, -TyrB1.2, -TyrB2.2, -TyrB5* and *P. margaritifera-TyrA2, -TyrB1, -TyrB2, -TyrB5.3*) were analysed by qPCR. Three reference genes (*ferritin, nascent polypeptide-associated complex alpha subunit (α-nac)* and *enoyl-CoA hydratase* (*enCOA*)*; P.*

maxima accession numbers: GT279936, GT279668, GT278168, and *P. margaritifera*: supplementary dataset S2) were selected as the most stably expressed genes from a number of potential candidates using the geNorm program [40]. All primer sequences are available upon request. PCR efficiencies for each primer set were determined by performing qPCR analysis on a serial dilution of a pooled cDNA sample.

qPCR was performed on triplicate samples in a reaction mix of SYBR Green PCR Master Mix (Roche) for amplification (55 cycles of 95°C for 15 sec, 58°C or 60°C for 5 sec and 72°C for 45 sec) on a Roche LightCycler® 480. Thermocycling was carried out in a final volume of 15 μL containing 3 μL cDNA sample (1:50 dilution), 0.5 μL of each primer (10 μM) and 7.5 μL of SYBR Green I Master mix (Roche). Absence of nonspecific products was confirmed by dissociation curve analysis (65-90°C). Quantification of tyrosinase gene expression in each sample relative to a standard were performed using the Roche LightCycler® 480 software. Normalisation of qPCR data to reference genes was performed using REST© [41], incorporating calculated primer efficiencies. All data were represented in terms of relative transcript abundance of the mean of the three replicates using $\log_{10}$ base scale.

3. Results

3.1 Identification of tyrosinase genes in molluscs

Profile HMM identification and sequence verification identified 88 tyrosinase genes from 9 bivalves, 4 gastropods and 2 cephalopods. No tyrosinase genes were detected in the mantle transcriptomes of the tropical abalone *Haliotis asinina* and the red abalone *H. rufescens*. These all encode tyrosinases

with a conserved pair of copper-binding domains. Genes and gene models lacking either the Cu(A) or Cu(B) domain were deemed to not be tyrosinases in this study. Although some of these may represent *bona fide* genes or pseudogenes, many of these appear to be incompletely or incorrectly assembled transcriptome or genome models.

Many bivalves have multiple tyrosinases (Table 1). The expansion of tyrosinase genes appears to be a common feature in bivalves, with more than ten gene family members present in *Pinctada* spp. and *Crassostrea gigas.* The freshwater mussel *Hyriopsis cumingii* has at least six genes, the green mussel *Perna virilis* has at least five genes, which have been previously identified to be expressed in the foot [42], and the saltwater clam *Laternula elliptica* minimally has two genes. Note that published transcriptomes are restricted to specific tissues and life cycles stages and thus might not include all tyrosinases in these bivalve genomes. Gastropods appear to have a limited number of genes encoding tyrosinases, with two genes present in the draft *Lottia gigantea* genome. We note that we found fewer than the recently reported 21 tyrosinase genes in the *P. fucata* genome [43], this is because several of these genes do not encode the six conserved histidine residues within the copper-binding domains that are essential for tyrosinase function ; these were not included in subsequent analyses.

3.2 Phylogenetic analyses of tyrosinase genes in molluscs

The most conserved regions in tyrosinase proteins correspond to the copper-binding sites [1, 2]. Using this region, we previously analysed the evolution of the entire type-3 copper protein superfamily [7]. Our analysis of

molluscan tyrosinases produce phylogenetic trees with very low support for many nodes (Fig. 1 and Fig. S1-S3), as was observed in the wider survey [7]. This may be because of the high level of conservation in the residues surrounding the copper-binding sites, resulting in a weak phylogenetic signal. Nonetheless, these analyses reveal two distinct clades of tyrosinase proteins (Fig. 1), one comprising bivalve, gastropod and cephalopod tyrosinases (clade A) and the other comprising only bivalve tyrosinases (clade B).

These analyses also demonstrate that the molluscan tyrosinase gene family has undergone independent lineage-specific gene expansions, with many of the tyrosinase genes present in *Pinctada* spp. and *C.* gigas restricted to these lineages (Fig. 1). This complex evolutionary history of molluscan tyrosinase genes required a naming scheme. First, genes falling into clade A or B are designated as TyrA or TyrB, respectively. These are then followed by an Arabic number to indicate different groups. In cases where two or more genes from the same species are part of a group a decimal number was added at the end of the name. For example, *C. gigas-TyrA1.1* and *C. gigas-TyrA1.2* are different genes that are part of the TyrA1 group (Fig. 1). Lineage-specific expansions are followed by a species-specific identifier and an Arabic number (e.g. *C. gigas-TyrACgig1* and *H. cumingii-TyrAHcum1*). The phylogenetic distribution of tyrosinases is consistent with a tyrosinase type A (TyrA) being ancestral and potentially present in the last common molluscan ancestor. This ancestral form likely duplicated and diverged before the diversification of bivalves surveyed in this study, giving rise to the tyrosinase type B (TyrB) (Figs 1, 2). TyrA and TyrB genes then underwent extensive expansions in the lineages leading to *C. gigas* and *Pinctada* spp., respectively (Figs 1, 2; Figs S1-S3).

The availability of genomic and transcriptomic resources for three closely-related pearl oyster species has allowed us to analyse the dynamics of tyrosinase gene family evolution in more detail. The phylogenetic relationships of the three species *P. fucata*, *P. maxima* and *P. margaritifera* are well understood, with latter two species diverging from *P. fucata* about 14 Mya and from each other approximately 8 Mya [14]. We identified at least six orthologous tyrosinase groups containing representatives from all *Pinctada* species, TyrA2, A3, TyrB1-4. TyrA1 may have been lost in the *P. maxima* + *P. margaritifera* lineage, although without genomic sequence this is difficult to ascertain, and TyrB5 appears to be an orthology group restricted to these two species (Figs 1, 2). In each of the conserved groups, there are cases of further lineage-specific gene duplications, such that there may species-specific paralogues within a given *Pinctada* orthology group (e.g. *P. fucata* and *P. margaritifera* have four and two paralogues respectively within orthology group TyrB4; Fig. 1; Figs S1-S3).

3.3 Linkage and syntenic relationship of tyrosinase genes in molluscan genomes

To further investigate the evolution of the tyrosinase gene family in molluscs, we examined the structure and organisation of tyrosinase genes in three molluscs whose genomes have been sequenced, assembled and annotated, *L. gigantea, C. gigas* and *P. fucata*. In the gastropod *L. gigantea*, two tyrosinase genes were located on different scaffolds. In *C. gigas*, there are five scaffolds with two or more tyrosinase genes. Only two of these scaffolds (337 and 867) possess a non-tyrosinase gene within the tyrosinase cluster (Fig. 3). In *P. fucata*, we found two tyrosinase gene clusters in the genome (Fig. 3), however the scaffolds for this species are relatively short and thus other tyrosinase clusters may exist.

In most clusters, one of the tyrosinase genes shows significantly higher expression (in terms of transcripts per million) than other genes located within that cluster (Table 2). Comparison of exon-intron architecture reveals that there is little conservation of tyrosinase gene structure within and between clusters. Two exceptions include *C. gigas* scaffolds 203, which contains closely related *TyrACgig3* and *TyrACgig4* with identical exon-intron organisation, and 867, which has two distantly related genes – *TyrA3.3* and *TyrB6* - with conserved architectures (Fig. 3).

We analysed five upstream and downstream genes that are adjacent to each tyrosinase and looked for synteny in *L. gigantea, C. gigas* and *P. fucata* genomes. Comparisons of *C. gigas* and *P. fucata* scaffolds identified two microsyntenic regions. Specifically *C. gigas* scaffold 867, which included *TyrB6*, *TyrA3.3*, *TyrA3.4* and *TyrA2.2* along with non-tyrosinase genes, is syntenic to *P. fucata* scaffolds 13287, 1286 and 19072, which house *TyrA3.1*, *TyrA3.2* and *TyrA2.2* and orthologous non-tyrosinase genes (Fig. 4; Fig. S4). *TyrA1.2* and *TyrA1* are adjacent to *Forkhead box gene*, *FOXP1*, in both *C. gigas* and *P. fucata* scaffolds (Fig. 4). No shared genes surrounding tyrosinase loci of *L. gigantea* and either bivalve species were identified. The exon-intron organisation of all syntenic tryosinase genes differed between *C. gigas* and *P. fucata* (Fig. 3), indicating that although synteny exists, the structure of these genes has evolved since the divergence of *Crassostrea* and *Pinctada* lineages.

3.4 Tyrosinase transcript abundance and gene expression across the mantle tissue of pearl oysters

De novo mantle transcriptome assembly for five bivalve species yielded a large number of putative single-copy genes, ranging from 25,135 to 224,965 unigenes (Table 3). Mantle RNA-seq data were used to evaluate tyrosinase transcript abundance in each species. Tyrosinase gene expression levels, as assessed by RNA-Seq read counts converted into TPM [38], vary markedly between genes and species (Fig. 5). Overall, pearl oysters had higher tyrosinase expression levels than the other bivalves, with few exceptions (Fig. 5). Many of these genes, at least in *P. maxima* and *P. margaritifera*, have significantly higher expression in the mantle tissue than other tissues (Table 4), which is consistent with previously reports of high tyrosinase expression levels in the mantle compared to other tissues in the Pacific oyster [16, 44]. Our qPCR analyses are consistent with transcript abundance estimations based on RNA-Seq data, lending further support to high tyrosinase transcript abundance in pearl oysters.

We assessed transcript abundance levels of nine of tyrosinase genes in different regions of the mantle in two species of pearl oyster (*P. maxima* and *P. margaritifera*) by qPCR; seven genes were deemed as orthologues based on phylogenetic analyses (Fig. 1) (*P. maxima-TyrA2* and *P. margaritifera-TyrA2* (group A2), *P. maxima-TyrB1.1* and -*TyrB1.2* and *P. margaritifera-TyrB1* (group B1), and *P. maxima-TyrB2.2* and *P. margaritifera-TyrB2* (group B2). *P. maxima-TyrB5* and *P. margaritifera-TyrB5.3* are also orthologues but were only found in these sister species. The mantle tissue was divided into different zones, distal, two central and proximal, with the distal zone in direct contact with the prismatic shell layer and the central and proximal zones with the nacreous shell layer (Fig. 6). Tyrosinase gene expression levels varied between regions of the

mantle and species and even between individuals within the same species (Fig. 6; Fig. S5). Most genes are more highly expressed at the distal mantle edge.

Among the genes analysed, the orthologous gene pair *P. maxima-TyrB2.2* and *P. margaritifera-TyrB2* were the most highly expressed at the distal mantle edge. Expression of these genes was approximately 1000-fold less in the central and proximal zones in both species. Although the expression profiles of these orthologues across the mantles of these two species are similar, the *P. maxima* gene is about 100 times more highly expressed (Fig. 6). The orthologous pairs *P. maxima-TyrB5* and *P. margaritifera-TyrB5.3* showed a decrease from the distal (outer part) to the proximal zone (mantle centre), however in this case the *P. margaritifera* gene is about 10 times more highly expressed in the distal mantle, but more lowly expressed in the other regions of the mantle. *P. maxima-TyrB1.1 and –TyrB1.2* and *P. margaritifera-TyrB1* are expressed at similar levels in the distal mantle but vary in other mantle territories. Likewise, the orthologous *TyrA2* genes display species-specific profiles across the mantle (Fig. 6).

4. Discussion

4.1 Independent large-scale expansions of the tyrosinase gene family in bivalves

The tyrosinase gene family has undergone multiple lineage-restricted expansions [7], including in the closely-related bivalve superfamilies Ostreoidea (containing *Crassostrea*) and Pterioidea (containing *Pinctada*) [45, 46]. In this study, we sought to reconstruct the evolution of this gene family in bivalves and other molluscs using existing and new genome and transcriptome data. Although this survey is far from exhaustive and largely relies on transcriptome data, phylogenetic analyses revealed that large tyrosinase gene expansions occurred

these taxa. Smaller lineage-restricted expansions were observed in other bivalves, including *P. viridis* and *H. cumingii*, leaving open the possibility that the tyrosinase gene family may have expanded in multiple mollusc lineages.

Phylogenetic analyses reveal that the ancestral molluscan tyrosinase gene duplicated early in bivalve evolution, giving rise to an ancestral clade (A) and bivalve-specific clade (B) (Fig. 1). Although it is difficult to further resolve the evolution to tyrosinase A genes, it is clear that the ancestral gene has undergone further independent duplication and divergence in both bivalves and gastropods. For example, there are three *TyrA* orthologues shared between *Pinctada* spp. and *C. gigas*. *C. gigas-TyrA1.2, -TyrA3.3 and -TyrA2.2* are orthologous to *P. fucata-TyrA1, -TyrA3.1 and -TyrA2.2*, respectively, indicating that these genes duplicated before the divergence of these two bivalve lineages (Fig. 1). These orthologues also display conserved synteny (Fig. 4). In addition to the expansion of TyrA genes prior to the divergence of *Crassostrea* and *Pinctada* lineages, there have been a number of separate *Crassostrea*-specific and *Pinctada*-specific expansions. In *C. gigas*, there has been a large *TyrA* expansion, leading to 12 paralogues and a number of other duplicates (24 genes total). There are only three *TryB* genes in *C. gigas*. In contrast, there appears to have been little further expansion of the *TyrA* genes in *Pinctada* after it diverged from the *C. gigas* lineage. Instead, *TyrB* genes have undergone continuous expansion during evolution of *Pinctada*, with shared and species-specific duplications evident (Figs 1, 2).

Duplicated *Crassostrea* and *Pinctada* tyrosinase genes can be found in clusters within the genomes of these species (Fig. 3), and likely arose via tandem duplication [47, 48]. In support of this hypothesis, many clusters consisted of genes that grouped closely within the phylogenetic tree and likely reflect more

recent duplicates (e.g., *C. gigas-TyrA3.5* and *C. gigas-TyrA3.6; C. gigas-TyrCgig3* and *C. gigas-TyrACgig4; P. fucata-TyrB3.2* and *P. fucata-TyrB3.3*, Fig. 3.). In some cases, however, clusters consisted of more distantly-related tyrosinase genes, for example, the cluster found on scaffold 867 of the *C. gigas* genome contains tyrosinase genes from Clades A and B. These genes also share a conserved exon-intron architecture, suggesting this may have been the organisation of the ancestral bivalve *TyrA and TyrB* genes. This cluster also displays synteny with the *P. fucata* genome, indicating that this arrangement has been maintained over evolutionary time. A number of reasons for the generation and/or maintenance of gene clusters have been proposed, including sharing of regulatory elements or the requirement for co-expressed genes to reside in a specifically-regulated region of chromatin [49, 50]. Genes from the same metabolic pathway are often found clustered within genomes [51]. The observation that one gene from each cluster is often much more highly expressed than the others may point towards a proximity-based shared enhancer, which may play a role in cluster maintenance [52].

4.2 Does functional divergence explain the retention of multiple tyrosinase duplicates?

The reason for the extensive tyrosinase gene duplication in *Crassostrea* and *Pinctada* lineages is difficult to ascertain. Retention of gene duplicates is often attributed to subfunctionalisation (division of ancestral roles between duplicated genes) or neofunctionalisation (attainment of a new functional role) of the duplicated genes, after which gene loss becomes detrimental [53-55]. We therefore investigated whether tyrosinase genes display differences in the

location or level of gene expression, as differences in gene expression are likely to reflect functional differences between the gene products. We analysed the gene expression profiles of nine tyrosinase genes in different tissues and across the mantle of two closely-related pearl oyster species, *P. maxima* and *P. margartifera*. Tissue-specific expression showed that tyrosinase transcripts are mostly expressed in the mantle, which contributes to the formation of the shell [56]. Within the mantle, tyrosinase genes were differentially expressed along the proximo-distal axis. In mollusc shells, the deposition of shell layers appears to be controlled by regionalised expression of genes within different zones of the mantle [57, 58]. Our qPCR results show high expression of several tyrosinase genes in the distal zone, suggesting roles in prismatic shell layer construction and/or periostracum formation. These results, and the detection of tyrosinase in different parts of the shell and at different ontological stages [9, 13, 59], indicates that tyrosinase duplicates may be retained because of their functional diversification in the mantle.

4.3 Substrate affinity and insights into the functionalities of tyrosinase genes in shell biomineralization

Although, the exact role of duplicated tyrosinase genes in shell formation is unknown, two lines of evidence suggest that they play key structural roles in shell formation. First, enzymatic assays and *in situ* hybridisation analyses reveal tyrosinase gene expression in the mantle cells of the middle fold, consistent with a role in periostracal layer formation [9]. Second, the spatial localisation of tyrosinase in the pigmented shell and mantle tissue suggest a role in shell pigmentation [13]. The enzymatic mechanism of tyrosinases in shell formation

and pigmentation is still under debate because of the presence of two catalytic activities and different substrate affinities. Nonetheless, the oxidation of monophenols to quinones [6, 12], and the subsequent reaction of quinones with nucleophilic amino acids can result in cross-linking accompanied by pigmentation [6]. This evidence suggests that tyrosinase has an important function in tanning periostracum proteins [11]. Different enzymatic inhibitors reveal differences in tyrosinase activity in various tissues in *C. gigas* [60], suggesting that new catalytic activities and metal binding properties may have evolved. This may be analogous to the vertebrate tyrosinase-related protein 2, which uses zinc instead of copper as cofactor [61]. These substrate affinities, in addition to the localisation and high level of expression of the genes, suggest that tyrosinases are important structural components of molluscan shells. It is therefore likely that the diversification of tyrosinase proteins in *C. gigas* and *Pinctada* spp. has contributed to the diversity of structure and patterning observed within these bivalve shells.

5. Conclusion

We show that the tyrosinase gene family has greatly expanded in two oyster lineages, with duplications occurring both before and after the divergence of Ostreoidea and Pterioidea. The majority of the tyrosinase genes in these groups are expressed at high levels in the mantle. However, there are noticeable differences in orthologue expression levels and profiles in this shell-fabricating organ between sister species, *P. maxima* and *P. margaritifera.* As these species diverged about 8 Mya [14], differences in expression levels are consistent with the rapid evolution of the regulatory architecture controlling expression of these

genes in mantle cells. These results are consistent with our previous suppositions that marked differences in the structure, colour and pattern of shells between closely related mollusc species, and sometimes individuals within a species, are underpinned by the rapid evolution of gene families that encode secreted proteins and are part of the mantle gene regulatory architecture [57, 62].

Acknowledgments

The authors thank Clipper Pearls/Autore Pearling for providing specimens of *P. maxima* for use in this study. This study was supported by funding from the Australian Research Council to BMD and a Becas Chile scholarship from Conicyt (Chile) to FA.

Appendix A. Supplementary data

Supplementary data associated with this article can be found with the online version.

References

[1] Decker H, Tuczek F. Tyrosinase/catecholoxidase activity of hemocyanins: structural basis and molecular mechanism. Trends Biochem Sci 2000;25:392-7.
[2] Decker H, Schweikardt T, Nillius D, Salzbrunn U, Jaenicke E, Tuczek F. Similar enzyme activation and catalysis in hemocyanins and tyrosinases. Gene 2007;398:183-91.
[3] Hofreiter M, Schoneberg T. The genetic and evolutionary bases of colour variation in vertebrates. Cell Mol Life Sci 2010;67:2591-603.
[4] Cieslak M, Ressmann M, Hofreiter M, Ludwig A. Colours of domestication. Biol Rev 2011;86:885-99.
[5] Cerenius L, Lee BL, Soderhall K. The proPO-system: pros and cons for its role in invertebrate immunity. Trends Immunol 2008;29:263-71.
[6] Andersen SO. Insect cuticular sclerotization: A review. Insect Biochem Mol Biol 2010;40:166-78.
[7] Aguilera F, McDougall C, Degnan BM. Origin, evolution and classification of type-3 copper proteins: lineage-specific gene expansions and losses across the Metazoa. BMC Evol Biol 2013;13:96.
[8] Sanchez-Ferrer A, Rodriguez-Lopez JN, Garcia-Canovas F, Garcia-Carmona F. Tyrosinase: a comprehensive review of its mechanism. Biochim Biophys Acta 1995;1247:1-11.
[9] Zhang C, Xie L, Huang J, Chen L, Zhang R. A novel putative tyrosinase involved in periostracum formation from the pearl oyster (*Pinctada fucata*). Biochem Biophy Res Commun 2006;342:632-9.
[10] Tsujii T. Studies on the mechanism of shell and pearl-formation. VIII. On the tyrosinase in the mantle. J Faculty of Fisheries University of Mie 1962;5:378-83.
[11] Timmermans LPM. Studies on shell formation in molluscs. Neth J Zool 1969;19:417-523.
[12] Waite JH. Quinone-tanned scleroproteins. In: Wilbur KM, Hochachka PW, editors. The Mollusca. New York: Academic Press; 1983. p. 467 - 504.
[13] Nagai K, Yano M, Morimoto K, Miyamoto H. Tyrosinase localization in mollusc shells. Comp Biochem Physiol B 2007;146:207-14.
[14] Cunha RL, Blanc F, Bonhomme F, Arnaud-Haond S. Evolutionary patterns in pearl oysters of the genus *Pinctada* (Bivalvia: Pteriidae). Mar Biotechn 2011;13:181-92.
[15] Simakov O, Marletaz F, Cho S-J, Edsinger-Gonzales E, Havlak P, Hellsten U, et al. Insights into bilaterial evolution from three spiralian genomes. Nature 2013;493:526-31.
[16] Zhang G, Fang X, Guo X, Li L, Luo R, Xu F, et al. The oyster genome reveals stress adaptation and complexity of shell formation. Nature 2012;490:49-54.
[17] Takeuchi T, Kawashima T, Koyanagi R, Gyoja F, Tanaka M, Ikuta T, et al. Draft genome of the pearl oyster *Pinctada fucata*: A platform for understanding bivalve biology. DNA Res 2012;19:117-30.
[18] Joubert C, Piquemal D, Marie B, Manchon L, Pierrat F, Zanella-Cleon I, et al. Transcriptome and proteome analysis of *Pinctada margaritifera* calcifying mantle and shell: focus on biomineralization. BMC Genomics 2010;11:613.
[19] Kinoshita S, Wang N, Inoue H, Maeyama K, Okamoto K, Nagai K, et al. Deep sequencing of ESTs from nacreous and prismatic layer producing tissues and a

screen for novel shell formation-related genes in the pearl oyster. PLoS ONE 2011;6:e21238.
[20] Freer A, Bridgett S, Jiang J, Cusack M. Biomineral proteins from *Mytulis edulis* mantle tissue transcriptome. Mar Biotechn 2014;16:34-45.
[21] Bai Z, Zheng H, Lin J, Wang G, Li J. Comparative analysis of the transcriptome in tissues secreting purple and white nacre in the pearl mussel *Hyriopsis cumingii*. PLoS ONE 2013;8:e53617.
[22] Clark MS, Thorne MAS, Vieira FA, Cardoso JCR, Power DM, Peck LS. Insights into shell deposition in the Antarctic bivalve *Laternula elliptica*: gene discovery in the mantle transcriptome using 454 pyrosequencing. BMC Genomics 2010;11:362.
[23] Werner GDA, Gemmell P, Grosser S, Hamer R, Shimeld SM. Analysis of a deep transcriptome from the mantle tissue of *Patella vulgata* Linnaeus (Mollusca: Gastropoda: Patellidae) reveals candidate biomineralising genes. Mar Biotechn 2013;15:230-43.
[24] Jackson DJ, McDougall C, Woodcroft B, Moase P, Rose RA, Kube M, et al. Parallel evolution of nacre building gene sets in molluscs. Mol Biol Evol 2010;27:591-608.
[25] de Wit P, Palumbi SR. Transcriptome-wide polymorphisms of red abalone (*Haliotis rufescens*) reveal patterns of gene flow and local adaptation. Mol Ecol 2012;22:2884-97.
[26] Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 2011;29:644-52.
[27] Huang X, Madan A. CAP3: A DNA sequence assembly program. Genome Res 1999;9:868-77.
[28] Katoh K, Kuma K-i, Toh H, Miyata T. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res 2005;33:511-8.
[29] Thompson JD, Thierry JC, Poch O. RASCAL: rapid scanning and correction of multiple sequence alignments. Bioinformatics 2003;19:1155-61.
[30] Castresama J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol 2000;17:540-52.
[31] Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 2011;28:2731-9.
[32] Jones DT, Taylor WR, Thornton JM. The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci 1992;8:275-82.
[33] Silvestro D, Michalak I. raxmlGUI: a graphical front-end for RAxML. Org Div Evol 2012;12:335-7.
[34] Whelan S, Goldman N. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol Biol Evol 2001;18:691-9.
[35] Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Hohna S, et al. MrBayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. Syst Biol 2012;61:539-42.
[36] Stanke M, Morgenstern B. AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. Nucleic Acids Res 2005;33:W465-W7.

[37] Gao J, Liu J, Li Z. Isolation and purification of functional total RNA from blue-grained wheat endosperm tissues containing high levels of starches and flavonoids. Plant Mol Biol Rep 2001;19:185a-i.
[38] Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 2011;12:323.
[39] Li B, Ruotti V, Stewart RM, Thomson JA, Dewey CN. RNA-Seq gene expression estimation with read mapping uncertainty. Bioinformatics 2010;26:493-500.
[40] Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol 2002;3:research0034.1-research.11.
[41] Pfaffl M, Horgan GW, Dempfle L. Relative expression software tool (REST©) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. Nucleic Acids Res 2002;30:e36.
[42] Guerette PA, Hoon S, Seow Y, Raida M, Masic A, Wong FT, et al. Accelerating the design of biomimetic materials by interacting RNA-seq with proteomics and materials science. Nat Biotechnol 2013;31:908-15.
[43] Miyamoto H, Endo H, Hashimoto N, Iimura K, Isowa Y, Kinoshita S, et al. The diversity of shell matrix proteins: Genome-wide investigation of the pearl oyster, *Pinctada fucata*. Zool Sci 2013;30:801-16.
[44] Wang X, Li L, Zhu Y, Du Y, Song X, Chen Y, et al. Oyster shell proteins originate from multiple organs and their probable transport pathway to the shell formation front. PLoS ONE 2013;8:e66522.
[45] Kocot KM, Cannon JT, Todt C, Citarella MR, Kohn AB, Meyer A, et al. Phylogenomics reveals depp molluscan relationships. Nature 2011;477:452-6.
[46] Smith SA, Wilson NG, Goetz FE, Feehery C, Andrade SC, Rouse GW, et al. Resolving the evolutionary relationships of molluscs with phylogenomics tools. Nature 2011;480:364-7.
[47] Hurles M. Gene duplication: The genomic trade in spare parts. PLoS Biol 2004;2:0900-4.
[48] Reams AB, Neidle EL. Selection for gene clustering by tandem duplication. Annu Rev Microbiol 2004;58:119-42.
[49] Hurst LD, Pál C, Lercher MJ. The evolutionary dynamics of eukaryotic gene order. Nat Rev Genet 2004;5:299-310.
[50] Kikuta H, Laplante M, Navratilova P, Komisarczuk A, Engström PG, Fredman D, et al. Genomic regulatory blocks encompass multiple neighboring genes and maintain conserved synteny in vertebrates. Genome Res 2007;17:545-55.
[51] Lee JM, Sonnhammer ELL. Genomic gene clustering analysis of pathways in eukaryotes. Genome Res 2003;13:875-82.
[52] Jiang S-Y, González JM, Ramachandran S. Comparative genomic and transcriptomic analysis of tandemly and segmentally duplicated genes in rice. PLoS ONE 2013;8:e63551.
[53] Rastogi S, Liberles DA. Subfunctionalization of duplicated genes as a transition state to neofunctionalization. BMC Evol Biol 2005;5:28.
[54] MacCarthy T, Bergman A. The limits of subfunctionalization. BMC Evol Biol 2007;7:213.

[55] Force A, Lynch M, Pickett FB, Amores A, Yan Y-l, Postlethwait J. Preservation of duplicate genes by complementary, degenerative mutations. Genetics 1999;151:1531-45.
[56] Marin F, Le Roy N, Marie B. The formation and mineralization of mollusk shell. Front Biosci 2012;4:1099-125.
[57] Jackson DJ, McDougall C, Green K, Simpson F, Worheide G, Degnan BM. A rapidly evolving secretome builds and patterns a sea shell. BMC Biol 2006;4:40.
[58] Jackson DJ, Worheide G, Degnan BM. Dynamic expression of ancient and novel molluscan shell genes during ecological transitions. BMC Evol Biol 2007;7:160.
[59] Huan P, Liu G, Wang H, Liu B. Identification of a tyrosinase gene potentially involved in early larval shell biogenesis of the Pacific oyster *Crassostrea gigas*. Dev Genes Evol 2013;223:389-91.
[60] Luna-Acosta A, Tomas-Guyon H, Amari M, Rosenfeld E, Bustamante P, Fruitier-Arnaudin I. Different tissue distribution and specificity of phenoloxidases from the Pacific oyster *Crassostrea gigas*. Comp Biochem Physiol B 2011;159:220-6.
[61] Olivares C, Solano F. New insights into the active site structure and catalytic mechanism of tyrosinase and its related proteins. Pigment Cell Melanoma Res 2009;22:750-60.
[62] McDougall C, Aguilera F, Degnan BM. Rapid evolution of pearl oyster shell matrix proteins with repetitive low-complexity domains. J R Soc Interface 2013;10:20130041.

Tables

Table 1. Minimal number of tyrosinase genes present in the genome or transcriptome of a variety of molluscs.

Organism	Class	Family	N° of genes
Pinctada maxima	Bivalvia	Pteriidae	11
Pinctada margaritifera	Bivalvia	Pteriidae	10
Pinctada fucata	Bivalvia	Pteriidae	19
Crassostrea gigas	Bivalvia	Ostreoidae	27
Azumapecten farreri	Bivalvia	Pectinidae	1
Mytilus edulis	Bivalvia	Mytilidae	0
Perna viridis	Bivalvia	Mytilidae	5
Hyriopsis cumingii	Bivalvia	Unionidae	6
Laternula elliptica	Bivalvia	Laternulidae	2
Lottia gigantea	Gastropoda	Lottidae	2
Patella vulgata	Gastropoda	Patellidae	2
Haliotis rufescens	Gastropoda	Haliotidae	0
Haliotis asinina	Gastropoda	Haliotidae	0
Illex argentinus	Cephalopoda	Ommastrephidae	2
Sepia officinalis	Cephalopoda	Sepiidae	1

Table 2. Expression of tyrosinase genes in the *C. gigas* and *P. fucata* mantle tissues, as transcripts per million (TPM), with their corresponding expected counts (EC).

C. gigas	TPM	EC	*P. fucata*	TPM	EC
scaffold203			**scaffold31287.1**		
C. gigas-TyrACgig3	8.65	372.22	*P. fucata-TyrA3.1*	0.87	2
C. gigas-TyrACgig4	67.4	3018.71	*P. fucata-TyrA3.2*	0	0
scaffold337			**scaffold1032.1**		
C. gigas-TyrA1.4	0.2	8	*P. fucata-TyrB3.3*	0	0
C. gigas-TyrA3.2	0.23	13	*P. fucata-TyrB3.2*	6.88	33.07
scaffold552			*P. fucata-TyrB3.4*	1	5.05
C. gigas-TyrA3.5	6.49	392.51			
C. gigas-TyrA3.6	0.9	53.3			
scaffold867					
C. gigas-TyrB6	0.02	1			
C. gigas-TyrA3.4	0.07	4			
C. gigas-TyrA3.3	12.98	643.25			
C. gigas-TyrA2.2	0.79	28.39			
scaffold43702					
C. gigas-TyrACgig6	0.09	4			
C. gigas-TyrACgig7	26.19	1199.7			
C. gigas-TyrACgig12	0.22	8.9			

Table 3. Summary of *de novo* assembled transcripts and open reading frame (ORFs) predictions of five bivalve species used to quantify tyrosinase transcript abundance.

Species[a]	Total raw reads	Total clean reads	Total transcripts	Transcript mean length	Transcript N50	Predicted ORFs
P. maxima	318,850	287,000	37,833	827.7	1,091	31,977
P. margaritifera	276,735	246,886	38,867	410.7	480	33,797
P. fucata	322,742	158,036	25,135	396.3	484	20,902
C. gigas	38,105,927	31,516,399	224,965	608.1	1,804	204,940
L. elliptica	1,033,522	804,965	69,256	438.6	551	54,093

[a] *P. maxima, P. margartifiera, P. fucata* and *L. elliptica* raw reads were obtained using 454 sequencing technology, and *C. gigas* raw reads were obtained using Illumina sequencing technology.

Table 4. Relative gene expression of nine tyrosinase genes in different tissues of the pearl oysters *P. maxima* and *P. margaritifera.*

	***P. maxima* tyrosinase gene expression**				
Tissues	***Pmax-TyrA2***	***Pmax-TyrB1.1***	***Pmax-TyrB1.2***	***Pmax-TyrB2.2***	***Pmax-TyrB5***
gill	3.70	0.00	4.41	0.00	2.97
foot	0.00	0.03	0.01	0.03	0.00
adductor muscle	0.56	0.13	0.52	0.00	9.80
mouth	0.01	0.05	0.49	11.89	0.07
labial palp	0.05	0.00	0.49	17.10	0.00
mantle edge	0.10	14.57	0.00	16977.88	25.77
mantle pallial	11.45	10.55	2.49	0.17	9.49
	***P. margaritifera* tyrosinase gene expression**				
Tissues	***Pmar-TyrA2***	***Pmar-TyrB1***	***Pmar-TyrB2***	***Pmar-TyrB5.3***	
gill	0.10	0.01	2.14	0.00	
foot	0.00	0.00	0.00	0.00	
adductor muscle	0.00	0.00	0.00	0.00	
mouth	5.13	8.66	3.61	6.49	
labial palp	0.06	0.00	0.00	0.00	
mantle edge	20.30	33.19	219.83	55.86	
mantle pallial	4.35	0.28	0.12	0.04	

Figure legends

Fig. 1. Phylogenetic analysis of tyrosinase proteins in molluscs. A consensus midpoint-rooted phylogenetic tree based on Maximum Likelihood topology is shown. Percentage bootstrap values (BV) are indicated at the nodes; first number NJ bootstrap support; second number, ML bootstrap support; third number, Bayesian posterior probabilities (BPP). Only statistical support values >50% and posterior probabilities >0.50 are shown. A black dot in the node indicates BV >90% and BPP >0.9. Bivalve and molluscan orthology TyrA groups are indicated by thick brackets and annotated A1-A3. *Pinctada*-specific TyrB orthology groups are bracketed and annotated B1- B4 and BPmax/Pmar5. Sequences used in this tree can be found in supplementary dataset S1. See supplementary Figures S1-S3 for trees of molluscan tyrosinase proteins generated using each phylogenetic method. Species are colour coded as follows: red: *Pinctada maxima*; blue: *P. margaritifera*; green: *P. fucata*; brown: *P. martensii*; black: *Hyriopsis cumingii*; orange: *Crassostrea gigas*; light green: *Perna viridis*; grey: *Laternula elliptica*; magenta: *Azumapecten farreri*; pink: *Lottia gigantea*; purple: *Patella vulgata*; sky blue: *Illex argentinus* and yellow: *Sepia officinalis*.

Fig 2. Evolution of bivalve and other molluscan tyrosinase genes. The phylogenetic relationship between the species is based on [45] and [46]. The origin of tyrosinase A groups (red dots) and B groups (blue dots) are shownand follows the nomenclature in figure 1. The number adjacent to the dots signify the minimal increase in gene number along a lineage. Circle with a slash represents

gene loss (A2 along gastropod lineage). Other gene losses may exist but can not be confirmed solely by comparing transcriptomes. Species are labelled according to colour code shown in Figure 1.

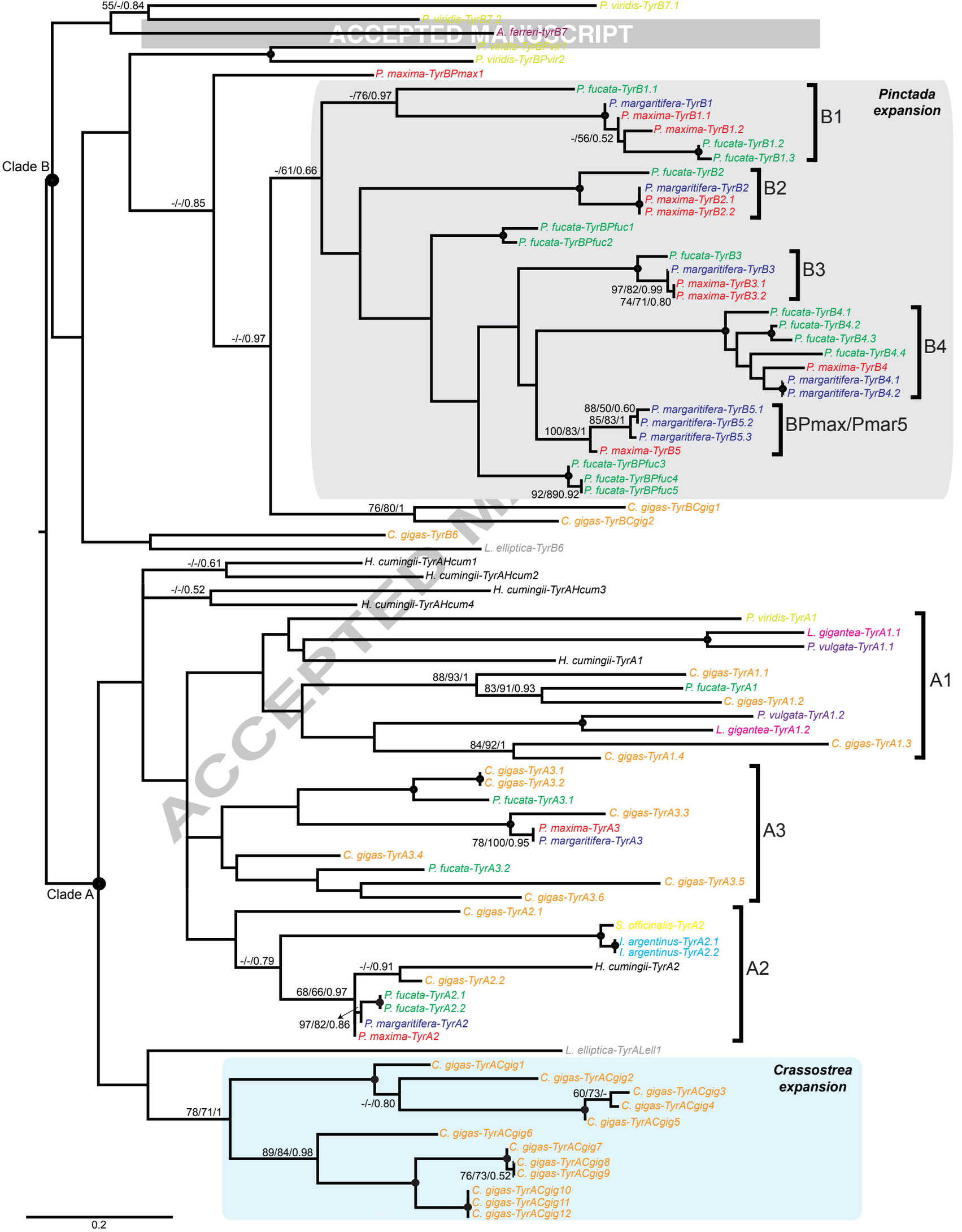
Fig. 3. Tyrosinase gene clusters in *C. gigas* and *P. fucata*. Scaffolds containing tyrosinase genes are to the left, with scaffold numbers corresponding to the original genome annotations for these species [16, 17]. Gene location and orientation are denoted by red arrows on the scaffolds. The distances between genes are shown, along with the location of this cluster from the ends of the scaffold. The exon-intron architecture of the *Tyr* genes are shown to the right. Exons are indicated by boxes and introns (not drawn to scale) are indicated by lines adjoining these. Scale bars presented for all gene models apply only to exons.

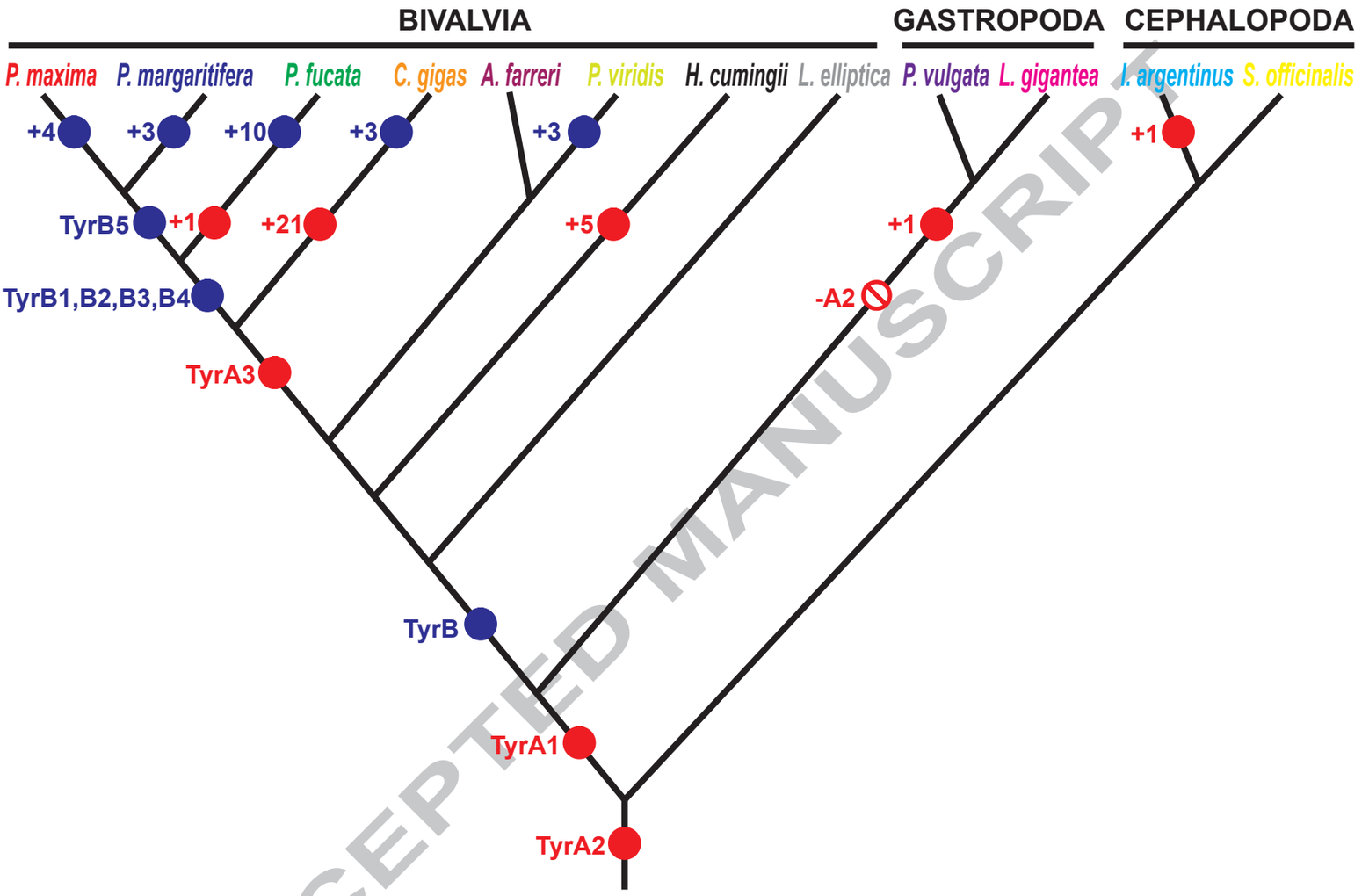
Fig. 4. Analysis of local synteny between the *C. gigas* and *P. fucata* genomes. Each *C. gigas* or *P. fucata* scaffold is represented as black bar and annotated as per Fig. 3. Predicted genes within each segment were identified by BLAST search similarity searching and are shown as rectangles. *C. gigas* or *P. fucata* orthologues are connected by a red line. Gene abbreviations are as follows: *APOD: Apolipoprotein D; SYF2: SYF2 pre-mRNA-splicing factor; KDM4B: Lysine (K)-specific demethylase 4b; SETBP1: SET binding protein 1; HDAC11: Histone deacetylase 11; PACSIN1: Protein kinase C and casein kinase substrate in neurons 1; LACC1: Laccase (multicopper oxidoreductase) domain containing 1; FOXP1: Forkhead box P1; HTR2B: 5-hydroxytryptamine (serotonin) receptor 2B, G protein-coupled; KLHL24: Kelch-like family member 24* and *ADRBK4*: *Adrenergic, beta,*

receptor kinase 4. For more details on hypothetical proteins that are adjacent to tyrosinase genes see Fig. S4.

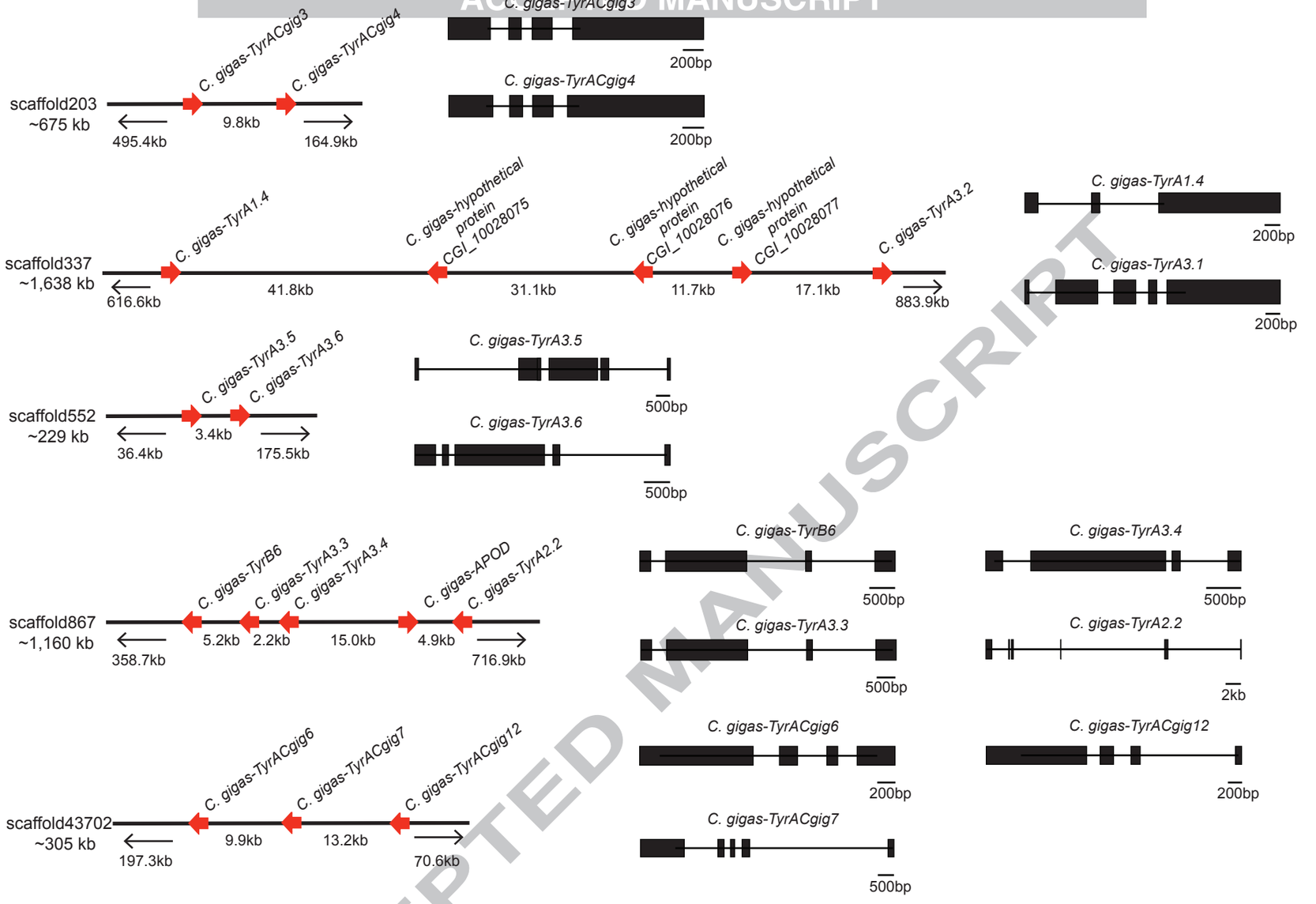
Fig. 5. Relative abundance of tyrosinase genes in the mantle tissue of five bivalves. A) *P. maxima*. B*) P. margaritifera*. C) *P. fucata*. D) *C. gigas*. E) *L. ellitica*. Values are expressed as transcripts per million (TPM) calculated using RSEM software [36] (From panel A to E).

Fig. 6. Comparison of quantitative PCR expression profiles between silver-lipped pearl oyster (*P. maxima*) and black-lipped pearl (*P. margaritifera*). A) Schematic of the internal anatomy of the pearl oyster. The region from which mantle tissue was extracted for qPCR analysis is indicated by red dotted line. (pl: prismatic layer; nl: nacreous layer; ma: mantle; gi: gill; by: byssus; fo: foot; lp: labial palp; dg: digestive gland; li: ligament; go: gonad; he: heart; am; adductor muscle; in: intestine). Adapted from http://journal.goingslowly.com/2010/12/peaceful-ride-on-phu-quoc. B) Schematic of the mantle tissue and shell of the pearl oyster to show sampling zones (of: outer mantle fold: mf: middle fold; if: inner fold). C) Relative expression (log scale) of nine tyrosinase genes. *P. maxima* mean expression is shown in red (N = 4 mantle zones/data points) for four individuals, *P. margaritifera* mean expression is shown in blue (N = 4 mantle zones/data points) for four individuals. See Fig. S5 for individual expression profiles.

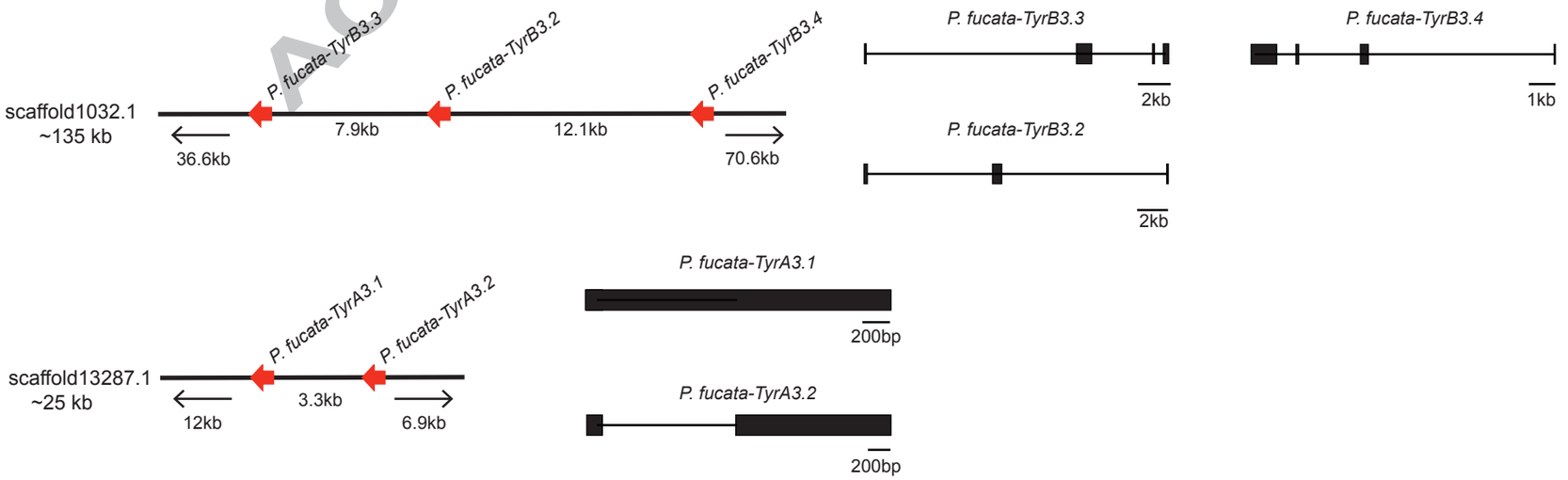

ACCEPTED MANUSCRIPT
Clade B
Clade A
Pinctada expansion
Crassostrea expansion
B1
B2
B3
B4
BPmax/Pmar5
A1
A2
A3
55/-/0.84
P. viridis-TyrB7.1
P. viridis-TyrB7.2
A. farreri-tyrB7
P. viridis-TyrBPvir1
P. viridis-TyrBPvir2
P. maxima-TyrBPmax1
-/76/0.97
P. fucata-TyrB1.1
P. margaritifera-TyrB1
P. maxima-TyrB1.1
P. maxima-TyrB1.2
-/56/0.52
P. fucata-TyrB1.2
P. fucata-TyrB1.3
-/61/0.66
P. fucata-TyrB2
P. margaritifera-TyrB2
P. maxima-TyrB2.1
P. maxima-TyrB2.2
-/-/0.85
P. fucata-TyrBPfuc1
P. fucata-TyrBPfuc2
P. fucata-TyrB3
P. margaritifera-TyrB3
97/82/0.99
P. maxima-TyrB3.1
74/71/0.80
P. maxima-TyrB3.2
P. fucata-TyrB4.1
P. fucata-TyrB4.2
P. fucata-TyrB4.3
P. fucata-TyrB4.4
P. maxima-TyrB4
P. margaritifera-TyrB4.1
P. margaritifera-TyrB4.2
-/-/0.97
88/50/0.60
P. margaritifera-TyrB5.1
85/83/1
P. margaritifera-TyrB5.2
100/83/1
P. margaritifera-TyrB5.3
P. maxima-TyrB5
P. fucata-TyrBPfuc3
P. fucata-TyrBPfuc4
92/890.92
P. fucata-TyrBPfuc5
76/80/1
C. gigas-TyrBCgig1
C. gigas-TyrBCgig2
C. gigas-TyrB6
L. elliptica-TyrB6
-/-/0.61
H. cumingii-TyrAHcum1
H. cumingii-TyrAHcum2
-/-/0.52
H. cumingii-TyrAHcum3
H. cumingii-TyrAHcum4
P. viridis-TyrA1
L. gigantea-TyrA1.1
P. vulgata-TyrA1.1
H. cumingii-TyrA1
88/93/1
C. gigas-TyrA1.1
83/91/0.93
P. fucata-TyrA1
C. gigas-TyrA1.2
P. vulgata-TyrA1.2
L. gigantea-TyrA1.2
84/92/1
C. gigas-TyrA1.3
C. gigas-TyrA1.4
C. gigas-TyrA3.1
C. gigas-TyrA3.2
P. fucata-TyrA3.1
C. gigas-TyrA3.3
P. maxima-TyrA3
78/100/0.95
P. margaritifera-TyrA3
C. gigas-TyrA3.4
P. fucata-TyrA3.2
C. gigas-TyrA3.5
C. gigas-TyrA3.6
C. gigas-TyrA2.1
S. officinalis-TyrA2
I. argentinus-TyrA2.1
I. argentinus-TyrA2.2
-/-/0.79
-/-/0.91
H. cumingii-TyrA2
C. gigas-TyrA2.2
68/66/0.97
P. fucata-TyrA2.1
P. fucata-TyrA2.2
97/82/0.86
P. margaritifera-TyrA2
P. maxima-TyrA2
L. elliptica-TyrALell1
C. gigas-TyrACgig1
C. gigas-TyrACgig2
60/73/-
C. gigas-TyrACgig3
-/-/0.80
C. gigas-TyrACgig4
C. gigas-TyrACgig5
78/71/1
C. gigas-TyrACgig6
89/84/0.98
C. gigas-TyrACgig7
C. gigas-TyrACgig8
76/73/0.52
C. gigas-TyrACgig9
C. gigas-TyrACgig10
C. gigas-TyrACgig11
C. gigas-TyrACgig12
0.2

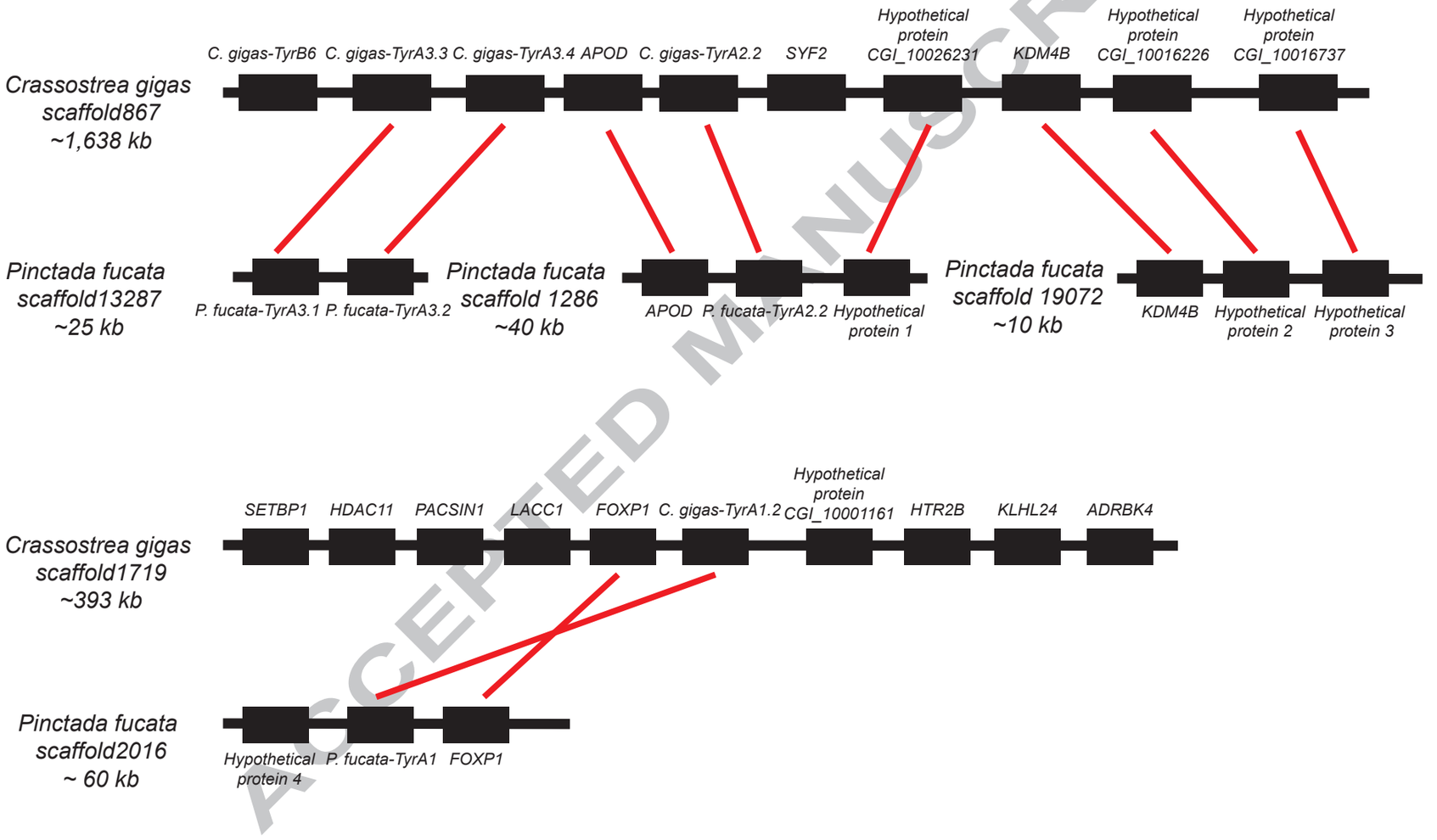


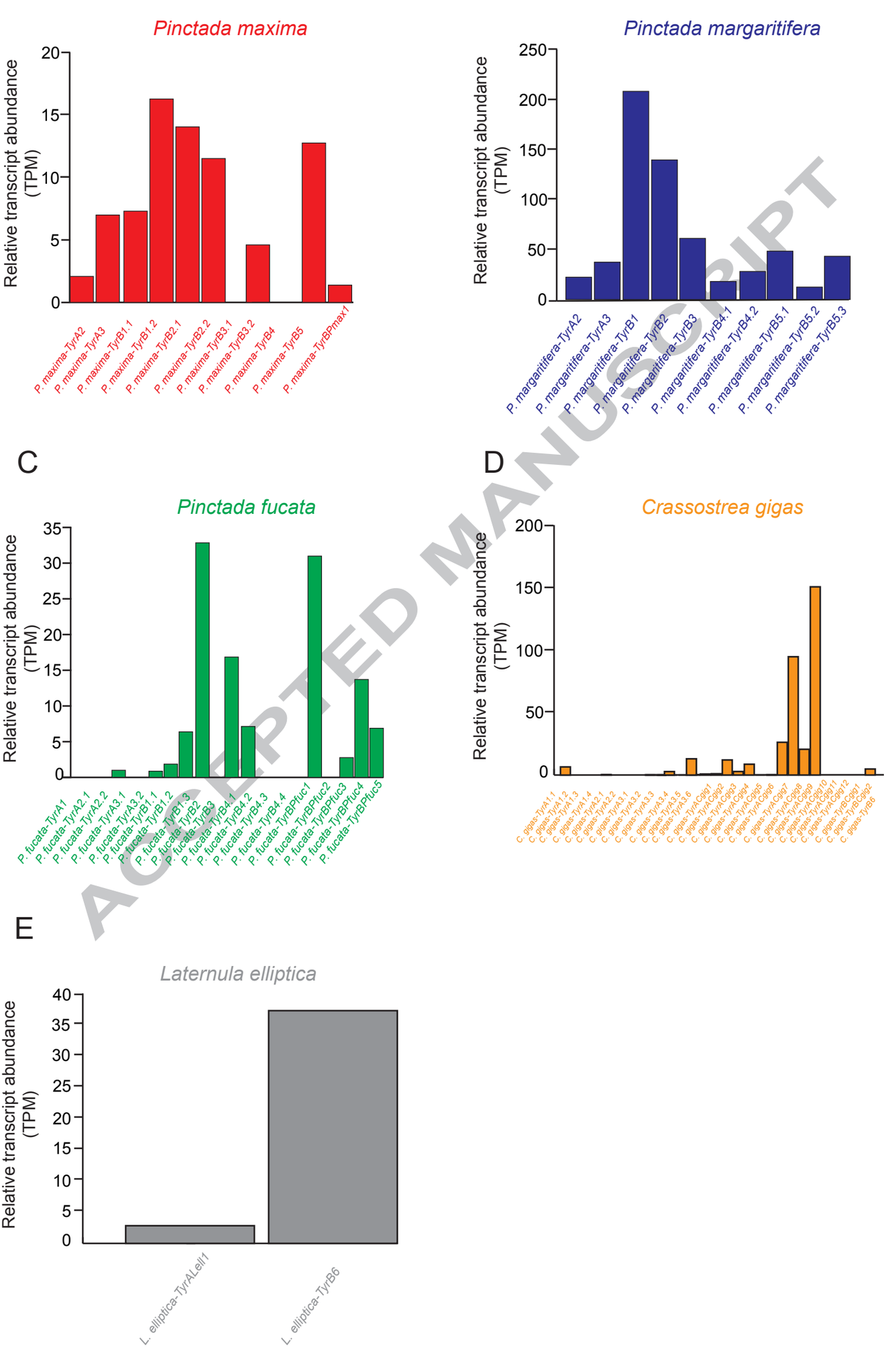
Crassostrea gigas

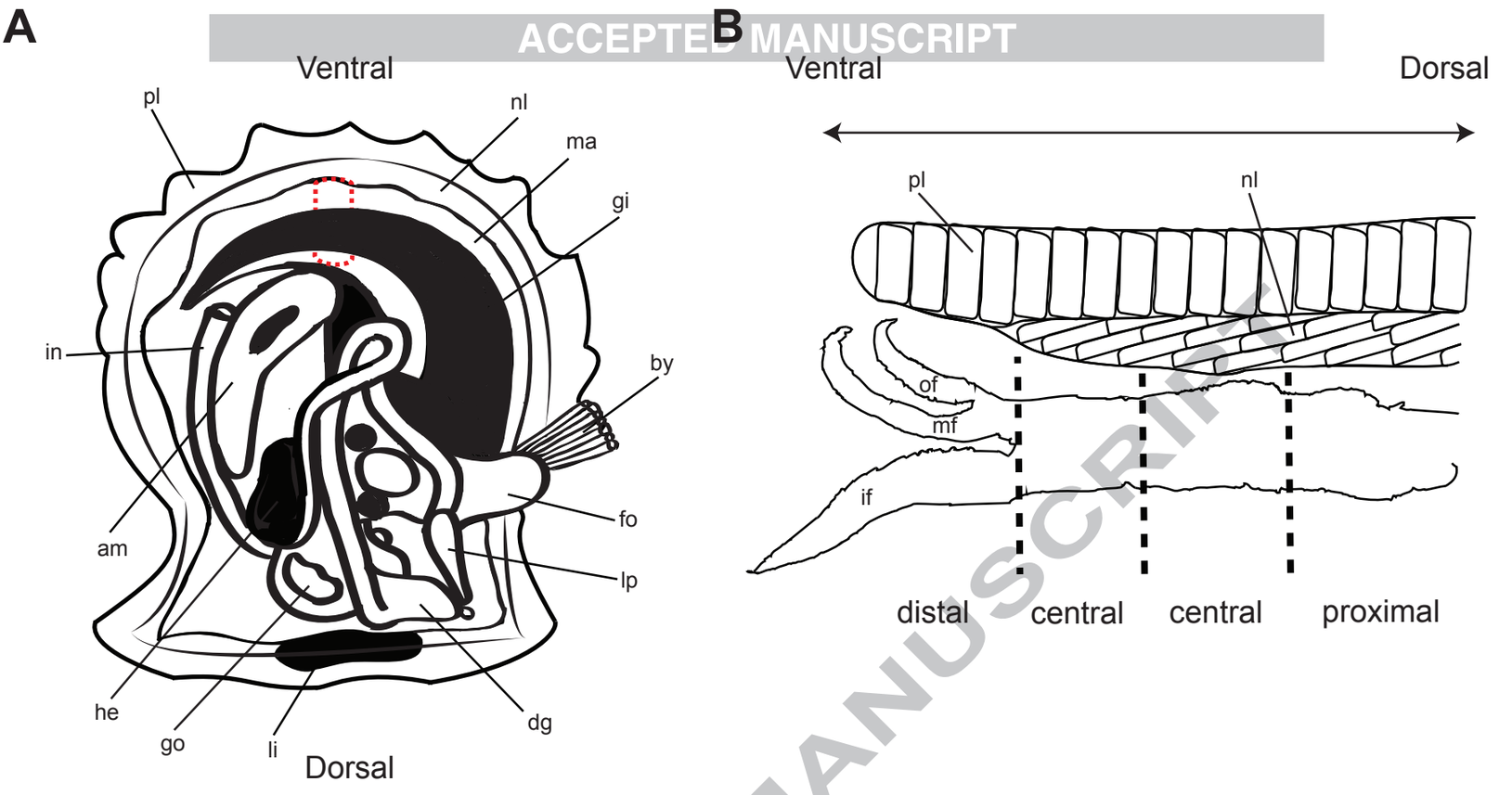


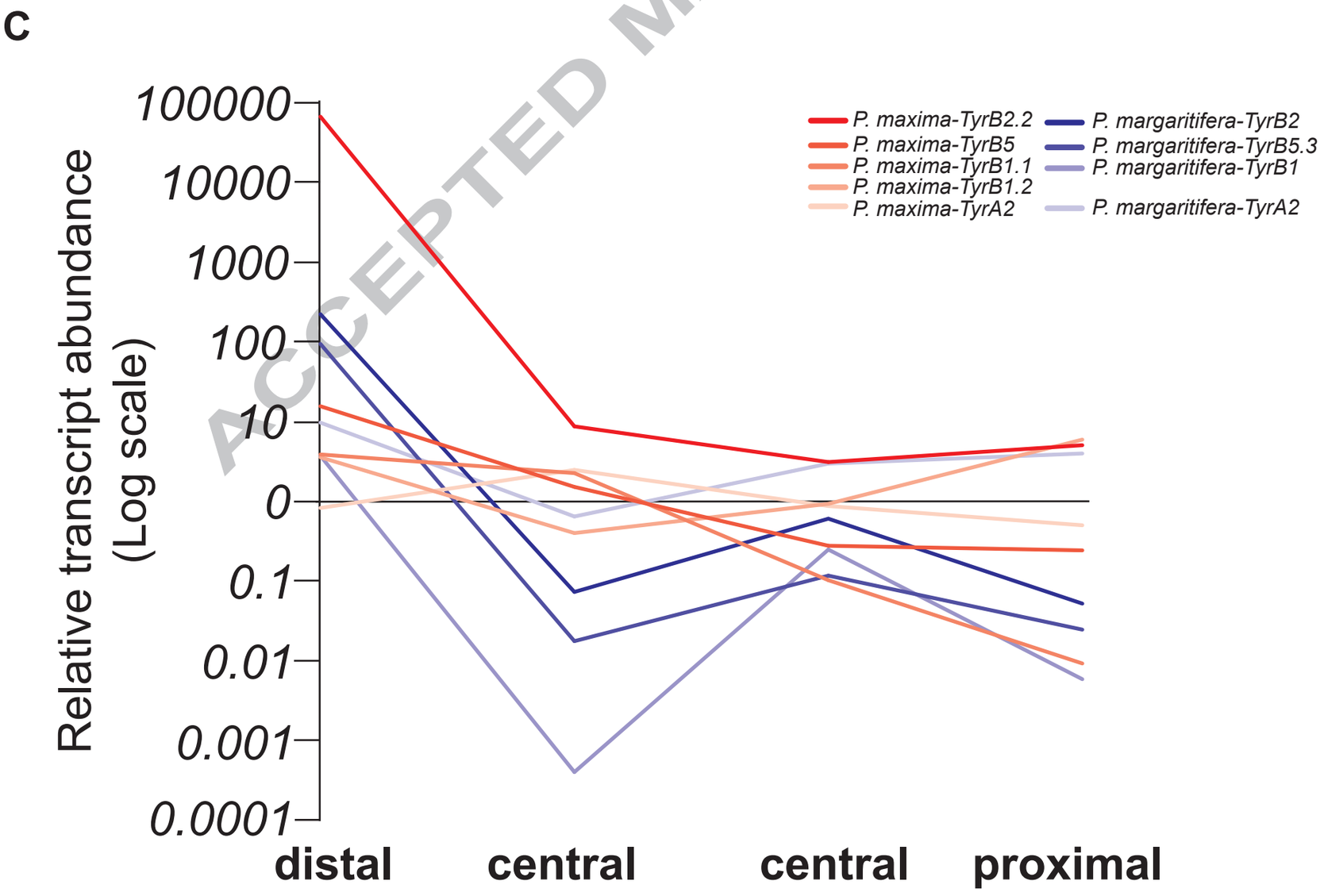
Pinctada fucata





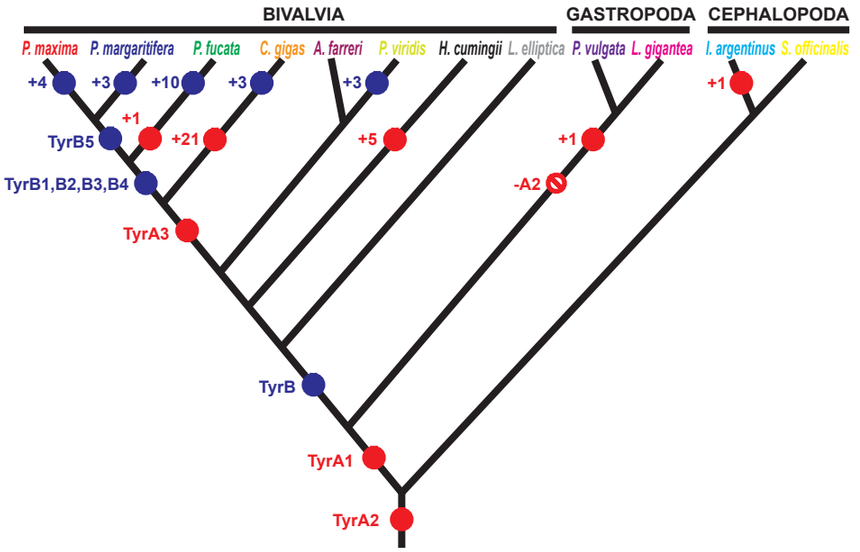






*Graphical Abstract

Evolutionary history of tyrosinase gene family in bivalves



Tyrosinase gene expression across mantle tissue in two pearl oyster species