Transcriptome-wide identification and evaluation of optimal reference genes for RT-qPCR expression analysis of *Saccharina latissima* responses to biotic and abiotic stress

Qikun Xing¹ · Sylvie Rousvoal¹ · Catherine Leblanc¹

Received: 14 July 2020 / Revised and accepted: 22 September 2020 / Published online: 12 November 2020
© The Author(s) 2020

Abstract

Saccharina latissima, known as sugar kelp, is a brown macroalga with huge ecological and economic values. In marine intertidal environment, *S. latissima* has to cope with both biotic and abiotic stress, which can cause the reduction of the yield during cultivation. To better understand the physiological responses of *S. latissima* under different stress conditions, large-scale transcriptomic analyses are useful to explore global metabolic pathway regulations. In addition, real-time quantitative polymerase chain reaction (RT-qPCR) is a powerful and rapid method for further quantifying changes in gene expression, and for targeting specific defense-related gene pathways. However, its level of accuracy is highly related to the expression stability of reference genes used for normalization and those still need to be evaluated in *S. latissima*. In this study, we therefore experimentally tested eight candidate reference genes identified from in silico screening of public transcriptomic datasets of *S. latissima* from different abiotic and biotic stress treatments. The stability analysis using complementary statistical approaches showed that *EIF5B* and *ATPase* are the most stable reference genes under biotic stress, whereas, under temperature and light stress, their combination with *NDH* gene is the best choice for RT-qPCR normalization. The validated reference genes were used to monitor the expression of target genes, related to oxidative responses, such as those involved in oxylipin pathways, in *S. latissima* plantlets submitted to different stress in laboratory-controlled conditions.

Keywords *Saccharina latissima* · Phaeophyceae · Reference gene · Real-time PCR · Environmental stress · Gene expression · RNA-Seq

Introduction

The brown alga *Saccharina latissima*, known as the sugar kelp, is a seaweed of great ecological and economic values (Bartsch et al. 2008; Lüning and Mortensen 2015). In the Northern hemisphere, the underwater forests formed by *S. latissima* and other kelp species provide habitats and food and play a significant role in the primary production of the coastal area. In addition, the cultivation of *S. latissima* is becoming an important research field in applied phycology, this species being the closest European relative to *S. japonica*, a

dominant species in Asian seaweed aquaculture (Buschmann et al. 2017). In their natural environment, kelps are exposed to biotic aggressions by a high number of potentially harmful organisms such as grazers, fungi, bacteria, or endophytic algae (Potin et al. 2002). In the intertidal area, kelps have also to cope with multiple abiotic stresses such as fluctuations of temperature and light intensity. The molecular mechanisms of acclimation and adaptation in kelp species are fundamental open questions. In this context, several studies have explored the global transcriptomic regulations in several kelp species under different stress conditions, including the combination of temperature and light stress (Heinrich et al. 2012, 2015) and osmotic and temperature stress (Li et al. 2020) and also along a geographical gradient in *S. latissima* (Monteiro et al. 2019), and under copper stress and ocean acidification in *S. japonica* (Xu et al. 2019, Zhang et al. 2019), during interactions with herbivores in *Laminaria digitata* and *Lessonia spicata* (Ritter et al. 2017) and under light, temperature, and nutrients stress in the giant kelp *Macrocystis pyrifera* (Konotchick et al. 2013).

Supplementary Information The online version of this article (<https://doi.org/10.1007/s10811-020-02279-x>) contains supplementary material, which is available to authorized users.

✉ Catherine Leblanc
catherine.leblanc@sb-roscoff.fr

¹ Sorbonne Université, CNRS, UMR 8227, Integrative Biology of Marine Models, Station Biologique, Roscoff, France

To further understand the stress-tolerance mechanisms in specific algae, targeted gene expression analysis by RT-qPCR is widely applied due to its advantage of sensitivity, specificity, and rapid execution (Pombo et al. 2017). However, given the potential biases introduced by reverse transcription step and PCR amplification, the relative quantification mode is based on the normalization by transcript level of reference genes, for which the expression is considered stable or constitutive. The choice of reference genes is therefore particularly important, as small variation in the expression of reference gene could lead to significant changes in target gene expression (Dheda et al. 2005). An ideal reference gene should have minimal or no fluctuation in different tissues, developmental stages, and diverse stress conditions. Housekeeping genes generally refer to a small set of highly conserved genes which are essential for basic biological functions in organisms (Fiume and Fletcher 2012) and were considered having a stable expression level through all conditions and tissues. However, some traditionally used housekeeping genes have shown variable expression in different physiological conditions. For example, in the plant model *Arabidopsis thaliana*, the genes coding for tubulin 6 (*TUB6*), elongation factor-1 α (*EF1 α*), actin 2 (*ACT2*), and glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*) display strongly reduced expressions in pollen samples, and the expressions of *ACT2* and *GAPDH* are also reduced in seed samples (Czechowski et al. 2005). In addition, *GAPDH* and α -tubulin (*TubA*) were proved to be the least stable reference genes in the fish *Odontesthes humensis* in different ages and tissues (Silveira et al. 2018). For each target species and each condition to be studied, it is therefore essential to validate the expression stability of reference genes in order to improve the reliability and accuracy of RT-qPCR expression analyses.

With the development of high-throughput sequencing technologies, RNA-seq has become an efficient method to obtain genome-wide gene expression data (Wang et al. 2009). This method is mainly applied for the analysis of global transcriptomic regulation and the identification of differentially expressed genes, when comparing different tissues, treatments, or stress conditions. The available transcriptomic data also constitute an interesting resource of constitutively expressed genes for identifying novel stable reference genes instead of only relying on available candidate genes. Recently, RNA-seq data have been used for the selection and validation of suitable reference genes in some land plant species such as *Lycoris aurea* (Ma et al. 2016) and *Oxytropis ochrocephala* Bunge (Fu et al. 2015). In marine macroalgae, some genes have been experimentally validated as reference genes in few model species such as *Ectocarpus siliculosus* (Le Bail et al. 2008), *Chondrus crispus* (Kowalczyk et al. 2014), *Heterosigma akashiwo* (Koeduka et al. 2015), and *Pyropia yezoensis* (Gao et al. 2018). However, up to now, there is no study about the identification and validation of RT-qPCR

reference genes in *S. latissima* whereas this kelp is becoming an important model species for studying molecular responses to environmental changes (Monteiro et al. 2019, Li et al. 2020).

In order to conduct studies on the expression regulation of targeted metabolic pathways, we looked for putative stable reference genes in RNA-seq datasets of this species, obtained from both biotic and abiotic stress experiments, using in silico analyses. Subsequently, the expression stability of selected genes was further validated using RT-qPCR under different light, temperature, and endophyte co-cultivation conditions as those used to generate RNA-seq data and following defense elicitation using endogenous elicitor, oligo-gulonates (GG) treatment. Indeed, GG applications have been shown to induce rapid oxidative bursts in kelps (Küpper et al. 2001, 2002) and later metabolic regulations (Goulitquer et al. 2009), including conserved defense and signaling pathways, such as the release of free fatty acids and oxylipins. In order to validate RT-qPCR normalization in *S. latissima* in a physiological context of both abiotic and biotic stress, all selected reference genes were then re-assessed to monitor the expression of two target genes encoding a lipoxygenase (*LOX*) and a glutathione s-transferase (*GST*), which are involved in oxylipin pathway and oxidative stress responses, respectively. The validated reference genes will be suitable internal standards to further explore gene expression regulations related to stress responses and resistance in *S. latissima* using RT-qPCR.

Material and methods

Algal material and treatments

Fertile sporophytes of the kelp *S. latissima* were collected at Roscoff (Perharidy, 48.73° N, 4.00° W), cleaned, and cut into small pieces for spore release using the hanging-drop technique (Wynne 1969). Lab-grown gametophytes were kept for several weeks in Provasoli-enriched, filtrated, and autoclaved natural seawater (FSW), under 12 °C and 20 $\mu\text{mol photons m}^{-2} \text{s}^{-1}$ with a 12-h light/dark cycle. After self-fertilization, the developing sporophytes were grown in Petri dishes in the same culture conditions (10 mL Provasoli solution L⁻¹ FSW (Provasoli 1968)). After 4 weeks, the sporophytes were detached from the coverslips and transferred to 10-L bottles connected to an aeration system. The culture medium (FSW enriched with Provasoli) was weekly changed. The kelp cultures were maintained under 12 °C and 20 $\mu\text{mol photons m}^{-2} \text{s}^{-1}$ with a 12-h light/dark cycle. The brown filamentous alga *Laminarionema elsbetiae* strain was maintained in Petri dishes in the same culture medium conditions, except the photon rate at 5 $\mu\text{mol photons m}^{-2} \text{s}^{-1}$.

The laboratory-grown sporophytes of *S. latissima* (~5 cm length) and its natural endophyte *L. elsbetiae* were co-

cultivated in 2-L bottles filled with 1.5-L sterile Provasoli-enriched FSW and connected to an aeration system. The co-cultivation experiment was performed for 2 days, and control cultures of *S. latissima* without endophytes were conducted in the same time, in the same conditions. In parallel, the GG elicitation was performed on *S. latissima* sporophytes, by adding $150 \mu\text{g mL}^{-1}$ of oligo-gulonates blocks (GG, prepared from *Laminaria hyperborea* according to Haug et al. (1974) in small glass bakerys filled with 30-mL FWS, under shaking for 1 h. All the co-cultivation tests, including GG treatments and corresponding control experiments, were conducted under the same conditions: 12°C and $20 \mu\text{mol photons m}^{-2} \text{s}^{-1}$ with a 12-h light/dark cycle.

For subjecting sporophytes to various abiotic stress, laboratory-grown sporophytes were transferred in 30-mL FSW glass bakerys and maintained for 2 days in four different temperature and light culture conditions as already conducted in Heinrich et al. (2012): high temperature and high light (HT&HL, 17°C and $107.8 \pm 5 \mu\text{mol photons m}^{-2} \text{s}^{-1}$), high temperature and low light (HT&LL, 17°C and $23.8 \pm 3.2 \mu\text{mol photons m}^{-2} \text{s}^{-1}$), low temperature and high light (LT&HL, 2°C and $107.8 \pm 5 \mu\text{mol photons m}^{-2} \text{s}^{-1}$), and low temperature and low light (LT&LL, 2°C and $23.8 \pm 3.2 \mu\text{mol photons m}^{-2} \text{s}^{-1}$). Three biological replicates were collected for each treatment. They were flash-frozen in liquid nitrogen and kept at -80°C until RNA extraction.

RNA isolation and cDNA synthesis

Total RNA was extracted using a combination of a classical CTAB-based method and of the RNeasy Mini kit (QIAGEN, Germany), as described in Heinrich et al. (2012). Compared to this protocol, before the use of the purification kit, an additional ethanol precipitation step was conducted to remove polysaccharides, and RNA was separated from DNA and major contaminants by precipitation in 3 M LiCl solution and 1% 2-mercaptoethanol, at -20°C overnight. After RNA extraction, a treatment with RNase-free DNase I (Turbo DNase, Ambion) was performed to eliminate residual genomic DNA. The quality and quantity of total RNA were tested on a NanoDrop™ spectrophotometer (Thermo Fisher Scientific, USA) and further checked on a 2% agarose gel. RNA samples with concentrations above $100 \text{ ng } \mu\text{L}^{-1}$ and A260/A280 ratio of 1.8–2.0 were used for cDNA synthesis. The RNA was reverse-transcribed to cDNA using the ImProm-II Reverse Transcription System (Promega).

Bioinformatic selection of candidate reference genes

Different RNA-seq data sets were used to screen for potential candidate reference genes. RNA sequencing of control and co-cultured samples were previously performed using Illumina Hi-Seq 3000 platform (Bernard, Xing et al.

unpublished data; ENA at the EBI; ID: PRJEB37483). After transcriptome assembly and annotation, the reads were mapped on the transcriptome using bowtie2 (Langmead and Salzberg 2012). The read counts were calculated by RSEM (Li and Dewey 2011) and converted into transcripts per million (TPM) for each unigene. The microarray data of *S. latissima* under abiotic stress (control, high temperature and high light, high temperature and low light, low temperature and high light, and low temperature and low light) were downloaded from the EMBL database (ArrayExpress at the EBI; <http://www.ebi.ac.uk/microarray-as/ae/>; ID: E-MEXP-3450) and the probes of each gene were mapped to the reference transcriptome to link the expression of probes with unigenes.

The expression stability of all unigenes in two datasets was calculated and ranked separately by the Pattern Gene Finder (PaGeFinder). PaGeFinder is a web-based server for the online detection of gene expression patterns from serial transcriptomic data generated by high-throughput technologies like microarray or next-generation sequencing. In PaGeFinder, the dispersion measure (DPM) was calculated to evaluate the variability of gene expression profiles. A value of DPM close to 0 suggests a similar expression level of a gene over all RNA-seq samples, and then a strong stability of its expression whatever the treatments. In the end, genes with putative and reliable protein annotation (nr databases), appropriate expression levels ($\text{TPM} > 10$), and a low dispersion measure ($\text{DPM} \leq 0.3$) were selected as candidate genes to be further tested as RT-qPCR reference genes.

RT-qPCR and data analysis

For each reference candidate gene and stress-related target genes, a pair of specific primers was designed in the 3' end coding region of sequences using Primer5 software (Table 1). Primers were designed to have a melting temperature around 60°C , a length between 18 and 26 nucleotides, a GC content between 40 and 60%, avoiding secondary structures and self- and cross-annealing. The specificity of each primer was tested in silico, using the BLASTN search on the reference transcriptome of *S. latissima*.

The RT-qPCR reactions were performed in 384-well plates in a LightCycler480 (Roche Molecular Biochemicals, Germany) with LightCycler480 SYBR Green I Master kit (Roche, Germany). The protocol was: 95°C for 5 min, followed by 55 cycles of 95°C for 10 s, 60°C for 10 s, and 72°C for 20 s. Four biological replicates were performed for each treatment. The specificity of the amplification was verified with a dissociation curve obtained by heating the samples from 65°C to 95°C . A 1:5 dilution series of the *S. latissima* cDNA was prepared and tested for each gene to determine the amplification efficiency of the selected primers pairs, which was calculated by the LightCycle480 gene scanning software (version

Table 1 Description of the candidate reference genes

Gene symbol	NR description	Accession ID	DPM* (abiotic stress)	DPM* (biotic stress)	Primer sequence: Forward Reverse	RT-qPCR efficiency ^a	Error ^b
60S	60S ribosomal protein L10A [<i>Ectocarpus siliculosus</i>]	CBN78308.1	0.2596	0.1423	TACAAAATAATAGCCACGACG CAGAGTTACATTACACCCAGGAG	1.99	0.01
ACT	beta-actin [<i>Fucus vesiculosus</i>]	Q39758.1	0.2919	0.0947	CACATTATTCTCGCACACAACAG TAAAGAACACGCACACTTAGCAG	2.006	0.167
ATM1	iron-sulfur clusters transporter atm1, mitochondrial precursor [<i>Ectocarpus siliculosus</i>]	CBN75449.1	0.1947	> 0.3	GTTTGTGTCGGATAGAGTGAATGG TCAAACAGACGCACGCTACG	2.039	0.0019
ATPase	F-type H-ATPase beta subunit [<i>Ectocarpus siliculosus</i>]	CBJ32298.1	0.2469	0.1222	TTGCTGTCGGTTCTGCTGG TGGCGGTGTGGAAAGTGG	1.963	0.108
EIF5B	EIF5B, eukaryotic translation initiation factor 5B [<i>Ectocarpus siliculosus</i>]	CBN80114.1	0.2807	0.1015	GAGGTGACGAGCGATTGA TCTTTCGGCGACGGGTT	2.002	0.0484
LISK	LISK family protein kinase [<i>Ectocarpus siliculosus</i>]	CBJ28172.1	0.2482	0.1149	CCGCTTCTCTCTGTTCTGG GCTTGATGGCTGCCTACGA	1.953	0.0255
NDH	ND11 homolog, NADH dehydrogenase (ubiquinone) sub- unit 11 [<i>Ectocarpus siliculosus</i>]	CBN74464.1	0.2593	0.0807	GGAGGAGATGGAGGAGCGTGTG GGCGTCCGTGCGGTAGAAAGT	1.962	0.116
TUBD	Tubulin folding cofactor D [<i>Ectocarpus siliculosus</i>]	CBN78333.1	0.1985	> 0.3	AGGTCAGCATCGGTTTAC TTTGGTTCGGCTGTCGTAG	1.961	0.046

*DPM: the dispersion measure

^aThe PCR amplification efficiency for each primers pair was determined by LightCycler480 gene scanning software (version 1.5)^bThe error value of amplification efficiency for each primers pair was determined by LightCycler480 gene scanning software (version 1.5) with an acceptable value of < 0.2

1.5), using the Fit-points method. These amplification tests on cDNA dilution series are also important to set the linearity limits of the amplification.

The gene expression level was determined by the number of amplification cycles (Cq), which was extracted from the raw fluorescence data of the RT-qPCR curve. Four statistical methods (NormFinder, geNorm, BestKeeper, and comparative delta Ct method) were used to evaluate the expression stability of selected reference genes. BestKeeper is an Excel-based tool that analyzes absolute raw Cq values directly using pairwise correlations. It calculates three variables related to the expression level of genes: the standard deviation (SD), the coefficient of correlation (r), and the coefficient of variance (CV). The gene with the lowest CV $\pm$ SD is considered the most stable transcript gene. In the other three methods, the Cq value is converted into relative quantities according to the formula: $2^{-\Delta Ct}$ (ΔCt = the corresponding Cq value–minimum Cq). In geNorm, the M values are calculated to evaluate the expression stability of genes. The most stable expressed gene exhibits the lowest M value. In NormFinder, the genes are ranked according to their stability value (SV) which is calculated using an ANOVA model based on ΔCt values. A lower SV is correlated with higher expression stability. The last method is based on the pairwise comparisons of ΔCt , ranking the genes from the most to the least stable transcripts, based on the average standard deviations, the lowest average SD corresponding to the most stable. Finally, the online software RefFinder (<http://leonxie.esy.es/RefFinder/>) was used to obtain the overall final ranking of all candidate genes based on the results of the four computational programs, NormFinder, BestKeeper, GeNorm, and comparative ΔCt .

For target genes, the raw data were divided by the Cq value of corresponding reference genes as a normalization step. And then the relative expression level of treated samples versus control samples was calculated by comparing the normalized Cq of treatment and control samples according to the formula: $2^{-\Delta\Delta Ct}$ ($\Delta\Delta Ct$ = normalized Cq value of treated sample – normalized Cq value of control sample).

Results

In silico *S. latissima* transcriptome-wide selection of candidate reference genes

A total of 23,049 unigenes from *S. latissima* was analyzed based on RNA-seq and microarray datasets: the dispersion measure (DPM) was calculated to evaluate their expression variability according to different experimental conditions, using PaGeFinder. The DPM value of 3,641 unigenes in biotic stress data, i.e., co-cultivation datasets including corresponding control datasets, and 2576 unigenes in abiotic stress data, i.e.,

temperature, light stress, and control conditions, were lower than 0.3 and were first selected as a pool of most stable transcripts. Among them, the unigenes without any function annotation or whose expression level was too low (TPM < 10) were removed, leading to 225 and 788 remaining candidate genes. Finally, eight genes were selected based on the combination of genes with low DPM values, either in biotic stress datasets only, or in abiotic stress datasets only or in all datasets.

PCR amplification and melting curve analysis were performed to confirm the specificity of the eight designed primer pairs, i.e., presenting a single band of amplification at the expected size (Sup. Fig. 1) and a single peak of the melting curve of the PCR product (Sup. Fig. 2). The mean amplification efficiency of each gene ranged from 1.95 to 2.03 (Table 1).

RT-qPCR expression profiles of candidate reference genes

The expression levels of the eight candidate genes were monitored by RT-qPCR in all *S. latissima* RNA samples, obtained from biotic and abiotic stress experiments, including corresponding control conditions. For each gene primer set amplification, the quantification cycle (Cq) value was reported in each RNA sample and the mean Cq value was calculated by pooling either all the Cq values derived from biotic conditions, or from abiotic conditions or from all tested conditions (Fig. 1). Globally, the eight tested genes featured a relatively wide expression range, from 19.62 to 30.65 Cq, and showed a very similar expression profile between the two treatment conditions. Since the Cq values are negatively correlated to the transcript levels, the *LISK* gene was the highest expressed gene and the 60S ribosomal protein-encoding gene was the lowest expressed gene. Whatever the analyzed datasets (biotic, abiotic, and all stress together), the *ACT* gene displayed the largest expression variation (Fig. 1). In biotic experimental conditions (Fig. 1a), the *LISK* gene exhibited the smallest expression variation, whereas in abiotic ones (Fig. 1b) and all stress conditions (Fig. 1c), it was the *ATPase* gene.

Computational analysis of candidate gene expression stability

The Cq values of the eight candidate genes were then used to evaluate their expression stability under different experimental conditions using four computational algorithms, namely geNorm, NormFinder, BestKeeper, and the comparative ΔCt method.

The geNorm analysis showed that the expression stability value (M) of all the candidate genes ranged from 0.16 to 0.80 under all conditions, demonstrating that their expression could be considered relatively stable, as M was always under 1.5

Fig. 1 Box-plots of Cq values obtained for the 8 candidate reference genes in biotic stress ($n = 16$) (**a**), temperature and light stress ($n = 20$) (**b**), and all tested conditions ($n = 36$) (**c**). Central lines in the boxes correspond to the median values

(Sup. Table 1). Under temperature and light stress, the *ACT* and *NDH* genes were predicted as the most stable reference genes, with the lowest M value, when the *ATPase* and *EIF5B* genes exhibited the lowest M value in biotic stress conditions.

Comprehensive analysis combining all the conditions together showed the M value-based ranking of all candidate genes, from the most (the lowest M value) to the least (the highest M value) constitutively expressed, was *EIF5B*, *NDH*, *ATPase*, *LISK*, *TUBD*, *ATM1*, *60S*, and *ACT* (Sup. Table 1). In addition, the pairwise variation analysis suggested that the combination of three reference genes was suitable for the light and temperature treatment group, as the V3/V4 value was the first one lower than 0.15 (Sup. Fig. 3). Using the co-cultivation and GG treatment group, the V2/V3 value was the first one lower than 0.15 (Sup. Fig. 3), indicating that two reference genes were sufficient for RT-qPCR normalization.

Based on NormFinder analysis, the most stable genes in terms of expression were *NDH* followed by *EIF5B*, both being considered suitable reference genes, in temperature and light stress conditions and when considering all conditions together (Sup. Table 2). However, the two most stable expressed genes were *ATPase* and *EIF5B*, in the biotic stress conditions.

In the BestKeeper analysis, only the SD value of the *ACT* gene in biotic stress conditions was bigger than one (Sup. Table 3), and its expression should be considered relatively unstable. Among all the other genes, BestKeeper identified two most stable genes, *ATPase* and *EIF5B*, considering all the sample sets or only the light and temperature stress-related samples. Under biotic stress, the *EIF5B* and *LISK* genes were predicted as the two most stable expressed genes.

The comparative delta Ct method identified the *ATPase* and *EIF5B* genes as the two most stable genes during biotic stress experiments (Sup. Table 4). When analyzing the light and temperature stress conditions or all conditions together, the expression stability was the strongest for the *NDH* and *EIF5B* genes.

Integrating computational methods for a comprehensive ranking of candidate reference genes

Based on these analyses, the prediction of the most stable genes revealed some variations according to experimental conditions tested and methods used (Fig. 2). Some overall trends appear such as the relatively bad ranking of the *ACT* gene as a potential reference gene, except for one method (delta Ct) in abiotic conditions (Fig. 2b). Indeed, the final ranks of the eight candidate reference genes were sometimes very different as observed for the *LISK* gene, classified from 1 to 4 in biotic conditions (Fig. 2a) or for the *ATPase* gene, ranked from 1 to 6 in abiotic conditions (Fig. 2b). Therefore, RefFinder, online software based on statistical approaches, was used to compare these data and establish a comprehensive ranking of all candidate

Fig. 2 Comparison of ranking of 8 candidate reference genes using geNorm, NormFinder, Delta Ct, and BestKeeper under biotic stress (a), light and temperature stress (b), and the combination of all stresses (c). The ordering of genes on X-axis indicates the comprehensive ranking obtained with RefFinder (see Sup. Table 5), from the least to the most stable genes

genes (Sup. Table 5). Our results showed that *ATPase* and *EIF5B* were the two most stable reference genes

Fig. 3 Relative expression levels of *LOX* gene, normalized by the most stable reference gene, *NDH*; the best reference gene combination, BRG (*NDH*, *ATPase*, and *EIF5B*); the least stable reference gene, *LISK*, in temperature and light stress compared to control condition. HT, high temperature; LT, low temperature; HL, high light; LL, low light. Error bars indicate standard errors ($n = 4$). The statistical significance between normalization modes is indicated by one or two asterisks (* P value < 0.05; ** P value < 0.01; t test)

under biotic stress, i.e., for co-cultivation or elicitation experimental conditions (Fig. 2). While under light and temperature stress, the predictive reference genes were *NDH* and *EIF5B*. In the end, the overall ranking of the eight genes considering all conditions from the most stable to the least stable expressed was *NDH* > *EIF5B* > *ATPase* > *TUBD* > *LISK* > *ATM1* > *60S* > *ACT* (Fig. 2; Sup. Table 5).

RT-qPCR reference gene validation

To further validate the reliability of the reference gene selection when studying gene regulation during either abiotic or biotic stress conditions in *S. latissima*, the expression levels of two target genes, *LOX* and *GST*, were monitored and normalized using the three best reference gene combination, the most stable reference gene or the least stable reference gene, predicted previously (Fig. 2a, b). Under light and temperature stress conditions, compared to control conditions, the *LOX* relative expression calculations appeared differently affected by the three normalization modes, depending on stress treatment (Fig. 3). For instance, no significant difference of the fold-change expression levels was detected under high temperature and low light treatment for the *LOX* gene. When comparing all conditions, the *LOX* expression always featured downregulated patterns, when using the least stable reference gene (*LISK*) or the most stable reference gene (*NDH*) for

normalization. Interestingly, this expression pattern was much more contrasted using the best reference gene combination (*NDH*, *EIF5B*, and *ATPase*) as calibrator and changed from 0.12-fold in high temperature and high light conditions to 1.53-fold, appearing upregulated in low temperature and low light conditions (Fig. 3). When comparing these results with the fold change value obtained from the RNA-seq data, the best reference gene combination as calibrators resulted in a similar expression pattern, with a downregulation in high-temperature stresses and an induction under low light and temperature stress conditions. In this latter condition, the relative expression levels of *LOX* seem to be significantly underestimated, when using the most stable reference gene (*NDH*). While using the least stable reference gene (*LISK*), the expression fold changes showed significant differences in two conditions out of four tested (Fig. 3), confirming the low-quality prediction of this gene as a reference gene in *S. latissima*.

In *S. latissima* samples cultivated with endophytes for 2 days, the *GST* expression levels showed strong upregulations, 5.73-fold and 9.41-fold changes, respectively, when using the most stable reference gene (*ATPase*) and the best reference gene combination (*ATPase* and *EIF5B*) as calibrators. These expression patterns agreed with that obtained from RNA-seq datasets (Fig. 4). The use of the least stable reference gene (*ACT*) for normalization significantly affects the calculation of the relative expression level, leading to a much lower value (2.39-fold change). A similar underestimation bias of the *GST* relative expression level was found in the GG-treated samples, where the best reference gene

combination (11.31-fold) and *ATPase* (6.66-fold) normalization showed the highest upregulation shifts, compared to the expression results normalized by *ACT* only. In that case, however, we could not compare these results with the corresponding GG RNA-seq datasets as a reference.

Discussion

RT-qPCR has been widely used for many purposes such as the quantification of target gene expression because of its specificity and high sensitivity of detection. However, the choice of internal standard genes has a huge effect on the normalization step and the final accuracy of RT-qPCR results. Using inappropriate reference genes as calibrators will result in significant differences in target transcript quantification as shown in the strong biases observed using different sets of genes for normalization (Figs. 3 and 4). Reference genes were generally selected from housekeeping genes, but recent studies have shown that some classic housekeeping genes do not present a stable expression under variable conditions, or not sufficiently stable to be used for RT-qPCR normalization (Czechowski et al. 2005, Silveira et al. 2018). In agreement with these results, our study also showed that a classic housekeeping gene coding for actin (*ACT*) did not exhibit high stability in *S. latissima* in any of the six experimental conditions analyzed, covering both abiotic and biotic stress conditions (Fig. 1). Therefore, it is not only necessary to select novel reference genes for *S. latissima*, but they also need to be adapted to each experimental setup, by validating the absence of variability of their expression in corresponding physiological conditions.

As an important source of transcriptomic information, RNA-seq analysis has been performed to explore suitable reference genes for specific experimental conditions in other model species (Liu et al. 2018, Zhan et al. 2019). In this study, we analyzed RNA-seq datasets related to biotic (co-cultivation with endophyte) and abiotic stress (light and temperature) in the brown alga *S. latissima*. Eight candidate reference genes were selected considering credible protein annotation (nr database), appropriate expression levels (TPM > 10), and a low dispersion measure (DPM ≤ 0.3), and their expression stability was further experimentally validated by RT-qPCR in *S. latissima* plantlets submitted to different stress treatments carried out in laboratory-controlled conditions. These treatments were chosen to cover a range of different physiological conditions, such as temperature and light environmental changing parameters, which could be encountered during intertidal exposures. For biotic stress validation, plantlets were cultivated with their natural endophyte *L. elsbetiae* (Bernard et al. 2019) using laboratory-controlled bioassays. In addition, oligo-guluronate (GG) elicitation was performed as previous studies indicated that this chemical treatment was able to induce an oxidative burst in *S. latissima* (Küpper et al. 2002) and transcriptomic and metabolic defense regulations in

Fig. 4 Relative expression levels of *GST* gene, normalized by the most stable reference gene, *ATPase*; the best reference gene combination, BRG (*ATPase* and *EIF5B*); the least stable reference gene, *ACT*, in biotic stress. GG, oligo-guluronate elicitation. Error bars indicate standard errors ($n = 4$). The statistical significance between normalization modes is indicated by one or two asterisks (* P value < 0.05; ** P value < 0.01; t test)

another closely related Laminariales, *L. digitata* (Cosse et al. 2009, Goulitquer et al. 2009). GG elicitation is therefore expected to mimic the damages of cell wall caused by numerous biotic agents and induced defense-related pathways in *S. latissima*.

Based on RT-qPCR expression data, the comprehensive ranking of candidate genes by RefFinder led to the identification of two optimal combinations of reference genes, according to experimental conditions (Fig. 2, Sup. Table 5). For biotic stress conditions, it was the two genes, *EIF5B* and *ATPase*. *ATPase* encodes the mitochondrial and chloroplast synthases, which use a proton gradient produced by electron transfer reactions to drive the synthesis of ATP. This gene has also been validated as the best internal standard in *Coleomegilla maculata* under different developmental stages (Yang et al. 2015). However, its use as a reference gene is not ubiquity, as in a dinoflagellate microalga, *Alexandrium catenella*, *ATPase* was predicted as the least stable gene across different nutritional conditions (Niaz et al. 2019). Unlike *ATPase*, the *EIF5B* gene is rarely used as a reference gene in other species. *EIF5B* belongs to the translation initiation factor family and is an essential initiation factor catalyzing the second GTP-dependent step in eukaryotic translation initiation (Lee et al. 2002). The other members of the translation initiation factor gene family have been widely used as constitutively reference genes in some macroalgal species. In the red alga *Chondrus crispus*, the expression stability of the gene coding for the translation initiation factor 4A-1 has been validated under light, osmotic, and heavy metal stress (Kowalczyk et al. 2014). In our study, *EIF5B* also exhibited a high stability of expression in all the experimental setup tested, including biotic, abiotic, and control conditions. For the light and temperature treatment group, three reference genes, namely *EIF5B*, *ATPase*, and *NDH*, were predicted to be the best combination in these physiological conditions. The *NDH* gene encodes NADH dehydrogenase located on the inner membrane of mitochondrial, transferring electron from NADH to coenzyme Q10. *NDH* was the best-ranked reference gene under temperature and light stress conditions (Fig. 2). However, it was not the case in the biotic treatment group (Fig. 2); therefore, its use should be adapted to physiological studies upon changing abiotic parameters or further validated using other experimental conditions. The actin (*ACT*) gene was first used as a reference gene for RT-qPCR normalization in the closely related species, *S. japonica*, under blue light induction (Deng et al. 2012), and later in different gene expression studies, such as a recent one related to copper stress (Zhang et al. 2019). This gene was experimentally validated in strawberry under light and temperature stress conditions according to its stable expression level (Zhang et al. 2018). However, in *S. latissima*, *ACT* was the least stable reference gene across all conditions we tested, as also observed in the brown alga *Ectocarpus siliculosus*, under multiple stress

conditions (Le Bail et al. 2008), suggesting that its use as reference gene is not adapted to these two brown algal species, and that its expression stability should be carefully checked before any application in brown algal gene expression normalization.

The *LOX* gene is involved in the oxylipin synthesis pathway, which plays an important role in the signal transduction under stress conditions (López et al. 2011, Lim et al. 2015). According to the RNA-Seq dataset analysis, the *LOX* gene was significantly upregulated under light and temperature stress conditions and *GST* was upregulated under co-cultivation stress. Therefore, these two target genes were used to validate the use of selected reference genes for normalizing their expression. In light and temperature stress, the results showed that the expression normalized by the least stable gene *LISK* was significantly underestimated comparing to the one normalized by the best reference gene combination (BRG), including the three genes *NDH*, *ATPase*, and *EIF5B*. It is worth noting that the normalization by a single reference gene, even if being the most stable one, *NDH*, also conducted to a significant underestimation compared to the normalization by the BRG, indicating the necessity of using multiple reference gene combination as a calibrator. The strong correlation observed with the expression pattern derived from the RNA-Seq data further supported the validity of this combination to be used in physiological experiments related to different abiotic conditions. Similarly, the *GST* expression analysis under biotic stress showed the same trend: the use of two reference gene combination (*ATPase* and *EIF5B*) led to accurate expression quantification, compared to RNA-seq results, at the contrary of the actin gene normalization, which has to be avoided in such conditions.

Conclusions

In this study, eight candidate genes with low expression variance were screened using large transcriptomic datasets generated in the kelp *S. latissima* under biotic and abiotic stress conditions. Among these genes, two combinations of reference genes were identified according to their expression stability under the different physiological conditions tested: *EIF5B* and *ATPase* are the best-adapted reference genes under biotic stress-related conditions, while *NDH*, *ATPase*, and *EIF5B* are the optimal combination for RT-qPCR normalization under temperature and light stress. Moreover, our study clearly showed that a commonly used reference gene, coding for Actin, is not suitable to normalize RT-qPCR gene expression under the different physiological conditions in *S. latissima*, its expression being too much variable. This result further illustrates the potential of large transcriptomic datasets for identifying good candidates and the importance of conducting specific experimental validation, before selecting accurate reference gene in RT-qPCR, especially in new algal models. The two validated reference gene combinations are expected to be further used as reliable calibrators

for targeted gene expression analysis and for monitoring gene biomarkers by RT-qPCR in *S. latissima* upon both abiotic and biotic environmental conditions.

Acknowledgments We would like to thank L. Darteville and E. Rolland from the LBI2M's Algal Culture facilities for technical support. This work has also benefited from the computational resources of the ABiMS bioinformatics platform (FR 2424, CNRS-Sorbonne Université, Roscoff), which are part of the Biogenouest core facility network.

Authors' contributions QX and CL conceived the experiments. QX and SR processed algal and RNA samples, and QX performed the in silico and RT-qPCR analyses and interpreted the data. QX wrote the first draft of the manuscript and SR and CL contributed substantially to the revisions. All authors read and approved the final manuscript.

Funding This project has received funding from the European Union's Horizon 2020 research and innovation program under the grant agreement N°727892, GENIALG (GENetic diversity exploitation for Innovative macro-228 ALGal biorefinery, <http://genialgproject.eu/>) for data analysis, from the Blue Granary company (QX's fellowship), and from the Centre National de la Recherche Scientifique (CNRS). It was also supported by the French National Research Agency via the "Investment for the Future" project IDEALG (no. ANR-10-BTBR-04).

Data availability The RNA-seq datasets analyzed during the current study are available in the EMBL databases under the accession ID: PRJEB37483 and E-MEXP-3450 (ArrayExpress at the EBI; <http://www.ebi.ac.uk/microarray-as/ae/>; ID: E-MEXP-3450). All other data generated during this study are either included in this published article and its supplementary information files, or available from the corresponding author on request.

Compliance with ethical standards

Conflict of interest The authors declare that they have no competing interests.

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

References

- Bartsch I, Wiencke C, Bischof K, Buchholz CM, Buck BH, Eggert A, Feuerpfeil P, Hanelt D, Jacobsen S, Karez R, Karsten U, Molis M, Roleda MY, Schubert H, Schumann R, Valentin K, Weinberger F, Wiese J (2008) The genus *Laminaria sensu lato*: recent insights and developments. *Eur J Phycol* 43:1–86
- Bernard MS, Strittmatter M, Murúa P, Heesch S, Cho GY, Leblanc C, Peters AF (2019) Diversity, biogeography and host specificity of kelp endophytes with a focus on the genera *Laminarionema* and *Laminariocolax* (Ectocarpales, Phaeophyceae). *Eur J Phycol* 54:39–51
- Buschmann AH, Camus C, Infante J, Neori A, Israel A, Hernández-González MC, Pereda SV, Gomez-Pinchetti JL, Golberg A, Tadmor-Shalev N, Critchley AT (2017) Seaweed production: overview of the global state of exploitation, farming and emerging research activity. *Eur J Phycol* 52:391–406
- Cosse A, Potin P, Leblanc C (2009) Patterns of gene expression induced by oligoguluronates reveal conserved and environment-specific molecular defense responses in the brown alga *Laminaria digitata*. *New Phytol* 182:239–250
- Czechowski T, Stitt M, Altmann T, Udvardi MK, Scheible WR (2005) Genome-wide identification and testing of superior reference genes for transcript normalization in *Arabidopsis*. *Plant Physiol* 139:5–17
- Deng Y, Yao J, Wang X, Guo H, Duan D (2012) Transcriptome sequencing and comparative analysis of *Saccharina japonica* (Laminariales, Phaeophyceae) under blue light induction. *PLoS One* 7:e39704
- Dheda K, Huggett J, Chang J, Kim L, Bustin S, Johnson M, Rook G, Zumla A (2005) The implications of using an inappropriate reference gene for real-time reverse transcription PCR data normalization. *Anal Biochem* 344:141–143
- Fiume E, Fletcher JC (2012) Regulation of *Arabidopsis* embryo and endosperm development by the polypeptide signaling molecule CLE8. *Plant Cell* 24:1000–1012
- Fu Y, He W, Wang L, Wei Y (2015) Selection of appropriate reference genes for quantitative real-time PCR in *Oxytropis ochrocephala* Bunge using transcriptome datasets under abiotic stress treatments. *Front Plant Sci* 6:475
- Gao D, Kong F, Sun P, Bi G, Mao Y (2018) Transcriptome-wide identification of optimal reference genes for expression analysis of *Pyropia yezoensis* responses to abiotic stress. *BMC Genomics* 19:251
- Goultier S, Ritter A, Thomas F, Ferec C, Salaün JP, Potin P (2009) Release of volatile aldehydes by the brown algal kelp *Laminaria digitata* in response to both biotic and abiotic stress. *ChemBioChem* 10:977–982
- Haug A, Larsen B, Smidsrød O (1974) Uronic acid sequence in alginate from different sources. *Carbohydrate Res* 32:217–225
- Heinrich S, Valentin K, Frickenhaus S, John U, Wiencke C (2012) Transcriptomic analysis of acclimation to temperature and light stress in *Saccharina latissima* (Phaeophyceae). *PLoS One* 7:e44342
- Heinrich S, Valentin K, Frickenhaus S, Wiencke C (2015) Temperature and light interactively modulate gene expression in *Saccharina latissima* (Phaeophyceae). *J Phycol* 51:93–108
- Koeduka T, Ishizaki K, Mwenda CM, Hori K, Sasaki-Sekimoto Y, Ohta H, Kohchi T, Matsui K (2015) Biochemical characterization of allene oxide synthases from the liverwort *Marchantia polymorpha* and green microalgae *Klebsormidium flaccidum* provides insight into the evolutionary divergence of the plant CYP74 family. *Planta* 242:1175–1186
- Konotchick T, Dupont CL, Valas RE, Badger JH, Allen AE (2013) Transcriptomic analysis of metabolic function in the giant kelp, *Macrocystis pyrifera*, across depth and season. *New Phytol* 198:398–407
- Kowalczyk N, Rousvoal S, Hervé C, Boyen C, Collén J (2014) RT-qPCR normalization genes in the red alga *Chondrus crispus*. *PLoS One* 9:e86574
- Küpper FC, Kloareg B, Guern J, Potin P (2001) Oligoguluronates elicit an oxidative burst in the brown algal kelp *Laminaria digitata*. *Plant Physiol* 125:278–291
- Küpper FC, Müller DG, Peters AF, Kloareg B, Potin P (2002) Oligoalginic acid recognition and oxidative burst play a key role in natural and induced resistance of sporophytes of Laminariales. *J Chem Ecol* 28:2057–2081
- Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357

- Le Bail A, Dittami SM, de Franco PO, Rousvoal S, Cock MJ, Tonon T, Charrier B (2008) Normalisation genes for expression analyses in the brown alga model *Ectocarpus siliculosus*. BMC Mol Biol 9:75
- Lee JH, Pestova TV, Shin BS, Cao C, Choi SK, Dever TE (2002) Initiation factor eIF5B catalyzes second GTP-dependent step in eukaryotic translation initiation. Proc Nat Acad Sci 99:16689–16694
- Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12:323
- Li H, Monteiro C, Heinrich S, Bartsch I, Valentin K, Harms L, Glöckner G, Corre E, Bischof K (2020) Responses of the kelp *Saccharina latissima* (Phaeophyceae) to the warming Arctic: from physiology to transcriptomics. Physiol Plant 168:5–26
- Lim CW, Han SW, Hwang IS, Kim DS, Hwang BK, Lee SC (2015) The pepper lipoxygenase CaLOX1 plays a role in osmotic, drought and high salinity stress response. Plant Cell Physiol 56:930–942
- Liu X, Guan H, Song M, Fu Y, Han X, Lei M, Ren J, Guo B, He W, Wei Y (2018) Reference gene selection for qRT-PCR assays in *Stellera chamaejasme* subjected to abiotic stresses and hormone treatments based on transcriptome datasets. PeerJ 6:e4535
- López MA, Vicente J, Kulasekaran S, Vellosillo T, Martínez M, Irigoyen ML, Cascón T, Bannenberg G, Hamberg M, Castresana C (2011) Antagonistic role of 9-lipoxygenase-derived oxylipins and ethylene in the control of oxidative stress, lipid peroxidation and plant defence. Plant J 67:447–458
- Lüning K, Mortensen L (2015) European aquaculture of sugar kelp *Saccharina latissima* for food industries: iodine content and epiphytic animals as major problems. Bot Mar 58:449–455
- Ma R, Xu S, Zhao Y, Xia B, Wang R (2016) Selection and validation of appropriate reference genes for quantitative real-time PCR analysis of gene expression in *Lycoris aurea*. Front Plant Sci 7:536
- Monteiro CMM, Li H, Bischof K, Bartsch I, Valentin KU, Corre E, Collén J, Harms L, Glöckner G, Heinrich S (2019) Is geographical variation driving the transcriptomic responses to multiple stressors in the kelp *Saccharina latissima*? BMC Plant Biol 19:1–15
- Niaz Z, Sui Z, Riaz S, Liu Y, Shang E, Xing Q, Khan S, Du Q, Zhou W (2019) Identification of valid reference genes for the normalization of RT-qPCR gene expression data in *Alexandrium catenella* under different nutritional conditions. J Appl Phycol 31:1819–1833
- Pombo MA, Zheng Y, Fei Z, Martin GB, Rosli HG (2017) Use of RNA-seq data to identify and validate RT-qPCR reference genes for studying the tomato-*Pseudomonas* pathosystem. Sci Rep 7:44905
- Potin P, Bouarab K, Salaün JP, Pohnert G, Kloareg B (2002) Biotic interactions of marine algae. Curr Opin Plant Biol 5:308–317
- Provasoli L (1968) Media and prospects for the cultivation of marine algae. Cultures and Collections of Algae. Proc US-Japan Conference, Hakone, September 1966, Japan Soc Plant Physiol
- Ritter A, Cabioch L, Brillet-Guéguen L, Corre E, Cosse A, Darteville L, Duruflé H, Fasshauer C, Goulitquer S, Thomas F, Correa JA, Potin P, Faugeron S, Leblanc C (2017) Herbivore-induced chemical and molecular responses of the kelps *Laminaria digitata* and *Lessonia spicata*. PloS One 12:e0173315
- Silveira TLR, Domingues WB, Remião MH, Santos L, Barreto B, Lessa IM, Varela Junior AS, Martins Pires D, Corcini C, Collares T, Seixas FK, Robaldo RB, Campos VF (2018) Evaluation of reference genes to analyze gene expression in silverside *Odontesthes humensis* under different environmental conditions. Front Genet 9:75
- Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. Nature Rev Genet 10:57
- Wynne M (1969) Life history and systematic studies of some Pacific North American Phaeophyceae (brown algae). Univ Calif Publ Bot 50:1–88
- Xu D, Brennan G, Xu L, Zhang XW, Fan X, Han WT, Mock T, McMin A, Hutchins DA, Ye N (2019) Ocean acidification increases iodine accumulation in kelp-based coastal food webs. Glob Chang Biol 25:629–639
- Yang C, Pan H, Noland JE, Zhang D, Zhang Z, Liu Y, Zhou X (2015) Selection of reference genes for RT-qPCR analysis in a predatory biological control agent, *Coleomegilla maculata* (Coleoptera: Coccinellidae). Sci Rep 5:18201
- Zhan A, Huang X, Li S (2019) Genome-wide identification and evaluation of new reference genes for gene expression analysis under temperature and salinity stresses in *Ciona savignyi*. Front Genet 10:71
- Zhang Y, Peng X, Liu Y, Li Y, Luo Y, Wang X, Tang H (2018) Evaluation of suitable reference genes for qRT-PCR normalization in strawberry (*Fragaria*×*Ananassa*) under different experimental conditions. BMC Mol Biol 19:8
- Zhang Y, Wang X, Shan T, Pang S, Xu N (2019) Transcriptome profiling of the meristem tissue of *Saccharina japonica* (Phaeophyceae, Laminariales) under severe stress of copper. Mar Genomics 47:100671

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