Heredity (2007) 99, 592–600

www.nature.com/hdy

ORIGINAL ARTICLE

Adaptive divergence in a high gene flow environment: *Hsc70* variation in the European flounder (*Platichthys flesus* L.)

J Hemmer-Hansen[1,2,3], EE Nielsen[3], J Frydenberg[1] and V Loeschcke[1]
[1]*Department of Ecology and Genetics, University of Aarhus, Aarhus C, Denmark;* [2]*Department of Marine Ecology, University of Aarhus, Finlandsgade, Aarhus N, Denmark and* [3]*Department of Inland Fisheries, Danish Institute for Fisheries Research, Vejlsøvej, Silkeborg, Denmark*

Little is known about local adaptations in marine fishes since population genetic surveys in these species have typically not applied genetic markers subject to selection. In this study, we used a candidate gene approach to investigate adaptive population divergence in the European flounder (*Platichthys flesus* L.) throughout the northeastern Atlantic. We contrasted patterns of genetic variation in a presumably neutral microsatellite baseline to patterns from a heat-shock cognate protein gene, *Hsc70*. Using two different neutrality tests we found that the microsatellite data set most likely represented a neutral baseline. In contrast, *Hsc70* strongly deviated from neutral expectations. Importantly, when estimating standardized levels of population divergence (F_{ST}), we also found a large discrepancy in the patterns of structuring in the two data sets. Thus, samples grouped according to geographical or historical proximity with regards to microsatellites, but according to environmental similarities with regards to *Hsc70*. The differences between the data sets were particularly pronounced in pairwise comparisons involving populations in the western and central Baltic Sea. For instance, the genetic differentiation between geographically close Baltic Sea and North Sea populations was found to be 0.02 and 0.45 for microsatellites and *Hsc70* respectively. Our results strongly suggest adaptive population divergence and indicate local adaptations at the DNA level in a background of high levels of gene flow, typically found in many marine fish species. Furthermore, this study highlights the usefulness of the candidate gene approach for demonstrating local selection in non-model organisms such as most marine fishes.
Heredity (2007) **99,** 592–600; doi:10.1038/sj.hdy.6801055; published online 12 September 2007

Keywords: local adaptation; adaptive evolution; candidate genes; marine fish; European flounder; heat shock cognate gene

Introduction

Marine fishes have been subject to an increasing number of population genetic studies applying neutral genetic markers in recent years. These studies have generally confirmed that gene flow between populations of many marine fishes is high (Waples, 1998). Since gene flow is expected to hamper adaptive population divergence, these results indicate that local adaptations may be rare or absent in marine fishes. On the other hand, the wide distributions over diverse environments and large effective population sizes of many marine fish species would tend to favor the effects of natural selection over the random effect of drift and the homogenizing effect of gene flow (see Conover *et al*., 2006 and references therein). Despite high levels of gene flow, a number of studies have identified significant, albeit small, levels of genetic structuring among populations of marine fishes, primarily with the aid of highly variable genetic markers such as microsatellites (for example Ruzzante *et al*., 1999; Nielsen *et al*., 2003; Bekkevold *et al*., 2005). However, because these markers are presumably neutral, any inferences about the evolutionary significance of the results in terms of adaptive population divergence have been based mostly on speculations regarding the potential for local adaptations to be present. Thus, evolutionary scenarios with temporally stable neutral genetic structuring among populations could indicate that strong local selection would be able to override the effects of drift and gene flow, resulting in temporally stable adaptive population divergence, that is local adaptations.

Only rarely have marine fishes been subject to studies of population differences at candidate genes believed to be directly affected by selection. One classical example of a selected gene in marine fishes is the hemoglobin locus in Atlantic cod (*Gadus morhua*). This locus shows clinal allele frequency variation along a latitudinal cline in the northeastern Atlantic as well as differences between the North Sea and the Baltic Sea (Sick, 1965a, b), and physiological experiments strongly suggest that natural selection is maintaining this variation (Brix *et al*., 1998). Few other examples of genes suggested to be under selection in marine fishes can be found (reviewed in Guinand *et al*., 2004) such as the vesicular membrane protein gene *Pan*I in Atlantic cod (Pogson, 2001; Pampoulie *et al*., 2006) and the *Ldh-B* gene in the killifish,

Correspondence: Dr J Hemmer-Hansen, Danish Institute for Fisheries Research, Department of Inland Fisheries, Vejlsøvej 39, Silkeborg 8600, Denmark.
E-mail: jhh@difres.dk
Received 19 February 2007; revised 14 July 2007; accepted 6 August 2007; published online 12 September 2007

Fundulus heteroclitus (Schulte *et al*., 2000). It has, however, rarely been attempted to separate the effects of natural selection from demographic effects, that is migration and genetic drift, by comparing population differentiation in markers under selection to a baseline generated from presumably neutral markers in these species (but see for example, Pampoulie *et al*., 2006).

In this study, we examine genetic variation in a candidate gene for adaptive divergence of European flounder (*Platichthys flesus* L.) populations in combination with a neutral baseline from genomic microsatellite loci. The European flounder is a wide-ranging euryhaline flatfish species inhabiting a large part of the northeastern Atlantic including the brackish Baltic Sea. As a coastal species inhabiting shallow areas the flounder is more exposed to spatial and temporal environmental differences in parameters such as temperature, salinity and light compared to other species found in deeper and more stable marine habitats. Furthermore, the flounder shows an unusual tolerance of low salinities; for example, it often stays in rivers for long periods before returning to the sea to spawn. Flounders are exposed to two major environmental gradients in European waters. One gradient is a gradual environmental transition with latitude in the Atlantic parts of the distributional area. The other is a more abrupt change between the marine North Sea and brackish Baltic Sea. Flounders in the innermost Baltic Sea are believed to have adapted to the local environment by changing from the normal pelagic spawning strategy to benthic spawning near the coast (Aro, 1989). This behavior is assumed to increase egg survivorship because eggs are not exposed to oxygen-poor water in the deeper parts of the Baltic Sea. Furthermore, and as seen in other marine fishes in the Baltic Sea, the flounders in the western Baltic Sea show distinct physiological characteristics. For instance, increased egg volumes and higher sperm mobility at lower salinities compared to neighboring marine populations in the North Sea are believed to reflect local adaptations to the brackish environment in the Baltic Sea (Nissling *et al*., 2002). Thus, several physiological and behavioral traits suggest adaptations to local environments in flounder populations. However, the genetic basis of these traits has never been investigated and hence the degree of plasticity of the traits remains unknown.

Due to the flounder's unique geographical and ecological distribution it is very well suited for studying the potential effects of environmental parameters on adaptive divergence among marine fish populations. In this study, we use a heat-shock protein gene as a candidate gene for local adaptations in the European flounder. Heat-shock cognate 70 (*Hsc70*, sometimes denoted *Hsc71*) is a member of the heat-shock protein 70 (*Hsp70*) gene family of molecular chaperones. These genes have been found in every organism from bacteria and plants to humans and play a central role in the cellular stress response system by assuring correct transport and folding of damaged proteins (Feder and Hofmann, 1999). *Hsp70* gene expression has been found to be induced by a large variety of stressors in fishes, such as elevated and lowered temperature, osmotic stress, radiation and heavy metals (reviewed in Iwama *et al*., 1998), illustrating their ubiquitous role in the cellular stress response. The function of *Hsc70* is believed to be similar to that of *Hsp70* except that the *Hsc70* genes are primarily constitutively expressed (Freeman and Morimoto, 1996, but see also Goldfarb *et al*., 2006). The few *Hsp70* gene sequences available from fishes show that the genes are highly conserved through fish evolution (Basu *et al*., 2002), which indicates that they are under heavy selective constraints.

Our aims were to study local adaptations in a marine fish, the European flounder, by investigating adaptive genetic divergence among populations covering a large part of the species' distribution. We compared patterns and levels of standardized genetic differentiation for a *Hsc70* linked indel and a set of neutral microsatellite markers. Despite the great potential for this candidate gene versus neutral variation approach to disclose the genetic basis of adaptive population divergence among natural populations, it has only rarely been applied in marine fish species (but see, for example, Pampoulie *et al*., 2006).

Materials and methods

Sampling and DNA extraction

Samples of European flounder were collected in 2003 and 2004 covering most of the species' distributional range (Figure 1 and Table 1). We collected samples from the Baltic Sea (Turku, Gotland, Bornholm and Ærø), Danish fjords (Ringkøbing Fjord and the Limfjord) and the Atlantic/North Sea (Thyborøn, Irish Sea, Bay of Biscay, the Faroe Islands and Trondheim). Finally, we collected flounders from Lake Pulmanki, which is connected to the Barents Sea via ~100 km of the river Teno, through which flounders presumably need to travel to the sea to spawn. DNA was extracted from ethanol stored fin or gill tissue by Chelex (Estoup *et al*., 1996), DNeasy (Qiagen, Valencia, CA, USA) and HotSHOT (Truett *et al*., 2000) methods.

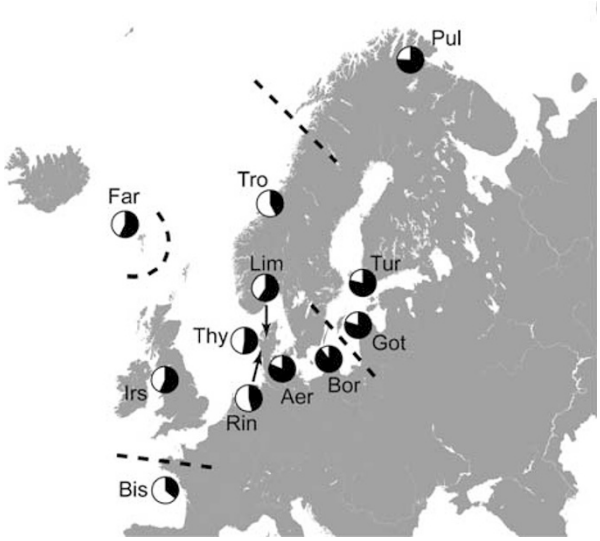

Figure 1 Map of sampling locations and *Hsc70* allele frequencies. Sample numbers refer to sample names in Table 1. Black proportion of pie charts denotes frequencies of the long allele. Major barriers to gene flow identified in the microsatellite data set are indicated by broken lines.

Table 1 Location and sample details of European flounder samples

	Locality	*Sample year*	*Approximate position*	*Spawning strategy*	*Proportion mature*	*Total sample size*
1	Turku (Tur)	2003 and 2004	22°E, 60°N	Benthic	39	79
2	Gotland (Got)	2003 and 2004	19°E, 57.5°N	Benthic	~24	94
3	Bornholm (Bor)	2003 and 2004	16°E, 55°N	Pelagic	100	108
4	Ærø (Aer)	2003 and 2004	10°E, 55°N	Pelagic	100	103
5	Thyborøn (Thy)	2003 and 2004	8°E, 57°N	Pelagic	35	80
6	Ringkøbing Fjord (Rin)	2004	8.3°E, 55.96°N	Pelagic	10	50
7	The Limfjord (Lim)	2003	8.59°E, 56.5°N	Pelagic	0	34
8	Irish Sea (Irs)	2003 and 2004	−4°E, 54°N	Pelagic	63	73
9	Bay of Biscay (Bis)	2003 and 2004	−2.3°E, 47.20°N	Pelagic	0	55
10	Trondheim (Tro)	2004	11°E, 65°N	Pelagic	Juveniles	49
11	Lake Pulmanki (Pul)	2004	28.02°E, 70.01°N	Pelagic	0	34
12	Faroe Islands (Far)	2003 and 2004	−6.45°E, 62°N	Pelagic	64	50

Genome walking and screen for genetic variation

A fragment of the gene coding the European flounder *Hsc70* gene was obtained by genome walking using the DNA walking SpeedUp kit (Seegene, Seoul, Korea). Primers were designed to walk upstream from aligned sequences of the 5′ end of other fish *Hsp70* genes. Obtained sequences were subsequently used to design primers for additional walks upstream and downstream from the known sequence.

Between two and four individuals from different populations were used for an initial screen for genetic variation. These individuals were chosen to represent populations from the entire geographical range sampled in the present study. All fragments were cloned in the TOPO TA Cloning kit (Invitrogen, Carlsbad, CA, USA) prior to sequencing. Several clones from each of a number of individuals were sequenced to verify that we were most likely amplifying a single locus.

PCR and genotyping

Nine highly polymorphic microsatellites developed for plaice (LIST1001, GenBank accession no. AF149831, Watts *et al*., 1999; PL142 and PL167, accession nos. AF406750 and AF406751, Hoarau *et al*., 2002) and European flounder (StPf1001, StPf1002, StPf1004, StPf1005, StPf1015 and StPf1022, accession nos. AJ315970, AJ315975, AJ315973, AJ315974, AJ538313, AJ538320, TJ Dixon *et al*., unpublished) were amplified under standard PCR conditions and genotyped on an ALFExpress automated sequencer (Amersham Biosciences, Uppsala, Sweden). Other primers were designed to amplify a 363 bp fragment containing a 21 bp indel in intron 1 of the flounder *Hsc70* gene (Forward 5′-GAGACATGTGAGG GATCCCTCC-3′; Reverse 5′-CATCATTCTTGCTGGAA ACAAGC-3′). These fragments were amplified under standard PCR conditions and genotyped on agarose gels stained with ethidium bromide.

Statistical analysis

Departures from Hardy–Weinberg equilibrium were tested with the exact test implemented in the software GENEPOP (Raymond and Rousset, 1995) for microsatellites and *Hsc70* separately. Temporal stability was assessed by estimating pairwise F_{ST} (θ, Weir and Cockerham, 1984), their confidence intervals (by bootstrapping over loci) and significance between temporal samples within localities in the program FSTAT (Goudet, 1995). We calculated standardized levels of population divergence (independent of levels of heterozygosity; Hedrick, 2005) to facilitate comparisons of population structuring between the microsatellite and *Hsc70* data sets because these markers are expected to differ markedly in levels of heterozygosity and hence in maximum levels of population divergence attainable (see Hedrick, 2005). We used the program RecodeData v. 1.0 (available from http://www.bentleydrummer.nl/software/index.html) to recode the data set so that each population had unique alleles. This recoded data set allowed the estimation of the maximum level of population differentiation. The actual estimate of population differentiation (F_{ST}) was then divided by this maximum level to estimate the standardized level of population subdivision (F'_{ST}). To compare patterns of structuring of the two marker types we constructed multidimensional scaling (MDS) plots of F'_{ST} with the program Vista 5.6.3 (Young, 1996) for microsatellites and *Hsc70* separately. Population structure of the microsatellite data set was also investigated by a landscape genetics approach using the program BARRIER (Manni *et al*., 2004) to find the largest breaks in genetic constitution based on genetic and geographic information from the samples. It should be noted that landscape genetics approaches are highly dependent on a sufficient spatial resolution of the sampling scheme. Thus, both the existence and geographical location of barriers in areas with large distances between samples (for example, between Lake Pulmanki and Trondheim in this study) can be difficult to determine. On the other hand, areas with extensive sample coverage (such as in the Baltic Sea) allow a more certain demonstration of major genetic breaks as well as of their geographical location. The landscape genetics analyses were run on 100 F'_{ST} matrices generated by bootstrapping over microsatellite loci as well as on each microsatellite locus separately to assess the robustness of barriers. Pairwise comparisons of F'_{ST} for microsatellites and *Hsc70* were conducted to examine which population pairs showed the largest discrepancy between the two marker types, that is where selection could be inferred. For this purpose, we calculated 95% confidence intervals for microsatellite F'_{ST} by bootstrapping over loci in 100 data sets. The outlier status of *Hsc70* could subsequently be assessed by evaluating estimates of F'_{ST} for *Hsc70* in relation to the microsatellite confidence intervals.

We used two different tests to assess the assumption of microsatellite neutrality. The simulation-based test by Beaumont and Nichols (1996) implemented in the program FDIST2 (available from Mark Beaumont's webpage at http://www.rubic.rdg.ac.uk/~mab/) and the LnRH test (Kauer *et al*., 2003 and references therein) were both developed to identify outlier loci in genome scans, but differ in their approaches and assumptions in a number of ways. The FDIST2 test is based on the assumption that outlier loci will show increased levels of population structuring if they are under diversifying selection or closely linked to a locus which is (or that is, genetic hitch-hiking) (Maynard Smith and Haigh, 1974). The method compares actual levels of differentiation at individual loci in relation to heterozygosity to a simulated distribution of loci generated from observed levels of population differentiation. We carried out the simulation under both the Stepwise Mutation Model (SMM), which should be well-suited for microsatellites, and the Infinite Alleles Model (IAM), which should describe the mutational process of an indel better. Beaumont and Nichols (1996) have shown that the type of marker applied (that is mutational mechanisms involved) has little effect on the simulated distribution under neutrality. Consequently, this method was used both to evaluate the neutrality of microsatellites as well as the outlier status of *Hsc70*. *Hsc70* was not, however, used to generate the expected neutral distribution, since we suspected this locus *a priori* to be under selection.

The LnRH test assumes that microsatellite loci which are linked to a gene of adaptive importance subject to a selective sweep will show reduced levels of diversity within the populations subject to selection (Kauer *et al*., 2003 and references therein). The LnRH test therefore compares relative levels of heterozygosity between loci in pairwise population comparisons. Since the LnRH test was specifically developed for microsatellite loci it has not been evaluated in situations where mutational mechanisms vary between loci and, accordingly, it was only used to assess the assumption of neutrality of the microsatellite baseline.

Results

Genome walking and screen for genetic variation

The genome walking produced a sequence of 1586 bp. This sequence showed the highest similarity to Japanese flounder (*Paralichthys olivaceus*) *Hsc70* mRNA followed by other *Hsc70/Hsc71* and *Hsp70* genes in a Blast search (Altschul *et al*., 1997). Based on an alignment with the Japanese flounder *Hsc70* mRNA, we assume that the European flounder sequence contains part of intron 1, the entire exon 1 and a part of intron 2 of a *Hsc70* gene. Other genomic sequences from fish contain an untranslated exon 0 upstream of the translational start site. This exon is located at least 1270 bp upstream in two *Hsc70* genes from *Rivulus marmoratus* (Lee, 2004) and ~1830 bp upstream in rainbow trout (Zafarullah *et al*., 1992). Based on comparisons with these few available fish genomic sequences, and the fact that gene prediction softwares failed to identify any transcriptional start sites in the European flounder sequence (results not shown), we find it unlikely that we have reached this potential untranslated exon 0 by the genome walking in European flounder. The partial sequence of European flounder *Hsc70* has been deposited in GenBank under accession no. EU029640. The initial screen for genetic variation identified several single-nucleotide polymorphisms and a 21 bp deletion in intron 1. The deletion was initially observed in individuals from the Irish Sea, Turku and Lake Pulmanki and was subsequently genotyped in all individuals.

Sequencing of several clones from single individuals confirmed that the primers used were most likely amplifying a single locus because only one or two alleles were found in each individual.

Genetic variation

Exact tests showed that no locus or sample exhibited consistent deviations from Hardy–Weinberg equilibrium with respect to microsatellites (see Supplementary Table S1). Likewise, *Hsc70* conformed to Hardy–Weinberg expectations in all populations (see Supplementary Table S1). Only one pairwise comparison between temporal samples was statistically significant (Thy03 versus Thy04 for microsatellites, $F_{ST} = 0.006$, $P = 0.04$). However, this is not more than expected by chance in 16 tests and it does not change the overall picture of temporal stability over the 2 years studied. Temporal samples from the same localities were therefore pooled for the remaining analyses.

Population structure

The overall F_{ST} for microsatellites was 0.025 (95% CI 0.019–0.032) and highly significant ($P < 0.0001$), while the overall F_{ST} for *Hsc70* was 0.14 and also highly significant ($P < 0.0001$). The overall standardized levels of population divergence (F'_{ST}) were 0.08 and 0.23 for microsatellites and *Hsc70*, respectively. As evidenced by Figure 1, the genetic variation at *Hsc70* seemed to be highly unevenly distributed throughout the distributional range. Samples in the Baltic Sea and Lake Pulmanki had very low frequencies of the shortest allele, while the remaining samples had higher and more variable frequencies of this allele. Examining the relationships between samples, it was evident that the two marker types showed highly discordant patterns of structuring.

The microsatellite MDS plot mirrored the geographical relationship between the samples which were separated along both dimension one, explaining 53% of the variance, and dimension two, explaining 24% of the variance. Thus, the microsatellite MDS plot revealed a clear separation of samples in the inner Baltic Sea (benthic spawners), Faroe Islands, the Bay of Biscay and Lake Pulmanki (Figure 2a). The Bornholm sample from the central Baltic Sea was positioned between the Atlantic and inner Baltic Sea samples, while the remaining samples were grouping together. This picture was largely supported by analyses on each microsatellite locus separately (results not shown). The MDS plot based on *Hsc70* (Figure 2b) identified the sample from Bornholm in the central Baltic Sea as distinct, reflecting the extremely low frequency of the short allele at this locality (see Figure 1). Furthermore, there was a clear separation along dimension one of samples in the Baltic Sea (western and innermost parts) and Lake Pulmanki in one group and all remaining samples in another and larger group. Dimension one explained 80% of the

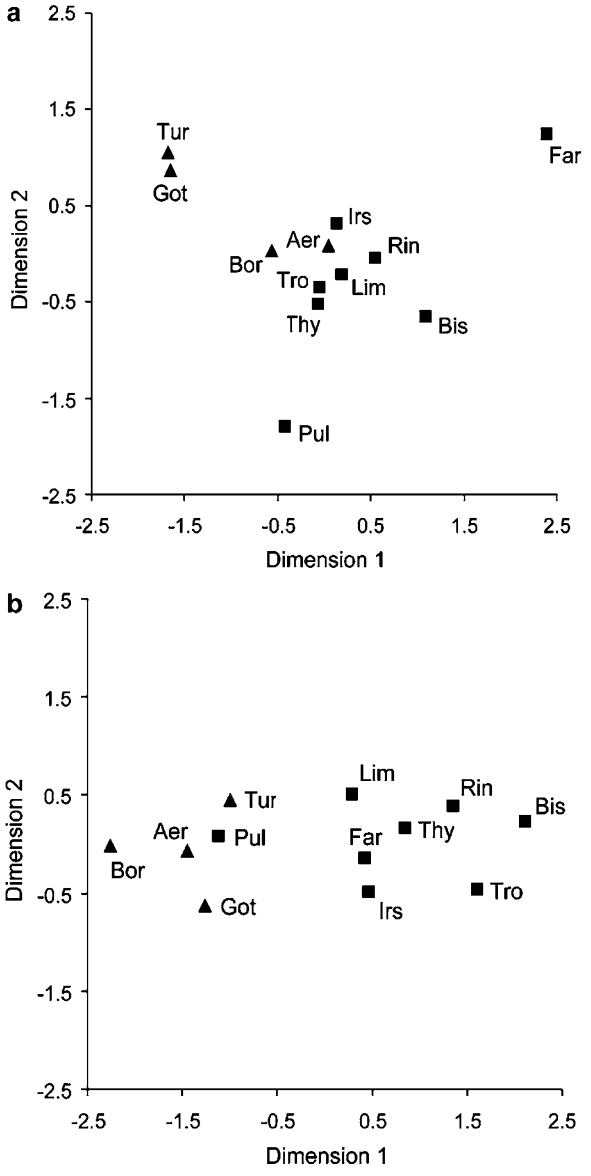

Figure 2 Multidimensional scaling plots of F'_{ST} (Hedrick, 2005) based on the microsatellite (**a**, stress 0.08) and *Hsc70* (**b**, stress 0.06) data sets. Sample names refer to Table 1. (▲) refer to Baltic Sea samples and (■) to Atlantic Sea and Lake Pulmanki samples.

variance while dimension two explained 12% of the variance.

The results from the landscape genetics analyses also provide support for a partitioning of the samples into distinct groups based on microsatellite data (Figure 1). Four barriers were identified based on the 100 bootstrapped data sets. These were positioned around the Faroe Islands, between Lake Pulmanki and Trondheim, between Gotland and Bornholm and to the north of the Bay of Biscay. A fifth barrier was initially placed between Thyborøn and the Limfjord/western Baltic Sea samples. However, this barrier resulted mainly from the Limfjord sample and not from a genetic break between the North Sea and western Baltic Sea, because the barrier disappeared when the Limfjord sample was removed from the data set. The distinctiveness of the Limfjord sample should be interpreted carefully, because this sample could contain a large fraction of released fish from a substantial supportive breeding program at this location. Another, but less well-supported barrier was placed south of Trondheim, but mainly when Trondheim was allowed a connection to the western Baltic Sea. The four main bootstrapped barriers were all supported by at least six individual microsatellite loci (except from the Bay of Biscay barrier, where four loci supported either no barrier or only a fraction of the barrier), supporting the robustness of these results.

The pairwise comparisons (F'_{ST}) in Figure 3 revealed the largest discrepancies between the two data sets in comparisons involving the western and central Baltic Sea samples. These showed particularly high levels of divergence from the Atlantic and North Sea samples for *Hsc70*, but high similarity with respect to microsatellites. For instance, the F'_{ST} estimates between Bornholm in the Baltic Sea and Thyborøn in the North Sea were 0.02 and 0.45 for microsatellites and *Hsc70* respectively, and the *Hsc70* estimate was markedly outside the 95% confidence interval for the microsatellite estimate (Figure 3). A number of pairwise comparisons showed a lower level of structure for *Hsc70* than for microsatellites. These results were almost exclusively from comparisons involving either the Faroe Islands or Lake Pulmanki samples, which both appeared quite isolated based on the microsatellite data set (see Figure 2a and Hemmer-Hansen *et al*., 2007).

When samples were divided into two major groups based on environmental similarity (Baltic Sea samples, that is Tur, Got, Bor and Aer and Lake Pulmanki versus all remaining Atlantic Sea samples, that is Tro, Thy, Rin, Lim, Irs, Bis and Far), there was a highly significant difference in *Hsc70* allele frequencies between the two groups (Mann–Whitney *U*-test, $P = 0.006$). Moreover, allele frequencies were more variable among the samples in the Atlantic group as evidenced by larger s.d. in this group (0.09 for the Atlantic group versus 0.05 for the Baltic Sea/Lake Pulmanki samples). However, this difference was only close to significant in Levene's test for variance homogeneity of arcsine transformed allele frequencies ($P = 0.08$).

Neutrality tests

We found little difference between the results obtained by simulations under the two mutation models (SMM versus IAM) as also noted by Beaumont and Nichols (1996) when analyzing simulated data. Consequently, we report only results obtained under the SMM. Results based on the estimated mean F_{ST} of 0.025 (see above) showed that only one microsatellite locus (Pl167) fell slightly outside the upper 95% confidence level ($P = 0.03$), while the *Hsc70* locus showed much higher levels of population differentiation given its level of heterozygosity ($P < 0.001$; Figure 4). We also conducted the simulations for the lower bound on F_{ST} (lower confidence level of 0.019), resulting in one additional locus (StPf1002) falling slightly outside the upper 95% confidence interval of the simulated loci. When the simulations were done with the upper bound on F_{ST} (0.032), no microsatellite loci were outliers while *Hsc70* was still identified as an extreme outlier (results not shown). The LnRH test on all 594 pairwise comparisons

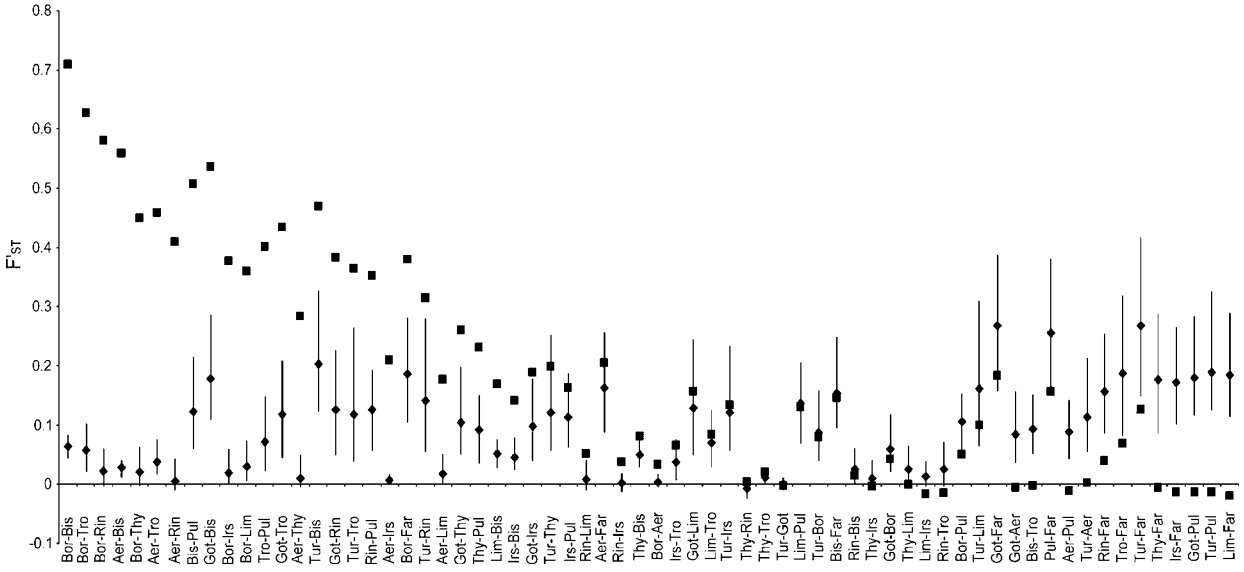

Figure 3 Pairwise point estimates of F'_{ST} for microsatellites (◆) with 95% confidence intervals calculated by bootstrapping over loci and point estimates of F'_{ST} for *Hsc70* (■). Population pairs are ranked according to the difference between the levels of divergence observed for the two marker types, that is leftmost population pairs have higher F'_{ST} for *Hsc70* than for microsatellites and vice versa for the rightmost pairs.

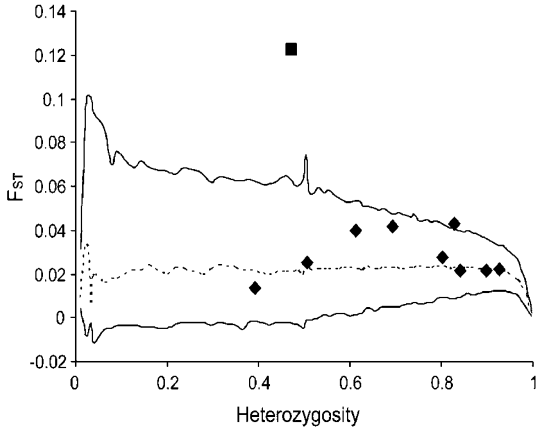

Figure 4 Results from the simulations with the FDIST2 program under the Stepwise Mutation Model. Shown is the distribution of F_{ST} values (median as broken line and 95% confidence intervals as solid lines) from 500 000 simulated loci as well as the estimated values of individual microsatellite loci (◆) and *Hsc70* (■).

(66 population comparisons in nine loci) revealed no loci as conspicuous outliers. As expected, some loci were in the outer 5% of the distribution of standardized LnRH values. These data points were primarily from loci StPf1004, StPf1001, Pl167 and StPf1005 in different pairwise sample comparisons (results not shown).

Discussion

Evidence for selection

We found no convincing evidence for non-neutrality of any of the microsatellites used to represent the neutral baseline in this study. In contrast, we found that the *Hsc70* linked marker was clearly identified as an outlier locus in terms of levels of population divergence. More importantly, there was a clear discrepancy with respect to patterns of structuring from the two data sets, that is the microsatellites reflected geographical relationships between samples while the *Hsc70* associated marker grouped samples with respect to environmental similarity (see below) rather than with geographical or historical proximity. These results are unlikely due to artifacts (see below) and hence they strongly suggest that natural selection is affecting the distribution of genetic variation at the *Hsc70* linked marker, implying adaptive population divergence of the *Hsc70* gene or at a closely linked locus in European flounder. Importantly, it is evident that such divergence is possible even in a background of very high levels of gene flow as inferred by the population similarities at microsatellite loci in population pairs where the signal of adaptive divergence is apparently largest.

Selective agents

Since it is not known if the indel polymorphism surveyed here affects transcription or function of the *Hsc70* gene in European flounder it is difficult to discriminate between a direct selection and a hitch-hiking scenario. Because there is no evidence for direct selection on the indel polymorphism (but also no evidence to suggest that there is not), we view the results in a hitch-hiking context. However, variation in a number of environmental parameters known to be associated with differences in expression or function of heat-shock protein genes can be found within the studied area, making these genes very plausible targets for local selection in this species.

In the narrow zone connecting the marine Atlantic and the brackish Baltic Sea there is a drastic environmental transition with respect to temperature and salinity in particular. Generally, both annual and intra-annual sea-surface temperature variations are larger and annual minimum temperatures are lower in the Baltic Sea than

in the nearby North Sea (for example Becker and Pauly, 1996; Siegel *et al*., 2006). Therefore, temperature fluctuations, or alternatively minimum temperatures, could be an important selective agent driving adaptive divergence at *Hsc70*, since both *Hsc70* expression and function have been found to be affected by temperature in fishes (Place and Hofmann, 2005; Fangue *et al*., 2006).

Hsc70 has also been found to be expressed in response to ambient salinity changes in fish (for example Deane and Wo, 2004), indicating a central role for *Hsc70* in the response to osmotic stress in fishes. Since this parameter is clearly distinguishing the brackish Baltic Sea and freshwater Lake Pulmanki populations from the remaining marine populations the results suggest that ambient salinity could also be involved in generating the observed *Hsc70* pattern.

A third potentially important environmental component is aquatic pollution, since heat-shock proteins have also been found to be induced in response to elevated levels of for example heavy metals (Iwama *et al*., 1998). However, while the Baltic Sea is quite heavily polluted, this is highly unlikely to apply to Lake Pulmanki because this lake is located far from any major urban areas. We therefore, regard it as more likely that temperature and/or salinity is the selective agent involved if *Hsc70* is the true target of selection. Since these environmental parameters are highly correlated among the samples in the present study, we are not able to differentiate between them and indeed they may not be mutually exclusive.

Variation in *Hsc70* allele frequencies is larger among the Atlantic Sea samples than among the Baltic Sea and Lake Pulmanki samples. Although this difference is only close to significant (probably because of a lack of statistical power with only 12 samples) this may indicate that selection is stronger at the latter localities, thereby reducing the overall variation in allele frequencies among populations.

Interestingly, the distribution of genetic variation at *Hsc70* is not perfectly correlated with temperature and salinity parameters; the benthic spawning populations in the innermost Baltic Sea (Gotland and Turku), which experience the most extreme environmental conditions (that is lowest temperatures and salinities) do not have the most extreme *Hsc70* allele frequencies in the Baltic Sea. These are found in the Bornholm sample, which represents the innermost Baltic Sea population with normal pelagic spawning. The low levels of microsatellite divergence between the samples in the western/central Baltic Sea and the North Sea could indicate that the former represent the most extreme Baltic Sea distribution of an Atlantic–Baltic Sea population complex connected through relatively high levels of gene flow. The fact that the signal of local selection at or near *Hsc70* is particularly strong in these pelagic spawning Baltic Sea populations could indicate that *Hsc70* acts as a 'first defence' against the stressful conditions, while more complex biochemical adaptations in addition to the known behavioral adaptations could be found in populations in the more extreme environments.

While the *Hsc70* allele frequencies in Lake Pulmanki are as extreme as in the Baltic Sea, it is more difficult to interpret this signal. Since these fish are most likely a component of the Barents Sea population, the signal could be reflecting allele frequencies in the Barents Sea or be specific to Lake Pulmanki, that is reflecting selective constraints on fish migrating to the lake. However, without a Barents Sea reference population it is not possible to know where the signal was generated.

Local selection in marine fishes

Recently, Pampoulie *et al*. (2006) compared genetic differentiation of microsatellites and the *Pan*I locus in Atlantic cod around Iceland in an analytical setup quite similar to the one applied here. The authors found similar patterns from the two marker types, but higher levels of structuring for the *Pan*I locus, and concluded that both oceanographic (that is demographic) and environmental (that is selection) factors were probably involved in structuring Atlantic cod around the island. In contrast, the *Pan*I locus has been found to exhibit a pattern of structuring different from neutral microsatellite markers in walleye pollock (*Theragra chalcogramma*) over broad geographical scales in the Pacific Ocean (Canino *et al*., 2005). Given the inherent difficulty of comparing levels of structuring across different marker types, the evidence of selection is considerably strengthened by the finding of diverging patterns as well as diverging levels of structuring (see also Lemaire *et al*., 2000 for an example comparing microsatellite and allozymes patterns in *Dicentrarchus labrax*).

In the present study, we found higher levels of standardized structuring for the marker supposedly under selection, but more importantly we also found highly divergent patterns of structuring for the two marker types over short geographical distances. The fact that adaptive divergence appears to be present in a background of very low levels of neutral structuring in European flounder suggest that local adaptations may not be uncommon in marine fishes, typically showing levels of neutral population structure comparable to the ones observed here.

Other species have been found to be genetically structured at neutral loci across the North Sea–Baltic Sea transition area (for example cod Nielsen *et al*., 2003; turbot Nielsen *et al*., 2004). The results from the present study could indicate that these signals are caused by reductions in gene flow due to local selection in the two areas. Future studies focusing on markers under selection in these species may clarify if this is indeed the case.

Potential influence of confounding factors

The use of different marker types in the present study could introduce statistical and interpretational problems, particularly related to downward bias of microsatellite divergence between the North and Baltic Seas caused by for example, size homoplasy (Estoup *et al*., 2002) and departure from migration-drift equilibrium (see for example Pogson *et al*., 2001). However, these mechanisms would be expected to affect all population comparisons and not just populations in the North Sea–Baltic Sea transition area. The fact that several populations of flounder appear to be neutrally structured (for example, pairwise microsatellite F'_{ST} values above 0.1, see Figure 3) suggests that the microsatellites applied in this study do not suffer seriously from size homoplasy and that time has been sufficient to allow neutral structuring to build up among populations. It therefore seems likely that the low microsatellite divergence

reflects high levels of gene flow rather than non-equilibrium situations or microsatellite mutational mechanisms. Finally, it should be noted that ascertainment bias could have influenced the observed pattern of variation in *Hsc70*. However, the *Hsc70* polymorphism was initially identified in several individuals from different populations and the deletion was originally identified in populations where it later turned out to be present at low frequencies. Furthermore, as evidenced by the microsatellite data set, the populations which are grouping with respect to *Hsc70* apparently belong to very different population components within the species. Thus, it seems unlikely that ascertainment bias should have influenced the results substantially.

Conclusions

The results from this study strongly suggest adaptive evolution even in a background of high levels of gene flow over relatively short geographical distances. Since many marine fishes demonstrate very low levels of neutral genetic differentiation, local adaptations may be much more widespread in the marine environment than previously believed from evaluating the distinctness of marine fish populations from neutral markers alone. Still, there is a need for further studies confirming the generality of the patterns observed here. Such knowledge of the scale and magnitude of local adaptations in marine fishes is central to improve our understanding of how evolution operates in the sea. Furthermore, knowledge of the genetic basis of local adaptation will add to our ability to manage biodiversity efficiently in relation to human impact such as exploitation and global warming by improving our ability to predict how the distribution and abundance of species will change in response to human-mediated evolutionary forces.

Acknowledgements

Tom Wiklund, Anders Nissling, Mike Armstrong, Richard Briggs, Yves Desaunay, Jarle Mork, Dag Altin, Hannu Sakari Mäkinen, Tarja Länsman, Ámundur Nolsøe, Palle Brogaard, Martin Sherfig, Henrik Degel, Erik Hansen, Hanne Nicolaisen, Ernst Nielson and Rene Mejlvang provided samples of flounders for which we are very grateful. We also thank Dorthe Meldrup and Karen-Lise Mensberg, who were very helpful in the laboratory. Lasse Fast Jensen, Michael Møller Hansen, Nina Poulsen and Anders Koed for help and discussions on statistics and Dorte Bekkevold for help with graphics. The comments from three reviewers greatly improved an earlier version of the article.

References

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W *et al.* (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389–3402.

Aro E (1989). A review of fish migration patterns in the Baltic. *Rapp P-v Reún Cons Int Explor Mer* **190**: 72–96.

Basu N, Todgham AE, Ackerman PA, Bibeau MR, Nakano K, Schulte PM *et al.* (2002). Heat shock protein genes and their functional significance in fish. *Gene* **295**: 173–183.

Beaumont MA, Nichols RA (1996). Evaluating loci for use in the genetic analysis of population structure. *Proc R Soc Lond B Biol Sci* **263**: 1619–1626.

Becker GA, Pauly M (1996). Sea surface temperature changes in the North Sea and their causes. *ICES J Mar Sci* **53**: 887–898.

Bekkevold D, Andre C, Dahlgren TG, Clausen LAW, Torstensen E, Mosegaard H *et al.* (2005). Environmental correlates of population differentiation in Atlantic herring. *Evolution* **59**: 2656–2668.

Brix O, Forås E, Strand I (1998). Genetic variation and functional properties of Atlantic cod hemoglobins: introducing a modified tonometric method for studying fragile hemoglobins. *Comp Biochem Physiol A* **119**: 575–583.

Canino MF, O'Reilly PT, Hauser L, Bentzen P (2005). Genetic differentiation in walleye pollock (*Theragra chalcogramma*) in response to selection at the pantophysin (*PanI*) locus. *Can J Fish Aquat Sci* **62**: 2519–2529.

Conover DO, Clarke LM, Munch SB, Wagner GN (2006). Spatial and temporal scales of adaptive divergence in marine fishes and the implications for conservation. *J Fish Biol* **69** (Suppl C): 21–47.

Deane EE, Wo NYS (2004). Differential gene expression associated with euryhalinity in sea bream (*Sparus sarba*). *Am J Physiol Regul Integr Comp Physiol* **287**: 1054–1063.

Estoup A, Jarne P, Cornuet JM (2002). Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Mol Ecol* **11**: 1591–1604.

Estoup A, Largiader CR, Perrot E, Chourrout D (1996). Rapid one-tube DNA extraction for reliable PCR detection of fish polymorphic markers and transgenes. *Mol Mar Biol Biotechnol* **5**: 295–298.

Fangue NA, Hofmeister M, Schulte PM (2006). Intraspecific variation in thermal tolerance and heat shock protein gene expression in common killifish, *Fundulus heteroclitus*. *J Exp Biol* **209**: 2859–2872.

Feder ME, Hofmann GE (1999). Heat-shock proteins, molecular chaperones, and the stress response: evolutionary and ecological physiology. *Annu Rev Physiol* **61**: 243–282.

Freeman BC, Morimoto RI (1996). The human cytosolic molecular chaperones *hsp90*, *hsp70* (*hsc70*) and *hdj-1* have distinct roles in recognition of a non-native protein and protein refolding. *EMBO J* **15**: 2969–2979.

Goldfarb SB, Kashlan OB, Watkins JN, Suaud L, Yan WS, Kleyman TR *et al.* (2006). Differential effects of *Hsc70* and *Hsp70* on the intracellular trafficking and functional expression of epithelial sodium channels. *Proc Natl Acad Sci USA* **103**: 5817–5822.

Goudet J (1995). FSTAT (Version 1.2): a computer program to calculate F-statistics. *J Hered* **86**: 485–486.

Guinand B, Lemaire C, Bonhomme F (2004). How to detect polymorphisms undergoing selection in marine fishes? A review of methods and case studies, including flatfishes. *J Sea Res* **51**: 167–182.

Hedrick PW (2005). A standardized genetic differentiation measure. *Evolution* **59**: 1633–1638.

Hemmer-Hansen J, Nielsen EE, Grønkjær P, Loeschcke V (2007). Evolutionary mechanisms shaping the genetic population structure of marine fishes; lessons from the European flounder (*Platichthys flesus* L.). *Mol Ecol* **16**: 3104–3118.

Hoarau G, Cook D, Stam WT, Olsen JL (2002). New microsatellites primers for plaice, *Pleuronectes platessa* L. (Teleostei: Pleuronectidae). *Mol Ecol Notes* **2**: 60–61.

Iwama GK, Thomas PT, Forsyth RB, Vijayan MM (1998). Heat shock protein expression in fish. *Rev Fish Biol Fish* **8**: 35–56.

Kauer MO, Dieringer D, Schlotterer C (2003). A microsatellite variability screen for positive selection associated with the 'Out of Africa' habitat expansion of *Drosophila melanogaster*. *Genetics* **165**: 1137–1148.

Lee JS (2004). Genomic cloning of a heat shock cognate 71-1 gene (*HSC71-1*) from the hermaphroditic fish *Rivulus marmoratus* (Cyprinodontiformes, Rivulidae). *DNA Seq* **15**: 33–38.

Lemaire C, Allegrucci G, Naciri M, Bahri-Sfar L, Kara H, Bonhomme F (2000). Do discrepancies between microsatellite and allozyme variation reveal differential selection between sea and lagoon in the sea bass. *Mol Ecol* **9**: 457–467.

Manni F, Guerard E, Heyer E (2004). Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by using Monmonier's algorithm. *Hum Biol* **76**: 173–190.

Maynard Smith J, Haigh J (1974). Hitch-hiking effect of a favorable gene. *Genet Res* **23**: 23–35.

Nielsen EE, Hansen MM, Ruzzante DE, Meldrup D, Gronkjaer P (2003). Evidence of a hybrid-zone in Atlantic cod (*Gadus morhua*) in the Baltic and the Danish Belt Sea revealed by individual admixture analysis. *Mol Ecol* **12**: 1497–1508.

Nielsen EE, Nielsen PH, Meldrup D, Hansen MM (2004). Genetic population structure of turbot (*Scophthalmus maximus* L.) supports the presence of multiple hybrid zones for marine fishes in the transition zone between the Baltic Sea and the North Sea. *Mol Ecol* **13**: 585–595.

Nissling A, Westin L, Hjerne O (2002). Reproductive success in relation to salinity for three flatfish species, dab (*Limanda limanda*), plaice (*Pleuronectes platessa*), and flounder (*Pleuronectes flesus*), in the brackish water Baltic Sea. *ICES J Mar Sci* **59**: 93–108.

Pampoulie C, Ruzzante DE, Chosson V, Jörundsdóttir TD, Taylor L, Thorsteinsson V *et al.* (2006). The genetic structure of Atlantic cod (*Gadus morhua*) around Iceland: insight from microsatellites, the *Pan I* locus, and tagging experiments. *Can J Fish Aquat Sci* **63**: 2660–2674.

Place SP, Hofmann GE (2005). Comparison of *Hsc70* orthologs from polar and temperate notothenioid fishes: differences in prevention of aggregation and folding of denatured proteins. *Am J Physiol Regul Integr Comp Physiol* **288**: 1195–1202.

Pogson GH (2001). Nucleotide polymorphism and natural selection at the pantophysin (*Pan I*) locus in the Atlantic cod, *Gadus morhua* (L.). *Genetics* **157**: 317–330.

Pogson GH, Taggart CT, Mesa KA, Boutilier RG (2001). Isolation by distance in the Atlantic cod, *Gadus morhua*, at large and small geographic scales. *Evolution* **55**: 131–146.

Raymond M, Rousset F (1995). Genepop (Version-1.2)—Population-genetics software for exact tests and ecumenicism. *J Hered* **86**: 248–249.

Ruzzante DE, Taggart CT, Cook D (1999). A review of the evidence for genetic structure of cod (*Gadus morhua*) populations in the NW Atlantic and population affinities of larval cod off Newfoundland and the Gulf of St Lawrence. *Fish Res* **43**: 79–97.

Schulte PM, Glémet HC, Fiebig AA, Powers DA (2000). Adaptive variation in lactate dehydrogenase-B gene expression: role of a stress-responsive regulatory element. *Proc Natl Acad Sci USA* **97**: 6597–6602.

Sick K (1965a). Haemoglobin polymorphism of cod in the Baltic and the Danish Belt Sea. *Hereditas* **54**: 19–48.

Sick K (1965b). Haemoglobin polymorphism of cod in the North Sea and the North Atlantic Ocean. *Hereditas* **54**: 49–69.

Siegel H, Gerth M, Tschersich G (2006). Sea surface temperature development of the Baltic Sea in the period 1990–2004. *Oceanologia* **48**: 119–131.

Truett GE, Heeger P, Mynatt RL, Truett AA, Walker JA, Warman ML (2000). Preparation of PCR-quality mouse genomic DNA with hot sodium hydroxide and TRIS (HotSHOT). *Biotechniques* **29**: 52–54.

Waples RS (1998). Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. *J Hered* **89**: 438–450.

Watts PC, Nash RDM, George SG, Kemp SJ (1999). Isolation and characterization of microsatellite loci in the European plaice, *Pleuronectes platessa* L. (Teleostei: Pleuronectidae). *Mol Ecol* **8**: 2151–2152.

Weir BS, Cockerham CC (1984). Estimating F-statistics for the analysis of population-structure. *Evolution* **38**: 1358–1370.

Young FW (1996). ViSta: the visual statistics system. *Research Memorandum 94-1(b)*, 2nd edn. L.L. Thursone Psychometric Laboratory, University of North Carolina: Chapel Hill, NC.

Zafarullah M, Wisniewski J, Schworak NW, Schieman S, Misra S, Gedamu L (1992). Molecular cloning and characterization of a constitutively expressed heat-shock-cognate *Hsc71* gene from rainbow trout. *Eur J Biochem* **204**: 893–900.

Supplementary Information accompanies the paper on Heredity website (http://www.nature.com/hdy)