Epigenetic divergence of key genes associated with water temperature and salinity in a highly invasive model ascidian

Chuanliang Pu · Aibin Zhan

Received: 22 May 2016 / Accepted: 7 March 2017 / Published online: 14 March 2017
© Springer International Publishing Switzerland 2017

Abstract Invasive species offer good models for studying mechanisms of response to rapid changing environments in the wild. DNA methylation is considered to be one of crucial drivers for rapid local adaptation. However, the extent of epigenetic variation and the factors driving such variation remain largely unexplored in biological invasions. Here we used direct bisulfite sequencing to investigate DNA methylation patterns of five key genes corresponding to two important environmental factors in marine ecosystems, water temperature and salinity, in a model invasive ascidian *Ciona robusta* (= *C. intestinalis* spA). Our results clearly showed that DNA methylation mainly occurred in gene bodies, rather than promoters, at regions with low values of CpG O/E. We detected significant variation of DNA methylation among populations in two genes (heat shock protein 90 and $\text{Na}^+\text{-K}^+\text{-2Cl}^-$ cotransporter). Interestingly,

significant correlation was detected between methylation levels and the two environmental factors at some CpGs in these two genes. When the data of all CpGs was subjected to principal component analysis, individuals were assigned back to their population origins. All the results suggest that environmental factors likely contribute, at least partially, to the observed DNA methylation variation. Such variation, either by some loci alone or through gene networks, might be involved in rapid local adaptation during biological invasions.

Keywords Ascidian · Biological invasion · DNA methylation · Environmental factor · Epigenetics · Invasive species · Local adaptation

Introduction

Biological invasions have been recognized as one of the major threats to the environment and economy worldwide (Pimentel et al. 2005). Understanding the mechanisms of invasion success is vital for both predicting future invasions and understanding how populations can evolve rapidly in response to novel environments (Zhan et al. 2015). Available evidence suggests that local adaptation may play one of the essential roles in determining invasion success (Sanford and Kelly 2011). Local adaptation is considered to be one of the major mechanisms to explain why some introduced species can establish, proliferate, and

Electronic supplementary material The online version of this article (doi:10.1007/s10530-017-1409-1) contains supplementary material, which is available to authorized users.

C. Pu · A. Zhan (✉)
Research Center for Eco-Environmental Sciences,
Chinese Academy of Sciences, 18 Shuangqing Road,
Haidian District, Beijing 100085, China
e-mail: zhanaibin@hotmail.com;
azhan@rcees.ac.cn

C. Pu · A. Zhan
University of Chinese Academy of Sciences, 19A Yuquan
Road, Shijingshan District, Beijing 100049, China

finally become invasive in dramatically different local environments (Colautti and Lau 2015; Kirkpatrick and Barrett 2015). In the process of local adaptation, mounting evidence supports that many factors can contribute to invasion success, including genetic attributes such as mutation, hybridization and genomic rearrangements (Ciancio et al. 2015; Colautti and Lau 2015; Kirkpatrick and Barrett 2015). Recently, studies successfully identified that epigenetic modifications, such as DNA methylation, could contribute substantially to rapid local adaptation during biological invasions (e.g. Xie et al. 2015).

DNA methylation, as one of the most common but important epigenetic modifications, is a process by which methyl groups are added to the 5th carbon of cytosine residues commonly situated adjacent to guanine residues (5mCpG; Feil 2006). Available evidence clearly showed that DNA methylation has played crucial roles in the regulation of gene expression in plants responding to environmental stresses, such as extreme temperature, salinity, and nutrition deficiency (Dyachenko et al. 2006; Kou et al. 2011; Pecinka et al. 2010). In animals, factors such as maternal diet and folic acid supplementation can alter de novo programming of DNA methylation during the development of offspring (Feil 2006; Lillycrop et al. 2005; Weaver et al. 2004). Collectively, epigenetic modification plays vital roles in response to environmental stresses and contributes to local adaptation via environmentally induced phenotypic variations by modifying gene expression without altering DNA sequences (Angers et al. 2010). Although many examples of environmentally induced phenotypic variations caused by DNA methylation have been described (Angers et al. 2010; Sumner et al. 2006), many issues remain largely unknown during the process of biological invasions, such as what epigenetic variation is left in populations and what factors cause such variation during population expansions in different habitats. As invasive species frequently encounter various environmental stresses, including both biotic (e.g. pathogens and species competition) and abiotic (e.g. temperature and salinity), during biological invasions, they offer great opportunities to study epigenetic mechanisms of rapid local adaptation.

In the past three decades, a number of ascidian species have spread across much of the globe (Zhan et al. 2015). The *Ciona* ascidians, including *C. intestinalis* spB (= *C. intestinalis*) and *C. intestinalis* spA (= *C. robusta*; Brunetti

et al. 2015) in the *C. intestinalis* species complex and *C. savignyi*, have long been recognized as model species in developmental and evolutionary biology (Procaccini et al. 2011). For the *C. intestinalis* species complex, the well sequenced and annotated small genome (Dehal et al. 2002; Vinson et al. 2005), coupled with detailed phylogenetic and population genetics analyses (Nydam and Harrison 2010; Zhan et al. 2010, 2012; Hudson et al. 2016), has expanded it as a promising model for evolutionary studies associated with biological invasions (Zhan et al. 2015). The highly invasive species *C. robusta* has successfully invaded coastal habitats with different environmental characteristics, especially water temperature and salinity, two of the most important environmental factors in marine ecosystems (Bates 2005; Rocha et al. 2009; Seo and Lee 2009). Owing to ambiguous taxonomy within the genus *Ciona*, the invasion history remains uncertain for many invaded areas. For example, the specimens detected in the benthic communities on the Pacific coast of North America in the early 1910s are probably *C. savignyi* (see the review by Zhan et al. 2015 and references therein). Since the late 1960s, new populations of *C. robusta* have been detected along the coasts of Oceania, South Africa and Europe (see review by Zhan et al. 2015 and references therein). The fast spread among habitats with dramatically different environmental factors in a short period of time makes *C. robusta* a good model to study both genetic and epigenetic mechanisms of rapid local adaptation (Zhan et al. 2015).

In this study, we used direct sequencing of bisulfite-converted DNA in four wild populations of *C. robusta* collected from different habitats with varied water temperature and salinity (Table 1; Fig. 1). By studying the epigenetic patterns of five key genes corresponding to two major environmental factors, water temperature and salinity (Gamba 2005; Han et al. 2006; Tsai and Lin 2007; Vabulas et al. 2010), we aim to understand what epigenetic variation is present in diverse populations and whether the two measured environmental factors may have contributed to such variation during biological invasions.

Materials and methods

Sample collection and environmental data acquisition

In order to explore the variation of DNA methylation, we selected four populations of *C. robusta* from

Table 1 Four selected *Ciona robusta* populations used in this study

Population	Region/city and country	No. of individuals tested	Invasion status
AM	Arenys de Mar, Spain	20	Cryptogenic
BL	Blanes, Spain	10	Cryptogenic
SA	Cape town, South Africa	20	Invasive
NZ	Nelson, New Zealand	12	Invasive

Fig. 1 Variation of water temperature and salinity at collection sites. *T* and *S* represent water temperature and salinity, respectively. Population names per Table 1

Europe [Arenys de Mar (AM) and Blanes (BL), Spain], South Africa [Cape town, South Africa (SA)] and Oceania [Nelson, New Zealand (NZ)] based on both the difference of the two environmental factors (i.e. water temperature and salinity; Table 1; Fig. 1) and population genetic diversity derived from our previous study (Zhan et al. 2010). A total of 62 individuals, including 20 individuals from AM, 10 individuals from BL, 20 individuals from SA and 12 individuals from NZ, were collected and stored in 95% ethanol at 4 °C and subsequently used for genomic DNA extraction (Table 1). For all sampling sites, monthly sea surface salinity and temperature were obtained at a quarter degree resolution from the National Oceanic and Atmospheric Administration (NOAA) World Ocean Data Atlas 2013 (www.nodc.noaa.gov/OC5/indprod.html; Fig. 1).

Selection of key genes

A total of five key genes associated with temperature [heat shock protein 90 (HSP90) and heat shock protein 70 (HSP70)] and salinity [$\text{Na}^+\text{-K}^+\text{-2Cl-cotransporter}$ (NKCC), $\text{Na}^+\text{-K}^+\text{-ATPase}$ (NKA) and $\text{Na}^+\text{-Cl-}\gamma$ -taurine transporter (TauT)] were selected to compare DNA methylation patterns in the four populations. Available evidence shows that the selected genes in this study play crucial roles in response to changes of water temperature and salinity in multiple species (e.g. Skou and Esmann 1992; Evans et al. 2005; Gamba 2005; Mayer and Bukau 2005; Yancey 2005; Han et al. 2006; Tsai and Lin 2007; Vabulas et al. 2010). The HSP90 and HSP70 genes, as molecular chaperones, have been reported to play vital roles in proper folding of proteins and adaptive responses to environmental

stresses in cells among numerous divergent organisms (Mayer and Bukau 2005; Vabulas et al. 2010). The NKCC gene, as a member of the chloride-cation cotransporter family, plays important roles in maintenance of the electrolyte content and transepithelial ion and water movement in polarized cells (Evans et al. 2005; Gamba 2005). The NKA gene, providing regulation of intracellular ionic gradients used for cellular homeostasis in all animal cells, is known to be critical for maintaining ion balance and homeostasis in relation to salinity acclimation and also to be important in adaptation to various salinity environments in many taxa (Skou and Esmann 1992; Tsai and Lin 2007). The TauT gene is found to be the most abundant organic osmolyte in numerous invertebrates, fishes and mammals (Yancey 2005; Han et al. 2006). Intracellular taurine content is mainly regulated via transmembrane transport and the accumulation of taurine is mediated by TauT (Yancey 2005; Han et al. 2006). Thus, TauT plays an important role in adjusting of the concentration of taurine in tissues in responses to altered salinity environments.

Bisulfite sequencing

Total genomic DNA was extracted from each animal's excurrent siphon using the Marine Animals DNA Kit (TIANGEN; Beijing, China). Subsequently, bisulfite conversion of genomic DNA (~1200 ng) and lambda DNA (unmethylated as an internal control, 100 ng) was conducted using the Epitect Bisulfite Kit (QIAGEN CHINA; Shanghai, China) based on the manufacturer's protocol. After bisulfite treatment, DNA was immediately used for Polymerase Chain Reaction (PCR). A total of 30 primer pairs targeting around 500-bp regions were designed by MethPrimer (Li and Dahiya 2002) to cover genomic regions of interest, including untreated DNA and both Watson and Crick strands of bisulfite-treated DNA in promoters and gene bodies of the selected key genes (Table 2). Promoters including -2000 to +200 bp around the transcriptional start site were obtained from NCBI GenBank for the five studied genes (Table 2). All PCRs were performed in a total volume of 25 µl containing 0.5 µM mix of forward and reverse primers, 2.0 mM MgCl₂, 0.2 mM dNTPs, and ~50 ng of converted DNA in 1× PCR buffer and 1 U *Taq* DNA polymerase (TAKARA CHINA; Dalian, China). The PCR program included a 5-min denaturation step at 95 °C,

followed by 36 cycles of 30 s at 95 °C, 30 s at an annealing temperature determined based on different primer pairs (Table 2), and 45 s at 72 °C; with a final extension at 72 °C for 7 min. Subsequently, a total of 1860 PCR products were directly sequenced using an ABI 3730 DNA analyzer.

Calculation of CpG O/E values

The value of CpG O/E is defined as $\text{CpG O/E} = P_{\text{CpG}} / P_{\text{C}} \times P_{\text{G}}$, where P_{CpG} , P_{C} , and P_{G} are the frequencies of CpG dinucleotides, C nucleotides and G nucleotides, respectively. The value of CpG O/E reflects the depletion of CpG dinucleotides for genomic regions (Bird 1980). It was calculated using sequences from promoters and gene bodies of each key gene.

Calculation and analysis for methylation level of CpGs

Genomic DNA from different populations was treated using bisulfite treatment: all unmethylated cytosines (Cs) were converted into uracils (Us) and replaced by thymines (Ts) after PCR amplification. The location of each CpG dinucleotide at all loci was determined by comparing the sequences obtained from PCR products of the bisulfite-treated DNA with untreated DNA sequences. Only data from bisulfite PCR whose lambda DNA showed complete C-to-T conversion by bisulfite treatment at Cs was used to calculate methylation levels for each CpG site. Because bisulfite treatment does not convert methylated Cs to Us, the C signal at each CpG corresponds to the degree of methylation at that site in a given sample, and the T signal corresponds to the degree of Cs that are unmethylated. Thus, the methylation level for each CpG can be quantified by measuring the ratio between the peak height values of Cs and Ts, yielding the basic equation: $\text{methylation percentage} = C / (C + T) \times 100\%$ or $\text{methylation percentage} = G / (G + A) \times 100\%$ (for the reverse complementary strand) (Jiang et al. 2010). Subsequently, data from each CpG was analyzed using one way analysis of variance (ANOVA) or wilcoxon rank sum test for data with heteroscedasticity to detect significant difference in the four populations. In order to explore the relationship between methylation and environmental factors, multiple correlation analyses (Pearson, Kendall's tau-b, Spearman's rho) between the DNA

Table 2 Gene regions and their corresponding primers used in this study

Accession no. and gene	Locus	N	Primer	Sequence (5' > 3')	T _a (°C)
NW_004190325. HSP90	Promoter	25	CGI F (−676/26)	CAGCAGCGGATGATGCCAGTTACAGT	54.5
			CGI R (+82/32)	CTTCGTGTTCCCTGTTCCATACCTTTGGTGTGA	55.0
			W F (−59/30)	AATTGGTTGATTTTAAATTTAAATTGATTG	57.0
			W R (−561/36)	CTTATCTAACCAATATATATAAAACAACAACTAC	56.7
			C F (−550/36)	AGTGTGTTAAAAATAGTAAGATTGTTTGTAGTG	58.5
			C R (−59/36)	AATTAACCTTTTAAATTTAACTAACTACCCAAC	59.2
			CGI F (+272/24)	TAATGTTACTGGTTAATGAAAAAG	58.5
			CGI R (+956/21)	TGTAGTTTACGACCCAGAAATG	57.7
			W F (+875/30)	TTAGGTTTAAAGTAAATGAAATTAAGTTAG	57.6
NW_004190451. HSP70	Promoter	25	W R (+305/34)	TTTACAAAATACTACTATATAATATATATCCAAC	57.0
			C F (+300/30)	TTTTATTATAAAATGTTGTTATGATGATG	54.9
			C R (+920/32)	TTTATTCAAATATAATTTAAACTCTAAATCC	55.4
			CGI F (−195/29)	TGAAAATACACGAACTACAGGAAACAT	55.5
			CGI R (+546/26)	ACGAAAACATAACAACTCTGCTCGGAAC	56.0
			W F (+509/36)	TTGGAATAGTATTAGTATTAAAGAGGTGTTAAAGA	56.8
			W R (−18/30)	AAAAAAAATTAACCCATCTAATCCATTTC	55.6
			C F (−37/36)	AAAAGTTGAGTAGAATAGAAAAAGAGATTAAAGTTA	58.6
			C R (+511/33)	CACATAAATACACCAACACCAAAAAATATACC	58.0
NW_004190330. NKCC	Promoter	39	CGI F (+464/22)	GAACTTTGATCTTTGGCACAC	58.0
			CGI R (+1149/23)	CGAAAAACAAAAAGGACATCTC	58.4
			W F (+1135/33)	GGATATTTTAAAGAGTAATAGGTAATTAGAAG	58.0
			W R (+473/34)	TCCTTAACACACCTCTTAATACTAATACTATTCC	58.8
			C F (+487/28)	TTTTGGTGTGTTGTTATTTAGTGAGG	58.5
			C R (+1143/32)	CACAAAAAACATCTCAAAAAATAACAAAAAC	59.0
			CGI F (−1048/27)	AGATAAGTGGGAAAAAATCACCCCTA	59.4
			CGI R (−322/28)	CGCCTTAAACAAACAGAAAATGTAATCGT	60.2
			W F (−517/28)	TTAGTTAAATTTGTTTAAATAATGAGTTG	55.0
NW_004190330. NKCC	Promoter	18	W R (−826/34)	TTTACTAAATAAACCTATATATTACTATAACATTAC	55.5
			C F (−826/33)	TTTATTGAATAAAATTTATATTATTATAGTATTG	54.9
			C R (−522/31)	TCAAACCTTACTTAATAATAAACTACATTAAAC	55.1
			CGI F (+1969/20)	AGGCTGTTTCAGGGATCAC	57.0
			CGI R (+2663/24)	AACAGAGGAATAGAATGCCTTGAT	56.6

Table 2 continued

Accession no. and gene	Locus	N	Primer	Sequence (5' > 3')	T _a (°C)
NW_004190401. NKA	Promoter	30	W F (+2617/30)	TGATTTAAATTTATTTTATTTTATTTAAGTTA	55.8
			W R (+1987/29)	ACATTTTCCACTAATAAAAATTTTAAAC	56.0
			C F (+1991/29)	TTTTTATGTAGGATTTTGTGTAAG	54.0
			C R (+2616/30)	AAATCAACTCATTTCTTTTCATTAATAATAC	55.4
			CGI F (−1261/26)	CCTACTAAGCCAGTGGATACTCGCAT	55.9
			CGI R (−560/28)	ACGATCAAATTAAGAATAAATGCCCCG	55.0
			W F (−619/35)	TTGTAGAAATTAATGAATTTTAGTAGAATA	58.4
			W R (−1114/35)	TATCTAAAAATTTCAATCTACTACAATTTATACAC	59.0
			C F (−1134/29)	AGAAATGTTGGGAGGTGTTTATTTGGGG	59.6
			C R (−627/37)	TCTAAATAAAACTTCTAACACAAAATATTACAAACAC	60.4
			CGI F (−2250/22)	GACTCCGAGGTCCTTACCAGCG	58.6
			CGI R (+2993/21)	CATTTTGGCGACGGGATTTG	58.0
			W F (+2917/28)	GTTAAATTTAAGAAAGTAAATAATAA	57.0
			W R (+2279/30)	TTTCAACCTATATACTAATAAAATATACTC	57.3
NW_004190564. TauT	Promoter	23	C F (+2272/34)	TTATGTTTTTAGTTTGTGTATTAGTAAATATA	58.6
			C R (+2916/28)	TTAAAAATTTAAAAAACAACAAACAAAC	57.8
			CGI F (−1188/30)	AAATTTTTTACATGTAACAAATCCGCTCA	62.2
			CGI R (−443/28)	AGAAATGAAAAGGTGTTCCATCTTCCCCC	62.0
			W F (−471/36)	ATTTATTATTGTATAGATTTATTAGTAAAGATGGGA	56.6
			W R (−1073/36)	TCCTAAAAAATAATCACCCACAAATAATACATA	55.8
			C F (−1072/36)	TTTAAAAAATAATTATTATATAAGTAATATATATGG	55.8
			C R (−511/27)	TTTAACACATAATTACAATAACCCCAAC	56.7
			CGI F (+872/24)	TTATTTTAGGGGAATATTTTCATT	56.4
			CGI R (+1471/22)	ATATTCATGCTTTGAGCAAAAC	55.8
			W F (+1428/30)	TAGAAATTTTAGAAAAGAAATATAGTTAAGG	55.0
			W R (+888/30)	ATTTCAATTTTACAAAAAATAACAAAA	54.2
			C F (+876/32)	TTTTAGGGGAATATTTTATTTTATAAAGAG	58.0
			C R (+1448/30)	TATTTTAAAAATTTAACCAACAAAAATCTC	57.7

N is the number of CpGs in a given region. Positions of CpGs in each gene region were annotated in Table S1. T_a is the optimized annealing temperature for different primer pairs. CGI = CpG island in each region, W = Watson strand, C = Crick strand, HSP90 = heat shock protein 90 gene; HSP70 = heat shock protein 70 gene; NKCC = Na⁺-K⁺-2Cl[−] cotransporter gene; NKA = Na⁺-K⁺-ATPase gene; TauT = Na⁺-Cl[−]-taurine transporter gene

methylation level and annual temperature and salinity were performed using SPSS version 22. Standard Bonferroni correction was used for all multiple tests. Finally, in order to understand the influence of environmental factors on DNA methylation patterns, principal component analysis (PCA) on the data of methylation level was also executed at the population level using the R software package ade4 (Dray and Dufour 2007).

Results

Relationship between CpG O/E values and methylation level

A total of five key genes were selected, including HSP90 and HSP70 associated with temperature, and NKCC, NKA and TauT associated with salinity (Table 2). By sequencing of PCR amplicons, a total of 1781 sequences including 620 unconverted sequences and 1161 converted sequences (591 Watson and 570 Crick strands) were obtained (GenBank accession no. KY247336–KY249116). Interestingly, we found that regions with lower values of CpG O/E showed a higher level of DNA methylation. For example, the gene bodies of HSP90 and NKCC had lower values of CpG O/E but higher methylation

levels than other studied regions (Fig. 2). More specifically, hypomethylated regions including the promoters of HSP90, HSP70, NKCC, NKA and TauT and gene bodies of HSP70, NKA, and TauT had CpG O/E values of >0.8 , but the CpG O/E values of gene bodies of HSP90 and NKCC with hypermethylation were <0.8 (Fig. 2).

Epigenetic patterns of key genes

Our results showed that DNA methylation largely varied among genes. For TauT, all CpGs examined in its promoter and gene body were completely unmethylated (Fig. S1). By contrast, HSP90 and NKCC had hypermethylated CpGs in their gene bodies, but almost all CpGs were unmethylated in the promoters, with only two hypomethylated CpGs detected in the promoter of HSP90 (Figs. S2 and S3). Similar to NKCC, all CpGs in NKA were completely unmethylated in the promoter, but 11 CpGs were hypomethylated in the gene body (Fig. S4). For HSP70, we detected three and 16 hypomethylated CpGs in the promoter and gene body, respectively (Fig. S5).

Despite different DNA methylation levels among genes, we detected similarity when comparing the methylation level of promoters with that in gene bodies. Collectively, DNA methylation mainly

Fig. 2 Distributions of CpG depletion (CpG O/E) and measured methylation levels for different regions of the five genes tested in the study. P and GB represent promoter and gene body, respectively. HSP90 = heat shock protein 90 gene; HSP70 = heat shock protein 70 gene; NKCC = $\text{Na}^+\text{-K}^+\text{-2Cl}^-$ cotransporter gene; NKA = $\text{Na}^+\text{-K}^+\text{-ATPase}$ gene; TauT = $\text{Na}^+\text{-Cl}^-$ taurine transporter gene

occurred in gene bodies, rather than promoters. Except TauT with complete unmethylation in both promoter and gene body (Fig. S1), all the other genes showed a higher level of DNA methylation in gene bodies than in promoters. For instance, HSP70 showed three and 16 hypomethylated CpGs in the promoter and gene body, respectively, and the average level of DNA methylation of gene body was higher than that in promoter at the population level (20 individuals versus four individuals, Fig. S5). For NKCC and NKA, no methylated CpGs were detected in their promoters, but methylated CpGs were detected in their gene bodies (Figs. S3 and S4). Especially in the gene bodies of HSP90 and NKCC, all CpGs were hypermethylated, however, only two hypomethylated CpGs were detected in their promoters (Figs. S2 and S3).

Variation of DNA methylation among populations

In order to understand the role of DNA methylation in response to varied temperature and salinity, the DNA methylation patterns of the five genes were compared in four populations. Although the results indicated that the DNA methylation patterns for both promoters and gene bodies of five analyzed genes were largely consistent among four populations (Figs. S1–S5), two hypermethylated regions—gene bodies of HSP90 and NKCC—significantly differed in the analyzed populations when the average levels of methylation for all CpGs were used for comparison (Fig. 3). When the data of methylation level for each CpG was analyzed to compare the difference among the four populations for both Watson and Crick strands, our results showed that the methylation level varied among each CpG in the gene bodies of these two genes (Fig. S6). For HSP90, the methylation levels at seven CpGs (CpG2, CpG3, CpG4, CpG5, CpG6, CpG7, CpG8) were significantly different among four populations in Watson strands (Fig. 4a). Similarly, the methylation levels at four CpGs (CpG2, CpG3, CpG6, CpG7) significantly varied among four populations in Crick strands (Fig. 4a). For NKCC, the significant difference was found at two CpGs (CpG6, CpG13) in both strands among four populations (Fig. 4b).

The correlation between DNA methylation and environmental factors

In order to explore the relationship between DNA methylation and the two environmental factors (water

temperature and salinity), multiple correlation tests between two environmental factors and DNA methylation level at each CpG was performed in the two genes with significant difference in four analyzed populations. Our results showed that significant correlation with temperature was detected at multiple CpGs in the gene body of HSP90 (Table 3). For the Crick strand, CpG7 was significantly correlated with temperature based on all tests, while for the Watson strand, CpG3, CpG7 and CpG8 were significantly correlated with temperature based on all tests (Table 3). In the gene body of NKCC, significant correlation with salinity was detected at CpG6 (Table 3). Especially for the Watson strand, CpG6 was significantly correlated with salinity based on all correlation tests (Table 3).

Moreover, when methylation levels at all CpGs in these two genes were included for principal component analysis (PCA), individuals were grouped into four clusters corresponding to their population origins (Fig. 5). For instance, individuals derived from Arenys de Mar, Spain (AM) and Capetown, South Africa (SA) were separated from individuals belonging to Blanes, Spain (BL) and Nelson, New Zealand (NZ) on PCA1 (22.5%). On PCA2 (14.7%), individuals belonging to SA were separated from individuals derived from AM.

Discussion

As invasive species can successfully colonize local habitats with dramatically different environmental conditions in a short period of time, they provide good models to answer many fundamental questions in microevolution, such as mechanisms of rapid local adaptation (Colautti and Lau 2015; Sanford and Kelly 2011). However, investigating possible mechanisms of rapid local adaptation can be challenging, mainly owing to ecological and genetic complexities generated during biological invasions (Colautti and Lau 2015). Therefore, it is necessary to identify common and/or unique genetic and epigenetic patterns caused by potential interactions between invasive species and local environmental factors, and subsequently elucidate how these genetic and epigenetic patterns contribute to invasion success. For the roles of epigenetic modifications, thus far it is known that DNA methylation plays a vital role in tuning gene

Fig. 3 Comparison of DNA methylation pattern at gene bodies of the heat shock protein 90 gene (HSP90) and $\text{Na}^+\text{-K}^+\text{-2Cl}^-$ cotransporter gene (NKCC) for both Watson and Crick strands in the four studied populations of *Ciona robusta*. The *left side* of the figure shows DNA methylation pattern of gene body of HSP90, the *right side* shows DNA methylation pattern of gene body of NKCC. The *upper part* shows the DNA methylation pattern of Watson strand of gene bodies in HSP90 and NKCC (GB-HSP90W and GB-NKCCW), the *lower part* shows the DNA methylation pattern of Crick strand of gene bodies in HSP90 and NKCC (GB-HSP90C and GB-NKCCC). Grey circles indicates no reliable data available for CpGs after sequencing mainly because those CpGs were too close to the beginning or end of each sequence. Population names per Table 1

Fig. 4 Comparison of DNA methylation level for gene bodies of the heat shock protein 90 gene (HSP90; **a**) and $\text{Na}^+\text{-K}^+\text{-2Cl}^-$ cotransporter gene (NKCC; **b**) by using one-way analysis of variance in the four studied populations of *Ciona robusta*. The *upper part* of each figure shows the DNA methylation level of Watson strand of gene bodies in HSP90 and NKCC (GB-

HSP90W and GB-NKCCW), the *lower part* shows the DNA methylation level of Crick strand of gene bodies in HSP90 and NKCC (GB-HSP90C and GB-NKCCC). The *bar chart plots* show the standard error of methylation level of each CpG in gene bodies of HSP90 and NKCC in the four studied populations. Population names per Table 1

expression in response to environmental stimuli. However, the ecological and evolutionary causes and consequences of DNA methylation in rapid local adaptation remains largely unknown during in the process of biological invasions. In order to understand epigenetic mechanisms responsible for rapid local adaptation, several epigenetic patterns should be studied, including what epigenetic variation is left in populations and how local environmental factors drive such variation during biological invasions. Here, we used direct sequencing of bisulfite-converted DNA to compare the epigenetic patterns of five key genes associated with two major environmental factors (temperature and salinity) in the populations of *C. robusta* collected from different habitats with varied water temperature and salinity.

Our results clearly showed that DNA methylation primarily occurred at regions with lower value of CpG O/E < 0.8 (Fig. 2). It is known that CpG methylation can influence the underlying sequence composition as methylated cytosines that spontaneously deaminate to thymine in genomes are gradually lost over evolutionary time (Elango et al. 2009). In other words, DNA methylation increases the frequencies of transition mutations from CpG to TpG, thus gradually depleting CpGs from DNA sequences. Several studies have

confirmed the negative correlation between the methylation levels and the CpG O/E ratio (Elango et al. 2009). In sea squirts, recent studies have found that approximately 54% of all genes belonged to the CpG-deficient fraction with values of CpG O/E < 0.8 (Zemach et al. 2010), and highly negatively correlation between methylation levels and the values of CpG O/E calculated for each locus was observed by genome-wide bisulfite sequencing (Zemach et al. 2010). Our experimental data was largely consistent with the observed CpG O/E pattern, but did not indicate that DNA methylation was entirely absent from hypomethylated regions as methylated CpGs were still detected in these analyzed genes (Figs. S1–S5).

We found that DNA methylation in *C. robusta* occurred mainly in gene bodies, rather than promoters. In vertebrates, regulation of transcription by DNA methylation is accomplished by differential patterns of methylation in intergenic regions, namely gene promoters (Kass et al. 1997). Therefore, studies of DNA methylation in vertebrates have generally focused on promoters, where hypermethylated promoters inhibit initiation of transcription (Mirouze and Paszkowski 2011). In contrast, invertebrate genomes do not show hypermethylated promoters. Most incidences of DNA

Table 3 The correlation analysis of water temperature (Tem.) and salinity (Sal.) with the methylation level at CpGs in gene bodies of heat shock protein 90 (GB-HSP90) and Na⁺-K⁺-2Cl⁻ cotransporter (GB-NKCC) genes for both Watson (W) and Crick (C) strands in the four populations of *Ciona robusta* (= *C. intestinalis* spA)

Tem.	Method	GB-HSP90	CpG2	CpG3	CpG4	CpG5	CpG6	CpG7	CpG8	Mean	
Crick strand	Pearson	Factor	−0.175	− 0.493**	−0.246	0.168	0.013	0.489**	−0.061	−0.183	
		<i>p</i> value	0.225	0.000	0.086	0.243	0.930	0.000	0.674	0.204	
	Kendall's tau-b	Factor	−0.031	−0.292	−0.166	0.179	0.182	0.451**	0.138	−0.030	
		<i>p</i> value	0.799	0.023	0.169	0.163	0.153	0.000	0.302	0.791	
	Spearman's rho	Factor	−0.037	−0.319	−0.191	0.199	0.207	0.539**	0.147	−0.064	
		<i>p</i> value	0.800	0.024	0.183	0.165	0.150	0.000	0.308	0.661	
	Watson strand	Pearson	Factor	0.167	− 0.972**	0.113	0.080	0.101	0.513**	0.351*	−0.003
			<i>p</i> value	0.206	0.000	0.393	0.548	0.448	0.000	0.006	0.979
Kendall's tau-b		Factor	0.329**	− 0.715**	0.326*	0.251	0.297*	0.479**	0.467**	0.240	
		<i>p</i> value	0.002	0.000	0.003	0.022	0.006	0.000	0.000	0.021	
Spearman's rho		Factor	0.422**	− 0.772**	0.409**	0.301	0.374*	0.562**	0.594**	0.301	
		<i>p</i> value	0.001	0.000	0.001	0.020	0.004	0.000	0.000	0.021	
Sal.	Method	GB-NKCC	CpG6 (C)	CpG13 (C)	Mean (C)	CpG6 (W)	CpG13 (W)	Mean (W)			
	Pearson	Factor	0.318	0.144	0.307	0.511**	0.156	0.438**			
		<i>p</i> value	0.013	0.267	0.016	0.000	0.241	0.001			
	Kendall's tau-b	Factor	0.196	0.051	0.127	0.356**	0.014	0.233			
		<i>p</i> value	0.062	0.606	0.195	0.001	0.893	0.021			
	Spearman's rho	Factor	0.227	0.059	0.170	0.433**	0.022	0.309			
		<i>p</i> value	0.078	0.654	0.190	0.001	0.872	0.018			

Notes The highlighted number in bold indicates significant correlation, * indicates significant correlation at $p < 0.0125$ level (after standard Bonferroni correction), ** indicates significant correlation at $p < 0.0025$ level (after standard Bonferroni correction). Standard Bonferroni correction was applied for all multiple tests

methylation in invertebrate genomes occur in the form of gene-body methylation (GBM), that is, methylation occurs in transcription units including exons and introns (Feng et al. 2010; Zemach et al. 2010). Recently, genome-wide bisulfite sequencing of multiple invertebrates, including *C. intestinalis*, *Apis mellifera*, *Bombyx mori* and *Nematostella vectensis*, confirmed that prominent GBM was a common feature in invertebrate genomes (Feng et al. 2010; Sarda et al. 2012; Xiang et al. 2010; Zemach et al. 2010). Unlike promoter methylation, GBM is not related simply to gene expression levels. Studies showed that a relatively high level of methylation in gene bodies tended to make methylated genes constitutively expressed at moderate to high levels (Suzuki et al. 2007; Elango et al. 2009; Sarda et al. 2012; Zemach et al. 2010), which may support the hypothesis that gene body DNA methylation functions to suppress spurious

transcriptional initiation within infrequently transcribed genes (Suzuki et al. 2007; Zemach et al. 2010).

Although the role of GBM remains enigmatic, DNA methylation in invertebrates is generally believed to be associated with the transcription level, alternative splicing and genome evolution (Jones 2012; Sarda et al. 2012; Takuno and Gaut 2012). Studies in invertebrates revealed that DNA methylation played roles in environmentally driven gene expression shifts (Gavery and Roberts 2010; Glastad et al. 2013; Wang et al. 2013). Processes associated with biological invasions, such as transfers of individuals to dramatically different environments and sudden changes in population demography, can lead to exceedingly rapid evolutionary changes (Prentis et al. 2008). DNA methylation is a primary mechanism of DNA modifications by which gene expression activities maybe modified in response to environmental

Fig. 5 Sample score plots generated by principal component analysis (PCA) on the data of methylation level based on all CpGs in gene bodies of the heat shock protein 90 gene (HSP90) and $\text{Na}^+\text{-K}^+\text{-2Cl}^-$ cotransporter gene (NKCC). Ellipses show the 95% confidence for each group. Population names per Table 1

stimuli, and in invertebrates DNA methylation appears to be targeted specifically to transcription units (Sarda et al. 2012; Zemach et al. 2010). Therefore, GBM may have an important role in local adaptation by regulating transcriptional level or improving the accuracy of splicing under temperature and salinity stresses in invertebrates. Our study found significant methylation difference at exons of HSP90 and NKCC in four populations of *C. robusta* (Fig. 4), suggesting that GBM may play roles in local adaptation to various environmental factors by regulating transcriptional level or the accuracy of splicing of HSP90 and NKCC expression in different environments. As a result, DNA methylation variation induced by environmental changes may be functional and allow a population to display phenotypic variability.

When the data from each CpG was analyzed to detect significant difference among populations, significant variation of DNA methylation level was detected at CpGs in gene bodies of two key genes. It is known that DNA methylation among populations can be affected by environmental factors and population genetic features (Dyachenko et al. 2006; Feil 2006; Kou et al. 2011; Lillycrop et al. 2005; Pecinka et al. 2010; Weaver et al. 2004). In animals,

environmental factors such as the maternal diet, rearing behavior and folic acid supplementation can alter de novo programming of DNA methylation during the development of offspring (Feil 2006; Lillycrop et al. 2005; Weaver et al. 2004). In addition, we used key genes directly associated with environmental factors here and found significant differences in DNA methylation in HSP90 and NKCC. All the results suggest that these two genes may play important roles in response to environmental stresses by changing DNA methylation patterns (Gamba 2005; Vabulas et al. 2010). However, genetic and epigenetic variation do not work absolutely independently in natural populations. Epigenetic variation can be affected by genetic variation and environment factors, as well as their complex interactions (Richards 2006). Therefore, the identification of independent epigenetic variation is extremely challenging in the wild, because any dependence on genotypes needs to be ruled out as a direct causal factor. It is particularly difficult in *Ciona* populations that contain extremely abundant genetic variation collected from the wild. Collectively, our results suggest that environmental factors analyzed here, as well as many unrecovered factors such as genetic features of different populations and many untested genes such as endoplasmic reticulum chaperones (Sato et al. 2015), could induce changes in DNA methylation, and in turn the changes of DNA methylation could play a role in rapid local adaptation to novel environmental conditions in biological invasions.

When multiple correlation analyses between the environmental factors and DNA methylation levels were performed, our results showed that some CpGs were significantly correlated with temperature and salinity in two key genes (Table 3). Many studies have also shown that environmental factors, such as salinity, drought and extreme temperature can lead to different epigenetic patterns and that such changes can play crucial roles in the regulation of gene expression responding to environmental stresses (Dyachenko et al. 2006; Feil 2006; Kou et al. 2011; Lillycrop et al. 2005; Pecinka et al. 2010; Weaver et al. 2004; Xie et al. 2015). However, it is a complicated molecular process to regulate DNA methylation level in responding to environmental stimuli, as well as many environmental factors (e.g. temperature and salinity) are fluctuant in different times at the same place, and it is difficult to detect significant correlation

between environmental factors and DNA methylation level at single CpG loci. Consequently, we performed principal component analysis (PCA) based on the data of methylation level at all CpGs in the two key genes with significant difference in the four populations. Our results showed that individuals were well assigned back to their population origins (Fig. 5). In contrast, population genetic analyses using both mitochondrial DNA and nuclear microsatellites did not find significant genetic differentiation in the population analyzed here (Zhan et al. 2010). Although the results obtained based on functional genes in this study cannot directly compared with those based on neutral molecular markers, multiple lines of evidence obtained here suggest that DNA methylation in key genes may play a vital role in response to environmental stresses either by single CpG loci or through gene networks.

In conclusion, direct sequencing of bisulfite-converted DNA was used to assess epigenetic patterns of five key genes corresponding to water temperature and salinity tolerance in four selected populations of a highly invasive model ascidian, *C. robusta*. The results clearly showed that DNA methylation mainly occurred in gene bodies, rather than promoters, at regions with low values of CpG O/E (<0.8). Significant DNA methylation variation was detected at CpGs in gene bodies of two of the five studied genes. More importantly, we found significant correlation between methylation level and environmental factors at CpGs in the gene body of HSP90 and NKCC. When methylation levels of all CpGs were used for PCA analysis, individuals were assigned back to their population origins. All the results suggest that environmental factors likely contribute to the variation of DNA methylation at the selected key genes, and such changes in DNA methylation, either by some loci alone or through gene networks, may be involved in rapid local adaptation during biological invasions.

Acknowledgements This work was supported by the National Natural Science Foundation of China (31622011; 31272665) and 100-Talent Program of the Chinese Academy of Sciences to A.Z.

References

- Angers B, Castonguay E, Massicotte R (2010) Environmentally induced phenotypes and DNA methylation: how to deal with unpredictable conditions until the next generation and after. *Mol Ecol* 19:1283–1295
- Bates WR (2005) Environmental factors affecting reproduction and development in ascidians and other Protochordates. *Can J Zool* 83:51–61
- Bird AP (1980) DNA methylation and the frequency of CpG in animal DNA. *Nucleic Acids Res* 8:1499–1504
- Brunetti R et al (2015) Morphological evidence that the molecularly determined *Ciona intestinalis* type A and type B are different species: *Ciona robusta* and *Ciona intestinalis*. *J Zoolog Syst Evol Res* 53:186–193
- Ciancio JE, Rossi CR, Pascual M, Anderson E, Garza JC (2015) The invasion of an Atlantic Ocean river basin in Patagonia by Chinook salmon: new insights from SNPs. *Biol Invasions* 17:2989–2998
- Colautti RI, Lau JA (2015) Contemporary evolution during invasion: evidence for differentiation, natural selection, and local adaptation. *Mol Ecol* 24:1999–2017
- Dehal P et al (2002) The draft genome of *Ciona intestinalis*: insights into chordate and vertebrate origins. *Science* 298:2157–2167
- Dray S, Dufour A (2007) The ade4 package: implementing the duality diagram for ecologists. *J Stat Softw* 22:1–20
- Dyachenko OV et al (2006) Effect of hypermethylation of CCWGG sequences in DNA of Mesembryanthemum crystallinum plants on their adaptation to salt stress. *Biochemistry (Mosc)* 71:461–465
- Elango N, Hunt BG, Goodisman MAD, Yi SV (2009) DNA methylation is widespread and associated with differential gene expression in castes of the honeybee, *Apis mellifera*. *Proc Natl Acad Sci U S A* 106:11206–11211
- Evans DH, Piermarini PM, Choe KP (2005) The multifunctional fish gill: dominant site of gas exchange, osmoregulation, acid-base regulation, and excretion of nitrogenous waste. *Physiol Rev* 85:97–177
- Feil R (2006) Environmental and nutritional effects on the epigenetic regulation of genes. *Mutat Res* 600:46–57
- Feng S et al (2010) Conservation and divergence of methylation patterning in plants and animals. *Proc Natl Acad Sci U S A* 107:8689–8694
- Gamba G (2005) Molecular physiology and pathophysiology of electroneutral cation chloride cotransporters. *Physiol Rev* 85:423–493
- Gavery MR, Roberts SB (2010) DNA methylation patterns provide insight into epigenetic regulation in the Pacific oyster (*Crassostrea gigas*). *BMC Genom* 11:483
- Glastad KM, Hunt BG, Goodisman MAD (2013) Evidence of a conserved functional role for DNA methylation intermites. *Insect Mol Biol* 22:143–154
- Han X, Patters AB, Jones DP, Zelickovic I, Chesney RW (2006) The taurine transporter: mechanisms of regulation. *Acta Physiol (Oxf)* 187:61–73
- Hudson J, Viard F, Roby C, Rius M (2016) Anthropogenic transport of species across native ranges: unpredictable genetic and evolutionary consequences. *Bio Lett* 12:20160620
- Jiang M et al (2010) Rapid quantification of DNA methylation by measuring relative peak heights in direct bisulfite-PCR sequencing traces. *Lab Invest* 90:282–290
- Jones PA (2012) Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet* 13:484–492

- Kass SU, Pruss D, Wolffe AP (1997) How does DNA methylation repress transcription? *Trends Genet* 13:444–449
- Kirkpatrick M, Barrett B (2015) Chromosome inversions, adaptive cassettes and the evolution of species' ranges. *Mol Ecol* 24:2046–2055
- Kou HP et al (2011) Heritable alteration in DNA methylation induced by nitrogen-deficiency stress accompanies enhanced tolerance by progenies to the stress in rice (*Oryza sativa* L.). *J Plant Physiol* 168:1685–1693
- Li LC, Dahiya R (2002) MethPrimer: designing primers for methylation PCRs. *Bioinformatics* 18:1427–1431
- Lillycrop KA, Phillips ES, Jackson AA, Hanson MA, Burdge GC (2005) Dietary protein restriction of pregnant rats induces and folic acid supplementation prevents epigenetic modification of hepatic gene expression in the offspring. *J Nutr* 135:1382–1386
- Mayer MP, Bukau B (2005) Hsp70 chaperones: cellular functions and molecular mechanism. *CMLS Cell Mol Life Sci* 62:670–684
- Mirouze M, Paszkowski J (2011) Epigenetic contribution to stress adaptation in plants. *Curr Opin Plant Biol* 14:267–274
- Nydam ML, Harrison RG (2010) Polymorphism and divergence within the ascidian genus *Ciona*. *Mol Phylogenet Evol* 56:718–726
- Pecinka A et al (2010) Epigenetic regulation of repetitive elements is attenuated by prolonged heat stress in *Arabidopsis*. *Plant Cell* 22:3118–3129
- Pimentel D, Zuniga R, Morrison D (2005) Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecol Econ* 52:273–288
- Prentis PJ, Wilson JR, Dormontt EE, Richardson DM, Lowe AJ (2008) Adaptive evolution in invasive species. *Trends Plant Sci* 13:288–294
- Procaccini G, Affinito O, Toscano F, Sordino P (2011) A new animal model for merging ecology and evolution. In: *Evolutionary biology—concepts, biodiversity, macroevolution and genome evolution*. Springer, Berlin Heidelberg, pp 91–106
- Richards EJ (2006) Inherited epigenetic variation—revisiting soft inheritance. *Nat Rev Genet* 7:395–401
- Rocha RM, Kremer LP, Baptista MS, Metri R (2009) Bivalve cultures provide habitat for exotic tunicates in southern Brazil. *Aquat Invasions* 4:195–205
- Sanford E, Kelly MW (2011) Local adaptation in marine invertebrates. *Annu Rev Mar Sci* 3:509–535
- Sarda S, Zeng J, Hunt BG, Yi SV (2012) The evolution of invertebrate gene body methylation. *Mol Biol Evol* 29:1907–1916
- Sato A, Kawashima T, Fujie M, Hughes S, Satoh N, Shimeld SM (2015) Molecular basis of canalization in an ascidian species complex adapted to different thermal conditions. *Scientific Reports* 5:16717–16725
- Seo KS, Lee Y (2009) A First Assessment of Invasive Marine Species on Chinese and Korean Coasts. *Biological Invasions in Marine Ecosystems* 204:577–585
- Skou JC, Esmann M (1992) The Na⁺, K⁺-ATPase. *J Bioenerg, Biomembrane* 24:249–261
- Sumner S, Pereboom JM, Jordan WC (2006) Differential gene expression and phenotypic plasticity in behavioural castes of the primitively eusocial wasp, *Polistes canadensis*. *Proceedings of the Royal Society B* 273:19–26
- Suzuki MM, Kerr ARW, Sousa DD, Bird A (2007) CpG methylation is targeted to transcription units in an invertebrate genome. *Genome Res* 17:625–631
- Takuno S, Gaut BS (2012) Body-methylated genes in *Arabidopsis thaliana* are functionally important and evolve slowly. *Mol Biol Evol* 29:219–227
- Tsai JR, Lin HC (2007) V-type H⁺-ATPase and Na⁺, K⁺-ATPase in the gills of 13 euryhaline crabs during salinity acclimation. *J Exp Biol* 210:620–627
- Vabulas RM, Raychaudhuri S, Hayer-Hartl M, Hartl FU (2010) Protein Folding in the Cytoplasm and the Heat Shock Response. *Cold Spring Harb Perspect Biol* 2:a004390
- Vinson JP et al (2005) Assembly of polymorphic genomes: algorithms and application to *Ciona savignyi*. *Genome Res* 15:1127–1135
- Wang X et al (2013) Function and evolution of DNA methylation in *nasonia vitripennis*. *PLoS Genet* 9:e1003872
- Weaver ICG et al (2004) Epigenetic programming by maternal behavior. *Nat Neurosci* 7:847–854
- Xiang H et al (2010) Single base resolution methylome of the silkworm reveals a sparse epigenomic map. *Nat Biotechnol* 28:516–520
- Xie HJ et al (2015) ICE1 demethylation drives the range expansion of a plant invader through cold tolerance divergence. *Mol Ecol* 24:835–850
- Yancey PH (2005) Organic osmolytes as compatible, metabolic and counteracting cytoprotectants in high osmolarity and other stresses. *J Exp Biol* 208:2819–2830
- Zemach A, McDaniel IE, Silva P, Zilberman D (2010) Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science* 328:916–919
- Zhan A, MacIsaac HJ, Cristescu ME (2010) Invasion genetics of the *Ciona intestinalis* species complex: from regional endemism to global homogeneity. *Mol Ecol* 19:4678–4694
- Zhan A, Darling JA, Bock DG, Lacoursière-Rousse A, MacIsaac HJ, Cristescu ME (2012) Complex genetic patterns in closely related colonizing invasive species. *Ecol Evol* 2:1331–1346
- Zhan A, Briski E, Bock DG, Ghabooli S, MacIsaac HJ (2015) Ascidiaceans as models for studying invasion success. *Mar Biol* 162:2449–2470