Historical DNA reveals the demographic history of Atlantic cod (Gadus morhua) in medieval and early modern Iceland

Guðbjörg Ásta Ölafsdóttir1, Kristen M. Westfall2, Ragnar Edvardsson1 and Snæbjörn Pálsson3

1Research Centre of the Westfjords, University of Iceland, Adalstræti 21, IS415 Bolungarvík, Iceland
2VOR Marine Research Centre, Nordurtangi, IS355 Snæfellsbær, Iceland
3Department of Environmental and Life Science, University of Iceland, Sturlugata 7, IS101 Reykjavík, Iceland

Atlantic cod (Gadus morhua) vertebrae from archaeological sites were used to study the history of the Icelandic Atlantic cod population in the time period of 1500–1990. Specifically, we used coalescence modelling to estimate population size and fluctuations from the sequence diversity at the cytochrome b (cytb) and Pantophysin I (PanI) loci. The models are consistent with an expanding population during the warm medieval period, large historical effective population size ($N_E$), a marked bottleneck event at 1400–1500 and a decrease in $N_E$ in early modern times. The model results are corroborated by the reduction of haplotype and nucleotide variation over time and pairwise population distance as a significant portion of nucleotide variation partitioned across the 1550 time mark. The mean age of the historical fished stock is high in medieval times with a truncation in age in early modern times. The population size crash coincides with a period of known cooling in the North Atlantic, and we conclude that the collapse may be related to climate or climate-induced ecosystem change.

1. Introduction

The population structure and demographic history of Atlantic cod have been subjects of numerous research projects in the past decades. The first studies on allozyme variation indicated complex population structuring at small geographical scales [1]. Whether such patterns resulted from population divergence or contemporary effects of natural selection, facilitated by high variation in individual reproductive success was later questioned [2], but research on rapidly evolving and functionally important genetic markers has again revealed significant population sub-structuring [3–6]. The differentiation of inshore and offshore stocks in Norway [7,8], Iceland [6,9] and Newfoundland [3] is an example of such sub-structuring. Recently found to have a broad genetic basis [10], genetic divergence by depth was first associated with allele frequencies at the PanI locus [8]. In contrast to this genetic structuring, mitochondrial DNA relationships in Atlantic cod are characterized by shallow gene genealogies, indicating low effective population sizes [11–13]. Disparity between effective ($N_E$) and census ($N_C$) population size is common in marine animals that are often highly fecund broadcast spawners with extreme mortality of sub-adults [14,15]. This mode of reproduction results in high variance in offspring number, i.e. reproductive sweepstakes [14] that may lead to low genetic diversity within populations and stochastic population structuring [14,16].

Archived material from fisheries surveys has revealed the history of many commercial fish populations, although the samples commonly include only few generations [17–19]. Atlantic cod vertebrae have accumulated at historical fishing sites in the North Atlantic since the beginning of commercial fishing [20,21] and in Iceland since the early settlement [22,23]. The quantity and temporal resolution of this material allow for both empirical estimates of population history and structure and improved model estimation of demographic variables. Estimation of
population size based on genetic modelling has improved rapidly in recent years [16,24] and the serial coalescent framework [25] has been successfully applied to study historical fluctuations [26–28].

Since the commencement of fisheries surveys in the twentieth century, the Icelandic cod stock has gone through a severe decline [29]. Historical records cite periods of low Atlantic cod fishing as far back as the fourteenth and fifteenth centuries in Iceland [30]. Intense fishing of Atlantic cod in Icelandic waters started in the thirteenth century as the European market for dried fish boomed, taking English and Dutch fishermen to new fishing grounds, particularly in Iceland and Newfoundland [31–33]. Historical fisheries may have impacted the Atlantic cod populations [34], including the Icelandic stock [35]. Commercial cod fishing in Icelandic waters collapsed in the seventeenth century, a period of both political [36] and climatic [37] upheaval. A cold period, ‘the little ice age’, followed ‘the medieval warm period’ in the sixteenth century and continued until the mid-nineteenth century. Recent evidence suggests an earlier drop of 2–3 °C in summer temperatures in the thirteenth century and subsequent decrease in sea temperatures in the fourteenth and fifteenth centuries [38–40]. Climate change and the concomitant ecosystem effects may have determined Atlantic cod catches both in historical [21] and modern times [41]; for example, temperatures warmer than at present are expected to be favourable for cod recruitment around Iceland [42].

Inference of historical population dynamics requires adequate sample sizes from known populations and time periods [43]. In this study, we used Atlantic cod vertebrae from historical fishing sites in Western Iceland, dating from 1500 to 1910, to examine genetic variation and demography in the Atlantic cod population and compare it to contemporary variation. Our first objective was to examine the effects of industrial fisheries on the Icelandic cod stock by comparing $N_E$ genetic variation and age composition in historical times and after the onset of modern fisheries. The effects of intense fisheries on fish stocks may include truncation in the mean age of fished stock, changes in population structure and the distribution of genetic variation [44,45]. Our second objective was to evaluate whether historical reports of abundant cod catches in the medieval warm period coincided with high estimates of $N_E$ and, similarly, if the early modern collapse of commercial fishing in Iceland and documented historical temperature fluctuations corresponded to changes in $N_E$, diversity or demography. First, we examined empirical variation at two genetic markers that have been extensively studied in modern Atlantic cod, the mitochondrial marker cytochrome $b$ (cytb) and the Pantophysin I (Pan-I) gene. Second, we modelled effective population size and fluctuation based on cytb variation. Finally, we estimated the age distribution of the fished stock within the study period.

2. Material and methods

(a) Historical sample selection, excavation and sample dating

Atlantic cod vertebrae were sampled from four historical fishing sites in Western Iceland; Breiðavík northwest Iceland (24°24′45.98″ W, 65°32′38.13″ N), Finnbogastaðir northwest Iceland (21°19′37.73″ W, 66°01′39″ N), Gufuskálar western Iceland (23°55′44.92″ W, 64°54′36.93″ N) and Gufuskálar in southwest Iceland (22°36′41.57″ W, 64°3′18.56″ N) (figure 1).

Stratigraphical layers were initially aged by context and finds, and ultimately by $^{14}$C dating (Scottish Universities Environment Research Centre). Two stratigraphical layers contained no material suitable for $^{14}$C dating, and dates were thus based on tephra deposits and context (see electronic supplementary material, table S1). All vertebrae from the excavations were aged by counting the growth rings with a dissection microscope. Temporal changes in age were tested with linear regression. Before calculation of descriptive statistics and haplotype networks, the samples were grouped into five temporal groups to minimize temporal overlap and to increase sample size for each time interval (see electronic supplementary material, table S1).

(b) DNA extraction, amplification and sequencing

DNA was extracted at Lakehead University Paleo-DNA Laboratory (Thunder Bay, Canada). A 340 base pair fragment of the mitochondrial cytb gene, with known variable sites in Atlantic cod [11], was targeted for PCR amplification with primers L14406 (5′-TTGGCTCTCTTCTTAGGCCCTTT-3′) and H14746 (5′-GCTCTC AGAATGATATTGTGCC-3′). A 210 base pair fragment of the Pantophysin I gene was amplified with primers (fwd 5′-TTGGCTCTCTTCTTAG GCCCTTT-3′) and (rev 5′-AAAATGCA CACCGACA GATTAAGC-3′). PCR of 2–20 ng per reaction was conducted including 2 mM MgCl$_2$, 3 mg bovine serum albumin (BSA), 1.5 μM each primer, 1.5 mM dNTP, 1 U Taq DNA polymerase and a thermostable buffer. Cycling conditions for both genes

Figure 1. Map of Western Iceland showing the excavation sites. Letters denote abbreviations for site names: GS, Gufuskálar in southwest Iceland; GW, Gufuskálar in Western Iceland; B, Breiðavík; F, Finnbogastaðir. Further information on sites and samples are found in electronic supplementary material, table S1.
were: 92°C, 4'; (92°C, 20'; 60°C/55°C, 1 Δ0.25°C, 45'; 70°C, 45') × 20; (92°C, 20'; 55°C, 45'; 70°C, 45') × 25; 12°C. PCR products were cleaned using Exo-SAP-IT and sequenced using LI4406 (cytb) and rev (Paml) on an ABI3500xl (Applied Biosystems). All PCR reactions were performed in a room where no previous work had been done on Atlantic cod DNA. A subset of individuals (n = 16) were re-sequenced for the cytb, using both PCR primers (see electronic supplementary material, table S2). The cytb sequences were easily aligned to known modern sequences covering positions 14494–14768 [46]. Previously published information on the modern Icelandic cod stock was used in all analyses (see electronic supplementary material, table S3).

(c) Population analyses
A haplotype network was generated using Pegas [47] in R v. 2.9.2 [48]. DnaSP v. 5.10 [49] was used to generate mismatch distributions and estimate haplotype (Hd) and nucleotide (π) diversities, Tajima’s D and DnaSP and pairwise FST estimates with 95% confidence intervals from 10 000 bootstrap replicates. Haplotypic richness (hr) [50] was calculated using the hierfstat package in R [51]. The linear relationship of haplotype richness with time was tested by comparing the slope of the observed data with slopes of regression lines obtained by permuting haplotypes across the temporal intervals. The permutation was conducted in R [48], using 1000 permutations.

(d) Coalescent simulations and demographic analyses
The Bayesian Skyride Plot (BSP) [52], implemented in BEAST v. 1.6.1 [53], was used to evaluate effective population size changes using time-stamped cytb sequences. The molecular substitution model (HKY) was selected using the program JModelTest v. 0.11 based on the Akaike Information Criterion (AIC) [54]. A range of mutation rates (1.8–2.2% Myr) was employed with a generation time of 5 years, based on previous research [11], while allowing for a level of uncertainty. Sample dates were rounded to the nearest 50 years BP. We used two replicate runs with 50 × 106 iterations after a burn-in of 5 × 106 iterations. Genealogical likelihood and model parameters were sampled every 5000th iteration. Mixing and convergence were assessed in TRACER v. 1.3, using a minimum effective sample size of 100 from combined runs for parameter estimates. We also modelled a constant size coalescent tree and calculated Bayes factor to compare the support for the BSP with the constant size model [55].

The genetic effect of three effective population size scenarios was modelled in a coalescence framework using approximate Bayesian computation (ABC) [56]: (i) constant population size, (ii) a single bottleneck between 0 and 600 years BP and (iii) two bottlenecks, 0–300 years BP and 300–600 years BP, as implemented in diy-ABC v. 1.0.4.37 [57]. We estimated N0 prior to any bottleneck (N0), after the first bottleneck (N1) and after the second bottleneck (N2), as well as the times of the first (t1) and second (t2) bottlenecks. One million simulations were obtained for each scenario and the number of haplotypes, pairwise number of haplotypes, number of segregating sites and pairwise FST values were recorded for each simulation. The different scenarios were compared by estimating the fit of 500 simulated datasets closest to the observed data (direct approach) and by logistic regression on the posterior probability of each scenario as predicted by the deviations between simulated and observed summary statistics, using 1% of the best-fitting datasets. Confidence in scenario choice was evaluated using the method of Cornuet et al. [58], using 500 simulated datasets under each scenario. Type I error was estimated as the proportion of instances when the supported true scenario had lower posterior probability than the alternative and type II error as the proportion of instances where the supported scenario was incorrectly chosen. The posterior distributions of population size and the timing of bottlenecks were estimated under the supported model using a local logistic regression on the 1% closest of 1.5 × 105 simulated datasets. To evaluate the performance of the estimation, we generated pseudo-observed datasets with known parameter values drawn from the posterior distribution of the supported scenario. The mean relative bias (MRB) for each parameter was estimated and averaged over the 500 datasets.

Table 1. The age composition of Atlantic cod vertebrae across the five temporal groups, from 1500 to 1910. Mean age and standard deviation (s.d.) are in years and n presents sample size.

<table>
<thead>
<tr>
<th>Year</th>
<th>n</th>
<th>Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1500</td>
<td>30</td>
<td>11.6</td>
<td>1.3</td>
</tr>
<tr>
<td>1600</td>
<td>50</td>
<td>12.8</td>
<td>1.5</td>
</tr>
<tr>
<td>1700</td>
<td>70</td>
<td>12.4</td>
<td>1.5</td>
</tr>
<tr>
<td>1800</td>
<td>50</td>
<td>12.9</td>
<td>1.5</td>
</tr>
<tr>
<td>1900</td>
<td>20</td>
<td>12.5</td>
<td>1.5</td>
</tr>
</tbody>
</table>

3. Results

(a) Age composition and genetic diversity
The mean age of the 1500–1550 sample was 13.14 (s.d. = 1.75) and 10.61 (s.d. = 2.11) years in the sample dated from 1910 (table 1 and figure 2). There was a significant decline of mean age over time (b = −0.458 (per 100 years), p < 0.001).

We successfully sequenced 258 bp of the 340 bp cytb target sequence of 156 historical samples (table 2). Ten haplotypes were observed from the historical samples. Two haplotypes had not been described in Atlantic cod from Iceland; a single individual with haplotype DO in the eighteenth century sample and a novel haplotype identified in three individuals from the 1500 to 1550 sample (GB: KP834568). The novel haplotype differed by a single base pair transition at site 14547 [46] from haplotype G. The G-clade [11] was more common in the sixteenth and seventeenth century samples than in more recent samples (figure 3). Haplotypic richness was highest in the earliest sample and decreased over time (p < 0.001, table 2). Pairwise FST values ranged from 0 to 0.053 and were significant among samples dated from before and after the sixteenth century (table 3). Population contraction/expansion was indicated by the mismatch distributions, except within the eighteenth century samples (see electronic supplementary material, figure S1). Tajima’s D values ranged from −0.8 to 0.6 and were not significant. Allele frequency at the Paml locus was determined for 55 historical specimens. The frequency of the Pan I B allele was low in all samples (table 2).

(b) Estimation of historical population size
The BSP suggested large changes in effective population size in historical times, most notably a rapidly growing medieval population with a severe bottleneck at around 1440 (figure 4). The N0 estimate before the bottleneck was high at 440 000 individuals (95% CI 191 000–1 037 000). Subsequent population size fluctuations were smaller and N0 estimates in modern times were low at 1955 individuals (95% CI 402–7550).
A Bayes factor of 271.7 supported the BSP model over a model of constant size. The estimated substitution rate was 5.12 x 10^{-2} (95% highest posterior density interval: 1.05 x 10^{-2} – 2.83 x 10^{-2}) substitutions per site per year.

The ABC estimation supported the scenario of two bottlenecks over the scenarios of either a single bottleneck or a constant population size with both the direct estimate and logistic regression (see electronic supplementary material, figure S2). The scenarios with bottleneck effects had high statistical support compared with constant population size, but the model with two bottlenecks was weakly supported. Estimated type I and type II error rates for the choice of a two bottleneck scenario compared with the constant size scenario were 44% and 46%, respectively, but much lower for single bottleneck scenario or 0.05% for both type I and type II errors. The median estimate of effective population size before the first bottleneck was 330 000 (95% quantile: 89 000–852 000) decreasing to 75 000 (95% quantile: 44 800–98 000) at the earlier bottleneck and to 21 300 (95% quantile: 3900–40 900) at the modern-day bottleneck. The time of the earlier bottleneck was estimated at 96.2 (95% quantile: 72.5–118.3) generations before the present (BP) (or between 1400 and 1630) and the modern bottleneck at 23.8 (95% quantile: 1.9–40.0) generations BP (or in the past 200 years), both dating from the samples collected at 1990. The bias (MRB) was low for all estimates of population size, N_s (1.002), N_1 (0.411) and N_2 (1.005), and for the bottleneck time estimates, t_1 (0.002) and t_2 (0.988).

### 4. Discussion
Our results indicate a large, growing Atlantic cod population during the medieval warm period and a marked population size bottleneck in the fifteenth or early sixteenth century.
century with concomitant loss of haplotype variation. Fish age declines over the study period, with a marked drop in the mid-seventeenth century.

The historical population size estimates are extreme compared with modern-day values but may be interpreted in the context of historical documents that claim vast numbers of cod in the North Atlantic at the height of the European fisheries expansion to Iceland and Newfoundland [31,32]. The scale of the early bottleneck observed in this study and the concomitant loss of genetic variation may be consistent with large-scale ecological change. Summer temperatures in Icelandic waters decreased by 2°C in the late fourteenth and early fifteenth centuries as minimal summer temperatures reached 4°C [38]. Written sources from the same period describe icebergs [37], unusually cold weather and low fish catches [30]. This period, the onset of 'the little ice age', immediately precedes the large-scale reduction in NE indicated by the current data and we conclude that the two may be linked, although without any assumptions of causality. Previous research on Atlantic cod growth using otoliths from archaeological excavations in Norway showed disruptions of growth patterns in the early sixteenth century, best explained by a significant drop in sea temperatures [60].

Although adult Atlantic cod are robust to a range of temperatures [61], early life stages are dependent on sea currents and the favourable timing of algae blooms and zooplankton availability [62,63], where the latter may be affected by even moderate changes in sea temperature [64–66]. The North Atlantic climate oscillations in medieval times provide the best current proxy for historical environmental change [67] and may have prompted unknown ecosystem changes, e.g. climatic events have been linked to severe changes in ocean circulation and currents [68,69].

Extensive fishing of the Icelandic cod stock started as early as the thirteenth century [23], and previous studies have suggested that historical fisheries affected the Atlantic cod stock [21,34,70]. Although the total catches of historical fishing fleets are difficult to estimate, historical documents refer to multiple fleets of up to 500 English and Dutch longliners fishing simultaneously in Icelandic waters in the sixteenth century [71]. Modern-day research finds that early signs of intensive fisheries included changes in the age composition of stock [44]. There was a truncation in mean age in our samples that started in the seventeenth century and continued into modern times (figure 2). Although these correlations may not imply causation, they correspond to

![Figure 3. Median joining haplotype network of all haplotypes in each temporal group and the common haplotypes in the modern sample (1990). The haplotypes are represented by circles, the area being proportional to their frequencies. Each trace corresponds to a single mutation. Letters identify different haplotypes (electronic supplementary material, table S3).](http://rspb.royalsocietypublishing.org/)

### Table 3. Differentiation among temporal samples estimated by pairwise $F_{ST}$ values based on haplotypic variation in cytb sequences. Asterisks denote a significant $p$-value, based on 10,000 permutations.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1910</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1700–1800</td>
<td>0.00</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1650–1700</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1600–1650</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.017</td>
<td></td>
</tr>
<tr>
<td>1500–1550</td>
<td>0.021*</td>
<td>0.01</td>
<td>0.053*</td>
<td>0.024*</td>
<td>0.015</td>
</tr>
</tbody>
</table>

* $p < 0.05$
modern studies on fisheries effects [44]. However, our samples represent the age of the fished stock rather than the population demography and the lack of younger individuals in our earliest samples does suggest that the early fisheries targeted older, larger fish (figure 2). As the decline in age occurred later than the early bottleneck, our data do not particularly support the hypothesis that historical fishing pressure influenced the population collapse.

Fisheries pose strong predation pressure on fish populations and may cause evolutionary change [72]. Two recent studies indicate fisheries effects on the relative frequencies of inshore (Pan1^A^) and offshore (Pan1^B^ and Pan1^M^) genotypes in the Icelandic Atlantic cod population [9,18]. The frequency of the offshore Pan1 B allele in the historical samples is low (table 2) in comparison with samples from modern-day inshore spawning stock [9]. These results could be biased by the development of fisheries and fishing gear throughout the period although, based on the geography and expected range of local fishing boats, it is probably that the current samples represent inshore cod. Thus, we tentatively conclude that the B-allele was at low frequency in the inshore stock during historical times, supporting a reduction in inshore allele frequencies in the twentieth century. Higher frequency of offshore alleles in the inshore stock after the onset of industrial fisheries may reflect changes in the target groups of the fishing industry but alternatively reflect a change in the relative proportions of inshore and offshore populations owing to greater fishing pressure on inshore stock. Given the low number of samples currently analysed, further studies are required to reach a firm conclusion on temporal changes in Pan1 allele frequencies.

Genetic estimates of population size assume an isolated population and are biased by gene flow or population admixture [73], which could effect the current results. The distribution range of Atlantic cod in the North Atlantic is dependent on temperature [13] and the genetic structure can be fluid: for example, some spawning populations in Greenland show genetic continuity over several decades, while others appear to have been replaced by neighbouring stock [74]. The Icelandic cod population is genetically structured by geography and depth [6], although the structure is weak or absent for neutral markers, and not detected by mtDNA markers [11]. Cod feeding migrations [75] are independent of genetic structure but could bias the historical samples. In our earlier samples, we observe a single individual with a haplotype currently found in low frequency in the Faroe Islands. This could be consistent with a historically panmictic population or increased structure at the present day. However, the reduction in haplotype richness, loss of rare haplotypes and concomitant increase in the frequency of intermediate haplotypes after 1550 concur with a population size bottleneck [76]. Distinguishing among these factors can only be speculative based on the current data but highlights the potential of archaeological material for understanding historical ecology.

The current results have both theoretical and applied implications irrespective of the cause of the population size changes. Genetic patterns of low nucleotide diversity and singleton polymorphism, characteristic of the cytb of Atlantic cod in modern times, indicate low N_e and low N_e/E ratios [11,77] that may have resulted from the severe bottleneck and loss of haplotypic variation in the sixteenth century. This may result in shallow genealogies with multiple merging coalescent events complicating reconstructions of population evolutionary history [78,79]. Moreover, the large discrepancy between current and historical N_e and the possible correlation with sea temperature raise questions about the mechanisms underlying fluctuations of N_e/E ratios in Atlantic cod. The N_e/E ratio affects harvesting tolerance, resilience to environmental change and adaptability [16] and has been found to depend on both environmental factors and population size in several other species [80–82]. To conclude, we estimate high historical N_e with a marked bottleneck in the fifteenth or early sixteenth century and continuing low N_e at modern times. Although causality is difficult to assume, we find support for both climate- and fisheries-induced change. The current results identify several lines of future studies, including: (i) the estimation of historical migration/gene flow and/or admixture to investigate the high historical effective population estimates from cytb by including older samples or samples from other regions, (ii) the changes in Pan1 frequency over time from inshore/offshore comparisons and (iii) concurrent shifts in population size in other Atlantic regions inferred from mtDNA or other genetic markers.

Acknowledgements. We thank Tom McGovern, Lilja Bjo¨rk Pa´lsdo´ttir and Frank Feeley for providing some of the samples used.

Data accessibility. Genbank accession numbers and haplotype frequencies can be found in electronic supplementary material, table S3.

Funding statement. This study was funded by grants from the Icelandic Ministry of Fisheries Research Fund to G.A.O., R.E. and S.P.

References


48. R Core Team. 2012


Correction to ‘Historical DNA reveals the demographic history of Atlantic cod (Gadus morhua) in medieval and early modern Iceland’

Guðbjörg Ásta Ölafs dóttir, Kristen M. Westfall, Ragnar Edvardsson and Snæbjörn Palsson


Concerns with the original article [1] were originally brought to light by Prof Einar Arnason and his concerns are outlined here [2]. Moreover, there were errors in the summary statistics presented in table 2. We corrected the errors in summary statistics (table 1) and repeated the analyses, including the rare haplotypes. The results strengthened the conclusions of Ölafs dóttir et al. [1].

1. Population analyses and estimates of effective population size ($N_e$) including rare modern haplotypes

All analyses followed the methods of Ölafs dóttir et al. [1], except that the Bayesian skyride model (BSM) implemented in BEAST [3] required a longer burn-in of $20.0 \times 10^6$. We used the complete haplotype frequencies of the modern Atlantic cod in Icelandic waters, reported in Arnason et al. [4]. The total modern sample was $n = 519$ (previously $n = 503$). As only 228 bp, within a 258 bp sequence, could be reliably used for the archaeological samples, not all modern haplotypes were distinguished. Specifically, haplotypes A, AJ and ZI; E and XI; G and LJ; NI and LI and EI, HI and UI could not be differentiated (table 2).

2. Results and conclusion

We report only results that depend on the modern Atlantic cod haplotypes, analyses relying solely on the archaeological samples have not changed. Including the rare modern haplotypes changed the summary statistic of the modern sample (table 1). The pair-wise $F_{ST}$ value between the new modern sample and the 1500–1550 temporal group was 0.022 ($p < 0.05$), (without rare haplotypes, $F_{ST} = 0.021$, $p < 0.05$), other pair-wise $F_{ST}$ values did not change.

When the rare haplotypes are included the BSM female effective population size (hereafter $N_e$) estimate before the bottleneck was high at 388 000 (95% CI $210 000–738 000$) (previously 440 000, 95% CI $191 000–1 037 000$) and $N_e$ estimate in modern times was low at 2680 (95% CI $658–9560$) (previously 1955, 95% CI $402–7550$; figure 1). The $N_e$ estimates when including the rare haplotypes are well within the error margins of the previous estimates. A Bayes factor of $2^{32.4}$ supported the BSM over a model of constant population size.

Mutation rate was estimated slightly higher than before at $8.22 \times 10^{-7}$ (95% CI $5.56 \times 10^{-7}–2.02 \times 10^{-6}$).

With the ABC estimation [5], logistic regression indicated a single bottleneck as the most likely scenario, but the direct estimate did not distinguish between one or two bottlenecks (two bottlenecks were indicated by both methods when excluding the rare haplotypes although two bottlenecks had negligible statistical support over a single bottleneck). The correct choice of the single bottleneck scenario was supported by a type I and type II errors of 11.2% and 12.0%. The
median estimates of effective population size including the rare haplotypes was considerably higher than when they were excluded; before the bottleneck $N_e = 1,420,000$ (95% quantile: 297,000–1,850,000) (was 330,000, 95% quantile: 89,000–852,000), and after, 160,000 (95% quantile: 46,200–382,000), (was 75,000, 95% quantile: 44,800–98,000 and 21,300, 95% quantile: 3,290–40,900). The time of the bottleneck was estimated at 74.6 generation BP (95% quantile: 40.8–102.2), (was 96.2, 95% quantile: 72.5–118.3 and 23.8, 95% quantile: 1.9–40.0). The bias (MRB) was low for both estimates of population size, $N_s (0.411)$, $N_1 (1.197)$ and the bottleneck time estimates, $t_1 (1.294)$.

All models supported a large-scale population size reduction in the late medieval period, with smaller subsequent changes in $N_e$, in line with the conclusions in Olafsdottir et al. [1]. This was further supported by consistently lower genetic variation in the modern and more recent archaeological samples and pair-wise $F_{ST}$ values that were only significant across the mid-sixteenth century, both with and without the rare modern haplotypes.

It is important to note that, because of different sample sizes and sampling ranges, the archaeological samples may not be directly comparable to the modern-day sample. Including all available sequences may increase the performance of the coalescence methods implemented in BEAST [3], but the difference in sample sizes may affect some of the summary statistics used in the Bayesian estimation of diyABC [5]. This could explain why including the rare haplotypes scaled up the $N_e$ estimates of the ABC analyses. For comparison, when excluding all the modern sequences, the BSM supported a considerably larger medieval $N_e$ of 601,000 individuals (95% quantile: 288,000–1,480,000) with a reduction to 310,000 individuals (95% quantile: 38,000–776,000) and the ABC analysis supported a large medieval $N_e$ of 1,110,000 (95% quantile: 750,000–1,500,000).
To conclude, all models estimate the female effective population size of Icelandic Atlantic cod to have been magnitudes higher before the sixteenth century, whereas population size reductions in modern times were more ambiguous. Importantly, for the conclusions of Ölafsdottr et al. [1], the relative scales of population size changes, $F_{ST}$ values and reduction in genetic variation remained consistent when the rare haplotypes were included.

Acknowledgements. We sincerely thank Professor Einar Árnason for pointing out these concerns and the editorial team of Proc. R. Soc. B for their assistance and patience while preparing this corrigendum.

References


Table 2. Modern haplotype frequencies and GenBank accession numbers. Frequencies of undistinguishable sequences are presented in parentheses.

<table>
<thead>
<tr>
<th>frequency</th>
<th>GB accession number</th>
<th>GB accession number</th>
<th>GB accession number</th>
</tr>
</thead>
<tbody>
<tr>
<td>243</td>
<td>A (238) U55961</td>
<td>AJ (4) AF205607.1</td>
<td>ZI (1) AF205614.1</td>
</tr>
<tr>
<td>80</td>
<td>D U55953</td>
<td></td>
<td></td>
</tr>
<tr>
<td>78</td>
<td>E (75) U55951</td>
<td>XI (3) AF205672.1</td>
<td></td>
</tr>
<tr>
<td>64</td>
<td>G (62) U55952</td>
<td>LI (2) AF020728.1</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>C U55956</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>NI (8) U55954</td>
<td>LI (1) AF20727.1</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>RI AF205609</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>BI AF205612.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>MJ AF205606</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>AI AF205613.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>DJ AF205631.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>DY AF205630.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>EI (1) AF205628.1</td>
<td>HI (1) AF205625.1</td>
<td>UI (1) AF020730.1</td>
</tr>
<tr>
<td>1</td>
<td>EM AF205627.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>HJ AF205624.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>O AF205621.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>OI AF205620.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>DI AF097466.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>H AF097468.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>MC AF205622.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>QI AF205618.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>TI AF205616.1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

quantile: 324,000–1,520,000), a reduction to 72,000 (95% quantile: 14,900–98,600) in the early-sixteenth century and to 51,300 (95% quantile: 10,900–93,600) in the late-eighteenth century (data not shown).