MOLECULAR ECOLOGY
RESOURCES

Molecular Ecology Resources (2015) **15**, 1205–1215 doi: 10.1111/1755-0998.12383

The red coral (*Corallium rubrum*) transcriptome: a new resource for population genetics and local adaptation studies

M. PRATLONG,*,† A. HAGUENAUER,* O. CHABROL,† C. KLOPP,‡ P. PONTAROTTI† and D. AURELLE*

*Aix Marseille Université, CNRS, IRD, Avignon Université, IMBE UMR 7263, 13397, Marseille, France, †Aix Marseille Université, CNRS, Centrale Marseille, I2M UMR 7373, Equipe Evolution Biologique et Modélisation, 13453, Marseille, France, ‡Plateforme Bioinformatique Toulouse Midi-Pyrénées, UR 875 UMIAT, INRA, Auzeville Castanet-Tolosan, France

Abstract

The question of species survival and evolution in heterogeneous environments has long been a subject for study. Indeed, it is often difficult to identify the molecular basis of adaptation to contrasted environments, and nongenetic effects increase the difficulty to disentangle fixed effects, such as genetic adaptation, from variable effects, such as individual phenotypic plasticity, in adaptation. Nevertheless, this question is also of great importance for understanding the evolution of species in a context of climate change. The red coral (*Corallium rubrum*) lives in the Mediterranean Sea, where at depths ranging from 5 to 600 m, it meets very contrasted thermal conditions. The shallowest populations of this species suffered from mortality events linked with thermal anomalies that have highlighted thermotolerance differences between individuals. We provide here a new transcriptomic resource, as well as candidate markers for the study of local adaptation. We sequenced the transcriptome of six individuals from 5 m and six individuals from 40 m depth at the same site of the Marseilles bay, after a period of common garden acclimatization. We found differential expression maintained between the two depths even after common garden acclimatization, and we analysed the polymorphism pattern of these samples. We highlighted contigs potentially implicated in the response to thermal stress, which could be good candidates for the study of thermal adaptation for the red coral. Some of these genes are also involved in the response to thermal stress in other corals. Our method enables the identification of candidate loci of local adaptation useful for other nonmodel organisms.

Keywords: cnidarians, gene expression, local adaptation, octocoral polymorphism, transcriptome

Received 11 August 2014; revision received 22 January 2015; accepted 26 January 2015

Introduction

Marine species usually deal with more or less pronounced gradients of environmental conditions (temperature, salinity, etc.) along their natural range. These species might have found adaptive responses to maximize their fitness in all environmental conditions they encounter. In this context, two types of mechanisms could occur. First, in the absence of other evolutionary pressures (migration, genetic drift) and if the environmental gradient is persistent for an extended period of time, each local population submitted to local selection could become genetically adapted to the corresponding local environmental conditions (Kawecki & Ebert 2004). On the other hand, during their life, individuals can develop physiological adaptations to their local environmental conditions. This acclimatization (or phenotypic plasticity) is often reversible at short term, but can be in some cases, maintained at medium to long term (Pespeni *et al.* 2013). In this context, we are interested in understanding the ability of individuals from contrasted environments to revert to a common basal state and in knowing whether they developed adaptive mechanisms stable at medium or long term. Two experimental approaches are mainly used for the study of local adaptation: reciprocal transplant and common garden experiments. Recent experimental studies of local adaptation in marine environments generally focus on the identification of differentially expressed genes between individuals from different environments and submitted to the same stress (Barshis *et al.* 2013; Bay *et al.* 2013; Haguenauer *et al.* 2013; Pespeni *et al.* 2013). In these cases, few candidates' loci are commonly used (Bay *et al.* 2013; Haguenauer *et al.* 2013). Nevertheless, transcriptomic patterns of individuals from different environments but acclimatized to a basal state in common garden condi-

Correspondence: Marine Pratlong, Fax: +33 4 91 04 16 35;
E-mail: marine.pratlong@imbe.fr

© 2015 John Wiley & Sons Ltd

tions remain poorly researched. If two individuals from different environmental conditions are kept in common garden conditions, expression differences due to phenotypic plasticity should tend to decrease in time, while fixed effects stemming from genetic adaptation, epigenetic or developmental modifications with medium to long-term stability will be maintained (Barshis *et al.* 2013; Palumbi *et al.* 2014). Furthermore, most studies treated separately the question of differential expression and sequence polymorphism, and these two aspects are rarely studied in the same framework (De Wit & Palumbi 2013). Differential expression analysis and expressed sequences polymorphism have different molecular and evolutionary origins. Indeed, these two aspects may bear traces of adaptation (broad sense) and taking them into account may contribute to our global understanding of this complex mechanism. Here, we propose to study local adaptation and acclimatization along an environmental gradient based on transcriptomic data for the red coral (*Corallium rubrum*). This species is an octocoral, an ecologically important clade but less studied than hexacorals. It lives in the Mediterranean Sea and eastern Atlantic between 5 m and 800 m depth (Costantini *et al.* 2011) where it meets very contrasted thermal conditions. Shallow populations are exposed to high maximum temperatures and to frequent and intense temperature fluctuations. These two parameters tend to decrease with increasing depth. Previous studies also showed that the red coral is an especially low disperser, with populations separated by fewer than 10 m detected as genetically different (Ledoux *et al.* 2010a,b). Furthermore, experiments have shown that individuals from different depths (20–40 m) of the same site have contrasted responses to thermal stress, in terms of respiration, growth, calcification and necrosis (Torrents *et al.* 2008; Ledoux *et al.* 2010a,b). The study of candidate genes in experimental conditions of thermal stress revealed the differential expression of HSP70 between individuals from different depths of the same site (Haguenauer *et al.* 2013). This study is a logical continuation of this previous research concerning the local adaptation of the red coral. As such, it is now necessary to use new molecular approaches to try to identify genes that could be good candidates for the study of local adaptation, because of their sequence polymorphisms or expression pattern, without preconceived ideas concerning candidate genes. To this end, we sequenced the transcriptome of individuals from two depths of the same site in the bay of Marseilles, after three months of in situ common thermal conditions and a one-month aquarium acclimatization. Our objective here was to evidence underlying expression differences stable at medium term between individuals from different depths of the same site. We also studied the patterns of sequence polymorphism according to depth. The general goal was to use these new transcriptomic resources to highlight candidate loci for the study of local adaptation in heterogeneous environment. This method could be extended to other nonmodel organisms. Additionally, this sequence database will be useful for the study of the response of other octocoral species to thermal stress.

Material and methods

Sampling and RNA extraction

Red coral (*Corallium rubrum*) colonies were collected from two areas near the city of Marseilles, the 5 m population of Figuier cave (FIG; 43°12.330′N, 5°26.790′E) and the 40 m population of Cap Morgiou (MOR; 43°12.060′N, 5°27.060′E) separated by 622 m of horizontal distance, with six individuals sampled at each depth. These two sampling depths are characterized by contrasted thermal conditions from April to the end of September (Haguenauer *et al.* 2013). The population from Figuier is the shallowest red coral population near Marseilles. These individuals are submitted to a large range of thermal conditions in spring and summer (from 15 to 26 °C), to fast and daily increase or decrease of temperature during this period, and are exposed to more days with temperatures above 20 °C than the deeper population from Morgiou (Haguenauer *et al.* 2013). Individuals from Morgiou are submitted to a narrower thermal range (from 14 to 20 °C), to less frequent thermal variations, and the temperature stays generally under 20 °C. From October to March, the temperature and others physical and chemical parameters are similar between these two depths because of a seasonal homogenization of the water column (D'Ortenzio *et al.* 2014). During this period, the thermal homogenization reaches a common minimal value of around 13 °C (Haguenauer *et al.* 2013). Furthermore, because the red coral does not bear photosynthetic symbionts, individuals from 5 m to 40 m should not be differentially impacted by light exposition differences. Sampling took place in December, when colonies from 5 m to 40 m had spent 3 months in common thermal conditions. Six individuals were sampled from each site and maintained during one additional month in common conditions in a semiclosed water circulation system with a temperature between 14 and 15 °C. After this acclimatization period, total RNA was extracted.

Library preparation and sequencing

Total RNA from each sample was purified as previously described in Haguenauer *et al.* (2013). Residual DNA was digested using TurboDNAse (Ambion) following the manufacturer's instructions. RNA samples were pooled in two duplicates for each depth (each duplicate

© 2015 John Wiley & Sons Ltd

containing the same quantity of RNA from three different individuals coming from the same depth). RNA-Seq libraries were generated using the TruSeq RNA-Seq Sample PREP KIT v2 according to the manufacturer's protocol. During this preparation, libraries were individually tagged to pool them before sequencing. Library sizes were controlled on analysis with a Bioanalyzer 2100 (Agilent Technologies) on high-sensitivity chips. The four libraries were quantified by qPCR following the manufacturer's protocol. Libraries were finally pooled before sequencing on one lane on the Illumina HiSeq2000 (Illumina Inc., San Diego, CA) as paired-end reads of length 100 bp with the TruSeq PE kit. Library preparation and sequencing were performed at the GENOTOUL platform (http://get.genotoul.fr/).

Assembly

One assembly was run for each condition (4 assemblies were performed in total). Raw sequence data were first filtered to remove unknown nucleotides. If a read contained Ns, it was split in subsequences without Ns and the longest of these subsequences was kept if its length exceeded half of the initial read length (a subsequence is a part of a read with no unknown nucleotides). Subsequently, because raw data produced are not filtered for low-quality reads, we applied the FASTQ_ILLUMINA_FILTER (v.0.1 available at http://cancan.cshl.edu/labmembers/gordon/fastq_illumina_filter/). Finally, duplicated reads were discarded and only unique reads were kept to perform assembly. The transcriptome de novo assembly was performed using OASES (v.0.2.06; Schulz *et al.* 2012). Nine assemblies using nine different k-mers (25, 31, 37, 43, 49, 55, 61, 65 and 69) were performed on preprocessed input data. We chose to keep only the best contig for each locus with a script developed by a Brown University team (available at https://sites.google.com/a/brown.edu/bioinformatics-in-biomed/velvet-and-oases-transcriptome). After that, contigs given by all k-mers were merged. Finally, antisense chimeras (accidentally produced by the assembly step) were cut. Then, because similar collection of contigs were produced by close k-mers, a CD-HIT-EST clustering step (v.4.6; Li & Godzik 2006) grouped similar contigs into clusters based on their sequence similarities (identity equal or greater than 0.95). TGICL (v.2.1; Pertea *et al.* 2003), an OLC (overlap layout consensus) assembler, clusterized sequences sharing significant fragments. After this assembly process, all input reads were mapped back to rebuilt contigs using BWA (v.0.7.0-r313; Li & Durbin 2009). For each contig, the longest ORF (region that is free of STOP codons) was extracted with the getorf EMBOSS TOOL (v.6.4.0.0; Rice *et al.* 2000) and contigs with longest ORF lower than 200 bp (66 aa) or very low coverage (less than 2 mapped reads for 1 million overall mapped reads) were discarded.

Meta-assembly

The second step of the assembly process was a meta-assembly resulting in the production of a single reference contig set. Contig fasta files coming from the four conditions were first concatenated. ORFs with sequence identity equal or greater than 0.9 were clusterized using CD-HIT (v.4.6; Li & Godzik 2006). Then, the contig with the longest ORF (or the longest contig if several contigs have an ORF of the maximal size) of each CD-HIT cluster was extracted. A CD-HIT-EST job was performed to clusterize remaining contigs with sequence identity equal or greater than 0.95. This step will be able to clusterize close contigs inside which potential frameshifts abort ORF detection. Finally, input reads from all conditions were mapped to selected contigs using BWA (v.0.7.0-r313; Li & Durbin 2009), and contigs with longest ORF lower than 200 bp (66 aa) or very low coverage (less than 2 mapped reads for 1 million overall mapped read) were discarded.

Comparative analysis

To compare the transcriptome of the red coral with those of other cnidarians, we performed a BLASTP (Altschul *et al.* 1990) search in the following cnidarian EST libraries: *Acropora palmata*, *Acropora millepora*, *Pocillopora damicornis*, *Stylopora Pistillata*, *Porites Astreoides*, *Monstastrae faveolata*, *Aiptasia Pallida*, *Anemonia Viridis*, *Metridium senile*, *Nematostella vectensis*, *Edwarsiella lineata*, *Paramuricea clavata*, *Eunicella verrucosa*, *Eunicella cavolinii*, *Gorgonia ventalina*, *Leptogorgia sarmentosa*, *Hydractinia echinata*, *Hydra vulgaris*, *Clytia hemisphaerica* (Table 1). We applied an e value cut-off of 10^{-10}.

Transcriptome annotation

All contigs were annotated using a two steps method which allowed to maximize the reliability of our annotation. For all contigs, we found the longest open reading frame using the getorf EMBOSS TOOL (Rice *et al.* 2000). First, we used orthoMCL (Li *et al.* 2003) to search for potential orthologous reciprocal relationships between the translated contigs, sequences of seven very well-annotated transcriptomes (*Adineta vaga*, *Arabidopsis thaliana*, *Caenorhabditis elegans*, *Drosophila melanogaster*, *Homo sapiens*, *Saccharomyces cerevisiae* and *Strongylocentrotus purpuratus*) and those of two transcriptomes of cnidarian species (*Hydra magnipapillata*, *Nematostella vectensis*). We applied an e value cut-off of 10^{-5} and we transferred the annotation

© 2015 John Wiley & Sons Ltd

Table 1 Number of ESTs used for comparative analysis of each species and number of BLASTP hits between the red coral transcriptome and the EST bank of each species (*e* value = 10^{-10})

Classification	Species	Source	Number of ESTs	Number of contigs
Cnidaria	*Acropora palmata*	NCBI	43 150	11 671
Anthozoa	*Acropora millepora*	Moya *et al.* (2012)	52 958	18 108
Hexacorallia	*Pocillopora damicornis*	Traylor-Knowles & Palumbi (2014)	70 786	14 716
Scleractinia	*Stylopora Pistillata*	Karako-Lampert *et al.* (2014)	21 810	16 664
	Porites Astreoides	NCBI	11 516	4753
	Monstastrae faveolata	NCBI	33 226	9606
Cnidaria	*Aiptasia Pallida*	NCBI	10 295	6391
Anthozoa	*Anemonia Viridis*	NCBI	39 939	10 535
Hexacorallia	*Metridium senile*	NCBI	29 412	9815
Actinia	*Nematostella vectensis*	NCBI	163 314	14 906
	Edwarsiella lineata	Stefanik *et al.* (2014)	90 440	17 166
Cnidaria	*Paramuricea clavata*	K. Mokthar-Jamaï, personal communication	6529	4341
Anthozoa	*Eunicella verrucosa*	Romiguier *et al.* (2014)	15 461	15 130
Octocorallia	*Eunicella cavolinii*	Romiguier *et al.* (2014)	20 731	17 982
	Gorgonia ventalina	Burge *et al.* (2013)	90 230	20 172
	Leptogorgia sarmentosa	Romiguier *et al.* (2014)	13 299	14 343
Cnidaria	*Hydractinia echinata*	NCBI	9464	2860
Medusozoa	*Hydra vulgaris*	NCBI	184 731	12 324
	Clytia hemisphaerica	NCBI	85 991	10 643
	Aurelia aurita	www.compagen.org	62 632	9836

from these orthologs, when available, to *C. rubrum* sequences. For all contigs which could not be annotated by this method, we performed BLASTP searches against the UniProt protein database (Altschul *et al.* 1990; The UniProt Consortium 2013). An e value cut-off of 10^{-5} was applied and we reported only the best hit of each query sequence. Among these UniProt hits, we listed the non-metazoan hits to avoid the bias of contaminations in our analysis. In all cases, we used the UniProt identifier to assign each contig to Gene Ontology biological process categories (Ashburner *et al.* 2000) associated with the predicted protein. Annotation results were summarized using Generic GO-Slim, a cut-down version of Gene Ontology containing a small number of relevant functional categories (http://www.geneontology.org/ontology/subsets/goslim_generic.obo).

Differentially expressed genes

To identify differentially expressed genes between the two depths, we compared the results of three packages in R (R Development Core Team 2008): EDGER (Robinson *et al.* 2010) and DESEQ (Anders & Huber 2010) using parametric statistical methods and NOISEQ (Tarazona *et al.* 2011) using nonparametric statistical method. In each case, we used parameters described as optimal by the authors. EdgeR was performed using the TMM (trimmed mean of M values) normalization method (Robinson & Oshlack 2010) and a common negative binomial dispersion parameter for the variance estimation. In the case of NOISeq, we applied the RPKM (reads per kilobase per million) normalization method (Mortazavi *et al.* 2008) and we used the noiseqbio function which is appropriate to treat biological replicates. In each case, we corrected p values for false discovery rate using the Benjamini–Hochberg procedure (Benjamini & Hochberg 1995). As our aim was to find the best candidates which could play a role in the depth differentiation, we chose to consider only genes detected as differentially expressed with an adjusted p value lower than 0.01 by these three statistical methods to be as stringent as possible. To identify functional categories enriched in our differentially expressed genes, we performed Fisher's exact tests and applied the false discovery rate correction (Benjamini & Hochberg 1995).

Detection and analysis of SNPs

The alignment files were first filtered to remove all PCR duplicates (samtools rmdup, Li *et al.* 2009), and all reads were aligned in multiple locations. The reads were then realigned and recalibrated using GATK (v.2.4-9, standard parameters; McKenna *et al.* 2010). Finally, all files were used to call high-quality SNPs and INDEL with GATK (same version, UnifiedGenotyper method; DePristo *et al.* 2011) using a minimum Phred quality score of 30 (Q30, corresponding to an error rate in base calling lower than 0.1 %). From this SNP file, we searched for polymorphism patterns potentially linked with depth. Our experimental framework did not allow us to precisely estimate SNP

© 2015 John Wiley & Sons Ltd

frequencies as individuals were pooled before sequencing and observed frequencies may also depend on expression levels. Therefore, we focused on the extreme case of differentially fixed SNPs (i.e. SNPs with the same allele for all individuals from the same depth and with an alternative allele for all individuals from the other depth). At this stage, to minimize errors in the detection of SNPs, we added a filter for the quality of each genotype. GATK variant calling generates the genotype quality (GQ) value that corresponds to the Phred-scaled confidence that the inferred genotype is true. We chose to apply the commonly used threshold of GQ = 20 and to discard low depth genotypes (DP < 12) in order to keep only high-quality genotypes. We tested for enriched functional categories within contigs containing differentially fixed SNPs and nonsynonymous SNPs using Fisher's exact test, and we applied the false discovery rate correction (Benjamini & Hochberg 1995).

Results

Transcriptome description

De novo assembly of reads resulted in 48 074 contigs with a mean length of 1 814 bp and N50 of 2 470 bp (Fig. 1). The Table 2 sums up the number of reads and alignments at different stages of the assembly process. The GC content of the *Corallium rubrum* transcriptome was 39.45 %. Our annotation method allowed us to successfully annotate 16 951 contigs (36 % of the red coral transcriptome). Among these contigs, 64 % could be annotated through orthoMCL. Among the 6187 annotated contigs which failed to be annotated through orthoMCL, the majority (73 %) had a hit with species we did not use in the orthoMCL analysis. The last 27 % could be sequences not present in the transcriptomes used in orthoMCL analysis or eliminated by the orthoMCL filters, as the orthoMCL algorithm is more stringent than a simple BLAST. Among contigs which failed to be annotated through orthoMCL, 426 had nonmetazoan UniProt hits, of which half (213) had similarities with bacteria. These nonmetazoan hits could be due to contaminations but also to horizontal transfers. In any cases, we chose not to include contigs with nonmetazoan hits in further analyses. Functional categories of annotated contigs were distributed into the 55 categories of biological processes of the Generic GO-Slim, but only eight categories contained more than 50 % of functions assigned to contigs (Fig. 2): cellular nitrogen compound metabolic process (12 %), biosynthetic processes (10 %), signal transduction (6 %), cellular protein modification process (6 %), transport (6 %), small molecule metabolic process (5 %), catabolic process (4 %), anatomical structure development (4 %). Twenty-nine categories were represented by less than 1% of functions attributed to contigs. The low number of data available on UniProt for species phylogenetically close to the red coral could partly explain the absence of significant hit for 31 123 sequences. Moreover, the mean length of these nonannotated sequences was lower than for the annotated ones ($p < 2.2 \times 10^{-16}$) and could correspond to noncoding RNA or technical artefacts (Fig. S1, Supporting information).

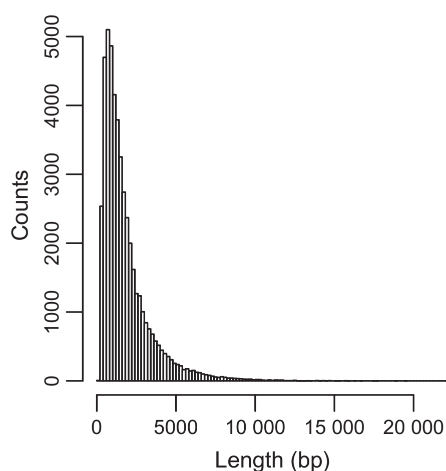

Fig. 1 Size distribution of contig lengths (bp).

Comparative analysis

Fourteen thousand one hundred and thirty-five of the red coral's contigs had homologues within the four clades of cnidarians (Fig. 3, Table 1). The highest number of clade-specific homologues was found in octocorals

Table 2 Number of reads and alignments at different stages of the assembly process. The Q30 (Phred quality score of 30) corresponds to an error rate in base calling lower than 0.1 %

Sample	Number of reads pairs	Number of alignments	Number of alignment after Q30 filtering	Number of alignment after PCR duplicates removing
FIG5A	57446789	114893578	47979281	15155126
FIG5B	51612319	103224638	43854431	14846834
MOR40A	69032155	138064310	58804098	17071186
MOR40B	62839889	125679778	55166310	16897159

© 2015 John Wiley & Sons Ltd

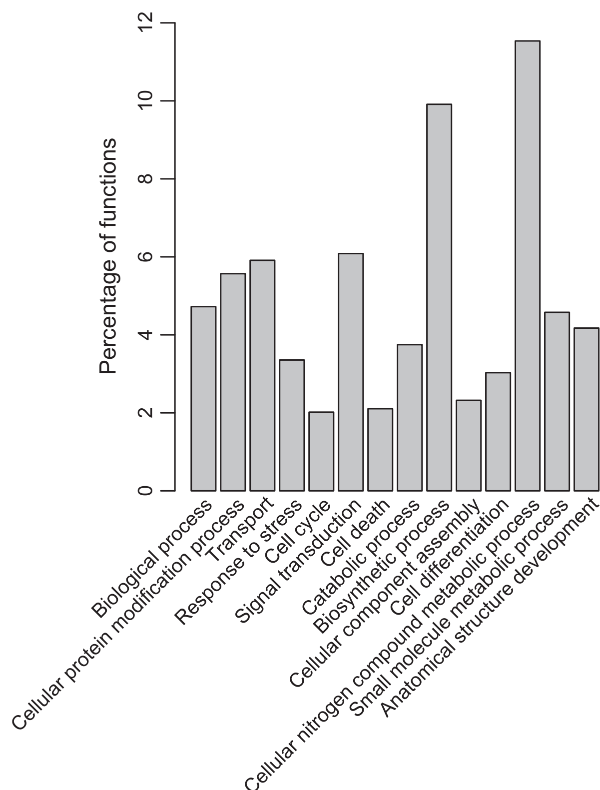

Fig. 2 Repartition of the functions attributed to contigs into the biological process GO functional categories. Functional categories represented by less than 2 % of functions attributed to contigs were not reported. The functional category 'biological process' represents the functions belonging to the biological process ontology, but which could not be classified in the categories presented here.

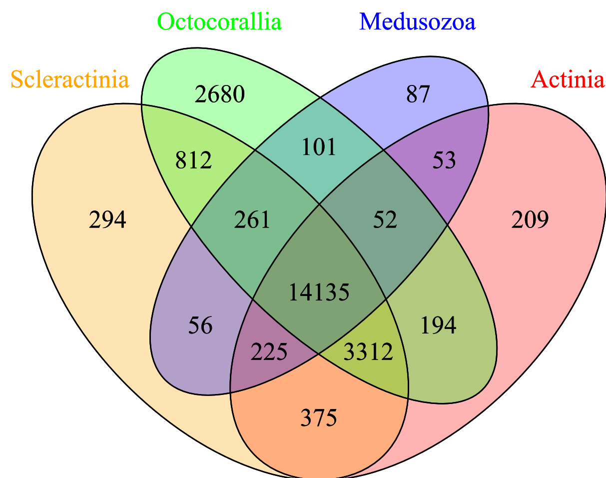

Fig. 3 Venn diagram of red coral's homologues with the cnidarians (Scleractinia, Actinia, Octocorallia, Medusozoa).

(2680 contigs), followed by scleractinians (294 contigs), actinians (209 contigs) and medusozoans (87 contigs) (Fig. 3, Table 1).

Analysis of depth specific expression patterns

Expression patterns of our 48 074 contigs were highly similar between the two pools of the same depth, but also between sets of different depths: the Pearson's correlation coefficient was ranged from 0.94 to 0.96 ($p < 0.0001$) for all comparisons (Table 3). Despite this correlation, 212 contigs were detected as differentially expressed with the three methods (FDR $\alpha = 0.01$, Table 4, Fig. S2, Supporting information). Among these 212 contigs, 92 had higher expression at 5 m and 120 had higher expression at 40 m (Table 5). The level of overexpression was greater at 40 m than at 5 m ($p = 1.37 \times 10^{-6}$, Table 5). We found UniProt hits for 77 of these 212 contigs (Table S1, Supporting information), but only 41 could be assigned to 18 Gene Ontology functional categories. Four of these categories contained more than 50 % of functions assigned to contigs: cellular protein modification process (16 %), biosynthetic process (13 %), signal transduction (11 %) and cellular nitrogen compound metabolic process (11 %). Among the functions assigned to the 41 differentially expressed contigs, two were detected as enriched by the functional enrichment analysis with a threshold of 0.05: cellular protein modification process ($p = 0.0054$) and cell death ($p = 0.048$). None of these functional categories remained significantly enriched after the false discovery rate correction (FDR $\alpha = 0.01$). The best informative UniProt hit for the 77 differentially expressed contigs annotated is listed in Table S1 (Supporting information). Among them, we can highlight several noteworthy functions because they are classified into the two Gene Ontology functional categories 'cellular protein modification process' and 'cell

Table 3 Pearson correlation coefficient of population expression patterns

	FIG 5A	FIG 5B	MOR40A	MOR40B
FIG 5A	1			
FIG 5B	0.96	1		
MOR40A	0.94	0.96	1	
MOR40B	0.94	0.95	0.96	1

Table 4 Number of contigs detected as differentially expressed between depths by each R package. The total row indicates contigs detected as differentially expressed by the three methods

	Differentially expressed contigs (FDR $\alpha = 0.01$)
DESeq	358
NOISeq	809
EdgeR	477
Total	212

© 2015 John Wiley & Sons Ltd

Table 5 Number of contigs overexpressed at each depth and their average fold change

Overexpression	Average fold change (x)				Total
	$x < 2$	$2 < x < 10$	$10 < x < 100$	$x > 100$	
Overexpressed at 5 m	4	58	24	6	92
Overexpressed at 40 m	0	39	63	18	120

death' mentioned before and potentially involved in corals adaptation (see discussion). Five contigs were homologous to tumour necrosis factor receptor-associated factors (TRAFs), and all were overexpressed at 5 m. A homologue of TNFAIP3 interacting protein was overexpressed at 40 m. Two homologues of E3 ubiquitin protein ligase were overexpressed at 40 m and one at 5 m. Finally, we can cite a lectin homologue and a F-type lectin homologue overexpressed at 5 m. Among the 134 contigs differentially expressed without UniProt hit, only 18 had homologues in the cnidarians data sets used for the comparative analysis (Table 1). Twelve of these contigs had hits only with octocorals.

Polymorphism analysis

The SNP detection allowed the identification of 691 940 SNPs within 43 474 contigs (5000 contigs were not polymorphic). Two of these SNPs presented four alleles, 1715 presented three alleles, and all others SNPs had two alleles. The frequency of SNPs in the transcriptome was one per 126 bp. Although the identified SNPs had very good quality, the genotype quality was poor (Fig. S3, Supporting information). Only 237 681 SNPs passed the genotype quality filter GQ = 20 and DP = 12. 73 % (173 825 SNPs) of these SNPs with high-quality genotypes were in the longest open frame of the corresponding contig and 31 % were nonsynonymous mutation (72 954 SNPs). Among these SNPs, we found 56 which were differentially fixed SNPs and which corresponded to 46 contigs. Thirty-six (63 %) of these differentially fixed SNPs were in the longest open reading frame of the corresponding contig. Among these 36 SNPs, 19 (distributed into 15 contigs) presented nonsynonymous mutations. There was no enrichment of nonsynonymous mutations among differentially fixed SNPs compared to the global set of SNPs with high-quality genotype ($p = 1$). The best informative UniProt hits (when it was available) for these contigs containing differentially fixed SNPs are listed in Table S2 (Supporting information). Only 19 of them could be assigned to 21 Gene Ontology functional categories. The functional categories the more represented were cellular nitrogen compound metabolic process (14 %), cellular protein modification process (11%) and transport (9%). No enrichment for any functional category was detected by the functional enrichment analysis. Twelve differentially expressed contigs (including a lectin homologue and two of the five TRAFs homologues previously mentioned) contained at least one differentially fixed SNP before the correction for genotype quality, but none of them were detected by the two approaches after that correction.

Discussion

Our results enabled us to explore the molecular basis of adaptation to local environments for the red coral and to complement previous results on adaptive diversity in this species (Haguenauer *et al.* 2013). Indeed, the transcriptome variability according to depth was yet to be studied for this species, in terms of both expression levels and sequence polymorphism. More generally, our study is the first to compare transcriptomic patterns between individuals from contrasted environmental conditions for a temperate octocoral. Extending such approaches to an octocoral in a domain where hexacorals are mostly studied is an important step in the study of the evolution of adaptive processes among cnidarians. The choice of a cnidarian that does not bear algal symbionts (*Symbiodinium*) also allowed us to overcome the effect of algal symbionts on stress response and on adaptation to local environment.

Transcriptome comparative analysis

Our comparative analysis allowed us to check for the congruence of the red coral transcriptome among other EST projects. As expected, the highest number of clade-specific homologues corresponded to octocorals. We found more clade-specific homologues with hexacorals (scleractinians and actinians) than with medusozoans. These relationships correspond to the most usual phylogeny of cnidarians (Collins 2002; Technau *et al.* 2005) although two recent papers uncovered a closer relationship of octocorals with medusozoans rather than with hexacorals (Park *et al.* 2012; Kayal *et al.* 2013). This analysis would therefore support the usual phylogenetic relationships within Cnidaria, but we could not exclude biases associated with the size and content of the EST libraries compared here.

© 2015 John Wiley & Sons Ltd

Biological functions involved in coral adaptation to thermal stress

We evidenced differentially expressed contigs between individuals from different depths in basal conditions (i.e. without thermal stress). These genes suggest an effect of environmental conditions at each depth acting on expression regulation either through genetic adaptation or acclimatization allowed by phenotypic plasticity at the transcriptomic level. Several functions identified here as differentially expressed between depths have been experimentally demonstrated as involved in other cnidarians' response to thermal stress, thereby supporting their potential involvement in adaptive response (broad sense) (Rodriguez-Lanetty *et al.* 2006; Wood-Charlson *et al.* 2006; Kvennefors *et al.* 2008; Davy *et al.* 2012; Barshis *et al.* 2013; De Wit & Palumbi 2013; Vidal-Dupiol *et al.* 2013). These genes could also constitute potential stress markers for experimental or in situ studies. Among these genes, we identified five homologues of tumour necrosis factor receptor-associated factors (TRAFs), all overexpressed at 5 m. These genes play a role in several signalling pathways in cell death, survival and mechanisms of cellular response to stress such as innate immunity or apoptosis in other metazoans (Arch *et al.* 1998; Bradley & Pober 2001). This gene family has been documented as a marker of acclimatization to thermal stress in the hexacoral *Acropora hyacinthus* by Barshis *et al.* (2013). These authors observed a higher basal level (i.e. before experimental stress) of expression of TRAFs homologues in *A. hyacinthus* individuals from thermally variable environments compared with more stable ones. Moreover, Palumbi *et al.* (2014), using reciprocal transplants, highlighted an overexpression of TRAF homologues in highly variable environments compared to moderately variable ones. Our results mirror these observations and extend them to temperate octocorals: the basal overexpression of these genes might be a conserved response of cnidarians to thermal stress.

An homologue gene of the TNFAIP3 interacting protein, also called ABIN, was overexpressed at 40 m in red coral. Overexpression of ABIN has been shown to inhibit NF-κB activation by tumour necrosis factor (Verstrepen *et al.* 2009). De Wit & Palumbi (2013) found a gastropod homologue of TNFAIP3 interacting protein as outlier based on sequence polymorphism. Nevertheless, we did not find any differentially fixed SNPs in the red coral homologue of TNFAIP3. This description of the TNFAIP3 as a potential candidate locus for the study of spatial adaptation, although interesting, requires further studies to be considered.

The lectins have been described in all metazoans and are involved in cell–cell interactions and pathogen recognition (Weis *et al.* 1998). Concerning cnidarians, lectins have been mainly cited for their role in the interaction between host and *Symbiodinium* for symbiotic species (Wood-Charlson *et al.* 2006; Davy *et al.* 2012). Additionally, lectin homologues from *Acropora millepora* and *Pocillopora damicornis* have been shown to be implicated in the bleaching process during thermal stress experiments (Rodriguez-Lanetty *et al.* 2006; Vidal-Dupiol *et al.* 2013). Nevertheless, Kvennefors *et al.* (2008) highlighted the ability of pathogens recognition for the lectin of the coral *Acropora millepora* and suggested a conservation of the function of lectin in the innate immunity between cnidarians and the bilaterians. Moreover, Martin *et al.* (2002) pointed that an increase in sea water temperature could favour the settlement of opportunistic pathogens or the increase of the sensibility of Mediterranean octocorals. Thus, the overexpression of the lectin red coral homologue at 5 m could be an adaptation to an environment where pathogens pressure would be higher: this remains to be investigated.

We thus identified new potential expression markers for the study of adaptation, but other candidate genes were not recovered as differentially expressed. For example, heat-shock protein homologues (three HSP60 and three HSP70) were expressed at the two depths, but without significant differential expression. This contrasts with experimental stress studies in cnidarians where HSPs have been shown as involved in stress response (DeSalvo *et al.* 2010; Barshis *et al.* 2013; Haguenauer *et al.* 2013; Kenkel *et al.* 2014). Here, HSPs would then correspond to an inducible response and the absence of expression in our samples could be considered as a negative control of the stress state of our colonies.

Phenotypic plasticity or genetic adaptation?

Tumour necrosis factor receptor-associated factor proteins seem to be consistent markers of the native environmental origin of individuals, but we still cannot discriminate the effects of phenotypic plasticity versus genetic adaptation. With the three-month period of common in situ temperature and a one-month period of acclimatization in common garden conditions, we expected to attenuate differences due to the environment of origin. Despite this, the effects of the original environment on individuals persisted. This has been much discussed recently and several nongenetic effects (plasticity, maternal effects, epigenetic processes) were probably maintained during our experiment and may cause the detection of some of the 212 differentially expressed genes (Kawecki & Ebert 2004; Pespeni *et al.* 2013). The effect of acclimatization on specific environmental conditions should decrease with the increase of the period of common acclimatization, but the time needed to reach the complete loss of the acclimatization effect is still

© 2015 John Wiley & Sons Ltd

unclear. Pespeni *et al.* (2013) found a signal of this environment condition memory even after 3 years of acclimatization in common conditions in the purple sea urchin *Strongylocentrotus purpuratus*. Indeed, we cannot characterize the proportion of genetic versus acclimatization effect at the transcriptome level, but with the combination of expression and annotation patterns, we highlight good candidates of the persistence of adaptation (broad sense) to thermally varying environments.

Polymorphism pattern

The frequency of SNPs in *C. rubrum* was 1 per 126 bp. This frequency is higher than those found for *Nematostella vectensis* (1 per 139 bp; Sullivan *et al.* 2008), *Acropora millepora* (1 per 207 bp; Meyer *et al.* 2009) and *Acropora palmata* (1 per 272 bp; Polato *et al.* 2011). However, much higher frequencies have been described for other marine organisms such as *Pinctata margaritifera* (1 per 23 bp; Teaniniuraitemoana *et al.* 2014), *Crassostrea gigas* (1 per 40 bp in noncoding regions and 1 per 60 bp in coding regions; Sauvage *et al.* 2007) and more particulary for the octocoral *Eunicella cavolinii* (1 per 63 bp; Romiguier *et al.* 2014). We chose to analyse the sequence polymorphism of our transcriptome to search for potential signals of genetic adaptation in these populations. Indeed, sequence polymorphism has already been shown to be an indicator of local adaptation for cnidarians by Lundgren *et al.* (2013) who highlighted correlations between allelic frequencies and environment for two hexacorals, *Acropora millepora* and *Pocillopora damicornis*. The conservative filter of genotype quality GQ = 20 and DP = 12 allows us to be confident about the quality of the remaining SNPs. The significant decrease caused by the application of this filter can be explained by the fact that individuals have been pooled and allelic frequencies drifted away from the expected values, although the red coral is probably diploid based on the observed microsatellite patterns (Ledoux *et al.* 2010b). However, our method of using pools of three individuals and the sequencing of expressed genes did not enable us to estimate precisely allele frequencies, and we focused on genes with major effects. Apart from the consequences of pooling, the analysis of six individuals per population probably induced an ascertainment bias leading to the nonidentification of SNPs with low-frequency alleles. We therefore focused on genes with strong differences in allelic frequencies between populations, the apparently differentially fixed loci. The 56 differentially fixed SNPs detected in our study are potential markers for the study of local adaptation in *C. rubrum*, but we cannot completely exclude other genes as potential markers of local adaptation. Indeed, one could also expect adaptation to be linked to various interacting loci with small effects. All candidate genes presented here still need to be individually tested, but they open the way to deeper investigations of the genetic adaptation to local conditions through the analysis of the polymorphism of these genes on more individuals with additional ecological replicates. This could be extended to other cnidarians as well to search for patterns of repeated evolution.

Conclusion

This study enabled the identification of the first candidates for the study of adaptation (broad sense, i.e. including genetic adaptation and acclimatization) to contrasted environmental conditions for the red coral. The consistence of some of these markers with other studies implicating hexacorals is a strong argument for the conservation of their adaptive function to thermal stress among cnidarian, whether symbiotic or not. Our results also support the key importance of studying the gene expression in basal conditions. Finally, we propose here SNP data for population genetic studies, as well as new resources for studying adaptation to contrasted environments in terms not only of expression pattern, but also of polymorphism.

Acknowledgements

This work is a contribution to the Labex OT-Med (no. ANR-11-LABX-0061) funded by the French Government 'Investissements d'Avenir' programme of the French National Research Agency (ANR) through the A*MIDEX project (no. ANR-11-IDEX-0001-02). This project has been funded by the ADACNI programme of the French National Research Agency (ANR) (project no. ANR-12-ADAP-0016; http://adacni.imbe.fr). We thank ECCOREV Research Federation (FR 3098) for the financial support of this study. We are grateful to the genotoul bioinformatics platform Toulouse Midi-Pyrénées. We thank Didier Zoccola and Sylvie Tambutté for their help on the transcriptome analysis. A particular thanks to the molecular biology staff of the IMBE laboratory and the Pythéas Institute's divers for their support and assistance.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *Journal of Molecular Biology*, **215**, 403–410.

Anders S, Huber W (2010) Differential expression analysis for sequence count data. *Genome Biology*, **11**, R106.

Arch RH, Gedrich RW, Thompson CB (1998) Tumor necrosis factor receptor-associated factors (TRAFs)—a family of adapter proteins that regulates life and death. *Genes & Development*, **12**, 2821–2830.

Ashburner M, Ball CA, Blake JA *et al.* (2000) Gene Ontology: tool for the unification of biology. *Nature Genetics*, **25**, 25–29.

Barshis DJ, Ladner JT, Oliver TA *et al.* (2013) Genomic basis for coral resilience to climate change. *Proceedings of the National Academy of Sciences, USA*, **110**, 1387–1392.

Bay LK, Guérécheau A, Andreakis N, Ulstrup KE, Matz MV (2013) Gene expression signatures of energetic acclimatisation in the reef building coral *Acropora millepora*. *PLoS ONE*, **8**, e61736.

© 2015 John Wiley & Sons Ltd

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, **57**, 289–300.

Bradley JR, Pober JS (2001) Tumor necrosis factor receptor-associated factors (TRAFs). *Oncogene*, **20**, 6482–6491.

Burge CA, Mouchka ME, Harvell CD, Roberts S (2013) Immune response of the Caribbean sea fan, *Gorgonia ventalina*, exposed to an Aplanochytrium parasite as revealed by transcriptome sequencing. *Frontiers in Physiology*, **4**, 1–9.

Collins AG (2002) Phylogeny of Medusozoa and the evolution of cnidarian life cycles. *Journal of Evolutionary Biology*, **15**, 418–432.

Costantini F, Rossi S, Pintus E *et al.* (2011) Low connectivity and declining genetic variability along a depth gradient in *Corallium rubrum* populations. *Coral Reefs*, **30**, 991–1003.

Davy SK, Allemand D, Weis VM (2012) Cell biology of cnidarian-dinoflagellate symbiosis. *Microbiology and Molecular Biology Reviews*, **76**, 229–261.

De Wit P, Palumbi SR (2013) Transcriptome-wide polymorphisms of red abalone (*Haliotis rufescens*) reveal patterns of gene flow and local adaptation. *Molecular Ecology*, **22**, 2884–2897.

DePristo MA, Banks E, Poplin R *et al.* (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics*, **43**, 491–498.

DeSalvo M, Sunagawa S, Voolstra C, Medina M (2010) Transcriptomic responses to heat stress and bleaching in the elkhorn coral *Acropora palmata*. *Marine Ecology Progress Series*, **402**, 97–113.

D'Ortenzio F, Lavigne H, Besson F *et al.* (2014) Observing mixed layer depth, nitrate and chlorophyll concentrations in the northwestern Mediterranean: a combined satellite and NO_3 profiling floats experiment. *Geophysical Research Letters*, **41**, 6443–6451.

Haguenauer A, Zuberer F, Ledoux J-B, Aurelle D (2013) Adaptive abilities of the Mediterranean red coral *Corallium rubrum* in a heterogeneous and changing environment: from population to functional genetics. *Journal of Experimental Marine Biology and Ecology*, **449**, 349–357.

Karako-Lampert S, Zoccola D, Salmon-Divon M *et al.* (2014) Transcriptome analysis of the Scleractinian coral *Stylophora pistillata*. *PLoS ONE*, **9**, e88615.

Kawecki TJ, Ebert D (2004) Conceptual issues in local adaptation. *Ecology Letters*, **7**, 1225–1241.

Kayal E, Roure B, Philippe H, Collins AG, Lavrov DV (2013) Cnidarian phylogenetic relationships as revealed by mitogenomics. *BMC Evolutionary Biology*, **13**, 5.

Kenkel CD, Sheridan C, Leal MC *et al.* (2014) Diagnostic gene expression biomarkers of coral thermal stress. *Molecular Ecology Resources*, **14**, 667–678.

Kvennefors ECE, Leggat W, Hoegh-Guldberg O, Degnan BM, Barnes AC (2008) An ancient and variable mannose-binding lectin from the coral *Acropora millepora* binds both pathogens and symbionts. *Developmental & Comparative Immunology*, **32**, 1582–1592.

Ledoux J-B, Garrabou J, Bianchimani O *et al.* (2010a) Fine-scale genetic structure and inferences on population biology in the threatened Mediterranean red coral, *Corallium rubrum*. *Molecular Ecology*, **19**, 4204–4216.

Ledoux J-B, Mokhtar-Jamaï K, Roby C *et al.* (2010b) Genetic survey of shallow populations of the Mediterranean red coral [*Corallium rubrum* (Linnaeus, 1758)]: new insights into evolutionary processes shaping nuclear diversity and implications for conservation. *Molecular Ecology*, **19**, 675–690.

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, **25**, 1754–1760.

Li W, Godzik A (2006) CD-HIT: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*, **22**, 1658–1659.

Li L, Stoeckert CJ, Roos DS (2003) OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Research*, **13**, 2178–2189.

Li H, Handsaker B, Wysoker A *et al.* (2009) The sequence alignment/map format and SAMTOOLS. *Bioinformatics (Oxford, England)*, **25**, 2078–2079.

Lundgren P, Vera JC, Peplow L, Manel S, van Oppen MJ (2013) Genotype–environment correlations in corals from the Great Barrier Reef. *BMC Genetics*, **14**, 9.

Martin Y, Bonnefont JL, Chancerelle L (2002) Gorgonians mass mortality during the 1999 late summer in French Mediterranean coastal waters: the bacterial hypothesis. *Water Research*, **36**, 779–782.

McKenna A, Hanna M, Banks E *et al.* (2010) The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, **20**, 1297–1303.

Meyer E, Aglyamova GV, Wang S *et al.* (2009) Sequencing and de novo analysis of a coral larval transcriptome using 454 GSFlx. *BMC Genomics*, **10**, 219.

Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods*, **5**, 621–628.

Moya A, Huisman L, Ball EE *et al.* (2012) Whole transcriptome analysis of the coral *Acropora millepora* reveals complex responses to CO_2-driven acidification during the initiation of calcification. *Molecular Ecology*, **21**, 2440–2454.

Palumbi SR, Barshis DJ, Traylor-Knowles N, Bay RA (2014) Mechanisms of reef coral resistance to future climate change. *Science*, **334**, 895–898.

Park E, Hwang D-S, Lee J-S *et al.* (2012) Estimation of divergence times in cnidarian evolution based on mitochondrial protein-coding genes and the fossil record. *Molecular Phylogenetics and Evolution*, **62**, 329–345.

Pertea G, Huang X, Liang F *et al.* (2003) TIGR gene Indices clustering tools (TGICL): a software system for fast clustering of large EST datasets. *Bioinformatics*, **19**, 651–652.

Pespeni MH, Barney BT, Palumbi SR (2013) Differences in the regulation of growth and biomineralization genes revealed through long-term common-garden acclimation and experimental genomics in the purple sea urchin. *Evolution*, **67**, 1901–1914.

Polato NR, Vera JC, Baums IB (2011) Gene discovery in the threatened elkhorn coral: 454 sequencing of the *Acropora palmata* transcriptome. *PLoS ONE*, **6**, e28634.

R Development Core Team (2008) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rice P, Longden I, Bleasby A (2000) EMBOSS: the European molecular biology open software suite. *Trends in Genetics*, **16**, 276–277.

Robinson MD, Oshlack A (2010) A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology*, **11**, R25.

Robinson MD, McCarthy DJ, Smyth GK (2010) EDGER: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, **26**, 139–140.

Rodriguez-Lanetty M, Phillips WS, Weis VM (2006) Transcriptome analysis of a cnidarian – dinoflagellate mutualism reveals complex modulation of host gene expression. *BMC Genomics*, **7**, 23.

Romiguier J, Gayral P, Ballenghien M *et al.* (2014) Comparative population genomics in animals uncovers the determinants of genetic diversity. *Nature*, **515**, 261–263.

Sauvage C, Bierne N, Lapègue S, Boudry P (2007) Single Nucleotide polymorphisms and their relationship to codon usage bias in the Pacific oyster *Crassostrea gigas*. *Gene*, **406**, 13–22.

Schulz MH, Zerbino DR, Vingron M, Birney E (2012) Oases: robust de novo RNA-seq assembly across the dynamic range of expression levels. *Bioinformatics*, **28**, 1086–1092.

Stefanik DJ, Lubinski TJ, Granger BR *et al.* (2014) Production of a reference transcriptome and transcriptomic database (EdwardsiellaBase) for the lined sea anemone, Edwardsiella lineata, a parasitic cnidarian. *BMC Genomics*, **15**, 71.

Sullivan JC, Reitzel AM, Finnerty JR (2008) Upgrades to StellaBase facilitate medical and genetic studies on the starlet sea anemone, *Nematostella vectensis*. *Nucleic Acids Research*, **36**, D607–D611.

© 2015 John Wiley & Sons Ltd

Tarazona S, García-Alcalde F, Dopazo J, Ferrer A, Conesa A (2011) Differential expression in RNA-seq: a matter of depth. *Genome Research*, **21**, 2213–2223.

Teaniniuraitemoana V, Huvet A, Levy P *et al*. (2014) Gonad transcriptome analysis of pearl oyster *Pinctada margaritifera*: identification of potential sex differentiation and sex determining genes. *BMC Genomics*, **15**, 491.

Technau U, Rudd S, Maxwell P *et al*. (2005) Maintenance of ancestral complexity and non- metazoan genes in two basal cnidarians. *Trends in Genetics*, **21**, 633–639.

The UniProt Consortium (2013) Activities at the universal protein resource (UniProt). *Nucleic Acids Research*, **42**, D191–D198.

Torrents O, Tambutté E, Caminiti N, Garrabou J (2008) Upper thermal thresholds of shallow vs. deep populations of the precious Mediterranean red coral *Corallium rubrum* (L.): assessing the potential effects of warming in the NW Mediterranean. *Journal of Experimental Marine Biology and Ecology*, **357**, 7–19.

Traylor-Knowles N, Palumbi SR (2014) Translational environmental biology: cell biology informing conservation. *Trends in Cell Biology*, **24**, 265–267.

Verstrepen L, Carpentier I, Verhelst K, Beyaert R (2009) ABINs: A20 binding inhibitors of NF-κB and apoptosis signaling. *Biochemical Pharmacology*, **78**, 105–114.

Vidal-Dupiol J, Zoccola D, Tambutté E *et al*. (2013) Genes related to ion-transport and energy production are upregulated in response to CO_2-driven pH decrease in corals: new insights from transcriptome analysis. *PLoS ONE*, **8**, e58652.

Weis WI, Taylor ME, Drickamer K (1998) The C-type lectin superfamily in the immune system. *Immunological Reviews*, **163**, 19–34.

Wood-Charlson EM, Hollingsworth LL, Krupp DA, Weis VM (2006) Lectin/glycan interactions play a role in recognition in a coral/dinoflagellate symbiosis. *Cellular Microbiology*, **8**, 1985–1993.

D.A., A.H. and P.P. conceived the project. A.H. and D.A. conceived experiments. A.H. performed the experiments. C.K. performed the assembly and the SNP calling. M.P. and O.C. performed the analyses. M.P. analysed the results. M.P. wrote the study. D.A. and P.P. supervised research.

Data accessibility

The raw DNA sequences are available in the Short Read Archive (SRA) database under the Accession no. SRX675792. All assembled contigs, SNP file and annotation file are available on DRYAD (doi: 10.5061/dryad.31f77).

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Fig. S1 Size distribution of contig lengths (bp) for annotated contigs (red) and non annotated ones (blue).

Fig. S2 Expression values for the two depths (FIG 5: Figuier – 5 m; MOR40: Morgiou – 40 m. Differentially expressed genes between the depth detected by each method are indicated (red).

Fig. S3 Effect of the increase of the Genotype Quality parameter on the number of SNPs. The red vertical line indicate the threshold applied.

Table S1 Best informative UniProt Match and for annotated differentially expressed contigs (contigs with non informative UniProt hits are not indicated).

Table S2 Best informative UniProt Match and for annotated contigs containing differentially fixed SNPs (contigs with non informative UniProt hits are not indicated).

© 2015 John Wiley & Sons Ltd