Selection of reliable reference genes for normalization of quantitative RT-PCR from different developmental stages and tissues in amphioxus

Qi-Lin Zhang1,*, Qian-Hua Zhu2,*, Xin Liao1,*, Xiu-Qiang Wang1, Tao Chen2, Han-Ting Xu3, Juan Wang3, Ming-Long Yuan3 & Jun-Yuan Chen1,4

Amphioxus is a closest living proxy to the ancestor of cephalochordates with vertebrates, and key animal for novel understanding in the evolutionary origin of vertebrate body plan, genome, tissues and immune system. Reliable analyses using quantitative real-time PCR (qRT-PCR) for answering these scientific questions is heavily dependent on reliable reference genes (RGs). In this study, we evaluated stability of thirteen candidate RGs in qRT-PCR for different developmental stages and tissues of amphioxus by four independent (geNorm, NormFinder, BestKeeper and deltaCt) and one comparative algorithms (RefFinder). The results showed that the top two stable RGs were the following: (1) S20 and 18S in thirteen developmental stages, (2) EF1A and ACT in seven normal tissues, (3) S20 and L13 in both intestine and hepatic caecum challenged with lipopolysaccharide (LPS), and (4) S20 and EF1A in gill challenged with LPS. The expression profiles of two target genes (EYA and HHEX) in thirteen developmental stages were used to confirm the reliability of chosen RGs. This study identified optimal RGs that can be used to accurately measure gene expression under these conditions, which will benefit evolutionary and functional genomics studies in amphioxus.
RGs were variable to some degree under different conditions\(^2\,\)\(^21\). Wang et al.\(^2\)\(^2\) used the BestKeeper and deltaCt algorithms to rank the expression stability for each RG, and two genes (EF1A and GAPDH) were the most stable genes, followed by EF1A and GAPDH (Supplementary Table S1). The results of comprehensive ranking with the ReFiFinder showed that the most stable genes were S20, 18S and EF1A, whereas four genes (GAPDH, HPRT, ACT and L13) were unstable (GM > 9.0).

### Determination of expression stability of reference genes in different tissues

For seven normal tissues, the three genes (EF1A, ACT and G6PDH) were the most stable by the geNorm analyses (Supplementary Fig. S5), which was consistent with the results of the deltaCt and Bestkeeper analyses. NormFinder consistently supported TBP, EF1A and G6PDH as the top three ranked RGs.
| Gene | gene ID | Primer Sequence | Amplicon Size (bp) | PCR Efficiency (E%)/Correlation Coefficient (R²) |
|------|---------|----------------|-------------------|-----------------------------------------------|
| 18S  | M97571.1* | F: CCTGAGAAGACGCTACTCC  |
|      |         | R: ATTAAAGTCTACCCATTTC  | 135               | 98.60/0.994                                    |
| ACT  | 249410R | F: ATCGTGCCGTGACATCAAGGA  |
|      |         | R: GGGAGGAAGCCTGGAAGAGA  | 178               | 97.35/0.999                                    |
| CYC  | 019450F | F: CTTCCGACATTACCCTGTAGG  |
|      |         | R: GGAGTGTCTCTGCGGTCTTG   | 120               | 101.78/0.998                                   |
| EF1A | 210870F | F: CTGTGCCGCTGCTGATTTGA  |
|      |         | R: GTTGAGATCTACTTGGTGAC   | 145               | 103.89/0.998                                   |
| GAPDH| 065110R | F: CGGCCAGTGTAAGCTGAAC   |
|      |         | R: CAGGATGATCTCTGCTGATGC  | 181               | 100.37/0.997                                   |
| HBP  | 238710R | F: CTTCTCTCTGGACTCTGACC  |
|      |         | R: TCCTGGAATATGGATCACAT  | 186               | 93.87/0.998                                    |
| L3   | 025620F | F: ACAAGGGTCTCGCTAAAGT   |
|      |         | R: CTCAGTGCGGTTGTTATACG  | 104               | 101.35/0.990                                   |
| L13  | 247350F | F: TGTTGGCTAGGAGTTGAAAGA  |
|      |         | R: ACTGGAAAGCCTGTTACTT    | 125               | 101.65/0.999                                   |
| S20  | 107260F | F: ACTCAGATCCACGCACTCA    |
|      |         | R: CCTTCACTTGGAGGTTTCTC   | 112               | 100.24/0.998                                   |
| TBP  | 122820R | F: TGGCCAGATGTTACAAGAT    |
|      |         | R: ATACGTCATGTCAGGCGAGG   | 176               | 100.04/0.993                                   |
| TUB  | 227130F | F: CCGTGTGATGTGTGAGTA    |
|      |         | R: ACCCTGCTAGTCTTCTC      | 184               | 100.71/0.996                                   |
| UBC  | 063240R | F: TGTCCATCTGCTACGCTCA    |
|      |         | R: TGGCGTATCTCTGGTCATC    | 130               | 99.66/0.995                                    |
| EYA* | 223870R | F: CCAAGCACCACAAACCACATAC |
|      |         | R: CAGGCCAGGTTGAAGATCCTC  | 306               | 95.72/0.995                                    |
| HHEX | 128560F | F: TACCCGACAGCTTCTCCTA    |
|      |         | R: TCTGGGACTGCAACTTCTT    | 106               | 100.50/0.990                                   |

Table 1. Description of thirteen reference genes in *B. belcheri* for qRT-PCR analysis. *DNA sequences download directly from Genbank. *Two target genes.

| Gene | RefFinder | geNorm | NormFinder | BestKeeper | deltaCt |
|------|-----------|--------|------------|------------|---------|
| S20  | 1         | 1.72   | 3          | 0.85       | 2       | 0.72 | 1 | 0.51 | 2 | 1.84 |
| 18S  | 2         | 2.23   | 1          | 0.76       | 3       | 0.76 | 2 | 0.55 | 5 | 1.89 |
| EF1A | 3         | 2.74   | 7          | 1.49       | 1       | 0.64 | 7 | 1.71 | 1 | 1.82 |
| CYC  | 4         | 3.11   | 1          | 0.76       | 5       | 0.84 | 3 | 0.63 | 4 | 1.88 |
| GAPDH| 5         | 4.79   | 4          | 0.95       | 9       | 1.03 | 4 | 0.67 | 3 | 1.86 |
| UBC  | 6         | 5.92   | 6          | 1.29       | 6       | 0.92 | 6 | 1.28 | 8 | 2.04 |
| L3   | 7         | 6.22   | 8          | 1.63       | 4       | 0.79 | 9 | 1.82 | 6 | 1.94 |
| TBP  | 8         | 7.04   | 5          | 1.15       | 10      | 1.07 | 5 | 0.84 | 10 | 2.12 |
| TUB  | 9         | 8.67   | 9          | 1.71       | 7       | 0.95 | 10 | 1.86 | 7 | 2.03 |
| GAPDH| 10        | 9.41   | 10         | 1.76       | 8       | 1.02 | 8 | 1.74 | 9 | 2.10 |
| HPRT | 11        | 11.37  | 11         | 1.85       | 12      | 1.46 | 11 | 1.89 | 12 | 2.59 |
| ACT  | 12        | 12.26  | 12         | 1.99       | 11      | 1.41 | 12 | 2.59 | 11 | 2.54 |
| L13  | 13        | 12.83  | 13         | 2.03       | 13      | 1.71 | 13 | 2.87 | 13 | 2.84 |

Table 2. Stability ranking of the candidate reference genes using five algorithms in different developmental stages of *B. belcheri*. Parameters of ranking were showed by using GM (geometric mean); SV, stability value. If no additional declarations, these abbreviations indicate the same means as above described in other tables.
comprehensive ranking with the RefFinder showed that three genes (EF1A, ACT and G6PDH) were the most stable, while four genes (18S, CYC, L3 and S20) exhibited unstable expressions (Table 3, Supplementary Tables S2 and S3).

**Table 3. Stability ranking of the candidate reference genes using five algorithms in different normal tissues of B. belcheri.**

| Gene | RefFinder Rank | geNorm Rank | NormFinder Rank | BestKeeper Rank | deltaCt Rank |
|------|----------------|-------------|----------------|-----------------|--------------|
| EF1A | 1 | 1.71 | 1 | 0.45 | 2 | 0.63 | 1 | 0.11 | 1 | 2.92 |
| ACT  | 2 | 1.84 | 1 | 0.45 | 6 | 1.51 | 2 | 0.59 | 3 | 3.18 |
| G6PDH| 3 | 2.83 | 3 | 0.90 | 3 | 1.08 | 3 | 0.85 | 2 | 3.14 |
| TUB  | 4 | 4.96 | 6 | 2.25 | 7 | 1.54 | 5 | 2.02 | 4 | 3.26 |
| HPRT | 5 | 5.32 | 4 | 1.44 | 5 | 1.43 | 4 | 1.55 | 7 | 3.48 |
| L13  | 6 | 5.96 | 7 | 2.47 | 4 | 1.39 | 7 | 2.29 | 5 | 3.40 |
| TBP  | 7 | 6.77 | 5 | 2.04 | 1 | 0.60 | 8 | 2.43 | 6 | 3.46 |
| UBC  | 8 | 8.11 | 9 | 2.78 | 9 | 1.94 | 6 | 2.23 | 9 | 3.91 |
| GAPDH| 9 | 8.64 | 8 | 2.64 | 8 | 1.80 | 10 | 2.45 | 8 | 3.86 |
| 18S  | 10 | 10.07 | 11 | 3.15 | 11 | 2.37 | 9 | 2.43 | 10 | 4.28 |
| CYC  | 11 | 11.35 | 10 | 2.93 | 10 | 2.27 | 11 | 3.06 | 11 | 4.33 |
| L3   | 12 | 12.04 | 12 | 3.43 | 12 | 3.16 | 12 | 3.70 | 12 | 4.76 |
| S20  | 13 | 12.87 | 13 | 3.79 | 13 | 3.70 | 13 | 3.88 | 13 | 5.74 |

**Table 4. Stability ranking of the candidate reference genes using five algorithms in intestine of B. belcheri challenged with LPS.**

| Gene | RefFinder Rank | geNorm Rank | NormFinder Rank | BestKeeper Rank | deltaCt Rank |
|------|----------------|-------------|----------------|-----------------|--------------|
| S20  | 1 | 2.81 | 1 | 0.28 | 1 | 0.10 | 7 | 1.62 | 1 | 1.32 |
| L13  | 2 | 3.30 | 1 | 0.28 | 2 | 0.23 | 5 | 1.61 | 2 | 1.37 |
| EF1A | 3 | 3.74 | 5 | 0.86 | 7 | 0.78 | 9 | 1.87 | 7 | 1.66 |
| UBC  | 4 | 4.01 | 3 | 0.64 | 3 | 0.61 | 10 | 2.08 | 3 | 1.58 |
| ACT  | 5 | 5.13 | 6 | 0.90 | 4 | 0.64 | 6 | 1.62 | 4 | 1.59 |
| GAPDH| 6 | 5.31 | 4 | 0.81 | 6 | 0.76 | 11 | 2.25 | 7 | 1.69 |
| HPRT | 7 | 5.84 | 8 | 1.17 | 5 | 0.67 | 4 | 1.35 | 5 | 1.63 |
| 18S  | 8 | 6.33 | 11 | 1.53 | 12 | 1.32 | 1 | 0.18 | 11 | 2.16 |
| G6PDH| 9 | 6.74 | 9 | 1.32 | 8 | 0.96 | 2 | 0.47 | 9 | 1.81 |
| TBP  | 10 | 7.52 | 10 | 1.42 | 10 | 1.06 | 3 | 0.57 | 10 | 1.90 |
| TUB  | 11 | 7.87 | 7 | 1.05 | 9 | 0.98 | 8 | 1.82 | 8 | 1.76 |
| CYC  | 12 | 12.24 | 12 | 1.62 | 11 | 1.24 | 12 | 2.35 | 12 | 2.34 |
| L3   | 13 | 12.89 | 13 | 1.81 | 13 | 1.80 | 13 | 2.45 | 13 | 2.60 |

**Determination for expression stability of reference genes in three different challenged tissues with LPS.** In the intestine challenged with LPS, three genes (S20, L13 and UBC) were the most stable analyzed by three independent algorithms (geNorm, NormFinder and deltaCt), whereas BestKeeper supported 18S, G6PDH and TBP as the top three RGs. In the gill challenged with LPS, three genes (S20, EF1A and UBC) were the most stable in all independent analyses, except for the BestKeeper analyses where 18S, ACT and G6PDH were the three top RGs. In the hepatic caecum challenged with LPS, geNorm and BestKeeper recommended S20, L13 and TBP as the top three stable genes, whereas three most stable genes were supported by NormFinder (S20, ACT and L13) and deltaCt (L13, ACT and S20) (Tables 4, 5 and 6, Supplementary Fig. S5, Supplementary Tables S2 and S3).

The results of comprehensive ranking with the RefFinder showed that S20 and L13 were the most stable genes in both challenged intestine and hepatic caecum, followed by EF1A in the former and TBP in the latter. In addition to S20, EF1A and UBC were the most stable genes in challenged gill. L3 was consistently recommended as the most unstable gene, followed by CYC in challenged intestine and gill, 18S, G6PDH and GAPDH in challenged hepatic caecum.

**Determinate optimal number of normalization factors under each experimental condition.** To determine optimal NFs number for normalization, the pairwise variation value \((V_n/V_{n+1}, V\text{-value})\) was calculated by geNorm. If V-value is firstly lower than the default value 0.15 or lowest value in all pairwise variation28, the number of gene pairings will be sufficient for the consistent normalization. In different developmental stages, V7/8 showed a minimum V-value, suggesting that the optimal number of RGs for normalization was seven. V8/9...
with the lowest V-value indicated that eight stable reference genes were reliable as NFs in the different normal tissues. For three challenged tissues with LPS, V-values of V4/5 were first lower than 0.15, suggesting that four RGs could be used for normalization in three tissues challenged with LPS (Fig. 1).

However, it is time-consuming and expensive for normalization by using excess RGs in the actual experiments, especially when large number of the target genes and the rare experimental samples are used. Generally, a reliable result could be obtained by using three or more RGs for normalization. To further identify the reliable and the least number of NFs, we calculated the correlation between NFn (n ranging from 3 to NFopt) and optimal number of NFs from geNorm (NFopt) for each experimental condition by Pearson correlation coefficient (r).

Then we considered NFn, which contained the minimum number of RGs and no significant difference with NFopt, as a target number of NFs. The results showed high correlation (r > 0.8) between the NF 5 and NFopt (r = 0.83, p < 0.01), indicating that normalization for expression level of target genes by combining the five most stable RGs (NF5) could obtain the same reliable results as normalization of NFopt in different normal tissues. The same is true for the normalization of the target genes in different developmental stages (r = 0.85, p < 0.01), challenged intestine (r = 0.99, p < 0.01) and gill (r = 0.97, p < 0.01) using top three stable RGs (NF3). However, four top RGs (NF4) may be necessary for normalization in challenged hepatic caecum due to low correlation between NF3 and NFopt though it reached the significant level (r = 0.69, p = 0.046) (Supplementary Fig. S6).

Determination of reference gene validation in different developmental stages of B. belcheri.
We performed expression profile analyses for two target genes (EYA and HHEX) using three [S20, 18S and EF1A; NF (1–3)], seven of the most stable genes [S20, 18S, EF1A, CYC, G6PDH, UBC and L3; NF (1–7)] as well as two of the least stable RGs [ACT and L13; NF (12–13)] as NFs for the normalization (Fig. 2 and Supplementary Fig. S7).

Comparing with middle gastrulae, a down-regulation of EYA expression was observed during later gastrulae stage by using NF (12–13) for normalization. Instead, when we used NF (1–3) and NF (1–7), we observed an

### Table 5. Stability ranking of the candidate reference genes using five algorithms in gill of B. belcheri challenged with LPS.

| Gene    | RefFinder | geNorm | NormFinder | BestKeeper | deltaCt |
|---------|-----------|--------|------------|------------|---------|
|         | Rank GM   | Rank SV | Rank SV    | Rank SV    | Rank SV |
| S20     | 1 1.21    | 1 0.51 | 1 0.30     | 7 1.04     | 1 1.40  |
| EF1A    | 2 2.95    | 3 0.78 | 2 0.41     | 8 1.10     | 3 1.44  |
| UBC     | 3 3.84    | 1 0.51 | 3 0.45     | 10 1.20    | 2 1.44  |
| ACT     | 4 3.91    | 5 0.97 | 4 0.49     | 2 0.64     | 5 1.51  |
| TUB     | 5 4.98    | 4 0.91 | 6 0.56     | 6 1.03     | 4 1.49  |
| 18S     | 6 5.61    | 10 1.31| 10 0.99    | 1 0.52     | 11 1.86 |
| HPRT    | 7 6.05    | 7 1.11 | 7 0.84     | 4 0.94     | 7 1.72  |
| L13     | 8 6.67    | 6 1.02 | 5 0.52     | 9 1.16     | 6 1.53  |
| G6PDH   | 9 6.77    | 9 1.27 | 9 0.90     | 3 0.72     | 8 1.79  |
| TBP     | 10 8.61   | 8 1.20 | 11 1.02    | 5 0.96     | 10 1.84 |
| GAPDH   | 11 9.85   | 11 1.38| 8 0.90     | 11 1.67    | 9 1.81  |
| CTC     | 12 11.83  | 12 1.59| 12 1.68    | 12 1.96    | 12 2.71 |
| L3      | 13 12.74  | 13 1.80| 13 1.91    | 13 2.49    | 13 2.97 |

### Table 6. Stability ranking of the candidate reference genes using five algorithms in hepatic caecum of B. belcheri challenged with LPS.

| Gene    | RefFinder | geNorm | NormFinder | BestKeeper | deltaCt |
|---------|-----------|--------|------------|------------|---------|
|         | Rank GM   | Rank SV | Rank SV    | Rank SV    | Rank SV |
| S20     | 1 0.94    | 1 0.76 | 1 0.41     | 1 0.42     | 3 1.40  |
| L13     | 2 1.91    | 1 0.76 | 3 0.49     | 2 0.45     | 1 1.35  |
| TBP     | 3 3.55    | 3 0.91 | 4 0.57     | 3 0.51     | 4 1.41  |
| ACT     | 4 3.92    | 5 1.02 | 2 0.45     | 7 0.84     | 2 1.35  |
| HPRT    | 5 4.82    | 4 0.93 | 6 0.62     | 4 0.56     | 5 1.45  |
| TUB     | 6 6.11    | 6 1.06 | 7 0.70     | 5 0.69     | 7 1.52  |
| EF1A    | 7 6.76    | 7 1.11 | 5 0.61     | 9 1.06     | 6 1.48  |
| CYC     | 8 7.72    | 8 1.16 | 8 0.81     | 6 0.82     | 8 1.61  |
| UBC     | 9 9.81    | 9 1.26 | 9 0.87     | 11 1.23    | 9 1.69  |
| GAPDH   | 10 10.04  | 10 1.33| 10 0.97    | 10 1.21    | 10 1.79 |
| G6PDH   | 11 10.32  | 11 1.41| 11 1.01    | 8 0.96     | 11 1.83 |
| 18S     | 12 11.91  | 12 1.46| 12 1.01    | 12 1.26    | 12 1.85 |
| L3      | 13 12.52  | 13 1.63| 13 1.62    | 13 2.34    | 13 2.55 |
up-regulation of the EYA expression. From hatching stage to adult stage of amphioxus, the highest expression level of EYA was observed during adult stage normalized with NF (12–13), but 2-gill arch stage and two week
after fertilization showed the highest expression level with NF (1–3) or NF (1–7) for normalization. The highest expression level of EYA (neurula stage) relative to lowest one (two month after fertilization) reached 694-fold when using NF (12–13) for normalization. However, using NF (1–3) or NF (1–7) for normalization of EYA, the highest expression level (neurula stage) relative to lowest one (8-cell stage) was 8- and 20-fold, respectively. These results showed that high expression of EYA mainly happened during embryonic stages. Expression level of EYA reached to peak of transcripts abundant at neural stage, and maintained the lowest expression level at 8-cell stage. Further, expression of EYA still kept the higher level from hatching to adult than 8-cell stage, indicating EYA had an important role in developmental process after embryonic stages. HHEX had a maximum expression level at later gastrula stage when using NF (1–3) or NF (1–7) for normalization, but it showed the highest expression level at neurula stage by normalization of NF (12–13). Although we investigated the down-regulation of these two target genes after neurula or hatching, medium expression levels were found after embryonic stages with slight up-regulation at later developmental stages. However, the expression profiles of EYA and HHEX evidently were altered and distorted when using NF (12–13) for normalization.

Discussion

Expression patterns of target genes by qRT-PCR analysis in amphioxus are important for exploring development homology, gene evolution and comparative immunology between vertebrates and cephalochordates. According to MIQE guidelines, the use of RGs as internal controls is the most appropriate normalization way for qRT-PCR analyses37. In this study, we firstly performed analyses for reliability evaluation of candidate RGs in different developmental stages (especially for embryonic stages), challenged intestine, gill and hepatic caecum with LPS, as well as further determined and improved reliable RGs in different normal tissues of amphioxus.

Comparative analyses showed a high consistency for the ranking of stability for thirteen candidate RGs under each experimental condition among different statistical methods. For example, the results from the deltaCt method were highly similar to that of geNorm and NormFinder calculations, because 2/3 or even all the three top stable RGs were consistent among these three statistical algorithms. However, there were also substantial discrepancies among the results from different algorithms for each experimental group, due to different statistical models in each algorithm, as found in other studies35,38. Compared to the other three independent algorithms, the BestKeeper exhibited the most discrepancies, as has been reported in previous studies37. In order to overcome differences among different algorithms, we performed overall ranking for thirteen candidates RGs based on the fifth algorithm (RefFinder) to obtain the final stability. In the different developmental stages, the three RGs (S20, 18S and EF1A) were identified as the most reliable, while the other three RGs (L13, HPRT and ACT) should be avoided in future study due to the exhibition of their low stability in all algorithms. In the different normal tissues of amphioxus, EF1A exhibited the most stability, consisting with the previous studies37. In contrast, the four RGs (S20, L3, CYC and 18S) should not be considered for normalization of tissue-specific gene expressions. Among all candidate RGs, S20 was ranked as a universal RG within three different tissues challenged with LPS, while L3 was one with the least stability. Besides, stability ranking of candidate RGs with GM value < 9 in three challenged tissues by LPS was also highly similar, suggesting that selecting the same RGs was reliable for normalization in homologous organs of amphioxus under immune-stimulation conditions.

S20 was a reliable RG in different tissues and immune relevant tissues of Atlantic salmon (Salmo salar)39, in different Sesamia inferens tissues and fifth instars larva treated by different temperatures38, but this gene exhibited the least stable expression by geNorm analysis in fruitfly by injury treatments39. EF1A was a stable RG in different developmental stages of Sesamia inferens and intestinal tissues of sea bass (Dicentrarchus labrax)38,40, but was the most variable gene in virus-infected plant hoppers41. In our study, no matter in different developmental stages, different tissues or three challenged tissues of amphioxus, S20 and EF1A had a good performance for their expression stability, which was very similar to that of Atlantic salmon35, but was discrepant with virus-insects pathosystems41. The three genes (18S, ACT and GAPDH) were not a good choice as RGs in developmental stages of Monopterus albus42, but were considered as reliable ones in virus-immune cell19 and human prostate cancer44. Our results showed that only in different developmental stages 18S was a reliable RG, whereas ACT and GAPDH exhibited the higher variations in different experimental conditions. These variations indicated that expression stability was different for the same RGs among multiple species and different RGs in the same experimental conditions, so the determination of RGs validation was essential for each specific condition in further experiments. L13 was not considered as a stable RG in avian species37; however, in our study stability ranking of this gene reached top three in intestine and hepatic caecum challenged with LPS. Therefore, unconventional RGs (e.g. L13) should not be ignored as candidate ones in further experiments of RGs selection.

According to previous studies, use of multiple RGs for normalization could obtain more accurate results than single RG28, because a biased normalization can be revised by a RG combination. An optimal number of NFs for normalization was evaluated by pairwise variation analysis in the geNorm. We further calculated correlation between NF and NFE to decrease number of NFs as soon as possible under precondition of no effect on accuracy of normalization. Our results will provide a practical number of NFs in common experiments.

The validation of selected RGs was confirmed by expression profiles of two target genes (EYA and HHEX). EYA gene family comprises of four members in vertebrates (i.e. EYA1, EYA2, EYA3 and EYA4) and modulates cell proliferation by phosphatase function to activate specific gene targets43. EYA also involves in innate immunity, DNA repair, cell migration, and cancer metastasis in adult vertebrates44. However, only one EYA was found in amphioxus and its indispensability for early development of amphioxus has been exhibited, potentially interacting with other gene functions44. Origin of EYA could be traced to fruitfly and it regulates eye development involved in cell proliferation, patterning, and neuronal information for invertebrates45. HHEX encodes an oligomeric homeodomain-containing transcription factor, and it was firstly cloned in hematopoietic tissues and highly conserved evolution in vertebrates46-48. HHEX is essential for embryonic development, especially for liver, thyroid and forebrain in mammal49. It is highly expressed in many kinds of hematopoietic cells, such as stem cells,
myeloid and lymphoid progenitors. Expression profiles of these two target genes showed that the transcript abundance of EYA and HHEX was strongly influenced with the development of amphioxus. Expression level of EYA and HHEX after normalization by NF (12–13) showed a huge difference from the results based on NF (1–3) or NF (1–7). Therefore, the normalization results based on NF (1–3) could not truthfully reflect the expression level of target genes in amphioxus. We found that high expression level of HHEX mainly concentrated at embryonic stages, which was also observed in endostyle of amphioxus by transcriptome of different amphioxus tissues (http://wcy.pkusulab.com/) and whole mount in situ hybridization (data not shown). Previous study reported that lymphocyte-like cells were found in endostyle of sea squirt and this tissue may be the germinal center of adult stem cells, implying that endostyle played a key role in amphioxus immunity and rudiment of hematopoietic cell had been formed in cephalochordate.

Overall, we obtained two RG sets (under development and adult tissues) for normalization of target genes. The RG sets under adult tissues included two treatment types (normal and immune-stimulation): one RG subset of normal tissues was used to normalize genes of tissue-specific expression for adult amphioxus, whereas the other three RG subsets under three tissues (intestine, gill and hepatic caecum) challenged with LPS were used to normalize immune-related gene expression for adult amphioxus. We found that there was no consensus in the RG sets to normalize data coming from adult tissues of amphioxus, and this was why we divided adult tissues into treated and untreated groups. In our present study, selection of RG sets for each of treated tissues was performed independently. We demonstrated that considerable variations of RG sets were found among adult tissues of amphioxus. These 13 candidate RGs were traditional RGs and stable RGs in other animals, so they could be potentially used for normalizing B. belcheri samples, particularly under development, normal tissues, challenged intestine, gill and hepatic caecum with LPS. The reliable RGs obtained here will be helpful for evolutionary and functional genomics studies in amphioxus. For other experimental conditions, however, it will be essential to evaluate the stability of 13 candidate RGs by standard process according to our manuscript.

Methods

Sample preparation. Adult specimens of B. belcheri were collected from the South China Sea (Maoming, Guangdong province, China) and reared in the cement pool with cuboid shape (1 m × 1 m × 1.2 m). We selected male and female amphioxus with a full gonad during the breeding season, and placed them into plastic cups (600 ml) with pre-filter sand and seawater in a dark place. Because most individuals were spawning during night, we collected their sperms and eggs every 30 minutes from 20:00 everyday till obtained enough experimental samples. Fertilization was performed by mixing sperms and eggs, and developmental stages were determined by using the microscope (Olympus DP71, Japan). Because of a small size for embryonic stages, we used self-made and small-bore bags consisting of stainless steel (400 mesh) to enrich experimental samples by filtering seawater that contained target stages. A total of thirteen different developmental stages were used in this study, including embryonic stages (i.e. 8-cell, morula, blastula, middle gastrula, late gastrula and neurula stages) (Supplementary Fig. S8), hatching stage, 10-somites stage, mouse-opening stage, 2-gill arch stage, two weeks after fertilization, two months after fertilization and adult stage. For each developmental stage, three biological replications were used.

Approximately 150 adult individuals were averagely put in three acrylic tanks (i.e. three biological replications) with filtered seawater, and were feed for several days to empty their contents in intestine and hepatic caecum. Seven different normal tissues (intestine, hepatic caecum, gill, skin, notochord, neural tube and muscle) were obtained from approximate 50 individuals in each tank.

We collected three tissues (intestine, gill and hepatic caecum) of B. belcheri which were challenged with 1 mg/ml LPS following the method of previous studies. For each tissue, samples with three biological replications were collected at nine timing points of immunostimulation (0 h, 2 h, 4 h, 6 h, 12 h, 24 h, 36 h, 48 h and 60 h). Each sample contained approximate 25 adult individuals. Before immunostimulation, adult individuals with empty contents of intestine and hepatic caecum were feed in several 1.5 L tanks that were filled with 1 L sterilized seawater.

All samples were put into a 1.5 ml RNase-free microcentrifuge tube containing 1 ml of TRIzol reagent (Invitrogen, Carlsbad, CA), stored overnight at 4 °C, and transferred to −20 °C till use.

Total RNA extraction and cDNA synthesis. Total RNA was isolated from each sample, as was previously described. Residual genomic DNA was digested by RNase-free DNase Set (Qiagen, Germany) according to the manufacturer’s instructions. The RNA concentration was quantified by measuring the absorbance at 260 nm using BioPhotometer Plus (Eppendorf, Germany). Quality of the total RNA was assessed by estimating the OD260/280 with expected values between 1.8 and 2.0. RNA structural integrity was verified on agarose gel electrophoresis. Single-stranded cDNA was synthesized using 1–5 μg of total RNA using an ReverTaid First Strand cDNA Synthesis Kit (Thermo Fisher Scientific, USA) with oligo(dT)18 primers according to the manufacturer’s protocol. Diluted cDNA (100 ng/μl) with free-RNase water was used for further experiments.

Identification of reference genes in the B. belcheri genome. We downloaded all target sequences annotated in GenBank and manually determined unannotated sequences based on the Chinese amphioxus genome (version B.belcheri_v18h27_r3). The coding sequence (CDS) set was downloaded from the website (http://mosas.sysu.edu.cn/) and annotated seed sequences of other model animals were hit to CDS set by using Blastn 2.2.24 (E-value 0.001). Candidate target sequences were further determined manually by using online BLAST in NCBI.
Reference genes selection, primer design and amplification efficiency test. We selected thirteen candidate RGs to investigate their robustness as internal controls for qRT-PCR and two target genes to verify the validation of recommended RGs (Table 1). Gene-specific primers were designed by Beacon Designer 7 software. The specific amplification of all candidate/target genes was confirmed by a single band of appropriate size in a 1.5% agarose gel electrophoresis. The qRT-PCR efficiency was determined for each gene by using slope analysis with a linear regression model. A pool of all of the cDNA samples was used to calculate the PCR efficiency and correlation coefficient ($R^2$) for each primer pair based on the standard curve method. Standard curves of five points were generated with serial dilutions of cDNA (1/10, 1/100, 1/1000, 1/10000, and 1/100000). The corresponding qRT-PCR efficiencies (E) were calculated according to the equation 54: $E(\%) = (10^{1(−1/\text{slope})} − 1) \times 100$.

Quantitative real-time PCR. qRT-PCR was performed using ABI 7300 real-time PCR system (Applied Biosystems, USA). cDNA was amplified in 96-well plates using the SYBR Premix Ex Taq (Takara, Japan) according to the manufacturer's protocol, with a final reaction volume of 20 μl in each well. 1 μl (about 100 ng) cDNA, 1 μl of each sense and anti-sense primer (10 μM), 10 μl of 2 × SYBR Green Premix and 7 μl ddH2O. The PCR reaction was conducted with 95 °C for 60 s, followed by 40 cycles of 95 °C for 10 s, 57 °C for 30 s and 72 °C for 35 s. In order to confirm the specificity of amplification, each reaction was performed with a dissociation curve. The reaction solution without cDNA template was used as negative controls to confirm template-specific amplification. The PCR reaction for each of three biological replicates was implemented according to above-described procedure, and the detection of each gene was performed with three technical replications in an independent sample.

Determining stability of candidate reference genes expression. Data analyses were performed independently for each of the five groups: developmental stages, normal tissues of adult amphioxus, and the other three groups (i.e. challenged intestine, gill and hepatic caecum with LPS). Average Ct value from three biological replicates was further analyzed according to previous studies55. The stability of thirteen RGs was analyzed by five algorithms: geNorm58, NormFinder29, BestKeeper60, deltaCt method31 and RefFinder42. The geNorm evaluates expression stability of each RG by calculating value (M-value) and excludes candidate one with the highest M-value (less stable) by stepwise cycles. This software also calculates pairwise variation between each RG and the other RGs to determine the optimal number of RGs required for normalization. The NormFinder ranks candidate RGs by calculating their stability value (SV) and standard error among samples in the given group29, and the higher expression stability of each gene shows a lower SV. The candidate RGs with a low SV are considered as the reliable ones in BestKeeper40. The deltaCt method calculates relative expression levels (REL) of gene pairs between one RG and the other RGs within each sample, and the candidate RGs with the smaller SD value of REL are more stable41. Finally, we comprehensively ranked candidates RGs based on the above results from four different statistical applets using a web-based analysis tool RefFinder (http://www.leonxie.com/referencegene.php). The RefFinder examines the stability of candidate RGs by calculating Geometric Mean (GM) values, and RGs with the lower GM values are considered as more stable ones.

Determination for validation of reference genes selection. To confirm the reliability of the RGs, the relative expression profiles of EYA and HHEX were determined in thirteen developmental stages and independently normalized with the three most stable RGs [S20, 18 S and EF1A; NF (1–3)], the seven top stable RGs [S20, 18 S, EF1A, CYC, G6PDH, UBC and RPL3; NF (1–7)] and the two most stable RGs [ACTIN and RPL13; NF (12–13)]. Relative quantification of these two target genes was calculated using the $2^{-ΔΔCt}$ method56. Statistical analysis of data was performed by using the IBM SPSS statistics 22 based on LSD test of one-way ANOVA. Products of statistical plots were performed by SigmaPlot 12.0.

References
1. Heid, C. A., Stevens, J., Livak, K. J. & Williams, P. M. Real-time quantitative PCR. *Genome Res.* 6, 986–994 (1996).
2. Güennin, S. et al. Normalization of qRT-PCR data: the necessity of adopting a systematic, experimental conditions-specific, validation of references. *J. Exp. Bot.* 60, 487–493 (2009).
3. Robledo, D. et al. Analysis of qPCR reference gene stability determination methods and a practical approach for efficiency calculation on a turbot (*Scophthalmus maximus*) gonad dataset. *BMC Genomics* 15, 648 (2014).
4. Czeckowski, T., Stitt, M., Altmann, T., Udvardi, M. K. & Scheible, W. R. Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. *Plant Physiol.* 139, 5–17 (2005).
5. Dragun, E. et al. A large nuclear US ribonucleoprotein required for 18S ribosomal RNA biogenesis. *Nature* 417, 967–970 (2002).
6. Crepe, C. et al. Elongator controls the migration and differentiation of cortical neurons through sirtylation of α-Tubulin. *Cell* 136, 551–564 (2009).
7. Holland, L. Z. et al. The amphioxus genome illuminates vertebrate origins and cephalochordate biology. *Genome Res.* 18, 1100–1111 (2008).
8. Putnam, N. H. et al. The amphioxus genome and the evolution of the chordate karyotype. *Nature* 453, 1064–1071 (2008).
9. Wang, Y. & Zhang, S. EF1a is a useful internal reference for studies of gene expression regulation in amphioxus Branchiostoma japonicum. *Fish Shellfish Immun.* 32, 1068–1073 (2012).
10. Huang, S. et al. Decelerated genome evolution in modern vertebrates revealed by analysis of multiple lancelet genomes. *Nat. Commun.* 5, 5896 (2014).
11. Li, G., Shu, Z. & Wang, Y. Year-round reproduction and induced spawning of Chinese amphioxus, *Branchiostoma belcheri*, in Laboratory. *PLoS ONE* 8, e75461 (2013).
12. Bustin, S. A. et al. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.* 55, 611–622 (2009).
13. Olia, P., Adam, I., Meyer, A., Scharff, C. & Gruber, A. D. Reference genes for quantitative gene expression studies in multiple avian species. *PLoS ONE* 9, e96678 (2014).
14. Mori, R., Wang, Q., Danenberg, K. D., Pinski, J. K. & Danenberg, P. V. Both 3’-actin and GAPDH are useful reference genes for normalization of quantitative RT-PCR in human FFPE tissue samples of prostate cancer. *Prostate* 68, 1555–1560 (2008).
15. Ingerslev, H. C., Petersen, E. F., Jakobsen, R. A., Petersen, C. B. & Wergeland, H. I. Expression profiling and validation of reference gene candidates in immune relevant tissues and cells from Atlantic salmon (*Salmo salar* L.). *Mol. Immunol.* 43, 1194–1201 (2006).
16. Dhorne-Pollet, S., Thélie, A. & Pollet, N. Validation of novel reference genes for RT-qPCR studies of gene expression in Xenopus tropicalis during embryonic and post-embryonic development. Dev. Dynam. 242, 709–717 (2013).
17. Wang, Y. & Zhang, S. Identification and expression of liver-specific genes after LPS challenge in amphioxus: the hepatic cecum as a liver-like organ and “pre-hepatic” acute phase response. Funct. Integr. Genomic. 11, 111–118 (2010).
18. Kuchipudi, S. V. et al. 18S rRNA is a reliable normalisation gene for real time PCR based on influenza virus infected cells. Virol. J. 9, 230 (2012).
19. Nakamura, A. M. et al. Reference genes for accessing differential expression among developmental stages and analysis of differential expression of OBP genes in Anastrepha obliqua. Sci. Rep. 6, 17480 (2016).
20. Xiao, X. et al. Validation of suitable reference genes for gene expression analysis in the halophyte Salicornia europaea by real-time quantitative PCR. Front. Plant Sci. 5, 788 (2014).
21. Thellin, O.; ElMoosayed, B.; Heinen, E. & Zorzzi, W. A decade of improvements in quantification of gene expression and internal standard selection. Biotechnol. Adv. 27, 323–333 (2009).
22. Niu, L. et al. Selection of reliable reference genes for gene expression studies of a promising oilseed crop, Plakentia volubilis, by real-time quantitative PCR. Int. J. Mol. Sci. 15, 12513–12530 (2015).
23. Jian, B. et al. Validation of internal control for gene expression study in soybean by quantitative real-time PCR. BMC Mol. Biol. 9, 1–14 (2008).
24. Chen, J. H. et al. Selection of suitable reference genes for normalization of quantitative RT-PCR in peripheral blood samples of bottlenose dolphins (Tursiops truncatus). Sci. Rep. 5, 15425 (2015).
25. Liu, M., Jiang, J., Han, X., Qiao, G. & Zhuo, R. Validation of reference genes aiming accurate normalization of qRT-PCR data in Dendrocalamus latiflorus Munro. PLoS ONE 9, e87417 (2014).
26. Quirós, F. G. et al. Housekeeping gene stability influences the quantification of osteogenic markers during stem cell differentiation to the osteogenic lineage. Cytotherapy 12, 109–120 (2010).
27. Tratwal, J., Follin, B., Ekblond, A., Kastrup, J. & Haack-Sørensen, M. Identification of a common reference gene pair for qPCR in human mesenchymal stromal cells from different tissue sources treated with VEGF. BMC Mol. Bio. 15, 1–11 (2014).
28. Vandesompele, J. et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol. 3, research0034. 1-research0034. 11 (2002).
29. Andersen, C. L., Jensen, J. L. & Ornstoft, T. F. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. Cancer Res. 64, 5245–5250 (2004).
30. Pfafl, M. W., Tischopad, A., Pregomet, C. & Neuvians, T. P. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper—excel-based tool using pair-wise correlations. Biotechnol. Lett. 26, 509–515 (2004).
31. Silver, N., Best, S., Jiang, J. & Thein, S. L. Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. BMC Mol. Bio. 7, 33 (2006).
32. Xie, F., Sun, G., Stiller, J. W. & Zhang, B. Genome-wide functional analysis of the cotton transcriptome by creating an integrated EST database. PLoS ONE 6, e26980 (2011).
33. Martinez Barbera, J. P. et al. The homeobox gene Hex is required in definitive endodermal tissues for normal forebrain, liver and thyroid formation. Development 127, 2433–2445 (2000).
34. Koznik, Z. et al. Pax–Six–Eya–Dach network during amphioxus development: Conservation in vitro but context specificity in vivo. Dev. Biol. 306, 143–159 (2007).
35. Mafra, V. et al. Reference genes for accurate transcript normalization in citrus genotypes under different experimental conditions. PLoS ONE 7, e31263 (2012).
36. Guo, J., Ling, H., Wu, Q., Xu, L. & Que, Y. The choice of reference genes for assessing gene expression in sugarcane under salinity and drought stresses. Sci. Rep. 4, 7042 (2014).
37. Petriccione, M., Mastroboumi, F., Zampella, L. & Scortichini, M. Reference gene selection for normalization of RT-qPCR gene expression data from Actinidia delicosa leaves infected with Pseudomonas syringae pv. actinidiae. Sci. Rep. 5, 16961 (2015).
38. Sun, M., Lu, M. X., Tang, X. T. & Du, Y. Z. Exploring valid reference genes for quantitative real-time PCR analysis in Sesamia inferens (Lepidoptera: Noctuidae). PLoS ONE 10, e0119979 (2015).
39. Ponton, F., Chapius, M. P., Pernrice, M., Sword, G. A. & Simpson, S. J. Evaluation of potential reference genes for reverse transcription–qPCR studies of physiological responses in Drosophila melanogaster. J. Insect Physiol. 57, 840–850 (2011).
40. Schaeck, M. et al. Laser capture microdissection of intestinal tissue from sea bass larvae using an optimized RNA integrity assay and validated reference genes. Sci. Rep. 6, 21092 (2016).
41. Madroniche, G., Sagadin, M., Mongelli, V. C., Truel, G. A. & del Vas, M. Reference gene selection for gene expression studies using RT-qPCR in virus-infected plantthoppers. Virol. J. 8, 1–8 (2011).
42. Hu, Q., Guo, W., Gao, Y., Tang, R. & Li, D. Reference gene selection for real-time RT-PCR normalization in rice field eld (Monopterus albus) during gonad development. Fish Physiol. Biochem. 40, 1721–1730 (2014).
43. Xue, L. et al. Eya protein phosphatase activity regulates Six1-Dach-Eya transcriptional effects in mammalian organogenesis. Nature 426, 247–254 (2003).
44. Tadjudie, E. & Hegde, R. S. The Eyes Absent proteins in development and disease. Cell. Mol. Life Sci. 70, 1897–1912 (2013).
45. Pigoni, F. et al. The eye-specification proteins so and eya form a complex and regulate multiple steps in eye development. Cell 91, 881–891 (1997).
46. Bedford, F. K., Ashworth, A., Enver, T. & Wiedemann, L. M. HEX: a novel homeobox gene expressed during haematopoiesis and conserved between mouse and human. Nucleic Acids Res. 21, 1245–1249 (1993).
47. Crompton, M. R. et al. Identification of a novel vertebrate homebox gene expressed in haematopoietic cells. Nucleic Acids Res. 20, 5661–5667 (1992).
48. Liao, W., Ho, C. Y., Yan, Y. L., Postlethwait, J. & Stainier, D. Y. Hhex and scl function in parallel to regulate early endothelial and blood differentiation in zebrafish. Development 127, 4303–4313 (2000).
49. Jayaraman, S. P., Frampton, J. & Goodwin, G. The homeodomain protein PRH influences the differentiation of haematopoietic cells. Leukemia Res. 24, 1023–1031 (2000).
50. Voskobowynik, A. et al. Identification of the endostyle as a stem cell niche in a colonial chordate. Nucleic Acids Res. 31, 856–862 (2003).
51. Yuan, S. et al. Identification of the endostyle as a stem cell niche in a colonial chordate. Nucleic Acids Res. 31, 856–862 (2003).
52. Yu, J. K. S. & Holland, L. Z. Extraction of RNA from amphioxus embryos or adult amphioxus tissue. Cold Spring Harb. Protoc. 2009, pdb.prot5288 (2009).
53. Radonic, A. et al. Guideline to reference gene selection for quantitative real-time PCR. Biochem. Biophys. Res. Co. 313, 856–862 (2004).
54. Yang, C., Pan, H., Liu, Y. & Zhou, X. Stably expressed housekeeping genes across developmental stages in the two-spotted spider mite, Tetranychus urticae. PLoS ONE 10, e0128033 (2015).
55. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2−ΔΔCT method. Methods 25, 402–408 (2001).
**Acknowledgements**
We are grateful to two anonymous reviewers for providing invaluable comments and suggestions. This work was supported by the 973 project of Ministry of Science and Technology of China (Grants No. 2013CB835300) and the Natural Science Foundation of China (41272008). We thank Professors X. D. Su and Dr. K. Yu at Peking University for providing data from database for transcriptomes of different amphioxus tissues and help from H. P. Li and X. Z. Li at Beihai Marine Station.

**Author Contributions**
Q.L.Z., M.L.Y. and J.Y.C. designed the study. Q.L.Z., X.Q.W., J.W. and X.L. performed the experiment. Q.L.Z., Q.H.Z., T.C. and H.T.X. analysed the data. J.Y.C. provided reagents. Q.L.Z., Q.H.Z. and X.L. drafted the manuscript. M.L.Y. and J.Y.C. revised the manuscript.

**Additional Information**
Supplementary information accompanies this paper at http://www.nature.com/srep

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Zhang, Q.-L. *et al.* Selection of reliable reference genes for normalization of quantitative RT-PCR from different developmental stages and tissues in amphioxus. *Sci. Rep.* **6**, 37549; doi: 10.1038/srep37549 (2016).

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

© The Author(s) 2016