ISBN 978-82-575-1019-0
ISSN 1503-1667

Hedmark University College
Department of Natural Sciences and Technology
NO-2306 Hamar, Norway
www.hihm.no, e-mail: postmottak@hihm.no

NOFIMA AS
Osloveien 1
NO-1430 Ås, Norway
Phone +47 64 97 01 00
www.nofima.no, e-mail: mat@nofima.no

Norwegian University of Life Sciences
NO–1432 Ås, Norway
Phone +47 64 96 50 00
www.umb.no, e-mail: postmottak@umb.no

Ola Frang Wetten

Norwegian University of Life Sciences • Universitetet for miljø- og biovitenskap
Department of Animal and Aquacultural Sciences
Philosophiae Doctor (PhD) Thesis 2011:56

Philosophiae Doctor (PhD) Thesis 2011:56

Molecular studies of Atlantic cod (*Gadus morhua*) globin genes – functional effects of hemoglobin polymorphisms and their geographical distribution

Molekylære studier av globingener hos Atlantisk torsk (*Gadus morhua*) – funksjonelle effekter av hemoglobin-polymorfier og deres geografiske utbredelse

Ola Frang Wetten

Molecular studies of Atlantic cod (*Gadus morhua*) globin genes – functional effects of hemoglobin polymorphisms and their geographical distribution

Molekylære studier av globingener hos Atlantisk torsk (*Gadus morhua*) – funksjonelle effekter av hemoglobin-polymorfier og deres geografiske utbredelse

Philosophiae Doctor (Ph.D.) Thesis

Ola Frang Wetten

Norwegian University of Life Sciences
Department of Animal and Aquacultural Sciences
P.O. Box 5003, NO-1432 Ås

Universitetet for Miljø- og Biovitenskap
Institutt for husdyr- og akvakulturvitenskap
Postboks 5003, NO- 1432 Ås

Hedmark University College
Department of Natural Sciences and Technology
P.O. Box 4010 Bedriftssenteret, N-2306 Hamar

Høgskolen i Hedmark
Institutt for naturvitenskap og teknologi
Postboks 4010 Bedriftssenteret, NO-2306 Hamar

Nofima Marin
P.O. Box 5010, NO-1432 Ås

Nofima Marin
Postboks 5010, NO-1432 Ås

Thesis number 2011:56

ISSN 1503-1667

ISBN 978-82-575-1019-0

Acknowledgements

This thesis was carried out at Hedmark University College (Hamar) in collaboration with the Norwegian University of Life Sciences (Ås) and Nofima Marin (Ås). The promoter analysis described in Paper III was carried out at the The Faculty of Mathematics and Natural Sciences, School of Pharmacy, University of Oslo. The research project was funded by the Norwegian Research Council, and the Ph. D. position by Hedmark University College.

My head supervisor, Øivind Andersen at Nofima Marin, deserves special thanks for taking me in on this fruitful project in the first place, for connecting me to his valuable network of research fellows, for always giving quick and useful responses, and for showing patience to a candidate of many focuses besides science. My always-positive supervisor at Hedmark University College, Robert C. Wilson, deserves great thanks for valuable help and discussions regarding the laboratory experiments, and for invaluable linguistic support.

I also would like to thank Tor Gjøen at the University of Oslo for sharing his knowledge and letting me work in his laboratory using necessary equipment and accessories during the promoter analysis (Paper III).

Thanks to all my colleagues in Hamar for creating a good and positive environment that makes going to work every morning a pleasure, and especially to Wenche, Fride and Trine for useful scientific talks, inspiration, support, and good laughs.

My extended family deserves big thanks for always helping us when my wife and I need you to take care of the children, do the cooking, or for all kinds of work at the farm. Especially I will call attention to my hardworking brother Espen who has rescued the inevitable seasonal farm duties from being neglected many seasons. Good luck with your own Ph. D. project that starts this fall!

My wonderful children Syver and Selma deserve thanks for taking me back to real life and making me happy every time I come home; now I look forward to spending more time with you! Finally, my love Marte, I thank you both for allowing me, and encouraging me to fulfill this thesis, for all love and support, and not least for all your patience with me and my duties.

Hamar, September 2011.

Ola Frang Wetten

Table of contents

Abstract

The phenotypic variation among Atlantic cod hemoglobins and their physiological properties has been of great interest to researchers for several decades, and many cod stocks have been "genotyped" according to the HbI nomenclature established by Knud Sick in 1961, stating that the HbI-1 and HbI-2 variants are related to warm and cold waters, respectively. Paper I unambiguously connects the HbI phenotypes to the two haplotypic variants Met55-Lys62 and Val55-Ala62 of the polymorphic β_1-globin gene. From three dimensional (3D) modeling of tetrameric $\alpha_1\beta_1\alpha_1\beta_1$ hemoglobin, we proposed that the polymorphisms at β_1-55 and β_1-62 are related to variable O_2 affinity and variation in temperature sensitivity, respectively. To haplotype these loci efficiently, a High Resolution Melting (HRM) assay was developed (Paper IV). Multiple variants, including recombinations between the β_1-55 and β_1-62 loci, were identified in trans-Atlantic populations.

The intergenic promoter of the head-to-head organized α_1-β_1 genes were shown to display an indel polymorphism of 73 bp strongly linked to the β_1-55 and β_1-62 polymorphisms (Paper III). Comparison of transcriptional activity revealed that the longer β_1-Val55-Ala62-linked promoter, mostly found in colder waters, generated twice the activity of the short one at temperatures of 15 and 20 °C. The increased activity at elevated temperatures is suggested to be a compensatory mechanism to counterbalance the low O_2 affinity at high temperatures of the β_1-Val55-Ala62 haplotype by increasing the Hb-β_1 concentration in erythrocytes.

Studies of the draft genome sequence of Atlantic cod showed that it harbors two unlinked globin clusters which host β_5-α_1-β_1-α_4 and β_3-β_4-α_2-α_3-β_2, respectively (Paper II). Expression analyses confirmed that several of the genes are mainly expressed during embryo development, putatively serving immunological functions, whilst α_1, α_2, β_1 and β_2 contributed polypeptides constituting the adult hemoglobins. The clusters are flanked by genes conserved in other genome-sequenced teleosts, and some flanking genes are also found in extant urochordate and cephalochordate species, linked or not to globins, leading to speculations about the origin of the ancestral vertebrate globin cluster.

Sammendrag

Fenotypisk variasjon mellom hemoglobin (Hb) varianter hos Atlantisk torsk og deres fysiologiske egenskaper, har vært av stor interesse for mange forskere i flere tiår, og mange torskestammer er "genotypet" i henhold til HbI nomenklaturen som ble etablert av Knud Sick i 1961, og som stadfester at HbI-1 og HbI-2 variantene henholdsvis er assosiert med varmt og kaldt vann. Artikkel I relaterer HbI fenotypene direkte til de to haplotypene Met55Val og Lys62Ala i det polymorfe β_1-globingenet. Basert på tredimensjonal (3D) modellering av hemoglobin $\alpha_1\beta_1\alpha_1\beta_1$ tetramerer foreslo vi at polymorfiene ved β_1-55 og β_1-62 er forbundet med henholdsvis økt O_2 affinitet og redusert temperatursensitivitet. For å kunne genotype disse to loci effektivt, ble det utviklet et "High Resolution Melting" (HRM) assay (Artikkel IV). Multiple varianter, inkludert rekombinanter mellom β_1-55 og β_1-62, ble avdekket i trans-Atlantiske populasjoner.

Det ble identifisert en 73 bp stor indel polymorfi i den intergeniske promoteren til de hode-mot-hode (5'-5') organiserte α_1-β_1 genene, som er sterkt koblet til β_1-55 og β_1-62 polymorfiene (Artikkel III). Sammenligning av transkripsjonsaktiviteten viste at den lange promotoren, assosiert til β_1-Val55-Ala62 allelet og kalde omgivelser, hadde mer enn dobbelt så høy aktivitet som den korte ved temperaturer på 15 and 20 °C. Den økte aktiviteten ved høye temperaturer antas å kunne være mekanisme som utligner og kompenserer for at β_1-Val55-Ala62 allelet har dårlig O_2 affinitet ved høye temperaturer, ved å øke konsentrasjonen av Hb-β_1 i erytrocyttene.

Studier av foreløpig genomsekvens fra Atlantisk torsk viste at genomet har to ikke-koblede globin-clustere som henholdsvis inneholder β_5-α_1-β_1-α_4 og β_3-β_4-α_2-α_3-β_2 (Artikkel II). Ekspresjonsanalyser viste at flere av globinene hovedsakelig uttrykkes under embryogenesen kanskje som et bidrag til immunforsvaret, mens α_1, α_2, β_1 og β_2 genene bidro med globin peptider til hemoglobin hos voksen fisk. Clusterne er flankert av gener som er konservert i andre genomsekvenserte teleoster, og noen flankerende gener er også påvist i nålevende urochordat- og cephalochordatarter, både koblet og ikke koblet til globiner, som legger grunnlag for spekulasjoner omkring dannelsen av det opprinnelige vertebrat globin clusteret.

List of papers

The present thesis is based on the following papers, which will be referred to in the text by their Roman numerals:

I: Andersen, Ø., Wetten, O.F., Rosa, M.C., Andre, C., Alinovi, C.C., Colafranceschi, M., Brix, O., Colosimo, A. (2009). Haemoglobin polymorphisms affect the oxygen binding properties in Atlantic cod populations. *Proc. R. Soc. B.* 276, 833-841.

II: Wetten, O.F., Nederbragt, A.J., Wilson, R.C., Jakobsen, K.S., Edvardsen, R.B., Andersen, Ø. (2010). Genomic organization and gene expression of the multiple globins in Atlantic cod: conservation of globin-flanking genes in chordates infers the origin of the vertebrate globin clusters. *BMC Evol Biol* 10, 315.

III: Star, B., Nederbragt, A.J., Jentoft, S., Grimholt, U., Malmstrøm, M., Gregers, T., Rounge, T., Paulsen, J., Sharma, A., Wetten, O.F., Lanzén, A., Winer, R., Knight, J., Vogel, J., Aken, B., Andersen, Ø., Lagesen, K., Solbakken, M., Tooming-Klunderud, A., Edvardsen, R.B., Tina, K., Espelund, M., Nepal, C., Previti, C., Karlsen, B., Moum, T., Skage, M., Berg, P., Gjøen, T., Kuhl, H., Thorsen, J., Malde, K., Reinhardt, R., Du, L., Johansen, S., Searle, S., Lien, S., Nilsen, F., Jonassen, I., Omholt, S., Stenseth, N.C. (2011). The genome sequence of Atlantic cod reveals a unique immune system. *Nature* 477, 207-210.

IV: Wetten, O.F., Wilson, R.C., Andersen, Ø. Recombinant subtypes of the polymorphic hemoglobin *β1* gene of Atlantic cod are sparsely distributed in trans-Atlantic populations as determined by High Resolution Melting analysis. *Under revision for resubmission to Canadian Journal of Fisheries and Aquatic Sciences*, October 2011.

General introduction

1. Atlantic cod, fishing industry and aquaculture

Atlantic cod (*Gadus morhua)* is a teleost of the family Gadidae inhabiting the North Atlantic Ocean on the east side from the Bay of Biscay to the Barents Sea, and on the west side from North Carolina to the southern part of Baffin Island and western Greenland. Fossil evidence of this species has been dated back to 65 million years before present (BP) (Kriwet and Hecht 2008). The largest registered size of an Atlantic cod is approximately 2 meters in length, and specimens can live up to 20 years (Cohen et al. 1990). The Norwegian commercial fishery harvest of wild Atlantic cod totaled 283 000 tons (live weight) in 2010, indicating a landed value of more than 2,93 billion NOK (http://www.ssb.no/fiskeri/). Aquaculture of this important species is still moderate compared to Atlantic salmon both in Norway and abroad, but reached a volume of 21,000 tons slaughtered fish for food in 2010, indicating a firsthand sales value of approximately 343 million NOK (http://www.ssb.no/fiskeoppdrett/). The aquaculture industry experienced a large increase in production during the years 2004-2008 (from 3000 to 18000 tons), but after this production has only slightly increased. Escape from net pens, high mortality in fry production, early sexual maturation, and diseases are some of the many reasons underlying this tendency, that poses a big challenge for the aquaculture industry in terms of the consequent loss of profits. Reflecting the high value of wild-caught cod, management of the fisheries has always been of great interest both to fishermen and government agencies aiming for a sustainable harvest in waters accessed by many nations, and the fishery industry has, during the last decades, shifted from a practice of unlimited exploitation of fish resources to a regulated industry of quotas and licenses (http://www.regjeringen.no/nb/dep/fkd/tema/fiske_og_fangst.html?id=1277). Decisions

allowing the search for oil and gas reservoirs near spawning locations for Arctic- and coastal cod populations near Lofoten are subject to passionate public and politic outcry and debate, as well as research to establish facts on risk factors and possible consequences (Karlsen et al. 2011). Considerable concerns are also pronounced regarding warnings of further increased sea temperatures as a consequence of global warming (Portner et al. 2001; Loeng 2008; Rogers et al. 2011).

2. Gas-exchange in teleost fish

2.1 Hemoglobin structure

Hemoglobin is found in all vertebrates except the Antarctic icefishes, and is an intracellular molecule of the erythrocytes that, through reversible ligand binding, is responsible for carrying O_2 between respiratory organs and metabolic tissues (Jensen et al. 1998). It is a tetrameric protein consisting of two α-globins and two β-globins arranged as two α-β dimers. Each of the four globin subunits has a heme (e.g. a protoporphyrin ring with a central Fe atom) which is responsible for oxygen binding, and is buried in a hydrophobic pocket (Figure 1). Three of the α-chains in Atlantic cod (*Gadus morhua*) are 143 residues in length (α_1, α_2 and α_4), while α_3 is 141 residues, and all five β-globins are 147 residues (Paper II). This is consistent with hemoglobins of other species, and the approximate molecular weight is 65,000 Da (Jensen et al. 1998). The secondary structure of the globin subunits represents a combination of α-helixes and intermediate non-helical segments. The helices are named from A through H from the N-terminal end, and the segments in between, AB through GH. Segments in the N- and C-terminal ends are non-helical and named NA and HC, respectively (Jensen et al. 1998). Reference to a specific residue can either be related to the total number of residues (e.g. β_1-55), or to the number of residues

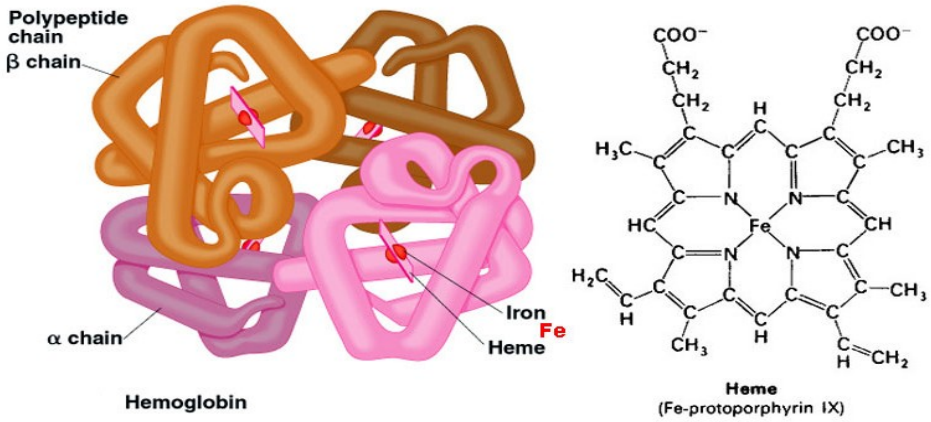

Figure 1: Schematic presentation of general hemoglobin structure. Two α-β dimers are interconnected to a tetramer, each subunit displaying one heme, which is able to carry one O_2 molecule.

in a specific segment (e.g. D6β_1). In all hemoglobins, the heme is attached to the protein by a covalent bond to the proximal histidine (His F8) (Figure 2). The distal histidine (His E7), which appears on the ligand side of the heme, close to Val E11, is also highly conserved. An important task for Val E11 is to sterically hinder endogenously produced CO from accessing the heme, instead favoring O_2 binding, because free heme originally has a far stronger affinity to CO than O_2 (Perutz 1990).

Hemoglobin has two alternative characteristic conformations. The tense (T) state is associated with deoxy-hemoglobin. This is the more thermodynamically stable structure because it harbors salt bridges and other non-covalent bonds between the two α-β dimers. These bonds are loosened as oxygen binds to hemoglobin, and simultaneously hemoglobin will shift to the relaxed (R) state, recognized as high-affinity oxy-hemoglobin (Perutz 1970). During the transition between T and R states, the dimers rotate by 15° respective to each other, and the transition occurs when at least one of the hemes in each dimer is oxygenated (Jensen et al. 1998).

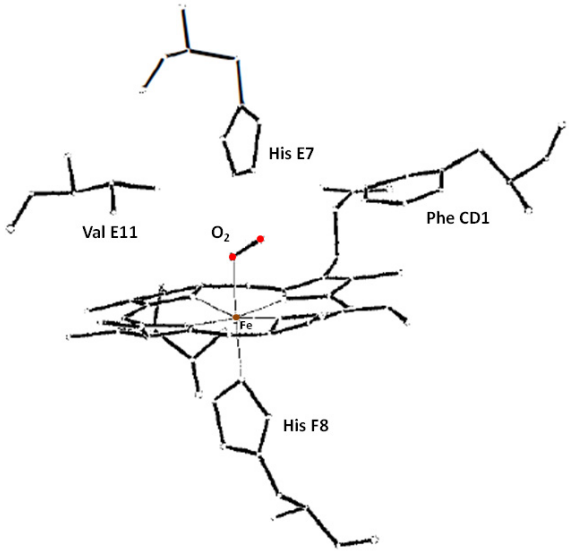

Figure 2: Heme and some of its surrounding well-conserved residues contributing to its anchoring (His F8), hydrophobic conditions (Phe CD1, Val E11) and O_2 binding (His E7). Modified from Perutz (1990).

2.2 Oxygen transport

Deoxygenated blood flows from the heart through the gill lamellas where it is loaded with oxygen, and further proceeds to metabolic tissues before it returns to the heart. Oxygen has to diffuse down its partial pressure gradient, moving from the surrounding water to the capillaries of the gills. Further, O_2 must diffuse into the erythrocytes and finally bind to hemoglobin. There are many variables determining how much oxygen that will be transported from the water flowing past the gill lamellas to hemoglobin. The concentration of erythrocytes in the blood and their hemoglobin content are of importance in addition to oxygen affinity of the hemoglobins and the partial pressure of oxygen in the water. As the oxygenated blood flow into the capillaries of oxygen-demanding tissues, the partial pressure of oxygen is decreased and oxygen diffuses from hemoglobin into the capillaries and further to oxygen-consuming sites like mitochondria (Nikinmaa and Salama 1998).

Hemoglobin must be able to operate under highly varying conditions, indeed a challenging task. The demand for oxygen will shift depending on activity level of the fish and water temperature (Gillooly et al. 2001; Seibel and Drazen 2007). In addition, hemoglobin always

has to *bind* O_2 efficiently in the gills, yet *release* O_2 with sufficient partial pressure in tissue capillaries to support further diffusion to target organelles. To obey these requirements, it is of high importance for an organism to be able to modify the O_2 affinity of hemoglobin quickly, and vertebrates are in the advantageous position of being able to do this by cellular mechanisms through the intracellular space of erythrocytes where hemoglobin is located (Tetens and Christensen 1987; Brauner et al. 2002). Invertebrates with intercellular pigments are prevented from such regulation.

2.3 Mechanisms for regulation of oxygen transport

The fractional oxygen saturation of hemoglobin is dependent on the partial pressure of O_2 surrounding it. Thus a fish can compensate for a decrease in O_2 saturation of hemoglobin by increased breathing or hyperventilation (Jensen et al. 1983; Jensen 1987). Further hemoglobin saturation depends on hemoglobin concentration in the erythrocyte, natural oxygen affinity of hemoglobin, the affinity of hemoglobin to ligands which interfere with O_2 affinity, and their concentration in the erythrocyte. The most important of such ligands in fish are protons and organic phosphates (mainly adenosine triphosphate (ATP) and guanosine triphosphate (GTP)) (Jensen et al. 1998). Both these ligands preferentially bind to the deoxygenated T-state of low oxygen affinity and stabilize it (Weber et al. 1988; Jensen et al. 1998). Thus an increase in organic phosphate concentration or increased proton concentration (decreased pH), will result in decreased O_2 affinity and support unloading of O_2 in capillaries where pH is decreased by metabolic CO_2. Acidification of blood is also utilized in combination with the Root effect (section 2.4) in teleosts upon unloading of O_2 to the retina and swim bladder (Fairbanks et al. 1969; Pelster and Scheid 1992).

2.4 The Root and Bohr effects

Metabolic activities in tissues produce either CO_2 or lactate both of which contribute to decrease pH levels in blood. The Root and Bohr effects both respond to such increased concentration of protons. The Root effect is named after the scientist R. W. Root (1931) who is recognized as the first to characterize it. From blood samples from toadfish (*Opsanus tau*), sea robin (*Prionotus carolinus*) and common mackerel (*Scomber scombrus*) exposed to various pressures of CO_2, he realized that shapes of oxygen dissociation curves appeared different from the well-known Bohr effects previously registered in mammals (Bohr et al. 1904; Root 1931). The Bohr effect is simply a negative shift in hemoglobin's affinity to oxygen induced by decreased pH in the alkali-to-neutral range (pH > 7), which, in theory, can be counterbalanced by increased oxygen tension (Bohr et al. 1904) (Figure 3). In the fish species examined by Root, he found a stronger and modulated variant of the Bohr effect in which the maximum oxygen carrying capacity of hemoglobin was reduced upon acid or CO_2 exposure, in addition to the decreased oxygen affinity (Root 1931; Root and Irving 1941) (Figure 3). In practical terms, the difference between the Root and Bohr effects is that Root-Hbs are able to unload oxygen at higher oxygen tensions (Root 1931; Pelster and Randall 1998). Much effort has been made by many researchers to show what the exact mechanism of the Root effect is, and many species have been evaluated according to Root effect, but still many questions remain unanswered (Brittain 2005). There are, however, several globin residues that have been confirmed to contribute to the Bohr and Root effects, including Asp95α, Asp99β, Asp101β, surface histidines in general and especially the C-terminal His of β-globins (Pelster and Randall 1998; Yokoyama et al. 2004; Berenbrink 2006; Mazzarella et al. 2006). Atlantic cod is among the species with the strongest reported Root effect, up to ~ 80 % decrease in oxygen saturation with strong pH reduction (Berenbrink et al. 2011).

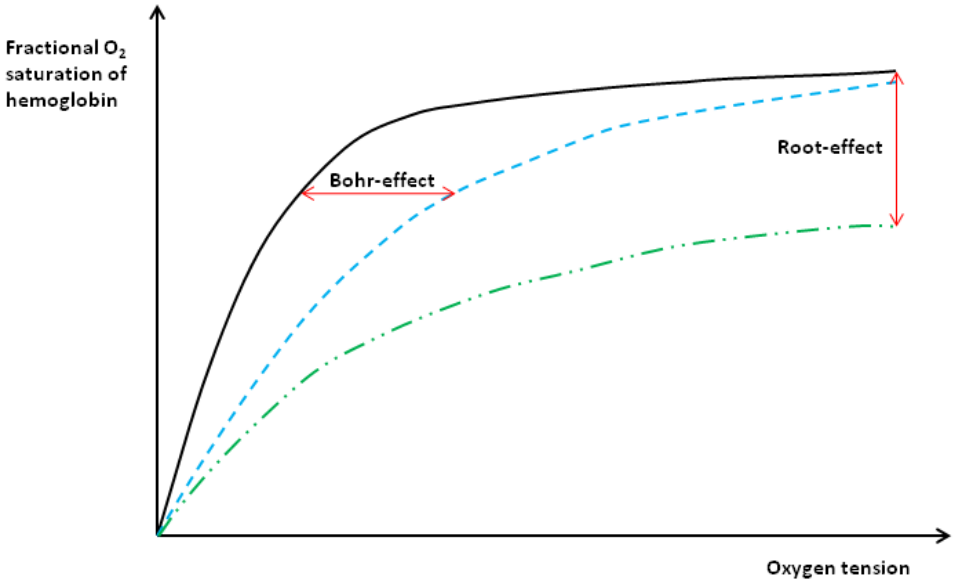

Figure 3: Schematic oxygen dissociation curves illustrating Bohr and Root effects. The Bohr effect reduces the oxygen affinity of hemoglobin upon decreased pH in the alkali-to-neutral range, shifting the oxygen dissociation curve to the right. The Root effect, in addition, decreases the oxygen carrying capacity of hemoglobin as pH decreases, shifting the dissociation both to the right and down.

The ability of Root-Hbs to unload oxygen at high oxygen tensions is of great importance regarding the supply of oxygen to the swim bladder for maintenance of demanded buoyancy. This involves secretion against high PO_2 in the swim bladder, and is assured through the Root-Hbs as protons are released in the erythrocytes to force oxygen unloading and thereby building a sufficiently high concentration of dissolved oxygen to allow its diffusion into the swim bladder. In addition, a peculiar organization of capillaries called *rete mirabile* (wonderful network) comes into play, where oxygen unloaded from Hb in veins is able to diffuse back to the arteries supplying oxygen to the swim bladder (Steen 1963; Kobayashi et al. 1990; Pelster and Randall 1998). A similar *rete* is found in the ocular connection of most fish species, called the choroid *rete*. This, again together with

Root-Hbs, contributes to a sufficient supply of oxygen to the retina, which is not vascularized in most teleosts, resulting in long diffusion distances (Wittenberg and Wittenberg 1962; Waser and Heisler 2005; Berenbrink 2007). Fish with Root-Hbs have also developed a reduced buffer capacity of their Hbs. This allows the intracellular pH of erythrocytes to be adjusted more rapidly, supporting the continuous shifts between loading and unloading of oxgygen (Jensen 1989). Finally, fish have developed a specialized capacity to regulate the proton level in erythrocytes to take full advantage of the Root effect, namely an isoform of the Na+/H+ exchanger family (βNHE) (Borgese et al. 1992; Nikinmaa 1992). Berenbrink et al. (2005) studied the evolutionary emergence of these physiological phenomena in fish, and concluded that the Root effect was introduced before any *rete mirable* occurred. The choroid *rete mirable* is estimated to have first existed 250 million years ago, while the swim bladder *rete mirable* did not arise until about 110-120 million years later. The reduced Hb-buffer values and increased βNHE activity evolved parallel to the Root effect and the choroid *rete mirable*, respectively (Berenbrink et al. 2005).

2.5 Temperature

The binding of oxygen to hemoglobin is, in general, an exothermic reaction producing heat, which in turn has a negative influence on the partial pressure of oxygen and O_2 saturation of hemoglobin (Barcroft and King 1909) (Figure 4). However, the negative effect of this fact is reduced as ligands like H^+ and organic phosphates that bind to the deoxygenated state are released in endothermic reactions upon oxygenation (Weber and Campbell 2011); additionally, increased temperature is beneficial for oxygen unloading to capillaries. Binding of oxygen is also subject to cooperativity in the sense of that oxygenation of one subunit increases the oxygen affinity of the remaining subunits. This capacity is responsible for the beneficial sigmoid curve, which results from a plot of PO_2 versus fractional Hb

saturation (Figure 4), and enables Hb to achieve low values for p50, which is defined as the PO_2 necessary to reach 50 % O_2 saturation of Hb.

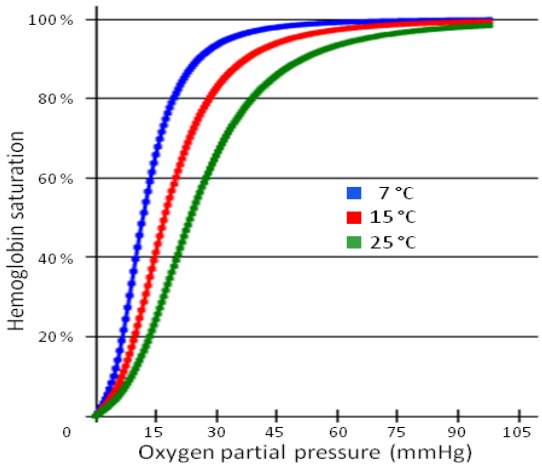

Figure 4: Schematic illustration of cooperativity (sigmoid curves) and temperature effects on human Hb saturation with O_2. Blue, red and green curves represents temperatures and p50 values of 7 °C / 11 mmHg, 15 °C / 18 mmHg and 25 °C / 24 mmHg, respectively.

It has also been observed that the concentration of organic phosphates in the erythrocytes increases with elevated temperatures, which will strengthen the effect of increased temperature on affinity of oxygen to hemoglobin. Anyway, if elevated temperatures cause insufficient oxygen loading of hemoglobin, an opposite reaction (decreasing organic phosphate levels in erythrocytes) can be effectuated (Nikinmaa et al. 1980; Albers et al. 1983; Laursen et al. 1985). Increasing temperature is known to coincide with decreased blood and erythrocyte pH (Heisler 1984), which, through the Bohr effect (section 2.4), also contributes to the reduced oxygen affinity of hemoglobin if pH drops low enough for protons to be accepted by hemoglobin histidines (Nikinmaa and Salama 1998).

3. Atlantic cod HbI polymorphisms

The Danish professor Knud Sick published a pioneering paper on hemoglobin polymorphisms in fish in Nature (1961). He described a pattern of bands from freshly prepared blood samples analyzed by modified agar electrophoresis. From Atlantic cod samples, his results showed an invariable component, HbII, and a possibly Mendelian inherited component HbI which turned out to harbor the isomers HbI-1, HbI-2, and in heterozygous combination HbI-1/2. Frequencies of the HbI-1 isomer seemed to vary in cod populations between Kattegat and the Baltic Sea, and in the following years, these Hb categories were used in studies of a broad scale of Atlantic cod stocks (Frydenberg et al. 1965; Sick 1965a; b; Moller 1966). As shown in Figure 5 reproduced from Petersen and Steffensen (2003), these studies showed that the HbI-2 isomer dominated in stocks in the north-western and north-eastern Atlantic Ocean, and in the Baltic and Barents Seas, while the HbI-1 isomers were found at highest frequencies in the temperate waters of Kattegat and the North Sea, and off the southern part of the Norwegian coast. Several of these studies also detected subtypes with deviate band patterns. Fyhn et al. (1994) first described a subtype named HbI-2b after analyzing samples on IEF gels, and Husebø et al. (2004) reported on the HbI-2b and several other bands in both agar and isoelectric focusing (IEF) gels.

The temperature-dependent geographical distribution of the HbI-1 and HbI-2 isomers tempted several research groups to investigate functional properties of the isomers related to oxygen affinity and temperature dependence. Karpov and Novikov (1980) showed that a HbI-2/2 suspension of erythrocytes had decreasing oxygen affinity in the temperature range from 0 to 20 °C, while as a HbI-1/1 suspension hardly showed any such effect. Brix et al. (1998; 2004) performed *in vitro* experiments with Hb hemolysates which showed that the

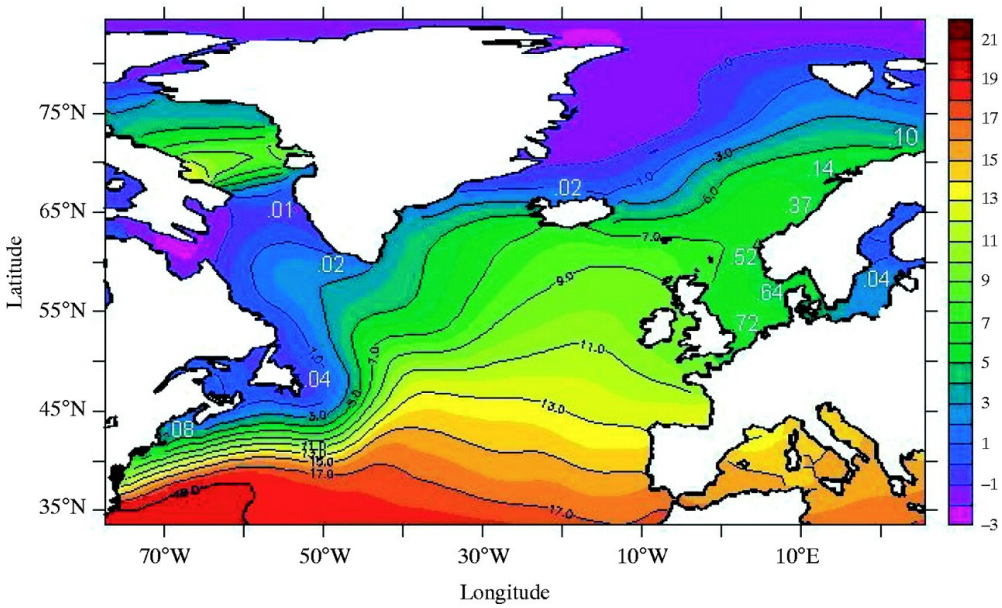

Figure 5: Sea surface temperatures (spring) and the frequency distribution of the HbI-1 allele (white numbers) throughout the range of the Atlantic cod (Frydenberg et al. 1965; Sick 1965a; b). The frequency of the HbI-1 allele is low in cold-water regions and becomes more dominant in warmer areas. A clear cline is found along the Norwegian coast (Petersen and Steffensen, 2003, reproduced with permission from The Journal of Experimental Biology).

HbI-2 isomer displayed higher oxygen affinity at temperatures below 12 °C than the HbI-1 phenotype, while the opposite situation occurred at temperatures between 12 and 20 °C. Petersen and Steffensen (2003) also demonstrated the difference between these phenotypes by showing that juvenile cod of the HbI-1/1 phenotype preferred higher temperatures (15.4 °C) than their HbI-2/2 conspecimens (8.2 °C) when given the opportunity to choose under normoxic conditions. Several studies have tried to verify a connection between the HbI isomers and growth rates, but contrasting results exist. In most cases it is stated that homozygous HbI-2 fish grow faster than HbI-1 fish (Mork et al. 1984; Nævdal et al. 1992; Imsland et al. 2004), while other studies fail to verify Hb dependency on growth (Jørstad and Nævdal 1994; Jordan et al. 2006), and finally Gamperl et al. (2009) found HbI-1 homozygote juveniles to grow faster than heterozygotes (Hb-1/2). In feeding trials

Salvanes and Hart (2000) found that homozygous HbI-2/2 fish captured the greater proportion of food given, and suggested that their behavior upon feeding reflects success in natural hunting for food. Feeding efficiency has also been related to the HbI genotype, and homozygous HbI fish exhibited higher feeding efficiency than heterozygotes while HbI-2/2 had the highest energy retention (Jordan et al. 2006). Johnston et al. (2006) studied relations between the HbI isomers and the development of muscle fibers, and showed that the HbI-2/2 specimens had a greater muscle fiber diameter than HbI-1/1 fish, but that the latter group had 15 % higher maximal number of muscle fibers.

Aims of the thesis

The fundamental goal of this project was to identify the *genetic* basis for the HbI polymorphisms in Atlantic cod, and to use this information to genotype different populations previously phenotyped according to the HbI isomers. The second goal was to propose an explanation of how any identified genetic variation could influence the functional properties of hemoglobin, as expected from knowledge regarding the prevalence of different isomers in different environments. A further aim was to study expression patterns of Atlantic cod globins to elucidate subunit composition of adult and embryonic variants, which necessitated the greater goal of identifying all cod globin genes and their organization in the genome. The search for globin genes also led to the discovery of a polymorphic promoter linked to genetic polymorphisms, which made us adopt the goal of describing the activity levels of this promoter under different temperature regimes. At last, we wanted to establish a PCR-based, effective, accurate and low cost method to genotype high numbers of Atlantic cod DNA samples with respect to the HbI-1/2 polymorphism.

Results and discussion

4. *Identification of the genetic origin of the HbI-1 and HbI-2 isomers (Paper I)*

To elucidate the genetic origin of the HbI isomers in Atlantic cod, we PCR amplified and sequenced globin genes from 35 specimens analyzed for the HbI isomer phenotypes by isoelectric focusing (IEF). An unambiguous relationship between the isomers and two alleles of the *β1*-globin gene was uncovered (Paper I). Namely, methionine at residue 55 and lysine at residue 62 followed the HbI-1 phenotype, whilst valine at residue 55 and alanine at residue 62 accompanied the HbI-2 phenotype. Specimens heterozygous from isomer analyses (HbI-1/2) were also heterozygous at residues 55 and 62 in the β_1-globin gene. It should be noted that Borza et al. (2009) found no relationship between the isomers and the β_1-Leu123Met polymorphism. The β_1-Met55Val polymorphism originated from a single non-synonymous a/g mutation, and the β_1-Lys62Ala resulted from two non-synonymous mutations aa/gc preceded by a synonymous c/t mutation (Paper I**,** Figure 6). Whereas no recombination was found between the neighboring mutations related to substitutions at residue 62, unexpectedly high levels of recombination were found between the codons of the 55 and 62 residues relative to their close physical proximity. Accordingly, the polymorphisms can represent four haplotypes, namely β_1-Met55-Lys62, β_1-Val55-Ala62, β_1-Val55-Lys62 and β_1-Met55-Ala62, which, for convenience, are referred to in this thesis as MK, VA, VK and MA, respectively. There is no evidence presenting how the recombinant VK and MA haplotypes relate to the HbI isomers, but in Paper I we suggest that these recombinants are connected to rare HbI subtypes that have been reported (Fyhn et al. 1994; Brix et al. 2004; Husebo et al. 2004; Imsland et al. 2007).

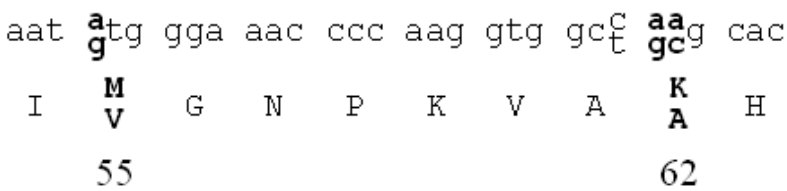

Figure 6: Extract of the β_1-globin gene nucleotide and amino acid sequences, indicating the polymorphic residues 55 and 62, and their respective DNA template. The MK and VA haplotypes corresponds to the HbI-1 and HbI-2 isomers, respectively.

5. HRM-genotyping assay (Paper IV)

The four haplotypes of the loci encoding residues 55 and 62 in β_1 globin (VA, MK, VK, MA) allow four homozygote genotypes (VA/VA, MK/MK, MA/MA, VK/VK) and six heterozygotes (MK/VA, MK/VK, MK/MA, VA/VK, VA/MA, MA/VK). Out of the latter six, two genotypes represent the coupling and repulsive double heterozygotes MK/VA and MA/VK, respectively. The fact that SNP genotyping (Sequenom MASSarray), which was employed for population analysis in Paper I, was not able to distinguish between coupling and repulsive double heterozygotes, encouraged us to develop an alternative method for haplotyping the Atlantic cod β_1 gene (Paper IV). As the High Resolution Melting (HRM) method has been adopted for similar purposes before (Tindall et al. 2009; Vossen et al. 2010), we amplified a 116 bp section of β_1-exon 2 harboring the mutations of interest (Figure 6), and found that melting curves of all the ten genotypes were distinguishable after optimizing the reaction and analysis conditions (Paper IV).

6. Hemoglobin functional characteristics influenced by polymorphisms at the β_1-55 and β_1-62 loci (Paper I)

The geographical distribution of the β_1 globin haplotypes represents a clear correlation with sea water temperature. In general, the correlation is recognized as dominance of the VA haplotype (HbI-2) in cold waters, and dominance of the MK haplotype (HbI-1) in more temperate waters (Paper I and IV), and is in consistence with earlier HbI distribution reports (Frydenberg et al. 1965; Sick 1965a; b; Moller 1966). To understand how variance in the primary structure of the β_1-peptide might explain the underlying genetic adaptation to an environmental factor like temperature, the three dimensional (3-D) structure of an Atlantic cod hemoglobin tetramer ($\alpha_1\alpha_1\beta_1\beta_1$) was modeled using crystallographic structures from four other teleost hemoglobins as templates. This 3D-model uncovered that the distance between residue 55 of β_1 and residue 120 of α_1 at the $\alpha_1\beta_1$ dimer interface increased by 0.77 Å replacing Met with the smaller Val at locus β_1-55 (Paper I). An increased distance at this interface might contribute to destabilization of the low O_2 affinity deoxy T-state, which in turn will favor transition to the more stable high-affinity oxy R-state (Abbasi and Lutfullah 2002; Shikama and Matsuoka 2003). Jessen et al. (1991) demonstrated that the destabilization of hemoglobin dimers resulting from similar mutations gave increased oxygen affinity in human hemoglobin, and distinct mutations at these same loci affecting the crucial dimer interface were uncovered in studies comparing hypoxia-tolerant Andean- and bar-headed geese to graylag geese inhabiting lower altitudes of higher PO_2 (Hiebl et al. 1987; Liang et al. 2001). Accordingly, convergent evolution between fish and birds seems to have exploited the exact same loci to achieve increased O_2 affinity, and is supported by the theory of Perutz (1983) that adaptations in hemoglobin have taken place at a few key positions. Together this implies that the higher oxygen affinity of the HbI-2 isomer (VA haplotype) compared to the HbI-1 (MK haplotype) observed by Brix et al. (1998) is partly due to increased distance between the subunits of the two dimers of the hemoglobin

tetramer. While β_1-Met55 and β_1-Val55 are found at significant frequencies in temperate waters, β_1-Met55 is a rare allele in the cold waters of the Barents Sea and off Greenland, indicating that β_1-Val55 may represent an adaptation to cold waters, rather than β_1-Met55 to temperate waters.

The Lys62Ala substitution of Atlantic cod β1 is located in close proximity to the heme and the highly conserved distal His63β, which regulates ligand access to the heme and affects the energetics of oxygen fixation (Olson et al. 1988; Marechal et al. 2006). We propose that the temperature-insensitive O_2 affinity of the MK variant as compared to VA (Brix et al. 2004) can be explained by a balance between two factors of opposite effects regarding O_2 affinity upon temperature change (Paper I). Subunits harboring β_1-Lys62 display both these factors leaving them insensitive to temperature change, while subunits that display β_1-Ala62 only host one of the factors, namely one that gives increased O_2 affinity with decreasing temperature. Conclusively, we imply that the VA haplotype represents strong oxygen affinity at low temperatures and high temperature sensitivity allowing it to be preferred in stable cold waters, whilst the MK haplotype shows weaker O_2 affinity but more temperature insensitivity allowing it to confer higher fitness in warmer and fluctuating waters, which is in good correspondence with the geographical distribution of both haplotypes (Paper I, Figure 1 and Table I in Paper IV). From this picture, the VK recombinant would represent high O_2 affinity independent of temperature, while the MA recombinant would represent low O_2 affinity that increases with decreasing temperatures.

7. Geographical distribution of β_1-globin genotypes in Atlantic cod populations (Paper I and IV)

In Paper I, Atlantic cod from eight selected populations were genotyped for β_1 single nucleotide polymorphisms using the established high-throughput Sequenom MASSarray platform, and in Paper IV an additional five populations were genotyped by HRM, also including unresolved double heterozygotes from the eight populations of Paper I. The 13 different sampling locations spanned the eastern and western North Atlantic Ocean, Kattegat, the North Sea, the Barents Sea and the Baltic Sea (Figure 1 in Paper IV). The main genotypes comprising more than 90 % of the genotyped fish were the homozygotes MK/MK (HbI-1/1) and VA/VA (HbI-2/2), and the heterozygote MK/VA (HbI-1/2) (Table 1 and Figure 1 in Paper IV). The VA haplotype dominated at most locations, except in the North Sea, Kattegat and the population of Molde, where the MK haplotype was most common. Along the Norwegian coast there was a clear cline from north to south with complete VA dominance in the Arctic cod population sampled at Bjørnøya and Båtsfjord, shifting to MK dominance in the population of Molde (Paper I). Both the broad VA dominance and the north/south cline are in accordance with previous analysis of the HbI isomers at similar Northeast Atlantic locations (Frydenberg et al. 1965; Sick 1965a; b; Moller 1966). Sick (1965a) also reported a week cline on the western side of the North Atlantic between Greenland and Maryland, but this was not apparent in the current study (Paper IV), as the VA haplotype was almost equally dominant in Sisimiut (Greenland) as in George Bank off the Canadian east coast. In any case, it should be noted that the recombinant VK haplotype was not accounted for in the early HbI isoform analysis by Sick and his collaborators. Additionally, they registered subtypes deviating from the general HbI variants that might have been encoded by genotypes possessing the VK haplotype. The VK haplotype was mostly found in cold waters, and was identified at allele frequencies up to 0.14 in the Canadian populations and the population of Molde, almost exclusively in terms

of VA/VK heterozygotes (Paper IV). Presence of the VK haplotype was higher in the Canadian populations than the Greenlandic, which might imply that it has an advantage over VA to cope with summer temperatures off the Canadian coast. Some rare specimens possessing this haplotype were also identified at Bjørnøya (Paper I), maybe contradicting the reported lack of a hemoglobin variant with double electrophoretic polypeptide bands in the Northeast Arctic population (Fyhn et al. 1994). The other recombinant MA, was hardly registered (Paper I and IV), probably due to low fitness in all populations. It is worth noting that β_2 globin, which, together with β_1, comprises the β globins in adult Hb tetramers (Verde et al. 2006), display MA at the same loci as in β_1 (Paper II). This strongly indicates proper functionality of β_1 MA, but also allows the speculation that β_1-MA cod might display less functional variability in their tetramer Hb when the β globin components β_1 and β_2 are identical at crucial loci.

According to the map of spring sea surface temperatures by Petersen and Steffensen (2003), reprinted in Figure 5, temperatures are similar in the Barents and Baltic Seas, and along the coast of Canada north to Greenland, supporting the abundance of the VA haplotype in these waters. In the more temperate waters of the North Sea, Kattegat and southern Norwegian coast, both the MK/MK homozygote and the coupled double heterozygote MK/VA were found at high frequencies (Paper I), indicating a physiological advantage of the MK haplotype in these waters, such as temperature insensitivity. Repulsion heterozygotes (MA/VK) were not detected in any populations, as expected from the very low frequency of the MA haplotype (Papers I and IV). A distinct partition in genotype frequencies was found between Kattegat and the Baltic Sea, where the VA haplotype exclusively dominated in the Baltic and MK was the most abundant haplotype of Kattegat (Paper I). A plausible theory to explain the total dominance of VA cod in the Baltic Sea is that they are remnants from the Arctic stock of the Barents Sea separated some 8000 to 4000 million years ago as

suggested by Sick (1965b). This theory is not supported by microsatellite studies that reveal closer relationships between cod of the Baltic and Kattegat than between cod of the Baltic and Barents Seas (Nielsen et al. 2001). Additionally, the temperature map (Figure 5) indicates low spring sea temperatures in the Baltic, and the fact that Baltic cod has to seek stagnant water of low oxygen content near the bottom to achieve correct salinity (Tomkiewicz et al. 1998) serves to justify the very high dominance of VA cod in the Baltic Sea. The rare MK/VK genotype was only identified in Kattegat and Helgoland (Paper I), in agreement with that a specific subtype has been identified in Danish waters (Husebo et al. 2004).

8. Globin gene clusters and expression patterns (Paper II)

The search for globin genes responsible for the phenotypic HbI isomers initiated a demand for detailed knowledge of the genomic content and composition of possible globin clusters in Atlantic cod. The Atlantic cod genome sequencing project (Paper III) revealed two unlinked clusters harboring altogether nine hemoglobin genes, namely *β5-α1-β1-α4* and *β3-β4-α2-α3-β2,* referred to as the MC (*MPG, c16orf35)* and LA (*LCMT, ARHGAP17)* loci, respectively (Paper II). Leftwards of the MC locus we identified seven conserved genes recognized as *c16orf33, POLR3K, Mgrn1, AANAT, RHBDF1, MPG* and *c16orf35,* and the three genes *ANKRD25, DOCK6* and *HuC* were identified rightwards of the cluster. The same flanking genes were found in the same transcriptional direction in the orthologous loci of zebrafish, medaka, stickleback and pufferfish. The LA locus was also associated with conserved flanking genes, on the leftward side duplicated *AQP8*, *LCMT* and *ARHGAP17,* on the rightward side *Foxj1* and a *RHBDF1*-like gene. The latter gene was coupled to the MC locus in stickleback and zebrafish, and an *ARHGAP17* duplicate was coupled to the MC locus in pufferfish, zebrafish and medaka. There is a wide range of variance between

species regarding the number of globins in each cluster, and recently, the genomic organization of Atlantic salmon globin genes was published revealing the highest numbers known for teleosts, 20 β and 16 α globins separated on chromosome 3 and 6 (Quinn et al. 2010). Teleosts, and salmonids in particular, are expected to harbor increased numbers of paralog globin genes because of the teleost- and salmonid-specific genome duplications. The high degree of similarity between globin clusters in different species indicates a common origin. If an original α-β globin cluster was duplicated to form the MC and LA clusters one would expect to find paralogous genes in the flanking region of both clusters. Such paralogs were, by *in silico* analysis, found linked to each of the globin loci for the *RHBDF1, ARHGAP17, Mgrn1, AQP8* and *FoxJ1* genes in several species, supporting a duplication theory (Paper II). From screening of the genomes of the tunicate *Ciona intestinalis* and the lancelet *Branchiostoma floridae* (amphioxus), we proposed that the man and chicken α-globin cluster originated from a fusion of chordate linkage group (CLG) 3, 15 and 17. The basis for this proposal is our finding of the flanking genes *RHBDF1, MPG, ARHGAP17* and *LCMT* localized to scaffold 17 of the amphioxus genome, which again has been localized to CLG 15 by FISH analysis (Putnam et al. 2008). The same flanking genes were localized to the α-globin containing chromosome 16 of man, which show syntenic association to both CLG 3, 5 and 17 (Putnam et al. 2008). No globin genes were found in scaffold 17 of amphioxus covering most of CLG 15, and we hypothesized that CLG 3 and 17 harbored globin genes. However, Ebner et al. (2010) later scanned these CLGs and found no globin genes, leaving our proposal α-globin cluster origin theory unsupported. Instead they found two amphioxus globin genes linked to *integrin-linked kinase (ILK),* which is a conserved flanking gene of the β-globin cluster in man, chicken and marsupials (Patel et al. 2008; Ebner et al. 2010). The β-globins in man are believed to be paralogs resulting from a more recent duplication of and/or transposition from the α-globin cluster to a region of olfactory receptor genes (Hardison 2008; Patel et al. 2008). On the other hand,

four amphioxus globins were identified in a scaffold corresponding to the region of β-globins in man, while olfactory receptor genes are dispersed in amphioxus (Churcher and Taylor 2009; Ebner et al. 2010). Deuterostome genome projects can contribute to find the globin loci of origin to vertebrate globins (Ebner et al. 2010).

Verde et al. (2006) reported three tetrameric variants of hemoglobin in adult Atlantic cod, namely Hb1, Hb2 and Hb3, consisting of $\alpha_1\alpha_1\beta_1\beta_1$, $\alpha_2\alpha_2\beta_2\beta_2$, $\alpha_1\alpha_1\beta_2\beta_2$, respectively. Expression patterns of the nine reported globins determined in Paper II are in agreement with this, confirming α_1, α_2, β_1 and β_2 as the sole highly expressed globins in adults. β_5 is expressed in early embryonic stages, while α_4 is most abundant around the hatching stage, both confirming phylogenetic studies grouping them with embryonic globins (Borza et al. 2009). Adult globins were detected in unfertilized eggs, and maternally inherited globins might serve the pelagic eggs immunological protection (Ullal et al. 2008). Globin expression during embryogenesis is in general low, consistent with Hall et al. (2004) who only found a transparent hemolymph circulating after the heart starts contracting when two-thirds of embryogenesis is completed. Putative key residues for possessing the Root effect, Asp95α, Asp99β and Asp101β, were detected in all the Atlantic cod globins, except from β1 and α3. This is consistent with Verde et al. (2006) who found the Hb3 ($\alpha_1\alpha_1\beta_2\beta_2$) component to exhibit the Root effect, which is important for specialized functions like delivery of oxygen to the swim bladder and to the retina (Berenbrink 2007). However, the Hb1 component has no β-globin subunits with confirmed Root effect, and we propose that this component unloads its oxygen in situations characterized by high O_2 consumption, when PO_2 in tissues is extra low.

9. Promoter polymorphism and temperature-dependent expression of β1

The MC-globin cluster harbors the *α1*- and *β1*-globin genes organized head-to-head separated by an intergenic promoter region of approximately 1.7 kb (Paper II). The promoter region was sequenced from a β_1 double heterozygote specimen, uncovering several promoter polymorphisms including a 73 bp indel, which is coupled to the functional β_1-polymorphisms (Paper III). In the Arctic population sampled at Bjørnøya and Båtsfjord, the frequency of triple homozygote cod with the long promoter variant and VA genotype was between 0.70 and 0.80, while the triple homozygote combination of short promoter and MK genotype was located with highest frequencies in the populations of Molde and Helgoland, 0.225 and 0.25, respectively (Paper III). The triple heterozygote representing the two latter haplotypes (long-VA and short-MK), but also recombinants between promoter and gene, was the most frequent genotype in Kattegat (0.387), but was also frequent in Molde and Helgoland (0.225 and 0.25, respectively). Activities of the two promoter allelic variants were examined by placing them in front of luciferase reporter genes, and transfecting the constructs into salmon kidney cells. Cells were incubated at 4, 15 or 20 °C, and the normalized results showed that the long promoter was twice as active compared to the short variant at the temperatures 15 and 20 °C, while they were similar at 4 °C. Considering the VA genotype displayed low oxygen affinity at elevated temperatures (Brix et al. 2004) (Paper I), we propose that the high promoter activity at such temperatures functions as compensation for low affinity by increased globin synthesis. Homozygous VA cod is mostly found in cold waters like the Barents and Baltic seas, and is only subject to elevated temperatures during short seasonal periods or migration, so the longer promoter variant contributes to sustain proper oxygen supply during such critical occasions.

Concluding remarks

The results presented in this thesis bring new insight and knowledge to the area of interactions between Atlantic cod and its fluctuating environment. The long sought link between the HbI blood phenotypes and their genetic origin has been unraveled, pinpointing two amino acid substitutions in the β_1 globin gene, and a possible structural based explanation of their different physiological properties in different environments is given. Two promoter variants of the β_1 gene and their transcriptional activities at different temperatures were documented, and together with the β_1 gene polymorphisms they represent convergent adaptation of expression level and functionality as responses to temperature. A cost-efficient and precise assay for genotyping the relevant globin loci in DNA samples was developed, enabling cod researchers to study stored tissue or DNA samples, and cod breeders to easily genotype their brood and recruits for these temperature-dependent and growth-related loci. Furthermore, the complete set of globin genes in the genome of Atlantic cod including their flanking regions is presented, representing valuable data both to future phylogenetic studies and to globin gene expression studies.

Further perspectives

There are still many unanswered questions regarding Atlantic cod and hemoglobin. One interesting question that should be addressed *in vivo* is to what extend cod is able to shift composition of their hemoglobin tetramers upon different environmental parameters such as hypoxia and temperature. Which genes or alleles are, in such case, preferred in different external conditions, and are, for example, double heterozygous MK/VA fish able to display overdominance in certain habitats? Why is the VK recombination widely distributed on the western side of the North Atlantic Ocean, and why is the other recombinant MA hardly found? An *in vivo* expression experiment should also include representatives of different $\alpha_1\beta_1$ promoter variants in combination with different β_1 alleles to give a more detailed picture of promoter effects. An *in vitro* deletion study of the intergenic $\alpha_1\beta_1$ promoter would also give a more complete picture of its relevance to globin expression, also considering other polymorphic sites than the 73 bp indel studied.

References

Abbasi, A., and Lutfullah, G. 2002. Molecular basis of bird respiration: primary hemoglobin structure component from tufted duck (Aythya fuligula, Anseriformes)--Role of [alpha]99Arg in formation of a complex salt bridge network. Biochem. Biophys. Res. Commun. **291**(1): 176-184. doi:10.1006/bbrc.2002.6399.

Albers, C., Goetz, K.H., and Hughes, G.M. 1983. Effect of acclimation temperature on intraerythrocytic acid-base-balance and nucleoside triphosphates in the carp, Cyprinus-carpio. Resp Physiol **54**(2): 145-159.

Barcroft, J., and King, W.O. 1909. The effect of temperature on the dissociation curve of blood. J Physiol **39**(5): 374-384.

Berenbrink, M. 2006. Evolution of vertebrate haemoglobins: Histidine side chains, specific buffer value and Bohr effect. Respir. Physiol. Neurobiol. **154**(1-2): 165-184. doi: 10.1016/j.resp.2006.01.002.

Berenbrink, M. 2007. Historical reconstructions of evolving physiological complexity: O2 secretion in the eye and swimbladder of fishes. J Exp Biol **210**(9): 1641-1652. doi: 10.1242/jeb.003319.

Berenbrink, M., Koldkjær, P., Hannah Wright, E., Kepp, O., and José da Silva, A. 2011. Magnitude of the Root effect in red blood cells and haemoglobin solutions of fishes: a tribute to August Krogh. Acta physiol **202**(3): 583-592. doi: 10.1111/j.1748-1716.2010.02243.x.

Berenbrink, M., Koldkjaer, P., Kepp, O., and Cossins, A.R. 2005. Evolution of oxygen secretion in fishes and the emergence of a complex physiological system. Science **307**(5716): 1752-1757.

Bohr, C., Hasselbalch, K., and Krogh, A. 1904. Über einen in biologischer Beziehung wichtigen Einfluss, den die Kohlensurespannung des Blutes auf dessen Sauerstoffbindung übt. Skand Arch Physiol **16**: 402-412.

Borgese, F., Sardet, C., Cappadoro, M., Pouyssegur, J., and Motais, R. 1992. Cloning and expression of a cAMP-activated Na+/H+ exchanger: evidence that the cytoplasmic domain mediates hormonal regulation. Proc Natl Acad Sci **89**(15): 6765-6769.

Borza, T., Stone, C., Gamperl, A.K., and Bowman, S. 2009. Atlantic cod (*Gadus morhua*) hemoglobin genes: multiplicity and polymorphism. BMC Genet. **10**: 51. doi:10.1186/1471-2156-10-51.

Brauner, C.J., Wang, T., and Jensen, F.B. 2002. Influence of hyperosmotic shrinkage and beta-adrenergic stimulation on red blood cell volume regulation and oxygen binding properties in rainbow trout and carp. J. Comp. Physiol., B **172**(3): 251-262. doi: 10.1007/s00360-001-0249-5.

Brittain, T. 2005. Root effect hemoglobins. J. Inorg. Biochem. **99**(1): 120-129. doi: 10.1016/j.jinorgbio.2004.09.025.

Brix, O., Foras, E., and 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**(2): 575-583. doi:10.1016/S1095-6433(97)00469-8.

Brix, O., Thorkildsen, S., and Colosimo, A. 2004. Temperature acclimation modulates the oxygen binding properties of the Atlantic cod (Gadus morhua L.) genotypes HbI*1/1, HbI*1/2, and HbI*2/2 by changing the concentrations of their major hemoglobin components (results from growth studies at different temperatures). Comp. Biochem. Physiol. A-Mol. Integr. Physiol. **138**(2): 241-251.

Churcher, A.M., and Taylor, J.S. 2009. Amphioxus (Branchiostoma floridae) has orthologs of vertebrate odorant receptors. BMC Evol Biol **9**: 242. doi: 10.1186/1471-2148-9-242.

Cohen, D.M., Inada, T., Iwamoto, T., and Scialabba, N. 1990. Gadiform fishes of the world (order Gadiformes) : an annotated and illustrated catalogue of cods, hakes, grenadiers, and other gadiform fishes known to date. FAO fisheries synopsis no. 125. Food and Agriculture Organization of the United Nations, Rome.

Ebner, B., Panopoulou, G., Vinogradov, S., Kiger, L., Marden, M., Burmester, T., and Hankeln, T. 2010. The globin gene family of the cephalochordate amphioxus: implications for chordate globin evolution. BMC Evol Biol **10**(1): 370. doi:10.1186/1471-2148-10-370.

Fairbanks, M.B., Hoffert, J.R., and Fromm, P.O. 1969. The dependence of the oxygen-concentrating mechanism of the teleost eye (Salmo gairdneri) on the enzyme carbonic anhydrase. J Gen Physiol **54**(2): 203-211.

Frydenberg, O., Moller, D., Naevdal, G., and Sick, K. 1965. Haemoglobin polymorphism in Norwegian cod populations. Hereditas **53**(1): 257-271.

Fyhn, U.E.H., Brix, O., Naevdal, G., and Johansen, T. 1994. New variants of the haemoglobins of Atlantic cod: a tool for discriminating between coastal and Arctic cod populations. ICES Mar. Sci. Symp. **198**: 666 -670.

Gamperl, A.K., Busby, C.D., Hori, T.S., Afonso, L.O., and Hall, J.R. 2009. Hemoglobin genotype has minimal influence on the physiological response of juvenile Atlantic cod (Gadus morhua) to environmental challenges. Physiol Biochem Zool **82**(5): 483-494. doi: 10.1086/603636.

Gillooly, J.F., Brown, J.H., West, G.B., Savage, V.M., and Charnov, E.L. 2001. Effects of size and temperature on metabolic rate. Science **293**(5538): 2248-2251. doi: 10.1126/science.1061967.

Hall, T.E., Smith, P., and Johnston, I.A. 2004. Stages of embryonic development in the Atlantic cod Gadus morhua. J Morphol **259**(3): 255-270. doi: 10.1002/jmor.10222.

Hardison, R.C. 2008. Globin genes on the move. J Biol **7**(9): 35. doi: 10.1186/jbiol92.

Heisler, N. 1984. Acid-base regulation in fishes. *In* Fish physiology. *Edited by* W.S. Hoar and D.J. Randall. Academic press, Inc. pp. 315-401.

Hiebl, I., Schneeganss, D., Grimm, F., Kosters, J., and Braunitzer, G. 1987. High altitude respiration of birds. The primary structures of the major and minor hemoglobin component of adult European black vulture (Aegypius monachus, Aegypiinae). Biol Chem Hoppe Seyler **368**(1): 11-18.

Husebo, A., Imsland, A.K., and Naevdal, G. 2004. Haemoglobin variation in cod: a description of new variants and their geographical distribution. Sarsia **89**(6): 369-378. doi:10.1080/00364820410002631.

Imsland, A.K., Foss, A., Naevdal, G., Johansen, T., Folkvord, A., Stefansson, S.O., and Jonassen, T.M. 2004. Variations in growth in haemoglobin genotypes of Atlantic cod. Fish Physiol. Biochem. **30**(1): 47-55. doi:10.1007/s10695-004-6787-5.

Imsland, A.K., Foss, A., Naevdal, G., Johansen, T., Stefansson, S.O., and Jonassen, T.M. 2007. New haemoglobin genotypes in Atlantic cod, Gadus morhua: possible relation with growth. Comp. Biochem. Physiol. A **147**(4): 955-960. doi:10.1016/j.cbpa.2007.03.001.

Jensen, F.B. 1987. Influences of exercise-stress and adrenaline upon intra- and extracellular acid-base status, electrolyte composition and respiratory properties of blood in tench (Tinca tinca) at different seasons. J Comp Physiol B **157**(1): 51-60.

Jensen, F.B. 1989. Hydrogen ion equilibria in fish haemoglobins. J Exp Biol **143**: 225-234.

Jensen, F.B., Fago, A., and Weber, R.E. 1998. Fish respiration. Haemoglobin structure and function. *In* Fish physiology. *Edited by* S.F. Perry and B.L. Tufts. Academic press Inc.

Jensen, F.B., Nikinmaa, M., and Weber, R.E. 1983. Effects of exercise stress on acid-base balance and respiratory function in blood of the teleost Tinca tinca. Respir Physiol **51**(3): 291-301.

Jessen, T.H., Weber, R.E., Fermi, G., Tame, J., and Braunitzer, G. 1991. Adaptation of bird hemoglobins to high-altitudes - demonstration of molecular mechanism by protein engineering. . Proc. Natl. Acad. Sci. U. S. A. **88**(15): 6519-6522.

Johnston, I.A., Abercromby, M., and Andersen, O. 2006. Muscle fibre number varies with haemoglobin phenotype in Atlantic cod as predicted by the optimal fibre number hypothesis. Biol. Lett. **2**(4): 590-592. doi:10.1098/rsbl.2006.0500.

Jordan, A.D., Lampe, J.F., Grisdale-Helland, B., Helland, S.J., Shearer, K.D., and Steffensen, J.F. 2006. Growth of Atlantic cod (Gadus morhua L.) with different haemoglobin subtypes when kept near their temperature preferenda. Aquaculture **257**(1-4): 44-52. doi: 10.1016/j.aquaculture.2006.03.015.

Jørstad, K.E., and Nævdal, G. 1994. Studies on associations between genotypes and growth rate in juvenile cod. ICES Mar. Sci. Symp. **198**: 671-675.

Karlsen, O.A., Bjørneklett, S., Berg, K., Brattås, M., Bohne-Kjersem, A., Grøsvik, B.E., and Goksøyr, A. 2011. Integrative Environmental Genomics of Cod (Gadus morhua): The Proteomics Approach. J Toxicol Env Heal A **74**(7-9): 494-507. doi: 10.1080/15287394.2011.550559.

Karpov, A.K., and Novikov, G.G. 1980. Hemoglobin alloforms in cod, *Gadus morhua* (Gadiformes, Gadidae), their functional characteristics and occurrence in populations. J. Ichthyol.(6): 45-50.

Kobayashi, H., Pelster, B., and Scheid, P. 1990. CO2 back-diffusion in the rete aids O2 secretion in the swimbladder of the eel. Respir Physiol **79**(3): 231-242.

Kriwet, J., and Hecht, T. 2008. A review of early gadiform evolution and diversification: first record of a rattail fish skull (Gadiformes, Macrouridae) from the Eocene of Antarctica, with otoliths preserved in situ. Naturwissenschaften **95**(10): 899-907. doi: 10.1007/s00114-008-0409-5.

Laursen, J.S., Andersen, N.A., and Lykkeboe, G. 1985. Temperature-acclimation and oxygen binding-properties of blood of the European eel, Anguilla-anguilla. Comp Biochem Phys A **81**(1): 79-86.

Liang, Y.H., Hua, Z.Q., Liang, X., Xu, Q., and Lu, G.Y. 2001. The crystal structure of bar-headed goose hemoglobin in deoxy form: The allosteric mechanism of a hemoglobin species with high oxygen affinity. J. Mol. Biol. **313**(1): 123-137.

Loeng, H. 2008. Klimaendringer i Barentshavet - Konsekvenser av økte CO2-nivåer i atmosfæren og havet. Spesialrapport, NorACIA, Tromsø.

Marechal, J.D., Maseras, F., Lledos, A., Mouawad, L., and Perahia, D. 2006. A DFT study on the relative affinity for oxygen of the alpha and beta subunits of hemoglobin. J. Comput. Chem. **27**(12): 1446-1453. doi:10.1002/jcc.20427.

Mazzarella, L., Vergara, A., Vitagliano, L., Merlino, A., Bonomi, G., Scala, S., Verde, C., and di Prisco, G. 2006. High resolution crystal structure of deoxy hemoglobin from Trematomus bernacchii at different pH values: the role of histidine residues in modulating the strength of the root effect. Proteins **65**(2): 490-498. doi: 10.1002/prot.21114.

Moller, D.A.G. 1966. Genetic differences between cod groups in the lofoten area. Nature **212**(5064): 824-824. doi:10.1038/212824a0.

Mork, J., Giskeødegård, R., and Sundnes, G. 1984. The haemoglobin polymorphism in Atlantic cod (Gadus morhua L.); genotypic differences in somatic growth and in maturing age in natural populations. Flødevigen rapportserie pp. 721–732, Skien.

Nævdal, G., Folkvord, A., Otterlei, E., and Thorkildsen, S. 1992. Growth-rate related to genotype of 0-group cod at 3 environmental temperatures. Sarsia **77**(1): 71-73.

Nielsen, E.E., Hansen, M.M., Schmidt, C., Meldrup, D., and Gronkjaer, P. 2001. Fisheries: Population of origin of Atlantic cod. Nature **413**(6853): 272. doi: 10.1038/35095112.

Nikinmaa, M. 1992. Membrane transport and control of hemoglobin-oxygen affinity in nucleated erythrocytes. Physiol Rev **72**(2): 301-321.

Nikinmaa, M., and Salama, A. 1998. Fish respiration. Oxygen Transport in Fish. *In* Fish Physiology. *Edited by* F.P. Steve and L.T. Bruce. Academic Press. pp. 141-184.

Nikinmaa, M., Tuurala, H., and Soivio, A. 1980. Thermoacclimatory changes in blood-oxygen binding-properties and gill secondary lamellar structure of Salmo-gairdneri. J Comp Physiol **140**(3): 255-260.

Olson, J.S., Mathews, A.J., Rohlfs, R.J., Springer, B.A., Egeberg, K.D., Sligar, S.G., Tame, J., Renaud, J.-P., and Nagai, K. 1988. The role of the distal histidine in myoglobin and haemoglobin. Nature **336**(6196): 265-266. doi:10.1038/336265a0.

Patel, V.S., Cooper, S.J., Deakin, J.E., Fulton, B., Graves, T., Warren, W.C., Wilson, R.K., and Graves, J.A. 2008. Platypus globin genes and flanking loci suggest a new insertional model for beta-globin evolution in birds and mammals. BMC Biol **6**: 34. doi: 10.1186/1741-7007-6-34.

Pelster, B., and Randall, D.J. 1998. Fish respiration. The physiology of the Root effect. *In* Fish physiology. *Edited by* B. Perry and B.L. Tufts. Academic press Inc. pp. 113-139.

Pelster, B., and Scheid, P. 1992. Countercurrent concentration and gas secretion in the fish swim bladder. Physiol. Zool. **65**(1): 1-16.

Perutz, M.F. 1970. Stereochemistry of cooperative effects in haemoglobin. Nature **228**(5273): 726-739.

Perutz, M.F. 1983. Species adaptation in a protein molecule. Mol Biol Evol **1**(1): 1-28.

Perutz, M.F. 1990. Mechanisms regulating the reactions of human hemoglobin with oxygen and carbon monoxide. Annu Rev Physiol **52**(1): 1-26. doi:10.1146/annurev.ph.52.030190.000245.

Petersen, M.F., and Steffensen, J.F. 2003. Preferred temperature of juvenile Atlantic cod Gadus morhua with different haemoglobin genotypes at normoxia and moderate hypoxia. J. Exp. Biol. **206**(Pt 2): 359-364. doi:10.1242/jeb.00111.

Portner, H.O., Berdal, B., Blust, R., Brix, O., Colosimo, A., De Wachter, B., Giuliani, A., Johansen, T., Fischer, T., Knust, R., Lannig, G., Naevdal, G., Nedenes, A., Nyhammer, G., Sartoris, F.J., Serendero, I., Sirabella, P., Thorkildsen, S., and Zakhartsev, M. 2001. Climate induced temperature effects on growth performance, fecundity and recruitment in marine fish: developing a hypothesis for cause and effect relationships in Atlantic cod (Gadus morhua) and common eelpout (Zoarces viviparus). Cont. Shelf Res. **21**(18-19): 1975-1997.

Putnam, N.H., Butts, T., Ferrier, D.E., Furlong, R.F., Hellsten, U., Kawashima, T., Robinson-Rechavi, M., Shoguchi, E., Terry, A., Yu, J.K., Benito-Gutierrez, E.L., Dubchak, I., Garcia-Fernandez, J., Gibson-Brown, J.J., Grigoriev, I.V., Horton, A.C., de Jong, P.J., Jurka, J., Kapitonov, V.V., Kohara, Y., Kuroki, Y., Lindquist, E., Lucas, S., Osoegawa, K., Pennacchio, L.A., Salamov, A.A., Satou, Y., Sauka-Spengler, T., Schmutz, J., Shin, I.T., Toyoda, A., Bronner-Fraser, M., Fujiyama, A., Holland, L.Z., Holland, P.W., Satoh, N., and Rokhsar, D.S. 2008. The amphioxus genome and the evolution of the chordate karyotype. Nature **453**(7198): 1064-1071. doi: 10.1038/nature06967.

Quinn, N.L., Boroevich, K.A., Lubieniecki, K.P., Chow, W., Davidson, E.A., Phillips, R.B., Koop, B.F., and Davidson, W.S. 2010. Genomic organization and evolution of the Atlantic salmon hemoglobin repertoire. BMC Genomics **11**: 539. doi: 10.1186/1471-2164-11-539.

Rogers, L.A., Stige, L.C., Olsen, E.M., Knutsen, H., Chan, K.-S., and Stenseth, N.C. 2011. Climate and population density drive changes in cod body size throughout a century on the Norwegian coast. Proc. R. Soc. B **108**(5): 1961-1966. doi: 10.1073/pnas.1010314108.

Root, R.W. 1931. The respiratory function of the blood of marine fishes. Biol Bull **61**(3): 427-456.

Root, R.W., and Irving, L. 1941. The equilibrium between hemoglobin and oxygen in whole and hemolyzed blood of the tautog, with a theory of the Haldane effect. Biol Bull **81**(3): 307-323.

Salvanes, A.G.V., and Hart, P.J.B. 2000. Is individual variation in competitive performance of reared juvenile cod influenced by haemoglobin genotype? Sarsia **85**(3): 265-274.

Seibel, B.A., and Drazen, J.C. 2007. The rate of metabolism in marine animals: environmental constraints, ecological demands and energetic opportunities. Philos Trans R Soc Lond B Biol Sci **362**(1487): 2061-2078. doi: 10.1098/rstb.2007.2101.

Shikama, K., and Matsuoka, A. 2003. Human haemoglobin - A new paradigm for oxygen binding involving two types of alpha beta contacts. Eur. J. Biochem. **270**(20): 4041-4051.

Sick, K. 1961. Haemoglobin polymorphism in fishes. Nature **192**(480): 894-896. doi:10.1038/192894a0.

Sick, K. 1965a. Haemoglobin polymorphism of cod in North Sea and north Atlantic Ocean. Hereditas **54**(1): 49-69.

Sick, K. 1965b. Haemoglobin polymorphism of cod in the Baltic and the Danish Belt Sea. Hereditas **54**(1): 19-48.

Steen, J.B. 1963. The physiology of the swimbladder in the eel Anguilla vulgaris. III. The mechanism of gas secretion. Acta Physiol Scand **59**: 221-241.

Tetens, V., and Christensen, N.J. 1987. Beta-adrenergic control of blood oxygen affinity in acutely hypoxia exposed rainbow trout. J Comp Physiol B **157**(5): 667-675.

Tindall, E.A., Petersen, D.C., Woodbridge, P., Schipany, K., and Hayes, V.M. 2009. Assessing high-resolution melt curve analysis for accurate detection of gene variants in complex DNA fragments. Hum. Mutat. **30**(6): 876-883. doi:10.1002/humu.20919.

Tomkiewicz, J., Lehmann, K.M., and St John, M.A. 1998. Oceanographic influences on the distribution of Baltic cod, Gadus morhua, during spawning in the Bornholm Basin of the Baltic Sea. Fish Oceanogr. **7**(1): 48-62.

Ullal, A.J., Litaker, R.W., and Noga, E.J. 2008. Antimicrobial peptides derived from hemoglobin are expressed in epithelium of channel catfish (Ictalurus punctatus, Rafinesque). Dev Comp Immunol **32**(11): 1301-1312. doi: 10.1016/j.dci.2008.04.005.

Verde, C., Balestrieri, M., de Pascale, D., Pagnozzi, D., Lecointre, G., and di Prisco, G. 2006. The oxygen transport system in three species of the boreal fish family gadidae - Molecular phylogeny of hemoglobin. J. Biol. Chem. **281**(31): 22073-22084. doi:10.1074/jbc.M513080200.

Vossen, R.H.A.M., van Duijn, M., Daha, M.R., den Dunnen, J.T., and Roos, A. 2010. High-throughput genotyping of mannose-binding lectin variants using high-resolution DNA-melting analysis. Hum. Mutat. **31**(4): E1286-E1293. doi:10.1002/Humu.21213.

Waser, W., and Heisler, N. 2005. Oxygen delivery to the fish eye: root effect as crucial factor for elevated retinal PO2. J Exp Biol **208**(Pt 21): 4035-4047. doi: 10.1242/jeb.01874.

Weber, R.E., and Campbell, K.L. 2011. Temperature dependence of haemoglobin-oxygen affinity in heterothermic vertebrates: mechanisms and biological significance. Acta Physiol (Oxf) **202**(3): 549-562. doi: 10.1111/j.1748-1716.2010.02204.x.

Weber, R.E., Lalthantluanga, R., and Braunitzer, G. 1988. Functional characterization of fetal and adult yak hemoglobins: An oxygen binding cascade and its molecular

basis. Arch Biochem Biophys **263**(1): 199-203. doi: 10.1016/0003-9861(88)90628-5.

Wittenberg, J.B., and Wittenberg, B.A. 1962. Active secretion of oxygen into the eye of fish. Nature **194**: 106-107.

Yokoyama, T., Chong, K.T., Miyazaki, G., Morimoto, H., Shih, D.T., Unzai, S., Tame, J.R., and Park, S.Y. 2004. Novel mechanisms of pH sensitivity in tuna hemoglobin: a structural explanation of the root effect. J Biol Chem **279**(27): 28632-28640. doi: 10.1074/jbc.M401740200.

Paper I Andersen et al. 2009

Paper II Wetten et al. 2010

Paper III Staar et al. 2011

Paper IV Wetten et al. 2011

Paper I

PROCEEDINGS OF THE ROYAL SOCIETY B

Proc. R. Soc. B (2009) **276**, 833–841
doi:10.1098/rspb.2008.1529
Published online 25 November 2008

Haemoglobin polymorphisms affect the oxygen-binding properties in Atlantic cod populations

Øivind Andersen[1,2,*], Ola Frang Wetten[3,4], Maria Cristina De Rosa[5], Carl Andre[6], Cristiana Carelli Alinovi[7], Mauro Colafranceschi[8,9], Ole Brix[10,11] and Alfredo Colosimo[8,9]

[1]*Nofima, PO Box 5010, 1430 Aas, Norway*
[2]*CIGENE—Centre of Integrative Genetics, 1430 Aas, Norway*
[3]*Department of Natural Science and Technology, Hedmark University College, 2318 Hamar, Norway*
[4]*Department of Animal and Aquaculture Sciences, Norwegian University of Life Sciences, 1430 Aas, Norway*
[5]*Institute of Chemistry of Molecular Recognition—CNR, 00168 Rome, Italy*
[6]*Department of Marine Ecology-Tjärnö, Göteborg University, 45296 Strömstad, Sweden*
[7]*Institute of Biochemistry and Clinical Biochemistry, Catholic University of Rome, 00168 Rome, Italy*
[8]*Department of Human Physiology and Pharmacology, University of Rome 'La Sapienza', 00185 Rome, Italy*
[9]*CISB Interdepartmental Centre, University of Rome 'La Sapienza', 00186 Rome, Italy*
[10]*Department of Biology, University of Bergen, 5020 Bergen, Norway*
[11]*The Michelsen Centre for Industrial Measurement Science and Technology, 5892 Bergen, Norway*

A major challenge in evolutionary biology is to identify the genes underlying adaptation. The oxygen-transporting haemoglobins directly link external conditions with metabolic needs and therefore represent a unique system for studying environmental effects on molecular evolution. We have discovered two haemoglobin polymorphisms in Atlantic cod populations inhabiting varying temperature and oxygen regimes in the North Atlantic. Three-dimensional modelling of the tetrameric haemoglobin structure demonstrated that the two amino acid replacements Met55β_1Val and Lys62β_1Ala are located at crucial positions of the $\alpha_1\beta_1$ subunit interface and haem pocket, respectively. The replacements are proposed to affect the oxygen-binding properties by modifying the haemoglobin quaternary structure and electrostatic feature. Intriguingly, the same molecular mechanism for facilitating oxygen binding is found in avian species adapted to high altitudes, illustrating convergent evolution in water- and air-breathing vertebrates to reduction in environmental oxygen availability. Cod populations inhabiting the cold Arctic waters and the low-oxygen Baltic Sea seem well adapted to these conditions by possessing the high oxygen affinity Val55–Ala62 haplotype, while the temperature-insensitive Met55–Lys62 haplotype predominates in the southern populations. The distinct distributions of the functionally different haemoglobin variants indicate that the present biogeography of this ecologically and economically important species might be seriously affected by global warming.

Keywords: haemoglobin; adaptation; polymorphism; *Gadus morhua*; oxygen affinity; convergent evolution

1. INTRODUCTION

Elucidating the molecular mechanisms of adaptation is a central aim in evolutionary biology, yet very few studies have identified the genes underlying adaptive variation in fitness-related traits in natural populations. The haemoglobins represent a unique system for studying adaptive changes because these oxygen-carrying proteins closely connect metabolic activities with external conditions. Hence, the haemoglobins have experienced a major evolutionary pressure and acquired a number of complex features to execute their primary function under extreme variable conditions (Perutz 1983). One remarkable instance of natural selection is the mutated haemoglobin of the Andean goose (*Chloephaga melanoptera*) and bar-headed goose (*Anser indicus*) enabling these species to tolerate low oxygen pressure at high elevations in the Andean and Himalayan mountains, respectively (Hiebl *et al.* 1987; Liang *et al.* 2001). The genetic basis of high-altitude adaptation in the deer mice (*Peromyscus maniculatus*) has recently been elucidated by identifying several mutations in the 5′α globin gene, which showed different genotypes in the elevation zones examined (Storz *et al.* 2007).

Temperature change strongly affects the oxygen consumption and haemoglobin–oxygen-binding affinity of poikilothermic organisms, as well as the ambient oxygen availability. This crucial environmental factor therefore sets the limits for life by influencing the maintenance and activity of all poikilothermic animals. Accordingly, the impact of global warming on seawater temperatures has been reported to seriously affect both the distribution and production of marine species (O'Brian *et al.* 2000; Perry *et al.* 2005; Brander 2007). Thus, the northward shift of

*Author and address for correspondence: Nofima, PO Box 5010, 1430 Aas, Norway (oivind.andersen@akvaforsk.no).

Electronic supplementary material is available at http://dx.doi.org/10.1098/rspb.2008.1529 or via http://journals.royalsociety.org.

Received 23 October 2008
Accepted 29 October 2008

the North Sea populations of Atlantic cod was reported to be due to climate-related changes (Perry *et al.* 2005), which, however, seems to contrast with the lack of selective behaviour to avoid high temperatures reported in a more recent study (Neat & Righton 2007). The tolerance of marine fishes to not only high, but also low, temperatures is characterized by a discrepancy between the demand for oxygen and the capacity of oxygen supply to the tissues (Lannig *et al.* 2003; Sartoris *et al.* 2003; Pörtner & Knust 2007). This emphasizes the temperature-dependent use of functionally different haemoglobin isoforms to optimize oxygen transport. The nucleated red blood cells of fishes are generally characterized by multiple haemoglobin components, which are functionally different and appear adapted for oxygen extraction in habitats that may vary in oxygen, temperature and carbon dioxide (Pérez *et al.* 1995; Berenbrink *et al.* 2007). The heterogeneity of fish haemoglobins is further increased in Atlantic cod through the display of population-specific haemoglobin phenotypes designated HbI-1/1, HbI-2/2 and HbI-1/2 (Sick 1961).

Atlantic cod is believed to comprise several more or less reproductively isolated populations of seasonal migrants and sedentary residents (Robichaud & Rose 2004; Pampoulie *et al.* 2008; Wennevik *et al.* 2008), which seem to be differentially adapted to the varying physico-chemical conditions in the Arctic and temperate regions of the North Atlantic (Nelson *et al.* 1996; Nissling & Westin 1997; Brix *et al.* 1998). The distribution of the cod haemoglobin phenotypes is extremely heterogeneous in the North Atlantic (Brix *et al.* 1998; Pörtner *et al.* 2001): the homozygote HbI-2/2 type predominates in the colder northern waters, whereas the homozygote HbI-1/1 type is prevalent at lower latitudes along the coast of Norway and in the North Sea. A similar, although less clear, cline can be seen along the North American East Coast (Sick 1965*a*). Accordingly, Petersen & Steffensen (2003) showed that temperature is a selective factor in the distribution of the haemoglobin phenotypes by demonstrating that HbI-2/2 cod prefer lower temperatures (8.2°C) than HbI-1/1 cod (15.4°C) under normoxic conditions. These results fit with oxygen-binding analyses demonstrating that the HbI-2/2 phenotype shows higher oxygen-binding affinity than HbI-1/1 fishes at low temperatures (less than 12°C), while the situation is reversed at higher temperatures up to 20°C (Karpov & Novikov 1980; Brix *et al.* 1998, 2004). The HbI-1/2 heterozygote seems to display intermediate properties of the two homozygotes.

In this study, we reveal the long-sought genetic basis of the haemoglobin polymorphism in Atlantic cod and present molecular mechanisms underlying the proposed adaptation of the cod populations to varying temperature and oxygen regimes. We first determine the relationship between the conventional haemoglobin phenotypes and sequence polymorphisms in the cod globins. Second, we infer from three-dimensional modelling of the tetrameric haemoglobin structure that the identified amino acid replacements influence the haemoglobin quaternary structure and electrostatic features, and thereby its oxygen-binding properties. Finally, we show that the allelic distribution of the haemoglobin polymorphisms in Atlantic cod populations is related to environmental temperature and oxygen saturation.

2. MATERIAL AND METHODS

(a) *Cloning of cod globin genes*

Templates for polymerase chain reaction (PCR) amplification of α and β globin genes were isolated from the spleen and blood of juvenile Atlantic cod caught in the sea off Bergen and Trondheim, Norway, and in Øresund, Denmark. The tissues were stored in RNAlater (Ambion, Foster City, CA, USA) until extraction of total RNA (Trizol, Gibco BRL, Gaithersburg, MD, USA) and cDNA synthesis (Pharmacia Biotech, Piscataway, NJ, USA). Design of PCR primers was based on gadoid globin sequences available in public databases, and primer sequences are available upon request. PCR was performed under standard protocols, the amplicons were inserted in pGEM-T Easy Vector (Promega, Fitchburg, WI, USA) and globin-encoding sequences were obtained by sequencing multiple clones in both directions with BigDye v. 3.1 sequencing kit on 3730 ABI DNA analyser (Applied Biosystems, Foster City, CA, USA).

Haemoglobin genotype–phenotype relationships were determined by studying globin sequence variation in a group of 35 juvenile fish caught in the sea off Bergen, which were phenotyped by isoelectric focusing (IEF) analysis of haemolysates (Fyhn *et al.* 1994). The fish were genotyped using erythrocyte cDNA and genomic DNA as templates for the PCR amplification and direct sequencing (Applied Biosystems) of polymorphic sites identified in the *Hb*-β_1, *Hb*-β_3 and *Hb*-α_2 genes (figure 1).

(b) *Globin genotyping*

A total of 363 Atlantic cod were sampled from different localities representing the stocks of northeast Arctic cod, Norwegian and Danish coastal cod, and Baltic cod. Genomic DNA was extracted from gill arches or muscle tissue, and genotyping was performed by single nucleotide polymorphism (SNP) analysis. PCR primers were designed using the software SpectroDESIGNER v. 3.0 (Sequenom) and hybridized to invariable sequences of the *Hb*-β_1 gene. The 5′ and 3′ capture primers flanked the polymorphic positions of Met55Val and Lys62Ala, and the sequences are 5′-ACTGTTGGATGATTTCGTCCATGTGGTCCAG-3′ and 5′-ACTGTTGGATGCTACTTCGGTAGCTTTGGCG-3′, respectively. The mutation A/G in the first position of the Met (ATG) or Val (GTG) codon was determined by a downstream extension primer (5′-GCACCGACGCCGCTATT-3′), while an upstream extension primer (5′-GGCCACGACGCCGTGC-3′) recognized the mutation A/C in the second position of the Lys (AAG) or Ala (GCG) codon. No other codon combinations were found by manual inspection of multiple DNA sequences. The Sequenom MassARRAY analyser was used for allele separation, and genotypes were assigned by the MassARRAY Spectrotyper RT v. 3.4 software based on the mass peaks present. The results were manually inspected using the MassARRAY TYPER v. 3.3 software.

(c) *Model building*

The predicted sequences of the cod globin chains were submitted to the NCBI–Blast server and aligned with homologous sequences using ClustalW. The α- and β-globins of Antarctic rock cod *Trematomus bernacchii* (Protein Data Bank (PDB) code 1HBH), the Dusky notothen *Trematomus newnesi* (PDB code 2AA1), bluefin tuna *Thunnus thynnus* (PDB code 1V4W), and rainbow trout *Oncorhynchus mykiss* (PDB code 1OUT), shared the

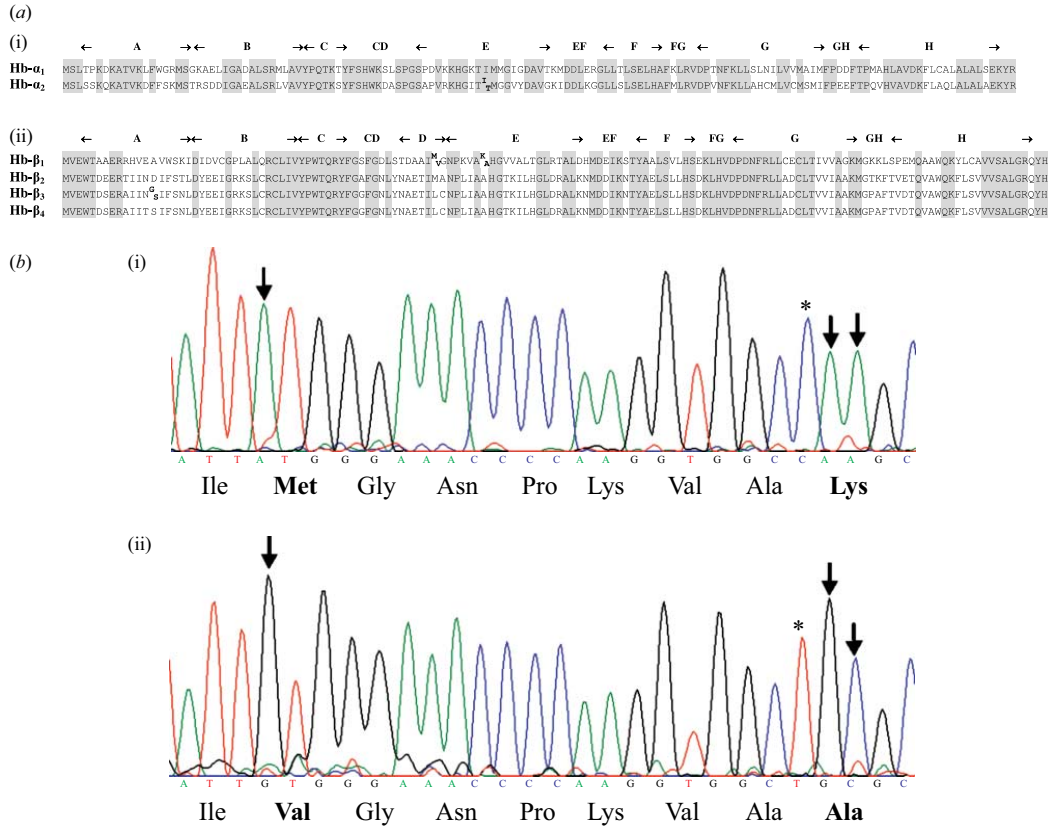

Figure 1. Amino acid polymorphisms in globin subunits of Atlantic cod. (*a*) Deduced amino acid sequences of the six isolated globin genes ((i) α-globins, (ii) β-globins), including the alternative amino acids at the polymorphic sites. The sequences are available in GenBank under the accession nos FJ392681–FJ392686. The helices are conventionally designated in bold letters. Highly conserved positions are shaded. (*b*) Sequence chromatogram of the mutated region of *Hb-β*$_1$ encoding the non-recombinant allelic globin variants (i) Met55–Lys62 and (ii) Val55–Ala62. Downward arrows, non-synonymous mutation; asterisks, synonymous mutation.

highest sequence identity with the cod globins and were selected as templates for homology modelling. Each chain model was built by comparative modelling using the MODELLER program v. 7 (Sali & Blundell 1993) as implemented in INSIGHTII (Accelrys Inc., San Diego, CA, USA) and using the crystallographic structures of the four fish haemoglobins. Three structural models were created for each globin chain in Atlantic cod, and the best model was selected based on the MODELLER objective function (*F*, molecular probability density function violation) and the stereochemical criteria of PROCHECK (Laskowski *et al.* 1996). The monomeric α and β chains obtained by homology modelling were superimposed onto the crystal structure of the *T. bernacchii* haemoglobin to keep the same relative orientation of the four subunits. The majority of the residues of the model structures were found to occupy the most favoured regions of Ramachandran plots, and the other residues occupied the additional allowed regions. Structural manipulations and energy minimization of the four tetrameric combinations of Met55βVal and Lys62βAla were performed using the molecular modelling program INSIGHTII (Accelrys) and the CHARMM force field (Brooks *et al.* 1983).

(d) *GRID analysis*

The hydrophobic and polar characteristics of the distal environment of the haem pocket were investigated with the GRID program. The interaction of a probe group with a protein of known structure is computed at sampled positions throughout and around the macromolecule, giving an array of energy values. GRID was used to predict the most favourable position for a water molecule near the distal His63β of the alternative pockets attributable to the Lys62βAla polymorphism. The probe used in the calculations was the water group (OH2), and the *grid* spacing was set to 0.25 Å. Using the flexibility option of the program, the flexible side chains of the protein can move in response to the probe and therefore mimic the adjustments that occur upon ligand binding. The results are displayed by INSIGHTII (Accelrys) as contour maps showing regions of the β haem pocket, which favourably interact with the hydrophilic probe.

(e) *Statistical analyses*

Departure from Hardy–Weinberg proportions within cod samples was estimated as F_{IS}, and population differentiation, F_{ST}, between sample pairs was evaluated using the estimator θ (Weir & Cockerham 1984). Calculations and statistical

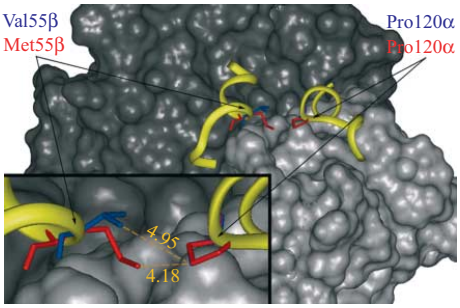

Figure 2. Met55βVal substitution in the $\alpha_1\beta_1$ subunit contact of Atlantic cod haemoglobin. The three-dimensional model shows the superimposition of Met55β (red) and Val55β (blue) at the interface of the α_1 (black) and β_1 (grey) subunits. The D helix of β_1 and the GH corner and initial H helix of α_1 chain are shown in ribbon representation (yellow). The inset image depicts the distances (Å) from the CG atom of Pro120α to the CE atom of Met55β and to the CG2 atom of Val55β.

testing of F_{IS}, F_{ST} and genotypic disequilibrium between loci were performed using GENEPOP on the web (Raymond & Rousset 1995).

3. RESULTS

(a) *Genotype–phenotype relationships*

We searched for polymorphic genes responsible for the conventional haemoglobin phenotypes in Atlantic cod by isolating six genetically distinct globins from reverse-transcribed erythrocyte mRNA using PCR. The identified globins comprise two α chains (Hb-α_1 and Hb-α_2) and four β chains (Hb-β_1, Hb-β_2, Hb-β_3 and Hb-β_4), and consist of 143 and 147 amino acids, respectively (figure 1*a*). Amino acid polymorphisms were identified in three of these globins by studying variation in the coding sequences. The *Hb-β_1* gene contains three non-synonymous mutations resulting in the replacement polymorphisms of Met55Val and Lys62Ala, in which the latter codon displays two mutations (figure 1*b*). We found strong association between the two haplotypes Met55–Lys62 and Val55–Ala62 and the two protein alleles HbI-1 and HbI-2 by comparing the genotypes and phenotypes of 35 individuals. All the fish displayed the classic IEF patterns (Sick 1961; Fyhn *et al.* 1994) representing either the HbI-1/1 ($n=12$), HbI-1/2 ($n=14$) or HbI-2/2 ($n=9$) phenotype. Only the non-recombinant haplotypes Met55–Lys62 and Val55–Ala62 were identified in this material by direct sequencing genomic DNA and erythrocyte cDNA. Cod homozygous for the Met55–Lys62 haplotype displayed the HbI-1/1 phenotype, whereas those homozygous for the Val55–Ala62 haplotype gave the HbI-2/2 phenotype. As expected, cod harbouring both haplotypes displayed the heterozygotic HbI-1/2 phenotype. Hence, unambiguous association was found between the polymorphic *Hb-β_1* gene, the expressed cDNA and the haemoglobin phenotypes, in contrast to the polymorphic *Hb-α_2* and *Hb-β_3* genes (figure 1*a*), which showed no relationship to the phenotypes.

(b) *Structure–function relationships*

We then addressed the functional impact of the Met55β Val and Lys62βAla replacements on the oxygen-binding

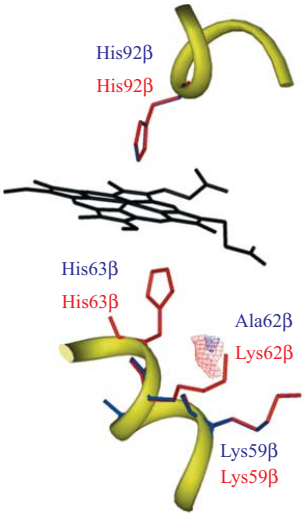

Figure 3. Lys62βAla substitution and water interaction in the haem pocket, showing the superimposition of GRID contour maps at -11.0 kcal mol^{-1} for the water probe calculated for the distal haem pocket containing Lys62β (red) or Ala62β (blue). His63β, Lys59β and Lys62β were identified as the key residues contributing to the hydrophilic contour maps.

properties of the cod haemoglobins. As the haemoglobins in Atlantic cod comprise three different tetrameric structures designated Hb1, Hb2 and Hb3 (Verde *et al.* 2006), the quaternary structure of the β_1-containing Hb1 tetramer $\alpha_1\alpha_1\beta_1\beta_1$ was built by three-dimensional homology modelling using the crystallographic structures of the α- and β-globins of four teleost species as templates. In the tetrameric cod haemoglobin, the position of Met55βVal is close to Pro120α at the $\alpha_1\beta_1$ subunit interface (figure 2). The calculated distances of the CG atom of Pro120α to the CE atom of Met55β and to the CG2 atom of Val55β in the two model structures are 4.18 and 4.95 Å, respectively (figure 2). Hence, the replacement of Met55 with the smaller Val residue increases the distance between the $\alpha_1\beta_1$ subunits that probably reduce the stability of the dimers.

Since the distal pocket residues primarily regulate oxygen binding through electrostatic interactions (Springer *et al.* 1994), we evaluated the effect of the Lys62βAla polymorphism by investigating the polar characteristics of the distal environment. In the de-oxygenated state, ligand access to the distal pocket is hindered by the presence of a water molecule stabilized by polar residues, whereas proton donors are required to stabilize the oxygenated state (Kachalova *et al.* 1999; Goldbeck *et al.* 2006). The relative magnitude of these two effects governs whether there is an increase or decrease in affinity (Springer *et al.* 1994). We estimated the binding energy of a water probe with the distal haem pocket of the Lys62β and Ala62β variants to -14.8 and -11.6 kcal mol^{-1}, respectively, demonstrating the stronger water interaction with the polar Lys compared with the Ala residue. Differences in the water interaction patterns in the two alternative pockets are evident in the calculated GRID contour maps (figure 3). The interaction area is significantly larger in the Lys62β-containing pocket due to

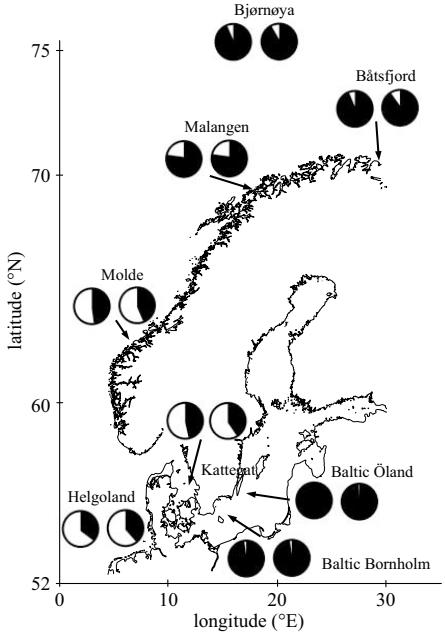

Figure 4. Relative allele frequencies of the Met55Val (left pie) and Lys62Ala (right pie) polymorphisms in the cod populations examined. The Met55 and Lys62 alleles are depicted in white, whereas the Val55 and Ala62 alleles are shown in black. Latitudes (ordinates) and longitudes (abscissas) are indicated.

the presence of the relatively large and flexible polar side chain. The smaller hydrophilic contour for the Ala62β variant is due to the interaction of the water probe limited to His63β and Lys59β.

(c) *Population genetic results*

The allele frequencies of the polymorphic *Hb-*β_1 locus showed profound differences in the cod populations examined (figure 4; table 1 in the electronic supplementary material). Both the Val55 and Ala62 alleles were completely dominant in the Baltic Sea, Barents Sea and in the northern part of the northeast Atlantic Sea. The frequencies of these alleles declined in both the north–south and east–west directions, and resulted in the predominance of the Met55 and Lys62 alleles in the southern population of the North Sea. The results are consistent with the cod population studies documenting the predominance of the protein allele HbI-2 in the waters of Greenland (99%), Iceland (98%), northern Norway (90%) and Baltic Sea (96%), decreasing to 28–36% in the warmer areas of the North Sea (Frydenberg *et al.* 1965; Sick 1965*a*,*b*). Allele frequencies within samples generally conformed to Hardy–Weinberg expectations, except at Malangen (table 2 in the electronic supplementary material). This sample showed a deficit of heterozygotes, possibly due to population mixing. The Val–Ala and Met–Lys haplotypes prevailed in the populations examined (table 1). Cod homozygous for the Val–Ala haplotype (corresponding to the HbI-2/2 phenotype) comprised 80–94% of the individuals in populations of the Baltic Sea and the northern waters, whereas homozygotic Met–Lys cod (HbI-1/1 phenotype) were completely lacking in these populations. Met–Lys homozygotes were mainly found in the mid- and southern populations together with the double heterozygotes. There was a significant linkage disequilibrium between the two loci (table 3 in the electronic supplementary material), but it is unclear to what extent heterozygotes are coupled (Met–Lys/Val–Ala) or repulsive (Met–Ala/Val–Lys). Intragenic recombination events have apparently occurred as indicated by the additional genotypes sparsely distributed in the populations examined (table 1). Consistently, sporadic subtypes of the cod haemoglobins have been reported along the Norwegian coast and in Danish waters, differing from the main types in IEF-banding patterns, oxygen-binding affinities and body growth rates (Fyhn *et al.* 1994; Brix *et al.* 2004; Husebø *et al.* 2004; Imsland *et al.* 2007).

4. DISCUSSION

(a) *Globin polymorphism and oxygen binding*

This study presents strong evidence for the crucial role played by the polymorphic β_1-globin in the adaptation of Atlantic cod to varying temperature and oxygen regimes. The unambiguous association between haemoglobin phenotypes and genotypes determined in both genomic DNA and expressed cDNA strongly supports *Hb-*β_1 as the active gene responsible for the protein alleles HbI-1 and HbI-2. This is further strengthened by the different isoelectric point of the Met55–Lys62 (pI 7.68) and the Val55–Ala62 (pI 6.99) variants, which is in agreement with the HbI-1-specific cathodic band as visualized by the IEF analysis (Sick 1961; Fyhn *et al.* 1994). Furthermore, the Met55Val replacement explains the appearance of an HbI-1-specific cathodic fragment in the fingerprinting analysis of chymotrypsin-digested cod haemoglobin (Rattazzi & Pik 1965). This His-containing peptide matches perfectly the acidic (pI 10.81) peptide 56GN-PKVAKH62 resulting from low-specific chymotrypsin cleavage at Met55, whereas treatment of the Val55 variant leaves the neutral (pI 7.55) peptide 49STDAA-IVGNPKVAAH62 uncleaved. The latter peptide was apparently not recognized by Rattazzi & Pik (1965).

The Met55Val and Lys62Ala replacements occur at the key positions of the $\alpha_1\beta_1$ subunit interface and haem pocket, respectively, and are therefore proposed to affect the oxygen-binding properties of cod haemoglobin. The $\alpha_1\beta_1$ contact is crucial to the stability of the dimers, and any gap loosening this $\alpha_1\beta_1$ interaction destabilizes the low-oxygen affinity deoxy state and facilitates transition to the high-affinity oxy state (Abbasi & Lutfullah 2002; Shikama & Matsuoka 2003). Consistently, increased oxygen affinity was induced in human haemoglobin by site-directed mutagenesis of Met55β→Ser or Pro119α→Ala (corresponding to cod Pro120α), which both excluded van der Waals contact at this subunit interface (Jessen *et al.* 1991). It is noteworthy that either of these crucial positions is mutated in the haemoglobin of the hypoxia-tolerant Andean goose (Leu55β→Ser) and bar-headed goose (Pro119α→Ala), when compared with the greylag goose (*Anser anser*) living in the plains and displaying normal oxygen affinity (Hiebl *et al.* 1987; Liang *et al.* 2001). In Atlantic cod, changing Met55 with the smaller Val residue increases the gap between the $\alpha_1\beta_1$ subunits that probably destabilizes the T-structure with a resultant

Table 1. Composite genotype frequencies (%) of the Met55βVal and Lys62βAla replacement polymorphisms in Atlantic cod populations. (Double heterozygotes represent both coupling and repulsive heterozygotes, which were not distinguished by the SNP analysis. Numbers of fish are given in parentheses. See figure 4 for collection localities of cod samples.)

genotype	Bjørnøya (40)	Båtsfjord (41)	Malangen (40)	Molde (39)	Helgoland (36)	Kattegat (74)	Baltic B (56)	Baltic Ö (37)
Val–Ala/Val–Ala	83	80	60	18	14	11	93	94
Met–Lys/Met–Lys	0	0	10	36	39	27	0	0
Met–Lys/Val–Ala Met–Ala/Val–Lys	15	15	18	31	44	52	7	3
Val–Ala/Val–Lys	2	5	8	13	0	5	0	3
Met–Lys/Val–Lys	0	0	0	0	3	4	0	0
Val–Lys/Val–Lys	0	0	2	0	0	1	0	0
Met–Lys/Met–Ala	0	0	0	2	0	0	0	0
Met–Ala/Met–Ala	0	0	2	0	0	0	0	0

increase in oxygen-binding affinity, as demonstrated in the human Met55β→Ser mutant (Jessen *et al.* 1991). We therefore propose that, during the process of convergent evolution, the water-breathing cod and air-breathing goose have acquired the same molecular mechanism for high-efficiency oxygen binding. The exploitation of the same position for tolerating hypoxic conditions in such diverse species strongly supports the hypothesis of Perutz (1983) that adaptive changes in haemoglobins have evolved by only a few amino acid substitutions in key positions.

The second polymorphism Lys62βAla in the cod β_1 globin is located close to the highly conserved His63β on the distal side of the haem. The distal His plays a key role in oxygen binding by regulating ligand access to the haem pocket and by affecting the energetics of the oxygen fixation (Olson *et al.* 1988; Maréchal *et al.* 2006). In Atlantic cod, the replacement of Ala62β with the polar Lys residue does not introduce any significant steric effects on the His63β position (figure 3), in contrast to the Ala62β→Pro substitution in the human Hb-Duarte mutant resulting in increased oxygen affinity (Ceccarelli *et al.* 2006). On the other hand, the mutation modifies the electrostatic feature of the haem pocket as shown by the stronger water interaction with Lys, compared with the neutral Ala, thus reducing the oxygen availability.

Binding of oxygen to haemoglobin is generally an exothermic reaction, and a decrease in temperature induces an increase in oxygen affinity (di Prisco *et al.* 2006). To evaluate the effect of temperature on the water interaction in the haem pocket, we considered the fact that the transfer of an amino group or an aliphatic group from a non-polar solvent to water is energetically more favourable for the amino group than the aliphatic group, with ΔG values of -132 and $+8$ cal mol^{-1} Å^{-2}, respectively (Ooi & Oobataka 1988). Apparently, this is essentially due to the more exothermic water interaction of the former as compared with the latter group, showing a ΔH of -192 and -26 cal mol^{-1} Å^{-2}, respectively (Ooi & Oobataka 1988). Consequently, at lower temperatures the oxygen path to the haem iron is restricted, being hindered by the stronger water interaction with Lys62β than with the Ala62β residue. The emerging mechanistic picture is consistent with the observation that the oxygen affinity of the HbI-1/1 (Met–Lys) form is much less temperature-sensitive than that of the HbI-2/2 (Val–Ala) form (Brix *et al.* 2004), due to the counterbalance between two exothermic reactions: the internal water interaction and oxygen binding to haem.

(b) *Cod in the Baltic Sea*

The strong shift in allele frequencies at the entrance to the Baltic coincides with the differentiation in cod at selectively neutral microsatellite DNA markers (Nielsen *et al.* 2003). However, the degree of differentiation found in the *Hb*-β_1 locus between the Kattegat and Baltic cod ($F_{ST}=0.480$, 0.472; table 1 in the electronic supplementary material) is more than one order of magnitude higher than that shown for neutral genetic markers ($F_{ST}=0.034$), and is thus consistent with the divergent selection between different environments. It might be speculated that the complete dominance of cod with the Val–Ala variant in the Baltic Sea could be the result of the colonization by Barents Sea cod *ca* 8000–4000 yr ago, as hypothesized by Sick (1965*b*). This is, however, not supported by microsatellite studies demonstrating a closer relationship between the Baltic and North Sea cod than that between the Baltic cod and cod from the Barents Sea (Nielsen *et al.* 2001).

The brackish water of the Baltic Sea restricts cod to deeper waters with higher salinity, but this strategy is compromised by low oxygen levels near the sea bottom due to water stagnation (Tomkiewicz *et al.* 1998). The Baltic cod are also challenged by long periods of low temperatures during winter and spring, which might, however, be an advantage by lowering the metabolic rate and increasing the oxygen solubility. In addition to the reported local adaptation of the Baltic cod eggs and sperm to low salinity (Nissling & Westin 1997), we propose that the Baltic cod stock is adapted to environmental hypoxia by possessing the high-affinity Val–Ala (HbI-2/2) form. Owing to the temperature sensitivity of this haemoglobin variant, its oxygen affinity would further increase with decreasing temperatures because of the strong exothermic character of oxygen binding.

(c) *Cod in the Arctic waters*

In the cold waters of the Northern Hemisphere, fishes are well known to possess high mitochondrial densities and elevated aerobic capacities, which are crucial traits in thermal adaptation (Guderley 2004; Pörtner *et al.* 2006). Accordingly, permanent cold adaptation in Atlantic cod inhabiting the Barents Sea and the Norwegian coast involves elevated aerobic metabolism in white muscle compared with the North Sea cod, in line with the concept of higher maintenance costs at low temperatures (Lannig *et al.* 2003). Increased oxygen delivery to high mitochondria densities is also supported by the

significantly higher cardiac expression of myoglobin in fish acclimatized to 4°C than 10°C, and the cold compensation was more pronounced in the northeast Arctic cod than North Sea cod (Lurman *et al.* 2007). HbI-2/2 cod acclimatized to 4°C showed significantly higher oxygen affinities than HbI-1/1 cod (P_{50} of 45.64 and 53.77 mmHg, respectively, and P_{80} of 130.79 and 177.28 mmHg, respectively), and larger arterial–venous differences were found in HbI-2/2 cod compared with the HbI-1/1 type when measured at 4°C (Brix *et al.* 2004). The importance of oxygen loading by high-affinity haemoglobins in water-breathing animals is supported by comparative studies of multiple water-breathing, bimodal and air-breathing fish species (Graham 2006). Thus, safeguarding post-branchial saturation seems to be given the first priority in a low-oxygen medium. Even though information about the ligands affecting both the arterial and venous site is limited in cod, the presented results strongly indicate that the high-affinity Val–Ala form is well adapted to the cold Arctic waters by ensuring high oxygen saturation of the arterial blood that is of importance for the high maintenance costs. The related ice cod (*Arctogadus glacialis*) and polar cod (*Boreogadus saida*) are strictly cold-water species, but the relatively stationary ice cod inhabits permanently ice-covered waters even farther north than the highly migratory polar cod. Physiological studies of a purified haemoglobin component (Hb3) revealed lower oxygen affinity, but higher Bohr effect, in the ice cod compared with both polar cod and Atlantic cod, which may be preferentially related to lifestyle (Verde *et al.* 2006). Intriguingly, the β_1 globin of ice cod and polar cod differs by displaying Met and Val, respectively, at position 55 (similar residues Gln and Asn at position 62; SwissProt accession nos. Q1AGS3, Q1AGS7). Based on the results herein, we suggest that the low-affinity Met variant probably transport sufficient amounts of oxygen in the sluggish ice cod, whereas the polar cod is probably better fitted to higher oxygen demands by possessing the destabilizing Val residue at the $\alpha_1\beta_1$ interface, similar to the Atlantic cod in these waters.

(d) *Cod in the North Sea*

Studying the effects of acclimation temperatures at 4 and 12°C, Brix *et al.* (2004) found a significant drop in oxygen affinity in HbI-2/2 cod compared with HbI-1/1 cod when acclimatized to and measured at 12°C. Cod possessing the temperature-sensitive Val–Ala form will therefore risk reduced oxygen uptake at higher temperatures. Accordingly, both oxygen-binding affinity and capacity decreased with acute temperature increases in cod from Newfoundland, which probably comprised much more than 90 per cent HbI-2/2 (Gollock *et al.* 2006). The large, exponential increase in blood flow measured in these Newfoundland cod differs from the hyperbolic increase in German Bight cod likely to be composed of more than 55 per cent HbI-1/1 (Lannig *et al.* 2004). Although different techniques may account for some of the observed differences between these studies, the unexpected resistance to warming waters specifically reported for the North Sea cod (Neat & Righton 2007) is highly suggestive of the possible great importance of the temperature insensitivity of the HbI-1/1 form. We therefore propose that cod possessing the Met–Lys form is better fitted for life in warmer waters than cod with the Val–Ala form, in agreement with the preference for higher temperatures by the HbI-1/1 cod than the HbI-2/2 fish (Petersen & Steffensen 2003). Interestingly, acclimation of heterozygotes HbI-1/2 to high temperature was shown to involve increased levels of the HbI-1 protein allele in the blood, whereas low temperature favoured the synthesis of the HbI-2 allele (Brix *et al.* 2004). Hence, the ability to change the ratio of the two haemoglobin forms by a hitherto unknown mechanism might be beneficial for the heterozygotic cod in these fluctuating environments.

(e) *Concluding remarks*

The distributions of both the exploited and unexploited North Sea fish populations have responded markedly to recent increases in sea temperatures as a consequence of global warming (O'Brian *et al.* 2000; Perry *et al.* 2005). Based on the strong relationship between the haemoglobin polymorphisms and temperature preference, we propose that the observed northward migration of Atlantic cod is due to increased temperatures exceeding that preferred by the HbI-1/1 (Met–Lys) cod in the North Sea. This implies further that a combination of increased water temperature (HELCOM 2007) and low oxygen levels would be even more unfavourable for the Baltic Sea cod, which predominantly possess the HbI-2 allele. The present frequency of the HbI-2 allele in the Baltic (approx. 97%) is, however, similar or even higher than that identified by Sick in the early 1960s (Sick 1965*b*), indicating either lack of environmental change or microevolutionary response during this time period.

We thank S. W. Omholt and R. Wilson for their discussions and reading of the manuscript. S. E. Fevolden is gratefully acknowledged for providing us with cod samples. This work was supported by the Norwegian Research Council and the Swedish Research Council FORMAS. The work was carried out under National Animal Board licences.

REFERENCES

Abbasi, A. & Lutfullah, G. 2002 Molecular basis of bird respiration: primary hemoglobin structure component from tufted duck (*Aythya fuligula*, Anseriformes)—role of α99Arg in formation of a complex salt bridge network. *Biochem. Biophys. Res. Commun.* **291**, 176–184. (doi:10.1006/bbrc.2002.6399)

Berenbrink, M., Koldkjær, P., Klepp, O. & Cossins, A. R. 2007 Evolution of oxygen secretion in fishes and the emergence of a complex physiological system. *Science* **307**, 1752–1757. (doi:10.1126/science.1107793)

Brander, K. M. 2007 Global fish production and climate change. *Proc. Natl Acad. Sci. USA* **104**, 19 709–19 714. (doi:10.1073/pnas.0702059104)

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. (doi:10.1016/S1095-6433(97)00469-8)

Brix, O., Thorkildsen, S. & Colosimo, A. 2004 Temperature acclimation modulates the oxygen binding properties of the Atlantic cod (*Gadus morhua* L.) genotypes-HbI*1/1, HbI*1/2, and HbI*2/2 by changing the concentrations of their major hemoglobin components (results from growth studies at different temperatures). *Comp. Biochem. Physiol. A* **138**, 241–251. (doi:10.1016/j.cbpb.2004.04.004)

Brooks, B. R., Bruccoleri, R. E., Olafsen, B. D., States, D. J., Swaminathan, S. & Karplus, M. 1983 CHARMM:

a program for macromolecular energy, minimization, and dynamics calculations. *J. Comp. Chem.* **4**, 187–217. (doi:10.1002/jcc.540040211)

Ceccarelli, M., Ruggerone, P., Anedda, R., Fais, A., Era, B., Sollaino, M. C., Corda, M. & Casu, M. 2006 Structure-function relationship in a variant hemoglobin: a combined computational–experimental approach. *Biophys. J.* **91**, 3529–3541. (doi:10.1529/biophysj.106.083170)

di Prisco, G., Condò, S. G., Tamburrini, M. & Giardina, B. 2006 Oxygen transport in extreme environments. *Trends Biochem. Sci.* **16**, 471–474. (doi:10.1016/0968-0004(91)90182-U)

Frydenberg, O., Møller, D., Nævdal, G. & Sick, K. 1965 Haemoglobin polymorphism in Norwegian cod populations. *Hereditas* **53**, 252–271.

Fyhn, U. E. H., Brix, O., Nævdal, G. & Johansen, T. 1994 New variants of the haemoglobins of Atlantic cod: a tool for discriminating between coastal and Arctic cod populations. *ICES Mar. Sci. Symp.* **198**, 666–670.

Goldbeck, R. A., Bhaskaran, S., Ortega, C., Mendoza, J. L., Olson, J. S., Soman, J., Gliger, D. S. & Esquerra, R. M. 2006 Water and ligand entry in myoglobin: assessing the speed and extent of heme pocket hydration after CO photodissociation. *Proc. Natl Acad. Sci. USA* **103**, 1254–1259. (doi:10.1073/pnas.0507840103)

Gollock, M. J., Currie, S., Petersen, L. H. & Gamper, A. K. 2006 Cardiovascular and haematological responses of Atlantic cod (*Gadus morhua*) to acute temperature increase. *J. Exp. Biol.* **209**, 2961–2970. (doi:10.1242/jeb.02319)

Graham, J. B. 2006 Aquatic and aerial respiration. In *The physiology of fishes* (eds D. H. Evans & J. B. Claiborne), pp. 85–117. 3rd edn. Boca Raton, FL: CRC Press.

Guderley, H. 2004 Metabolic responses to low temperature in fish muscle. *Biol. Rev.* **79**, 409–427. (doi:10.1017/S1464793103006328)

HELCOM 2007 Climate change in the Baltic Sea area. *Balt. Sea Environ. Proc.* **111**, 1–49.

Hiebl, I., Braunitzer, G. & Schneeganss, D. 1987 The primary structures of the major and minor hemoglobin-components of adult Andean goose (*Chloephaga melanoptera*, Anatidae): the mutation Leu–Ser in position 55 of the β-chains. *Biol. Chem. Hoppe-Seyler* **368**, 1559–1569.

Husebø, Å., Imsland, A. K. & Nævdal, G. 2004 Haemoglobin variation in cod: a description of new variants and their geographical distribution. *Sarsia* **90**, 1–11.

Imsland, A. K., Foss, A., Nævdal, G., Johansen, T., Folkvord, A., Stefansson, S. O. & Jonassen, T. M. 2007 Variations in growth in haemoglobin genotypes of Atlantic cod. *Fish. Physiol. Biochem.* **30**, 47–55. (doi:10.1007/s10695-004-6787-5)

Jessen, T. H., Weber, R. E., Fermi, G., Tame, J. & Braunitzer, G. 1991 Adaptation of bird hemoglobins to high altitudes: demonstration of molecular mechanism by protein engineering. *Proc. Natl Acad. Sci. USA* **88**, 6519–6522. (doi:10.1073/pnas.88.15.6519)

Kachalova, G. S., Popov, A. N. & Bartunik, H. D. 1999 A steric mechanism for inhibition of CO binding to heme proteins. *Science* **284**, 473–476. (doi:10.1126/science.284.5413.473)

Karpov, A. K. & Novikov, G. G. 1980 The hemoglobin aloforms in cod (*Gadus morhua* L.), their functional characteristics and distribution in the populations. *J. Ichthyol.* **6**, 45–50.

Lannig, G., Eckerle, L. G., Serendo, I., Sartoris, F. J., Fischer, T., Knust, R., Johansen, T. & Pörtner, H. O. 2003 Temperature adaptation in eurythermal cod (*Gadus morhua*): a comparison of mitochondrial enzyme capacities in boreal and Arctic populations. *Mar. Biol.* **142**, 589–599.

Lannig, G., Bock, C., Sartoris, F. J. & Pörtner, H. O. 2004 Oxygen limitation of thermal tolerance in cod, *Gadus morhua* L., studied by magnetic resonance imaging and on-line venous oxygen monitoring. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **287**, R902–R910.

Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R. & Thornton, J. M. 1996 AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **8**, 477–486. (doi:10.1007/BF00228148)

Liang, Y., Hua, Z., Liang, X., Xu, Q. & Lu, G. 2001 The crystal structure of bar-headed goose hemoglobin in deoxy form: the allosteric mechanism of a hemoglobin species with high oxygen affinity. *J. Mol. Biol.* **313**, 123–137. (doi:10.1006/jmbi.2001.5028)

Lurman, G. J., Koschnick, N., Pörtner, H. O. & Lucassen, M. 2007 Molecular characterisation and expression of Atlantic cod (*Gadus morhua*) myoglobin from two populations held at two different acclimation temperatures. *Comp. Biochem. Physiol. A* **148**, 681–689. (doi:10.1016/j.cbpa.2007.08.021)

Maréchal, J. D., Maseras, F., Lledós, A., Mouawad, L. & Perahia, D. 2006 A DFT study of the relative affinity for oxygen of the α and β subunits of hemoglobin. *J. Comput. Chem.* **27**, 1446–1453. (doi:10.1002/jcc.20427)

Neat, F. & Righton, D. 2007 Warm water occupancy by North Sea cod. *Proc. R. Soc. B* **274**, 789–798. (doi:10.1098/rspb.2006.0212)

Nelson, J. A., Tang, Y. & Boutilier, R. G. 1996 The effects of salinity changes on the exercise performance of two Atlantic cod (*Gadus morhua*) populations inhabiting different environments. *J. Exp. Biol.* **199**, 1295–1309.

Nielsen, E. E., Hansen, M. M., Schmidt, C., Meldrup, D. & Grønkjær, P. 2001 Determining the population of origin of individual cod in the Northeast Atlantic. *Nature* **413**, 272. (doi:10.1038/35095112)

Nielsen, E. E., Hansen, M. M., Ruzzante, D. E., Meldrup, D. & Grønkjær, 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. (doi:10.1046/j.1365-294X.2003.01819.x)

Nissling, A. & Westin, L. 1997 Salinity requirements for successful spawning of Baltic and Belt Sea cod and the potential for cod stock interactions in the Baltic Sea. *Mar. Ecol. Progr. Ser.* **152**, 261–271. (doi:10.3354/meps152261)

O'Brian, C. M., Fox, C. J., Planque, B. & Casey, J. 2000 Climate variability and North Sea cod. *Nature* **404**, 142. (doi:10.1038/35004654)

Olson, J. S., Mathews, A. J., Rohlfs, R. J., Springer, B. A., Egeberg, K. D., Sligar, S. G., Tame, J., Renaud, J. P. & Nagai, K. 1988 The role of the distal histidine in myoglobin and haemoglobin. *Nature* **336**, 265–266. (doi:10.1038/336265a0)

Ooi, T. & Oobataka, M. 1988 Effects of hydrated water on protein unfolding. *J. Biochem.* **103**, 114–120.

Pampoulie, C., Örn Stefansson, M., Jörundsdottir, T. D., Danilowicz, B. S. & Danielsdottir, A. K. 2008 Recolonization history and large-scale dispersal in the open sea: the case study of the North Atlantic cod, *Gadus morhua* L. *Biol. J. Linn. Soc.* **94**, 315–329. (doi:10.1111/j.1095-8312.2008.00995.x)

Pérez, J., Rylander, K. & Nirchio, M. 1995 The evolution of multiple haemoglobins in fishes. *Rev. Fish Biol. Fish.* **5**, 304–319. (doi:10.1007/BF00043004)

Perry, A. L., Low, P. J., Ellis, J. R. & Reynolds, J. D. 2005 Climate change and distribution shifts in marine fishes. *Science* **308**, 1912–1915. (doi:10.1126/science.1111322)

Perutz, M. 1983 Species adaptation in a protein molecule. *Mol. Biol. Evol.* **1**, 1–28.

Petersen, M. F. & Steffensen, J. F. 2003 Preferred temperature of juvenile Atlantic cod *Gadus morhua* with different haemoglobin genotypes at normoxia and moderate hypoxia. *J. Exp. Biol.* **206**, 359–364. (doi:10.1242/jeb.00111)

Pörtner, H. O. & Knust, R. 2007 Climate change affects marine fishes through the oxygen limitation of thermal tolerance. *Science* **315**, 95–97. (doi:10.1126/science.1135471)

Pörtner, H. O. *et al.* 2001 Climate induced temperature effects on growth performance, fecundity and recruitment in marine fish: developing a hypothesis for cause and effect relationships in Atlantic cod (*Gadus morhua*) and common eelpout (*Zoarces viviparus*). *Cont. Shelf Res.* **21**, 1975–1997. (doi:10.1016/S0278-4343(01)00038-3)

Pörtner, H. O., Bennett, A. F., Bozinovic, F., Clarke, A., Lardies, M. A., Lucassen, M., Pelster, B., Schiemer, F. & Stillman, J. H. 2006 Trade-offs in thermal adaptation: the need for a molecular to ecological integration. *Physiol. Biochem. Zool.* **79**, 295–313. (doi:10.1086/499986)

Rattazzi, M. C. & Pik, C. 1965 Haemoglobin polymorphism in the cod (*Gadus morhua*): a single peptide difference. *Nature* **208**, 489–491. (doi:10.1038/208489a0)

Raymond, M. & Rousset, F. 1995 Genepop on the web (v. 1.2): population genetic software for exact test and ecumenism. *J. Hered.* **86**, 248–249.

Robichaud, D. & Rose, G. A. 2004 Migratory behaviour and range in Atlantic cod: inference from a century of tagging. *Fish Fish.* **5**, 185–214.

Sali, A. & Blundell, T. L. 1993 Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815. (doi:10.1006/jmbi.1993.1626)

Sartoris, F. J., Bock, C., Serendo, I., Lannig, G. & Pörtner, H. O. 2003 Temperature-dependent changes in energy metabolism, intracellular pH and blood oxygen tension in the Atlantic cod. *J. Fish Biol.* **62**, 1239–1253. (doi:10.1046/j.1095-8649.2003.00099.x)

Shikama, K. & Matsuoka, A. 2003 Human haemoglobin—a new paradigm for oxygen binding involving two types of αβ contacts. *Eur. J. Biochem.* **270**, 4041–4051. (doi:10.1046/j.1432-1033.2003.03791.x)

Sick, K. 1961 Haemoglobin polymorphism in fishes. *Nature* **192**, 894–896. (doi:10.1038/192894a0)

Sick, K. 1965*a* Haemoglobin polymorphism in the North Sea and North Atlantic ocean. *Hereditas* **54**, 49–69.

Sick, K. 1965*b* Haemoglobin polymorphism of cod in the Baltic and the Danish Belt Sea. *Hereditas* **54**, 19–48.

Springer, B. A., Sligar, S. G., Olson, J. S. & Phillips, G. N. 1994 Mechanisms of ligand recognition in myoglobin. *Chem. Rev.* **94**, 699–714. (doi:10.1021/cr00027a007)

Storz, J. F., Sabatino, S. J., Hoffman, F. G., Gering, E. J., Moriyama, H., Ferrand, N., Monteiro, B. & Nachman, M. W. 2007 The molecular basis of high-altitude adaptation in deer mice. *PLoS Genet.* **3**, 448–459. (doi:10.1371/journal.pgen.0030045)

Tomkiewicz, J., Lehmann, K. M. & St John, M. A. 1998 Oceanographic influences on the distribution of Baltic cod, *Gadus morhua*, during spawning in the Bornholm Basin of the Baltic Sea. *Fish Oceanogr.* **7**, 48–62. (doi:10.1046/j.1365-2419.1998.00051.x)

Verde, C., Balesrieri, M., de Pascale, D., Pagnozzi, D., Lecointre, G. & di Prisco, G. 2006 The oxygen transport system in three species of the boreal fish family Gadidae. *J. Biol. Chem.* **281**, 22 073–22 084. (doi:10.1074/jbc.M513080200)

Weir, B. S. & Cockerham, C. C. 1984 Estimating *F*-statistics for the analysis of population structure. *Evolution* **38**, 1358–1370. (doi:10.2307/2408641)

Wennevik, V., Jørstad, K. E., Dahle, G. & Fevolden, S.-E. 2008 Mixed stock analysis and the power of different classes of molecular markers in discriminating coastal and oceanic Atlantic cod (*Gadus morhua* L.) on the Lofoten spawning grounds. *Hydrobiologia* **606**, 7–25. (doi:10.1007/s10750-008-9349-5)

Supplementary table 1. Genetic hemoglobin differentiation among pairs of samples of Atlantic cod. a**)** Met55Val, b) Lys62Ala, c) both loci combined. Below diagonal are pair-wise F_{ST}-values, and above diagonal significance values. *P*-values < 0.05 are in italics, and *P*-values significant after correction for multiple testing ($\alpha' = 0.0018$) are denoted in italics and bold.

F_{ST}*P*-value	Båtsfjord	Bjørnøya	Malangen	Molde	Helgoland	Kattegat	Baltic B	Baltic Ö
a. Met55Val								
Båtsfjord	-	1.000	*0.033*	***0.000***	***0.000***	***0.000***	0.314	0.111
Bjørnøya	-0.012	-	*0.049*	***0.000***	***0.000***	***0.000***	0.314	0.111
Malangen	0.061	0.058	-	***0.001***	***0.000***	***0.000***	*0.002*	***0.001***
Molde	0.398	0.393	0.188	-	0.241	0.893	***0.000***	***0.000***
Helgoland	0.523	0.518	0.310	0.008	-	0.186	***0.000***	***0.000***
Kattegat	0.375	0.371	0.195	-0.010	0.010	-	***0.000***	***0.000***
Baltic B	0.004	0.005	0.136	0.498	0.617	0.453	-	0.407
Baltic O	0.030	0.032	0.162	0.498	0.617	0.454	-0.001	-
b. Lys62Ala								
Båtsfjord	-	0.785	*0.025*	***0.000***	***0.000***	***0.000***	0.075	0.056
Bjørnøya	-0.010	-	*0.021*	***0.000***	***0.000***	***0.000***	0.193	0.156
Malangen	0.060	0.076	-	***0.000***	***0.000***	***0.000***	***0.000***	***0.001***
Molde	0.391	0.409	0.181	-	0.637	0.771	***0.000***	***0.000***
Helgoland	0.450	0.468	0.232	-0.009	-	0.766	***0.000***	***0.000***
Kattegat	0.394	0.409	0.208	-0.007	-0.008	-	***0.000***	***0.000***
Baltic B	0.027	0.015	0.178	0.530	0.590	0.506	-	1.000
Baltic Ö	0.034	0.022	0.173	0.507	0.567	0.489	-0.010	-
c. both loci								
Båtsfjord	-	0.975	*0.007*	***0.000***	***0.000***	***0.000***	0.112	*0.038*
Bjørnøya	-0.011	-	*0.008*	***0.000***	***0.000***	***0.000***	0.230	0.088
Malangen	0.061	0.068	-	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***
Molde	0.394	0.401	0.184	-	0.442	0.946	***0.000***	***0.000***
Helgoland	0.487	0.494	0.272	0.000	-	0.421	***0.000***	***0.000***
Kattegat	0.385	0.391	0.202	-0.009	0.001	-	***0.000***	***0.000***
Baltic B	0.017	0.010	0.159	0.515	0.604	0.480	-	0.773
Baltic Ö	0.033	0.027	0.168	0.503	0.593	0.472	-0.006	-

Supplementary table 2. Deviation from Hardy-Weinberg expectations, F_{IS}, in cod samples at the SNP-loci Met55Val and Lys62Ala. *P* indicates the probability to obtain the observed F_{IS} under the null hypothesis: $F_{IS} = 0$. Corrected critical *P*-value for eight multiple test is 0.0064 at alpha=0.05. There is an indication of heterozygote deficiency in the sample from Malangen.

	Met55Val		**Lys62Ala**	
Sample	F_{IS}	***P***	F_{IS}	***P***
Båtsfjord	-0.067	1.000	-0.103	1.000
Bjørnøya	-0.068	1.000	-0.083	1.000
Malangen	0.487	0.005	0.345	0.040
Molde	0.357	0.027	0.040	1.000
Helgoland	0.037	1.000	0.065	0.736
Kattegat	-0.111	0.361	-0.171	0.158
Baltic B	-0.028	1.000	-0.028	1.000
Baltic Ö	-	-	-0.014	1.000

Supplementary table 3. Contingency test for genotypic linkage disequilibrium. *P*-values for the null hypothesis: "genotypes at locus Met55Val are independent from genotypes at locus Lys62Ala".

Sample	***P***
Båtsfjord	0.000
Bjørnøya	0.000
Malangen	0.000
Molde	0.000
Helgoland	0.000
Kattegat	0.000
Baltic B	0.000
Baltic Ö	0.053

Paper II

BMC Evolutionary Biology

RESEARCH ARTICLE Open Access

Genomic organization and gene expression of the multiple globins in Atlantic cod: conservation of globin-flanking genes in chordates infers the origin of the vertebrate globin clusters

Ola F Wetten[1,2], Alexander J Nederbragt[3], Robert C Wilson[2], Kjetill S Jakobsen[3], Rolf B Edvardsen[4], Øivind Andersen[1,5*]

Abstract

Background: The vertebrate globin genes encoding the α- and β-subunits of the tetrameric hemoglobins are clustered at two unlinked loci. The highly conserved linear order of the genes flanking the hemoglobins provides a strong anchor for inferring common ancestry of the globin clusters. In fish, the number of *α-β*-linked globin genes varies considerably between different sublineages and seems to be related to prevailing physico-chemical conditions. Draft sequences of the Atlantic cod genome enabled us to determine the genomic organization of the globin repertoire in this marine species that copes with fluctuating environments of the temperate and Arctic regions.

Results: The Atlantic cod genome was shown to contain 14 globin genes, including nine hemoglobin genes organized in two unlinked clusters designated *β5-α1-β1-α4* and *β3-β4-α2-α3-β2*. The diverged cod hemoglobin genes displayed different expression levels in adult fish, and tetrameric hemoglobins with or without a Root effect were predicted. The novel finding of maternally inherited hemoglobin mRNAs is consistent with a potential role played by fish hemoglobins in the non-specific immune response. *In silico* analysis of the six teleost genomes available showed that the two *α-β* globin clusters are flanked by paralogs of five duplicated genes, in agreement with the proposed teleost-specific duplication of the ancestral vertebrate globin cluster. Screening the genome of extant urochordate and cephalochordate species for conserved globin-flanking genes revealed linkage of *RHBDF1*, *MPG* and *ARHGAP17* to globin genes in the tunicate *Ciona intestinalis*, while these genes together with *LCMT* are closely positioned in amphioxus (*Branchiostoma floridae*), but seem to be unlinked to the multiple globin genes identified in this species.

Conclusion: The plasticity of Atlantic cod to variable environmental conditions probably involves the expression of multiple globins with potentially different properties. The interspecific difference in number of fish hemoglobin genes contrasts with the highly conserved synteny of the flanking genes. The proximity of globin-flanking genes in the tunicate and amphioxus genomes resembles the *RHBDF1-MPG-α-globin-ARHGAP17-LCMT* linked genes in man and chicken. We hypothesize that the fusion of the three chordate linkage groups 3, 15 and 17 more than 800 MYA led to the ancestral vertebrate globin cluster during a geological period of increased atmospheric oxygen content.

* Correspondence: oivind.andersen@nofima.no
[1]Department of Animal and Aquacultural Sciences, University of Life Sciences, P.O. Box 5003, N-1430 Aas, Norway
Full list of author information is available at the end of the article

Background

Hemoglobin plays a critical role in both terrestrial and aquatic animals by transporting oxygen from the respiratory surface to the inner organs. The functional complexity and evolutionary adaptation of this heme-containing molecule to different environments has therefore attracted researchers for more than a half-century. In jawed vertebrates, or gnathostomes, the hemoglobin tetramer consists of two pairs of α- and β-globins, which probably arose by duplication of a single primordial globin gene about 500-570 million years ago (MYA) [1,2]. Whereas α- and β-globin genes are juxtaposed in teleost fish, birds and mammals are characterized by unlinked clusters of α- and β-globin genes, which in mammals are arranged in the order of their expression during ontogeny [3,4]. Based on the conservation of the globin-flanking genes, including *MPG* and *c16orf35*, all gnathostomes examined share a common globin cluster referred to as the MC locus [5] corresponding to the α-globin cluster in placental mammals and chicken. Silencing of the β genes in the ancestral MC-α-β cluster has apparently also occurred in non-amniotic species, such as pufferfish, whereas a single β-like ϖ-globin is retained in the α cluster of marsupials and monotremes [6-8]. The teleost-specific genome duplication event 350-400 MYA probably gave rise to the second fish α-β globin cluster flanked by *ARHGAP17*, *LCMT* and *AQP8* [5,8]. It should be noted that this LA locus lacks globin genes in tetrapods, but is positioned on the α-containing chromosome 16 and 14 in man and chicken, respectively [5]. The amniotic β-globin cluster is thought to have originated from the transposition of a β gene copy into a region of olfactory receptor genes in their ancestor [8-10].

In contrast to the linked α-β globin pairs identified in *Xenopus*, the fish α-β pairs are commonly organized head-to-head or tail-to-tail with respect to transcriptional polarity [11-16]. These configurations probably arose from an inversion of one of the paired α-β genes in an ancestral ray-finned fish, thus resembling the reported case of gene inversion within the human β-globin cluster [17]. The structural and functional diversity of the multiple hemoglobins in teleosts strongly indicates that they have experienced a major evolutionary pressure to execute their oxygen-transporting function under highly variable physico-chemical conditions [18-20]. The selective forces have apparently resulted in the loss of hemoglobin genes in the white-blooded Antarctic icefishes (*Channichthyidae*) to reduce the blood viscosity at stable subzero temperatures [21-23].

The genomic organization of the fish α-β globin clusters has only been investigated in the model species pufferfish, zebrafish and medaka [5,6,8,10,15,24]. Atlantic cod is a marine cold water species being widely distributed from the sea surface to depths of 600 m in the Arctic and temperate regions of the North Atlantic Ocean, including the low saline Baltic Sea. Adaptation of the different cod populations to the varying physico-chemical conditions seems to involve hemoglobins with highly pH-sensitive oxygen affinities (Root effect) to adjust the swimming bladder to variable pressure during vertical migrations [25,26], together with the novel feature of expressing polymorphic variants with different oxygen-binding properties [27]. A variable number of cod hemoglobin genes and allelic variants have been reported in Norwegian, Icelandic and Canadian populations [27-29]. Here, we screened the draft cod genome [30] and identified nine α- and β-globin genes, which are organized in two unlinked clusters flanked by highly conserved syntenic regions. We document close linkage between the conserved globin-flanking genes in extant cephalochordate and urochordate species, and hypothesize that the fusion of three chordate chromosomes formed the ancestral vertebrate globin cluster more than 800 MYA.

Results

Identification of cod globin clusters

PCR primers were designed and employed to identify α-β-linked globin genes from genomic DNA, but this strategy resulted only in the amplification of the head-to-head organized *α1-β1* pair (Figure 1). We therefore screened an Atlantic cod Bacterial Artificial Chromosome (BAC) library for α- and β-globin genes. Pyrosequencing of two positive BAC clones respectively resulted in 33,889 and 32,029 reads, which were assembled into 60 (BAC1) and 46 contigs (BAC2). Multiple hemoglobin genes and conserved globin-flanking genes were identified by performing BLAST searches of the contigs using the pufferfish globin loci as query sequences. Finally, the most updated draft sequences from the cod genome project (http://www.codgenome.no) were screened to confirm the gene sequences identified in the BAC clones, and the genomic organization of the hemoglobin loci was determined. The presented sequence information therefore represents the north-east Arctic population of Atlantic cod.

Cod MC locus

Four hemoglobin genes designated *β5*, *α1*, *β1* and *α4* were identified within a 7-kb region of a scaffold spanning 1.7 Mb in the draft cod genome (Figure 1). The hemoglobin genes show the characteristic structure of three exons and two introns encoding the predicted α- and β-globins of 143 and 147 amino acids (aa), respectively (Figure 2). The

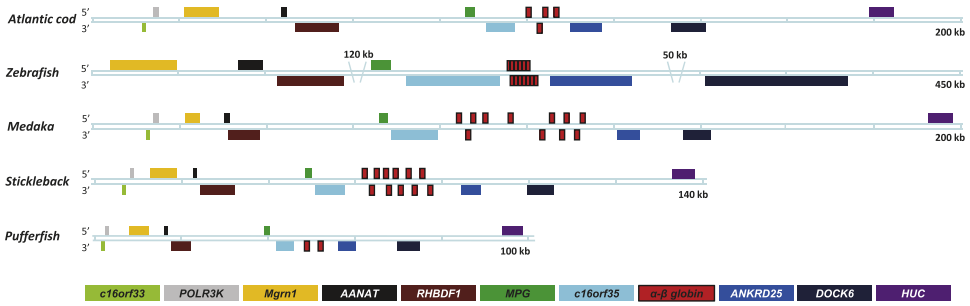

Figure 1 Conserved synteny at the MC locus in teleosts. Genomic organization of the Atlantic cod *β5-α1-β1-α4* globin cluster compared with the orthologous locus in zebrafish, medaka, stickleback and pufferfish. Omitted regions in the zebrafish loci are shown. The boxed genes are shown above or below the doubled line to indicate rightward and leftward transcriptional direction, respectively. The color codes for gene names are shown below.

paired α-β genes are organized tail-to-tail (*β5-α1*), head-to-head (*α1-β1*) or tail-to-head (*β1-α4*), and the *α1* gene is transcribed in the opposite direction of the others. Seven conserved genes, *c16orf33*, *POLR3K*, *Mgrn1*, *AANAT*, *RHBDF1*, *MPG* and *c16orf35*, were identified within a 90 kb region leftwards of the *β5-α1-β1-α4* cluster, while the rightwards flanking region of 80 kb harbors *ANKRD25*, *DOCK6* and *HuC* (Figure 1). A single major regulatory element (MRE; YGCTGASTCAY) was identified as a reversed motif (ATGACTCAGCA) in intron 5 of *RHBDF1* close to a putative GATA binding site. Whereas paired MREs are located at this position in other vertebrates examined, a second single MRE motif was found in intron 9 of the cod *Mgrn1* gene. In zebrafish, two

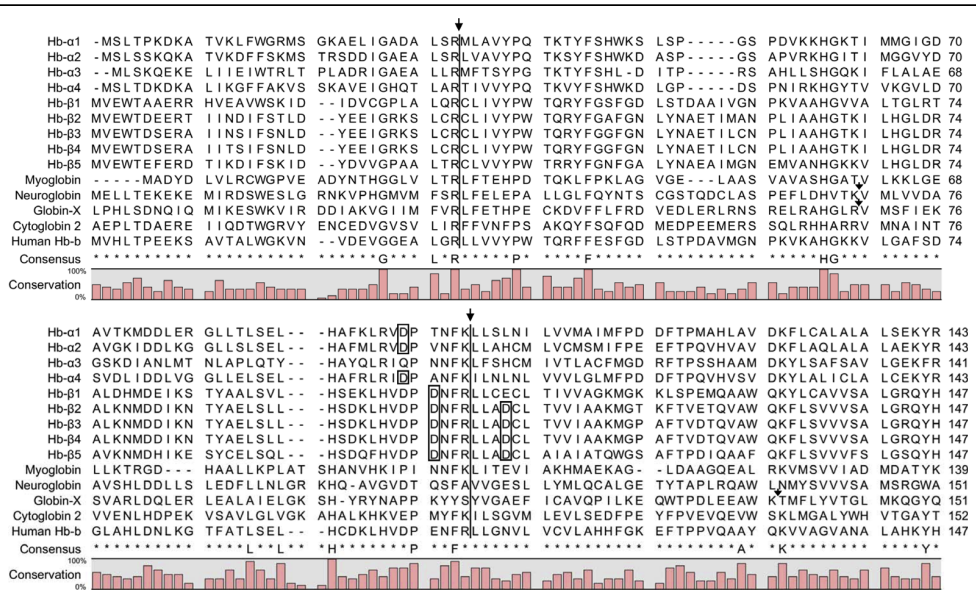

Figure 2 Sequence alignment of the Atlantic cod α-β globins, myoglobin, neuroglobin globin-X and cytoglobin 2. The sequences are based on the draft genome of the northeast Arctic population of Atlantic cod. Human β-globin is included for comparison. The alignment was optimized by omitting the N-and/or C-terminal sequences of the non-hemoglobins, and numbers refer to the residues presented. The consensus sequence shows residues with > 80% identity. Putative residues required for Root effect are boxed. GenBank accession numbers: α1 (ACJ66341), α2 (ACJ66342), α3 (ACV69832), α4 (ACV69835), β1 (ACV69840), β2 (ACJ66344), β3 (ACJ66345), β4 (ACJ66346), β5 (ACV69854). Introns are indicated by arrows.

additional *Mgrn1* genes (ENSARG00000018347, ENSDARG00000057481) are also linked to the LA globin locus.

Cod LA locus

The second cod globin cluster was shown to contain five hemoglobin genes in the order *β3-β4-α2-α3-β2* positioned within a region of about 12 kb in a scaffold spanning 381 kb (Figure 3). The tail-to-head organized pairs *β4-α2* and *α3-β2* are transcribed in opposite directions. The three exons encode the 147-aa long β-globins, while the predicted α3 contains only 141 aa compared to the other α-globins of 143 aa (Figure 2). The globin cluster is flanked on the leftward side by a 70-kb region harboring duplicated *AQP8* genes similar to the zebrafish locus, and the adjacent *ARHGAP17* and *LCMT* genes are conserved in the teleosts examined (Figure 3). A *RHBDF1*-like gene is juxtaposed to *FoxJ1* in the LA locus of only cod and pufferfish, whereas we found a *FoxJ1* gene coupled to the MC locus in stickleback (ENSGACG00000014879) and zebrafish (ENSDARG00000059545). We also identified paralogs of stickleback *RHBDF1* (ENSGACG00000004462), *ARHGAP17* (ENSGACG00000009145) and *FoxJ1 (*ENSGACG00000014879) linked on chromosome 5, which, however, contains no globin genes, whereas an *ARHGAP17* duplicate is coupled to the MC locus in pufferfish (ENSTING00000017988), zebrafish (ENSDARG00000075341) and medaka (ENSORLG00000009090). A second *ARHGAP17* gene was also identified in the cod genome, but we presently lack information about any linkage to the globin loci.

Other cod globin genes

Five additional globin genes encoding myoglobin, neuroglobin, globin-X and two cytoglobins were identified in the cod genome (Figure 2). The gene encoding the predicted cod myoglobin of 145 aa is organized as the α-β globins, while neuroglobin and globin-X of 159 and 197 aa, respectively, are encoded by four and five exons. The three exons of the cytoglobin-2 gene encode a protein of 202 aa, while the draft genome sequences contained only a partial cytoglobin-1 gene. The four α-globins are less similar (35-67% identity) than the five β-globins (57-99%) of which β2, β3 and β4 show high sequence identity. The α-globins share only 25-33% identity with the β-globins, compared to sequence identities of about 20% between the cod α-β globins and the other globins, except for the very low similarity with globin-X. Despite this low overall identity, highly conserved positions were identified throughout the aligned sequences, including human β-globin (Figure 2). Rare mutations in almost all these positions have been reported to affect the functionality of human hemoglobin [31], and suggest the importance of these residues for the proper structure and/or function of different oxygen-binding molecules in diverse vertebrate species.

Globin gene mapping and expression

The cod α-β globin clusters were mapped to different linkage groups by genotyping multiple single nucleotide polymorphic (SNP) markers, including the globin SNPs underlying the Metβ1Val and Thrα2Ile polymorphisms [27]. The segregation of the SNPs in full-sib cod families localized the MC and LA loci to linkage groups 17 and 16, respectively, among the total of 24 linkage groups [32].

The nine α-β globin genes were shown to be transcriptionally active by quantifying the mRNA levels throughout the life cycle of Atlantic cod using real-time qPCR (Figure 4). Intriguingly, *α1*, *α2*, *β1* and *β2* mRNAs were identified in unfertilized eggs, whereas fertilized eggs and early embryos contained mainly the *β5* transcript. The later stages of embryogenesis showed very low hemoglobin mRNA levels prior to the larval expression of several

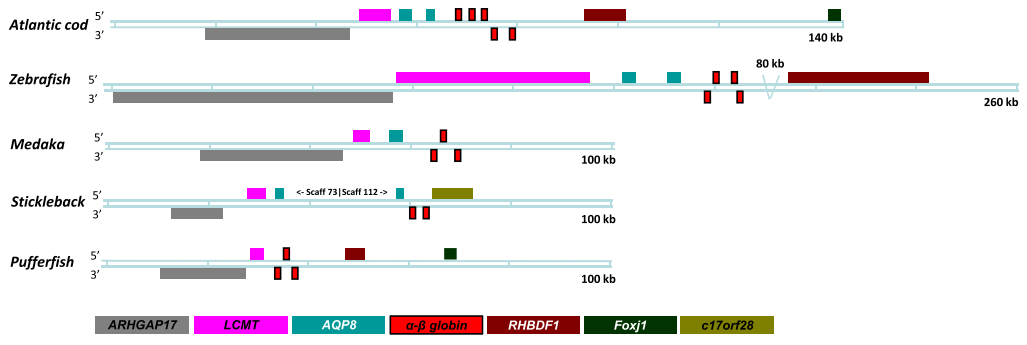

Figure 3 Conserved synteny at the LA locus in teleosts. Genomic organization of the Atlantic cod *β3-β4-α2-α3-β2* globin cluster compared with the counterpart in zebrafish, medaka, stickleback and pufferfish. Further details are as given in Figure 1.

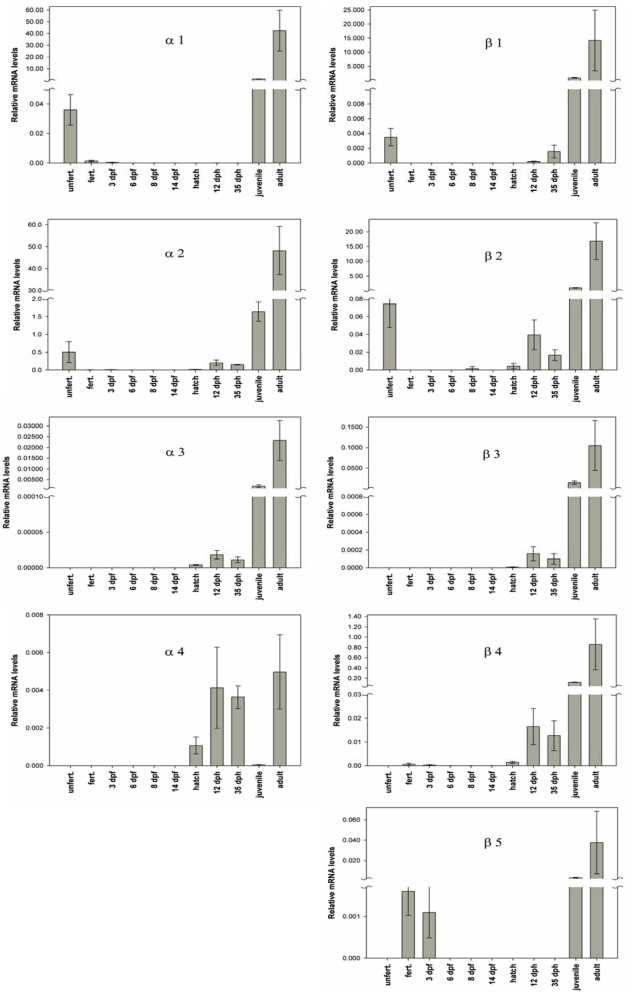

Figure 4 Real-time PCR analysis of globin gene expression in Atlantic cod. The globin mRNA levels are presented relative to the level of ubiquitin mRNA at each developmental stage examined. The juvenile and adult expression profiles include spleen and blood mRNAs, respectively. dpf, days post fertilization; dph, days post hatching.

α- and β-globin genes, and all hemoglobin genes were expressed in the juvenile and adult fish. Abundant expression of *α1*, *α2*, *β1* and *β2* was measured in the adult fish, while the other genes showed low mRNA levels (Figure 5).

Discussion

The Atlantic cod genome was shown to harbor altogether nine α- and β-globin genes organized in two unlinked clusters similar to the other teleost genomes available. The expression of many hemoglobin genes in adult cod is consistent with the multiple tetrameric hemoglobin types and subtypes identified by gel electrophoresis of blood proteins [33,34]. The cod hemoglobin repertoire is further extended by the polymorphic α1, β1, β3 and β4 globins [27,29] of which the functionally different variants of β1 are differentially distributed in cod populations [27,35,36]. The dominant expression of *α1*, *α2*, *β1* and *β2* in adult fish is in agreement with the isolation of three major tetramers designated Hb1, Hb2

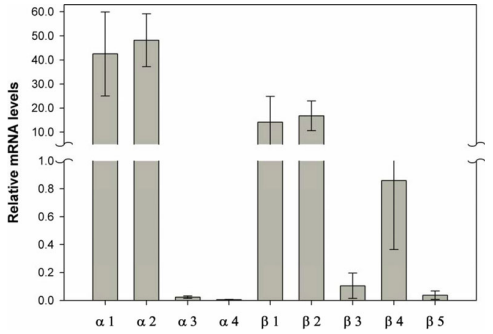

Figure 5 Comparison of globin mRNA levels quantified in adult Atlantic cod. See Figure 4 for details.

and Hb3, which comprise different combinations of these four subunits [37]. The tetrameric Hb3 (α1-α1-β2-β2) was shown to exhibit a marked Root effect of importance for the delivery of oxygen to the swim bladder for neutral buoyancy and to the retina for enhanced visual acuity via the highly specialized vascular structures [25,38]. The structural basis for this extreme acid-induced reduction in oxygen affinity is far from understood, but the putative key residues, including Asp95α, Asp99β and Asp101β [39,40], are conserved in the cod hemoglobins, except for β1 and α3. We therefore suggest that the β1-containing Hb1 tetramer (α1-α1-β1-β1) has no Root effect and might function as an emergency oxygen supplier when fish exercise vigorously.

The detection of hemoglobin mRNAs in unfertilized cod eggs is the first evidence of maternally inherited α-β globins, while Vlecken et al. [41] recently reported maternal transfer of myoglobin mRNA in zebrafish. The function of these oxygen-binding molecules in the early fish embryo is uncertain, as aerobic processes have been shown to continue in the zebrafish embryo after functional ablation of hemoglobin [42]. Hemoglobin-derived antimicrobial peptides expressed in the fish epithelium have been suggested to play a significant role in the non-specific immune response [43], together with maternally transferred transcripts encoding lysozyme and cathelicidin [44]. The very low embryonic expression of globin genes is consistent with the transparent hemolymph flowing through the heart, which starts contracting after embryogenesis is two-thirds completed [45]. Thus, the early larval expression of hemoglobins probably represents the initial stage of hemoglobin oxygen binding and coincides with gill development. The embryonic expression of *β5* and the dominant mRNA levels of *α4* at hatching are in agreement with the phylogenetic analysis grouping these genes together with other fish globins expressed in embryonic stages [29].

Duplication and loss of hemoglobin genes have apparently occurred within specific teleost sublineages and have resulted in a variable number of α- and β-globins as summarized in Figure 6. The LA locus comprises from two

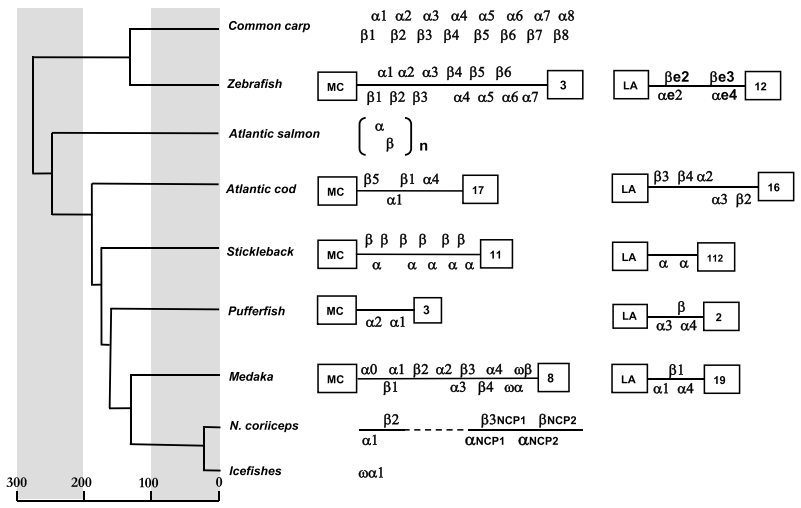

Figure 6 Comparison of the α-β globin genes in the MC and LA loci of diverse teleosts. Linkage between globin genes is represented by solid line (broken line indicates uncertainty). The genes are transcribed in the rightward (upper) or leftward (lower) direction. The linkage groups are numbered. The number of paired α-β genes in salmon is unknown (n). The estimated divergence times (MYA) are based on mitochondrial DNA sequences [64].

(stickleback) to five globins (cod), and the phylogenetic analysis of the highly similar β2, β3 and β4 globins in the cod cluster indicated a relatively recent gene duplication event in gadids [29]. Whereas the cod MC locus contains four globins, this cluster harbors up to 13 globins in zebrafish, stickleback and medaka. Maruyama et al. [24] suggested that the latter globin cluster originated from a subcluster duplication, while subsequent gene silencing is evidenced by the ϖβ-ϖα pseudogene pair in medaka. In pufferfish, the MC locus has been reduced to only two α-globins [6], while only remnants of an α-globin gene are found in icefishes inhabiting the cold Polar Ocean saturated with oxygen [46,47]. The metabolic functions are maintained in the hemoglobin-less icefishes by the elevated cardiac output of blood of low viscosity through the highly vascularized gills and skin [48]. Although the Arctic variant of cod β1 and a major β globin component of the pelagic Antarctic teleosts *Pagothenia borchgrevinki* and *Trematonus newnesi* share only 58% sequence identity, similar functional features of these globins were recently hypothesized based on their close position in the PC (principal component) plane in the hydrophobicity analysis of multiple fish globins [49].

The highly conserved linear order of the globin-flanking genes provides a strong anchor for inferring common ancestry of the vertebrate globin clusters. The proposed teleost-specific duplication of an ancient α-β globin cluster implies that paralogs of the flanking genes should still be present in both loci. *In silico* analysis of the teleost genomes available revealed linkage of *RHBDF1*, *ARHGAP17*, *Mgrn1*, *AQP8* and *FoxJ1* paralogs to the MC and LA loci in several species. Consistent with these findings, comparative gene mapping of medaka, zebrafish, pufferfish and human genomes demonstrated large conserved syntenic segments in paired fish chromosomes, including the globin-containing pairs of linkage groups 8 and 19 (medaka), 3 and 12 (zebrafish), and 2 and 3 (pufferfish) [50,51] (see Figure 6). Furthermore, we found evidence for the origin of the *RHBDF1-MPG-α-globin-ARHGAP17-LCMT1* syntenic region in man and chicken by screening the genomes of the tunicate *Ciona intestinalis* and the lancelet *Branchiostoma floridae* (amphioxus). Four *Ciona* globin genes designated *CinHb1-4* were shown to form a monophyletic group basal to the vertebrate hemoglobin, myoglobin and cytoglobin [52]. We recognized *CinHb3* (ENSCING00000006495) linked to *MPG* and *ARHGAP17* on chromosome 3q, while an additional *Ciona* globin gene (ENSCING00000002015) is coupled to *RHBDF1* on chromosome 1q (Figure 7). In amphioxus, we identified *RHBDF1* (position 17_000132), *MPG* (17_000133), *ARHGAP17* (17_000183) and *LCMT* (17_000184 and 17_000191) on the 4.2-Mb long scaffold 17, which has been localized to the chordate linkage group (CLG) 15 by FISH analysis [53]. We were, however, unable to position any of the multiple globin genes to the 16 scaffolds

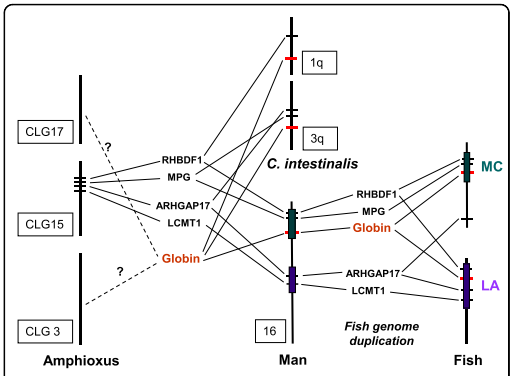

Figure 7 Proposed model for the evolution of the ancient globin cluster. Conserved synteny of globin-flanking genes in extant cephalochordate (amphioxus) and urochordate (*C. intestinalis*) species together with the MC (green) and LA (violet) globin loci in man and fish. The fusion of three chordate linkage groups (CLG) to form the homolog of the α-containing human chromosome 16 is based on the reported chromosomal rearrangements of the ancestral chordate genome [53].

spanning almost the complete CLG15. Based on conserved chromosomal segments of the amphioxus and human genomes, Putnam et al. [53] reconstructed a total of 17 ancestral CLGs of which CLG3, CLG15 and CLG17 showed syntenic association with the α-containing human chromosome 16. Although we presently lack information about any coupling of the amphioxus globin-like genes to these linkage groups, we propose that the fusion of CLG15 to CLG3 and CLG17 resulted in the linkage of the *RHBDF1-MPG-ARHGAP17-LCMT* region to globin gene (s) as illustrated in Figure 7. The identification of remnants of this globin linkage in the *Ciona* genome indicates that the proposed chromosomal rearrangement occurred prior to the divergence of the vertebrates and urochordates about 800 MYA [54]. Thus, the formation of this syntenic region seems to have coincided with a period of Earth history characterized by a rise in atmospheric oxygen from 0.02-0.04 atm 850 MYA to present day levels of 0.2 atm 540 MYA [55]. The increased oxygen content would be expected to have a strong impact on the regulation and structure of H_2S-binding globins. In sulfide-rich environments, the unusual sulfide-binding function is found in annelid globins containing key cysteine residues, which are absent in annelid globins from sulfide-free environments [56]. Concomitant with increased atmospheric oxygen, the role of globins as oxygen scavengers would probably be lost in oxygen-tolerant organisms to function as oxygen-transporting hemoglobins. Based on the close phylogenetic relation of cyclostome hemoglobins to gnathostome cytoglobins, the ancestors of cyclostome and gnathostome vertebrates were recently stated to have independently

invented erythroid-specific oxygen-transporting hemoglobins about 450-600 MYA [57]. The transcriptional regulation of the hemoglobins in extant vertebrates involves both proximal promoters and distant enhancers [58]. In mouse erythroid cells, the active *α1* and *α2* genes are in close spatial proximity of the flanking *RHBDF1*, *MPG* and *c16orf35*, including the *cis*-regulatory MREs, as the result of erythroid-specific changes in the chromatin conformation [59]. The chromosomal rearrangements forming this highly conserved syntenic region seem to have occurred more than 800 MYA, and we therefore propose that the molecular mechanism underlying the oxygen-dependent regulation of globin expression evolved prior to the structural changes in the duplicated ancestral globins.

Conclusions

In contrast to the low number of globin genes reported in Antarctic teleosts [23], the adaptation of Atlantic cod to fluctuating environmental conditions probably involved the evolution of multiple globins with potentially different oxygen binding properties. The unlinked globin pairs *α1-β1* and *α2-β2* are abundantly expressed in the adult fish and form three major hemoglobin tetramers with different Root effect. The identification of paralogous genes in the flanking regions of the two globin clusters in diverse teleosts supports the proposed teleost-specific duplication of the vertebrate globin cluster. Based on the conserved synteny of globin-flanking genes in extant urochordate and cephalochordate species, we hypothesize that the ancestral globin cluster contained both the MC and LA loci, and was formed by the fusion of three chordate chromosomes. We propose that these chromosomal rearrangements facilitated the transcriptional regulation of globin synthesis to cope with increased atmospheric oxygen content about 850 MYA. Thus, these regulatory changes probably preceded the convergent evolution of different ancestral globins to function as erythroid-specific oxygen transporting hemoglobins.

Methods

Identification of globin clusters

PCR

Forward and reverse PCR primers were designed to amplify pairs of α-β genes using cod globin gene sequences available in GenBank (Table. 1). PCR was performed under standard conditions (Applied Biosystems 2720 thermal cycler) using genomic DNA as template (Qiagen DNeasy blood & tissue kit). The amplified products were ligated into the pGEM®-T easy vector (Promega) and sequenced in both directions (Applied Biosystems 3130xl genetic analyzer).

BAC library screening

A cod BAC library consisting of 92,000 clones with average insert size of 125 kb was screened for globin genes by PCR using gene specific primers (Table 2) on pools and super-pools of BAC clones. Positive BAC clones were purified (NucleoBond BAC 100), and sequenced using the 454 GS FLX instrumentation at the Norwegian Sequencing Center (http://www.sequencing.uio.no). The resulting reads were assembled using Newbler v. 2.0 (gsAssembler) [60], using default settings and filtering of the reads against contaminating *E. coli* genomic sequences. The pufferfish globin loci (AY016023, *Sphoeroides nephelus;* AY016024, *Takifugu rubripes*) were utilized as query sequences in BLAST searches of the assembled contigs.

Table 1 PCR primers for amplification of cod α1-β1 gene pair and for screening BAC library

Gene	Name	Sequence (5' to 3')
α1-β1	A1-B1F	GCAAATTGTTCAAGTTATTCCCCCTAAC
	A1-B1R	TAAAGACTGACCTGCAACGCGAGTGGT
α1	A1-bacF	CAGACCAAGACTTACTTCAGCC
	A1-bacR	GCTCGCTCAGAGTGAGAAGAC
α2	A2-bacF	CCGATGATATCGGAGCTGAGG
	A2-bacR	CTAAGGCTGAGGAGTCCTCC
β1	B1-bacF	ATGGTTGAATGGACAGCTGC
	B1-bacR	GTCGACGTGCAGTTTCTC
β2	B2-bacF	TGGACAGATAGTGAGCGCG
	B2-bacR	AGTGGAGCAGAGACAGCTC

Cod genome BLAST

The Atlantic cod genome project (http://www.codgenome.no) is based on the genome sequences of the north-east Arctic cod population. Scaffold sequences harboring globin genes were identified among the assemblies of the cod genome project [30] using the BLAST search tool at http://www.bioportal.uio.no. Annotation of genes located on the scaffolds was completed based on results from TBLASTN searches of known protein sequences from related species, using the bioinformatics software CLC genomics workbench (CLC bio).

Chordate genome BLAST

Conserved globin and globin-flanking genes were identified in cephalochordate and urochordate species by BLAST searching the genomes of *Branchiostoma floridae* (version 1.0, http://genome.jgi-psf.org/Brafl1/Brafl1.home.html) and *Ciona intestinalis* (release 43, http://www.ensembl.org/Ciona_intestinalis/Info/Index).

Real-time qPCR

Fish

Spleen and blood were sampled from juvenile (n = 5) and adult (n = 12) fish kept at the National Cod Breeding Centre (Kraknes, Tromsø, Norway) and the University of Bergen, respectively. Sexually mature fish were hand-stripped, and eggs were fertilized *in vitro*. The

Table 2 Primers for real-time qPCR, amplification efficiency (%) and amplicon size (bp)

Gene	Name	Sequence (5' to 3')	Efficiency	Size
α1	A1F	GACTTACTTCAGCCACTGGAAGAGCCT C	96	153
	A1R	TTGAAGGCGTGCAGCTCGCTCAGAG		
α2	A2F	GTCCTATTTCTCTCACTGGAAGGACGCG	85	153
	A2R	ATGAACGCGTGCAGCTCGCTAAGGC		
α3	A3F	CACATCATACCCTGGCACCAAGAC	95	172
	A3R	CTGGTAGGCGTGGTAGGTTTGAAGAG		
α4	A4F	TTCTCCCACTGGAAAGACCTCGG	70	138
	A4R	ATGGAGCTCACTGAGCTCGAGAAG		
β1 allele A	B1FA	TTATGGGAAACCCCAAGGTGGCCAA	91	131
	B1R	GTGCAGTTTCTCGGAGTGCAGCACGC		
β1 allele B	B1FB	TTGTGGGAAACCCCAAGGTGGCTGC	98	131
	B1R	GTGCAGTTTCTCGGAGTGCAGCACGC		
β2	B2F	CCTGTACAATGCAGAGACCATCATGGC	84	151
	B2R	GTGCAGCTTGTCAGAGTGGAGCAGAG		
β3	B3F	ACAGATAGTGAGCGCGCCATCATTAA	86	176
	B34R	GCGGCGATCAGGGGGTTGCACAG		
β4	B4F	ACAGATAGTGAGCGCGCCATCATTAC	95	176
	B34R	GCGGCGATCAGGGGGTTGCACAG		
β5	B5F	GTGGACTCGGAGGTACTTTGGAAAC	89	168
	B5R	TGCAGCTGACTGAGCTCGCAATAG		
Ubiquitin	UbiF	GGCCGCAAAGATGCAGAT	81	69
	UbiR	CTGGGCTCGACCTCAAGAGT		

incubation of embryos and feeding of larvae were carried out as described [44]. Sampling of unfertilized eggs, fertilized eggs and larvae was performed during 10 weeks. All samples were rapidly submerged in RNAlater (Ambion, Austin, TX, USA) and incubated at 4°C overnight, then stored at -20°C.

RNA isolation and cDNA synthesis

5-10 eggs/embryos or 3-5 larvae were pooled and homogenized in 1.5 ml microcentrifuge tubes containing lysis buffer (Qiagen RNeasy mini kit) using a plastic pestle. After centrifugation through a QiaShredder column (Qiagen, Hilden, Germany), RNA was isolated according to the manufacturer's protocol (Qiagen RNeasy mini kit), and followed by the recommended on-column DNase treatment. The Qiagen RNeasy mini kit was also used for the spleen and blood samples from juvenile and adult fish, respectively. cDNA was synthesized from 1 μg total RNA using TaqMan® Reverse Transcription Reagents (Applied Biosystems) and oligo-dT primer in 20 μl reactions using the conditions of: 25°C for 10 min, 48°C for 30 min and 95°C for 5 min. Primers used for real-time qPCR were adopted from Borza et al. [29] for the globins, while *ubiquitin* primers were taken from Olsvik et al. [61] (Table 2). For the *β1* gene, two allele-specific primer sets were used on all samples, and relative expression was calculated dependent on the actual genotype of each sample. Ten-fold dilution series were prepared to generate standard curves, and PCR efficiencies and relative quantification results were calculated according to Ståhlberg et al. [62] using ubiquitin as the reference transcript [63]. Cycling parameters were 50°C for 2 min, 95°C for 10 min, 40 cycles of 95°C for 15 sec, 61°C for 1 min, including a final dissociation stage to yield melting curves. Reactions of 25 μl consisted of 12.5 μl 2× Power SYBR®Green PCR Master Mix (Applied Biosystems), 0.5 μl each of sense and antisense primers (10 μM) and 11.5 μl of 50× diluted cDNA.

List of abbreviations

MPG: N-methylpurine-DNA glycosylase; ARHGAP17: Rho GTPase activating protein 17; RHBDF1: rhomboid 5 homolog 1; LCMT: leucine carboxyl methyltransferase; c16orf35: human chromosome 16 open reading frame 35; POLR3K: DNA-directed RNA polymerase III subunit RPC10; Mgrn1: mahogunin Ring Finger 1; AANAT: arylalkylamine N-acetyltransferase; DOCK: dedicator of cytokinesis; ANKRD: ankyrin repeat domain; AQP: aquaporin; FoxJ1: fork head J1; PC: principial component; FISH: fluorescence *in situ* hybridization; CLG: chordate linkage group; dpf: days post fertilization; dph: days post hatching.

Acknowledgements

We thank Carl Andrè and two anonymous reviewers for helpful comments and suggestions. This work was funded by grants to Ø.A., A.N. and K.S.J. from the Norwegian Research Council and by a PhD stipend to O.F.W financed by The Norwegian Ministry of Education and Research.

Author details

[1]Department of Animal and Aquacultural Sciences, University of Life Sciences, P.O. Box 5003, N-1430 Aas, Norway. [2]Department of Natural Sciences and Technology, Hedmark University College, P.O. Box 4010 Bedriftsenteret, N-2306 Hamar, Norway. [3]Centre for Ecological and

Evolutionary Synthesis (CEES), Department of Biology, University of Oslo, P.O. Box 1066 Blindern, N-0316 Oslo, Norway. [4]Institute of Marine Research, P.O. Box 1870 Nordnes, N-5817 Bergen, Norway. [5]Nofima Marine, P.O. Box 5010, N-1430 Aas, Norway.

Authors' contributions

OFW carried out the majority of the analyses. AJN performed the sequencing and assembly of the reads. RW participated in the real-time PCR analysis. KSJ participated in the design of the study and the sequencing. RBE screened the BAC library. ØA conceived and designed the study, and wrote the manuscript. All authors critically read the manuscript drafts and approved the final version of the manuscript.

Received: 18 June 2010 Accepted: 20 October 2010
Published: 20 October 2010

References

1. Goodman M, Moore GW, Matsuda G: **Darwinian evolution in the genealogy of haemoglobin.** *Nature* 1975, **253**:603-608.
2. Jeffreys AJ, Wilson V, Wood D, Simons JP, Kay RM, Williams JG: **Linkage of adult α- and β-globin genes in X. laevis and gene duplication by tetraploidization.** *Cell* 1980, **21**:555-64.
3. Efstratiadis A, Posakony JW, Maniatis T, Lawn RM, O'Connell C, Spritz RA, Deriel JK, Forget BG, Weissman SM, Slightom JL, Blechl AE, Smithies O, Baralle FE, Shoulders CC, Proudfoot NJ: **The structure and evolution of the human β-globin gene family.** *Cell* 1980, **21**:653-668.
4. Lauer J, Shen CK, Maniatis T: **The chromosomal arrangement of human alpha-like globin genes: sequence homology and alpha-globin gene deletions.** *Cell* 1980, **20**:119-130.
5. Hardison RC: **Globin genes on the move.** *J Biol* 2008, **7**:35.
6. Flint J, Tufarelli C, Peden J, Clark K, Daniels RJ, Hardison R, Miller W, Philipsen S, Tan-Un KC, McMorrow T, Frampton J, Alter BP, Frischauf AM, Higgs DR: **Comparative genome analysis delimits a chromosomal domain and identifies key regulatory elements in the alpha globin cluster.** *Hum Mol Genet* 2001, **10**:371-382.
7. Wheeler D, Hope RM, Cooper SJ, Gooley AA, Holland RA: **Linkage of the beta-like omega-globin gene to alpha-like globin genes in an Australian marsupial supports the chromosome duplication model for separation of globin gene clusters.** *J Mol Evol* 2004, **58**:642-52.
8. Patel VS, Cooper SJ, Deakin JE, Fulton B, Graves T, Warren WC, Wilson RK, Graves JA: **Platypus globin genes and flanking loci suggest a new insertional model for beta-globin evolution in birds and mammals.** *BMC Biol* 2008, **6**:34.
9. Bulger M, van Doorninck JH, Saitoh N, Telling A, Farrell C, Bender MA, Felsenfeld G, Axel R, Groudine M: **Conservation of sequence and structure flanking the mouse and human beta-globin loci: the beta-globin genes are embedded within an array of odorant receptor genes.** *Proc Natl Acad Sci USA* 1999, **96**:5129-5134.
10. Gillemans N, McMorrow T, Tewari R, Wai AW, Burgtorf C, Drabek D, Ventress N, Langeveld A, Higgs D, Tan-Un K, Grosveld F, Philipsen S: **Functional and comparative analysis of globin loci in pufferfish and humans.** *Blood* 2003, **101**:2842-2849.
11. Wagner A, Deryckere F, McMorrow T, Gannon F: **Tail-to-tail orientation of the Atlantic salmon alpha- and beta-globin genes.** *J Mol Evol* 1994, **38**:28-35.
12. McMorrow T, Wagner A, Deryckere F, Gannon F: **Structural organization and sequence analysis of the globin locus in Atlantic salmon.** *DNA and Cell Biol* 1996, **15**:407-414.
13. Miyata M, Aoki T: **Head-to-head linkage of carp α- and β-globin genes.** *Biochim Biophys Acta* 1997, **1354**:127-133.
14. Lau DT, Saeed-Kothe A, Parker SK, Detrich HW: **Adaptive evolution of gene expression in Antarctic fishes: Divergent transcription of the 5'-to-5' linked adult α1- and β-globin genes of the Antarctic teleost Notothenia coriiceps is controlled by dual promoters and intergenic enhancers.** *Amer Zool* 2001, **41**:113-132.
15. Maruyama K, Yasumasu S, Naruse K, Mitani H, Shima A, Iuchi I: **Genomic organization and developmental expression of globin genes in the teleost *Oryzias latipes*.** *Gene* 2004, **335**:89-100.
16. Fuchs A, Burmester T, Hankeln T: **The amphibian globin gene repertoire as revealed by the *Xenopus* genome.** *Cytogenet Genome Res* 2006, **112**:296-306.
17. Jennings MW, Jones RW, Wood WG, Weatherall DJ: **Analysis of an inversion within the human beta globin gene cluster.** *Nucl Acids Res* 1985, **13**:2897-2906.
18. Perutz MF: **Species adaptation in a protein molecule.** *Mol Biol Evol* 1983, **1**:1-28.
19. Weber RE, Fago A: **Functional adaptation and its molecular basis in vertebrate hemoglobins, neuroglobins and cytoglobins.** *Resp Physiol Neurobiol* 2004, **144**:141-159.
20. Weber RE: **Adaptations for oxygen transport: lessons from fish hemoglobins.** In *Hemoglobin Function in Vertebrates, Molecular Adaptation in Extreme and Temperate Environments.* Edited by: Di Prisco G, Giardina B, Weber RE. Milano: Springer-Verlag; 2000:23-37.
21. Ruud JT: **Vertebrates without erythrocytes and blood pigment.** *Nature* 1954, **173**:848-850.
22. Bargelloni L, Marcato S, Patarnello T: **Antarctic fish hemoglobins: Evidence for adaptive evolution at subzero temperature.** *Proc Natl Acad Sci USA* 1998, **95**:8670-8675.
23. di Prisco G, Ennio C, Parker SK, Detrich HW III: **Tracking the evolutionary loss of hemoglobin expression by the white-blooded Antarctic icefishes.** *Gene* 2002, **295**:185-191.
24. Maruyama K, Yasumasu S, Iuchi I: **Evolution of globin genes of the medaka Oryzias latipes (Euteleostei; Beloniformes; Oryziinae).** *Mech Dev* 2004, **121**:753-769.
25. Berenbrink M, Koldkjær P, Kepp O, Cossins AR: **Evolution of oxygen secretion in fishes and the emergence of a complex physiological system.** *Science* 2005, **307**:1752-1757.
26. van der Kooij J, Righton D, Strand E, Michalsen K, Thorsteinsson V, Svedäng H, Neat FC, Neuenfeldt S: **Life under pressure: insights from electronic data-storage tags into cod swimbladder function.** *ICES J Mar Sci* 2007, **64**:1293-1301.
27. Andersen Ø, Wetten OF, De Rosa MC, Andre C, Carelli Alinovi C, Colafranceschi M, Brix O, Colosimo A: **Hemoglobin polymorphisms affect the oxygen binding properties in Atlantic cod populations.** *Proc Royal Soc B* 2009, **276**:833-841.
28. Halldorsdottir K, Arnason E: **Organization of a β and α globin gene set in the teleost Atlantic cod, *Gadus morhua*.** *Biochem Genet* 2009, **47**:817-830.
29. Borza T, Stone C, Gamperl AK, Bowman S: **Atlantic cod (*Gadus morhua*) hemoglobin genes: multiplicity and polymorphism.** *BMC Genetics* 2009, **10**:51.
30. Johansen SD, Coucheron DH, Andreassen M, Karlsen BO, Furmanek T, Jorgensen TE, Emblem A, Breines R, Nordeide JT, Moum T, Nederbragt AJ, Stenseth NC, Jakobsen KS: **Large-scale sequence analyses of Atlantic cod.** *New Biotechnology* 2009, **25**:263-271.
31. Hardison RC, Chui DH, Riemer C, Giardine B, Lehväslaiho H, Wajcman H, Miller W: **Databases of human hemoglobin variants and other resources at the globin gene server.** *Hemoglobin* 2001, **25**:183-190.
32. Moen T, Delghandi M, Wesmajervi MS, Westgaard JI, Fjalestad KT: **A SNP/microsatellite genetic linkage map of the Atlantic cod (*Gadus morhua*).** *Animal Genetics* 2009, **40**:993-996.
33. Fyhn UEH, Brix O, Nævdal G, Johansen T: **New variants of the haemoglobins of Atlantic cod: a tool for discriminating between coastal and Arctic cod populations.** *ICES Mar Sci Symp* 1994, **198**:666-670.
34. Husebø Å, Imsland AK, Nævdal G: **Haemoglobin variation in cod: a description of new variants and their geographical distribution.** *Sarsia* 2004, **90**:1-11.
35. Karpov AK, Novikov GG: **The hemoglobin aloforms in cod (*Gadus morhua* L.), their functional characteristics and distribution in the populations.** *J Ichthyol* 1980, **6**:45-50.
36. Brix O, Thorkildsen S, Colosimo A: **Temperature acclimation modulates the oxygen binding properties of the Atlantic cod (*Gadus morhua* L.) genotypes-Hbl*1/1, Hbl*1/2, and Hbl*2/2 by changing the concentrations of their major hemoglobin components (results from growth studies at different temperatures).** *Comp Biochem Physiol* 2004, **138A**:241-251.
37. Verde C, Balesrieri M, de Pascale D, Pagnozzi D, Lecointre G, di Prisco G: **The oxygen transport system in three species of the boreal fish family Gadidae.** *J Biol Chem* 2006, **281**:22073-22084.
38. Scholander PF, van Dam L, Enns T: **The source of oxygen secreted into the swimbladder of cod.** *J Cell Comp Physiol* 2005, **48**:517-522.

39. Yokoyama T, Chong KT, Miyazaki G, Morimoto H, Shih DT, Unzai S, Tame JR, Park SY: **Novel mechanisms of pH sensitivity in tuna hemoglobin: a structural explanation of the Root effect.** *J Biol Chem* 2004, **279**:28632-29640.
40. Mazzarella L, Vergara A, Vitagliano L, Merlino A, Bonomi G, Scala S, Verde C, di Prisco G: **High resolution crystal structure of deoxy hemoglobin from Trematomus bernacchii at different pH values: the role of histidine residues in modulating the strength of the root effect.** *Proteins* 2006, **65**:490-498.
41. Vlecken DH, Teserink J, Ott EB, Sakalis PA, Jaspers RT, Bagowski CP: **A critical role for myoglobin in zebrafish development.** *I J Dev Biol* 2009, **53**:517-524.
42. Pelster B, Burggren WW: **Disruption of hemoglobin oxygen transport does not impact oxygen-dependent physiological processes in developing embryos of zebrafish (*Danio rerio*).** *Circulation Res* 1996, **79**:358-262.
43. Ullal AJ, Litaker RW, Noga EJ: **Antimicrobial peptides derived from haemoglobin are expressed in epithelium of channel catfish (*Ictalurus punctatus*, Rafinesque).** *Dev Comp Immun* 2008, **32**:1301-1312.
44. Seppola M, Johnsen H, Mennen S, Myrnes B, Tveiten H: **Maternal transfer and transcriptional onset of immune genes during ontogenesis in Atlantic cod.** *Dev Comp Immunol* 2009, **33**:1205-1211.
45. Hall TE, Smoth P, Johnston IA: **Stages of embryonic development in the Atlantic cod *Gadus morhua*.** *J Morphology* 2004, **259**:255-270.
46. Cocca E, Ratnayake-Lecamwasam M, Parker SK, Camardella L, Ciaramella M, di Prisco G, Detrich HW: **Genomic remnants of alpha-globin genes in the hemoglobinless antarctic icefishes.** *Proc Natl Acad Sci USA* 1995, **92**:1817-21.
47. di Prisco G, Eastman JT, Giordano D, Parisi E, Verde C: **Biogeography and adaptation of Notothenioid fish: hemoglobin function and globin-gene evolution.** *Gene* 2007, **398**:143-155.
48. Kock KH: **Antarctic icefishes (Channichthyidae): a unique family of fishes. A review, Part II.** *Polar Biol* 2005, **28**:897-909.
49. Colafranceschi M, Giuliani A, Andersen Ø, Brix O, De Rosa MC, Giardina B, Colosimo A: **Hydrophobicity patterns and biological adaptation: an exemplary case from fish hemoglobins.** *OMICS* 2010, **14**:275-281.
50. Kasahara M, Naruse K, Sasaki S, Nakatani Y, Qu W, Ahsan B, Yamada T, Nagayasu Y, Doi K, Kasai Y, Jindo T, Kobayashi D, Shimada A, Toyoda A, Kuroki Y, Fujiyama A, Sasaki T, Shimizu A, Asakawa S, Shimizu N, Hashimoto S, Yang J, Lee Y, Matsushima K, Sugano S, Sakaizumi M, Narita T, Ohishi K, Haga S, Ohta F, Nomoto H, Nogata K, Morishita T, Endo T, Shin-I T, Takeda H, Morishita S, Kohara Y: **The medaka draft genome and insights into vertebrate genome evolution.** *Nature* 2007, **447**:714-719.
51. Naruse K, Tanaka M, Mita K, Shima A, Postlethwaite J, Mitani H: **A medaka gene map: the trace of ancestral vertebrate proto-chromosomes revealed by comparative gene mapping.** *Genome Res* 2004, **14**:820-828.
52. Ebner B, Burmester T, Hankeln T: **Globin genes are present in *Ciona intestinalis*.** *Mol Biol Evol* 2003, **20**:1521-1525.
53. Putnam NH, Butts T, Ferrier DE, Furlong RF, Hellsten U, Kawashima T, Robinson-Rechavi M, Shoguchi E, Terry A, Yu JK, Benito-Gutiérrez EL, Dubchak I, Garcia-Fernàndez J, Gibson-Brown JJ, Grigoriev IV, Horton AC, de Jong PJ, Jurka J, Kapitonov VV, Kohara Y, Kuroki Y, Lindquist E, Lucas S, Osoegawa K, Pennacchio LA, Salamov AA, Satou Y, Sauka-Spengler T, Schmutz J, Shin-I T, Toyoda A, Bronner-Fraser M, Fujiyama A, Holland LZ, Holland PW, Satoh N, Rokhsar DSP: **The amphioxus genome and evolution of the chordate karyotype.** *Nature* 2008, **453**:1064-1072.
54. Blair JE, Hedges SB: **Molecular phylogeny and divergence times of deuterostome animals.** *Mol Biol Evol* 2005, **22**:2275-2284.
55. Holland H: **The oxygenation of the atmosphere and oceans.** *Phil Trans Roy Soc B* 2006, **361**:903-910.
56. Bailly X, Leroy R, Carney S, Collin O, Zal F, Toulmond A, Jollivet D: **The loss of the hemoglobin H_2S-binding function in annelids from sulfide-free habitats reveals molecular adaptation driven by Darwinian positive selection.** *Proc Natl Acad Sci USA* 2003, **100**:5885-5890.
57. Hoffmann FG, Opazo JC, Storz JF: **Gene cooption and convergent evolution of oxygen transport hemoglobins in jawed and jawless vertebrates.** *Proc Natl Acad Sci USA* 2010, **107**:14274-14279.
58. Hardison R: **Hemoglobins from bacteria to man: evolution of different patterns of gene expression.** *J Exp Biol* 1998, **201**:1099-1117.
59. Zhou GL, Xin L, Song W, Di LJ, Liu G, Wu XS, Liu DP, Liang CC: **Active chromatin hub of the mouse alpha-globin locus forms in a transcription factory of clustered housekeeping genes.** *Mol Cell Biol* 2006, **26**:5096-5105.
60. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer ML, Jarvie TP, Jirage KB, Kim JB, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF, Rothberg JM: **Genome sequencing in microfabricated high-density picolitre reactors.** *Nature* 2005, **437**:376-380.
61. Olsvik PA, Søfteland L, Lie KK: **Selection of reference genes for qRT-PCR examination of wild populations of Atlantic cod *Gadus morhua*.** *BMC Res Notes* 2008, **1**:47.
62. Ståhlberg A, Åman P, Ridell B, Mostad P, Kubista M: **Quantitative real-time PCR method for detection of B-lymphocyte monoclonality by comparison of κ and λ immunoglobulin light chain expression.** *Clin Chem* 2003, **49**:51-59.
63. Sæle Ø, Nordgreen A, Hamre K, Olsvik PA: **Evaluation of candidate reference genes in Q-PCR studies of Atlantic cod (*Gadus morhua*) ontogeny, with emphasis on the gastrointestinal tract.** *Comp Biochem Physiol B* 2009, **152**:94-101.
64. Azuma Y, Kumazawa Y, Miya M, Mabuchi K, Nishida M: **Mitogenomic evaluation of the historical biogeography of cichlids toward reliable dating of teleostean divergences.** *BMC Evol Biol* 2008, **8**:215.

doi:10.1186/1471-2148-10-315
Cite this article as: Wetten *et al.*: Genomic organization and gene expression of the multiple globins in Atlantic cod: conservation of globin-flanking genes in chordates infers the origin of the vertebrate globin clusters. *BMC Evolutionary Biology* 2010 **10**:315.

Paper III

LETTER

doi:10.1038/nature10342

The genome sequence of Atlantic cod reveals a unique immune system

Bastiaan Star[1], Alexander J. Nederbragt[1], Sissel Jentoft[1], Unni Grimholt[1], Martin Malmstrøm[1], Tone F. Gregers[2], Trine B. Rounge[1], Jonas Paulsen[1,3], Monica H. Solbakken[1], Animesh Sharma[4], Ola F. Wetten[5,6], Anders Lanzén[7,8], Roger Winer[9], James Knight[9], Jan-Hinnerk Vogel[10], Bronwen Aken[10], Øivind Andersen[11], Karin Lagesen[1], Ave Tooming-Klunderud[1], Rolf B. Edvardsen[12], Kirubakaran G. Tina[1,13], Mari Espelund[1], Chirag Nepal[4,8], Christopher Previti[8], Bård Ove Karlsen[14], Truls Moum[14], Morten Skage[1], Paul R. Berg[1], Tor Gjøen[15], Heiner Kuhl[16], Jim Thorsen[17], Ketil Malde[12], Richard Reinhardt[16], Lei Du[9], Steinar D. Johansen[14,18], Steve Searle[10], Sigbjørn Lien[13], Frank Nilsen[19], Inge Jonassen[4,8], Stig W. Omholt[1,13], Nils Chr. Stenseth[1] & Kjetill S. Jakobsen[1]

Atlantic cod (*Gadus morhua*) is a large, cold-adapted teleost that sustains long-standing commercial fisheries and incipient aquaculture[1,2]. Here we present the genome sequence of Atlantic cod, showing evidence for complex thermal adaptations in its haemoglobin gene cluster and an unusual immune architecture compared to other sequenced vertebrates. The genome assembly was obtained exclusively by 454 sequencing of shotgun and paired-end libraries, and automated annotation identified 22,154 genes. The major histocompatibility complex (MHC) II is a conserved feature of the adaptive immune system of jawed vertebrates[3,4], but we show that Atlantic cod has lost the genes for MHC II, CD4 and invariant chain (Ii) that are essential for the function of this pathway. Nevertheless, Atlantic cod is not exceptionally susceptible to disease under natural conditions[5]. We find a highly expanded number of MHC I genes and a unique composition of its Toll-like receptor (TLR) families. This indicates how the Atlantic cod immune system has evolved compensatory mechanisms in both adaptive and innate immunity in the absence of MHC II. These observations affect fundamental assumptions about the evolution of the adaptive immune system and its components in vertebrates.

We sequenced the genome of a heterozygous male Atlantic cod (NEAC_001, Supplementary Notes 1 and 2), applying a whole-genome shotgun approach to 40× coverage (estimated genome size of 830 megabases (Mb), Supplementary Note 4 and Supplementary Fig. 2) using 454 technology (Supplementary Note 3). Two programs (Newbler[6] and Celera[7], Supplementary Notes 5 and 6) produced assemblies with short contigs, yet with scaffolds of comparable size to those of Sanger-sequenced teleost genomes (Supplementary Note 10 and Supplementary Fig. 8). Although fragmentation due to short tandem repeats is difficult to address (Supplementary Note 7), we resolved numerous gaps attributable to heterozygosity (Supplementary Note 8). The assemblies differ in scaffold and contig length (Table 1), although their scaffolds align to a large extent (Supplementary Note 9 and Supplementary Fig. 7). We obtained about one million single nucleotide polymorphisms (SNPs) by mapping 454 and Illumina reads from the sequenced individual to the Newbler assembly (Supplementary Note 11). Both assemblies cover more than 98% of the reads from an extensive transcriptome data set, indicating that the proteome is well represented (Supplementary Note 13). The assemblies are consistent with four independently assembled bacterial artificial chromosome (BAC) insert clones (Supplementary Note 14 and Supplementary Fig. 9), and with the expected insert size of paired BAC-end reads (Supplementary Note 15 and Supplementary Fig. 10).

A standard annotation approach based on protein evidence was complemented by a whole-genome alignment of the Atlantic cod with the stickleback (*Gasterosteus aculeatus*), after repeat-masking 25.4% of the Newbler assembly (Supplementary Note 16 and Supplementary Table 6). In this way, 17,920 out of 20,787 protein-coding stickleback genes were mapped onto reorganized scaffolds (Supplementary Note 17). Additional protein-coding genes, pseudogenes and non-coding RNAs were annotated using the standard Ensembl pipeline. These approaches resulted in a final gene set of 22,154 genes (Supplementary Table 7). Comparative analysis of gene ontology classes indicates that the major functional pathways are represented in the annotated gene set (Supplementary Note 18 and Supplementary Fig. 11). We anchored 332 Mb of the Newbler assembly to 23 linkage groups of an existing Atlantic cod linkage map using 924 SNPs[8] (Supplementary Note 19 and Supplementary Table 8). These linkage groups have distinct orthology to chromosomes of other teleosts, on the basis of the number of co-occurring genes, showing that the whole-genome shotgun assembly reflects the expected chromosomal ancestry (Fig. 1, Supplementary Note 20 and Supplementary Table 9).

Table 1 | Assembly statistics

	Number	Bases (Mb)	N50L (bp)*	N50 (*n*)†	ML (bp)‡
Newbler					
Contigs§	284,239	536	2,778	50,237	76,504
Scaffolds	6,467	611	687,709	218	4,999,318
Entire assembly‖	157,887	753	459,495	344	4,999,318
Celera					
Contigs§	135,024	555	7,128	19,938	117,463
Scaffolds	3,832	608	488,312	373	2,810,583
Entire assembly‖	17,039	629	469,840	395	2,810,583

*Minimum sequence length in which half of the assembled bases occur.
†Number of sequences with lengths of N50L or longer.
‡Maximum length.
§Contigs longer than 500 bp.
‖Scaffolds and unplaced contigs.

[1]Centre for Ecological and Evolutionary Synthesis (CEES), Department of Biology, University of Oslo, PO Box 1066, Blindern, N-0316 Oslo, Norway. [2]Department of Molecular Biosciences, Centre for Immune Regulation, University of Oslo, Blindern, N-0316 Oslo, Norway. [3]Bioinformatics Core Facility, Institute for Medical Informatics, Oslo University Hospital, Montebello, N-0310 Oslo, Norway. [4]Department of Informatics, University of Bergen, N-5020 Bergen, Norway. [5]Department of Natural Sciences and Technology, Hedmark University College, P.O. Box 4010, Bedriftsenteret, N-2306 Hamar, Norway. [6]Department of Animal and Aquacultural Sciences, University of Life Sciences, P.O. Box 5003, N-1432 Ås, Norway. [7]Department of Biology, Centre for Geobiology, University of Bergen, N-5020 Bergen, Norway. [8]Computational Biology Unit, Uni Computing, Uni Research AS, N-5020 Bergen, Norway. [9]454 Life Sciences, 15 Commercial Street, Branford, Connecticut 06405, USA. [10]Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK. [11]Nofima Marine, P.O. Box 5010, N-1430 Ås, Norway. [12]Institute of Marine Research, P.O., Box 1870, Nordnes, N-5817 Bergen, Norway. [13]Department of Animal and Aquacultural Sciences, CIGENE, Centre for Integrative Genetics, Norwegian University of Life Sciences, PO Box 5003, 1432 Ås, Norway. [14]Faculty of Biosciences and Aquaculture, University of Nordland, N-8049 Bodø, Norway. [15]Department of Pharmaceutical Biosciences, School of Pharmacy, University of Oslo, P.O. Box 1068, Blindern, N-0316 Oslo, Norway. [16]Max Planck Institute for Molecular Genetics, Ihnestrasse 63-73, D-14195 Berlin-Dahlem, Germany. [17]Institute for Basic Sciences and Aquatic Medicine, School of Veterinary Sciences, N-0033 Oslo, Norway. [18]Department of Medical Biology, Faculty of Health Sciences, University of Tromsø, N-9037 Tromsø, Norway. [19]Department of Biology, PO Box 7803, University of Bergen, N-5020 Bergen, Norway.

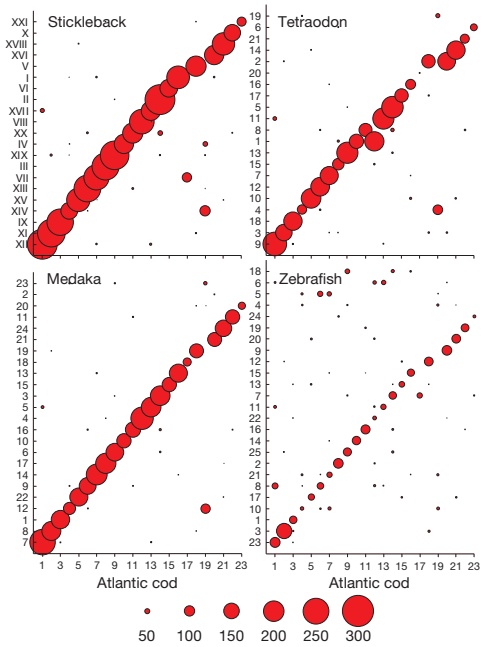

Figure 1 | Synteny between Atlantic cod and selected teleosts. The co-occurrence of orthologous genes (with a minimum of 50% sequence identity over 50% of the alignment, sphere size indicates the numbers of syntenic genes) in 23 Atlantic cod linkage groups[8] (*x*-axis) reveals synteny with the chromosomes of four teleosts (*y*-axis). Several genes located on the stickleback chromosome XIV, tetraodon chromosome 4 and medaka chromosome 12 indicate a lineage-specific chromosomal rearrangement in Atlantic cod.

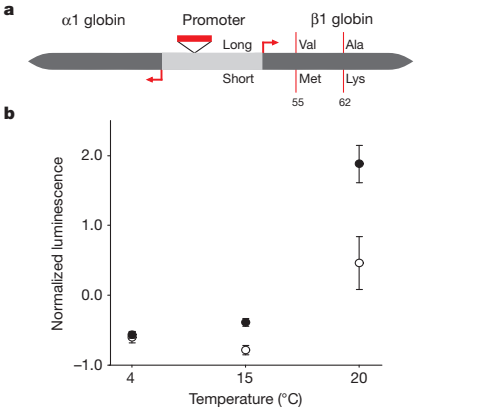

Figure 2 | Functional haemoglobin polymorphisms in Atlantic cod. **a**, Schematic of the head-to-head organized α1 and β1 globin genes, the intergenic promoter region and transcription start sites (red arrows). A promoter polymorphism consisting of a 73-bp indel (red box) segregates in linkage disequilibrium with two amino-acid-substitution polymorphisms (vertical lines) at positions 55 and 62 in β1 globin that affect its oxygen-binding affinity. This linkage disequilibrium results in two predominant haplotypes, long–Val–Ala and short–Met–Lys. **b**, Normalized luciferase luminescence ratios in salmon kidney cells. Cells were transfected using the long promoter (black circles) or the short promoter (white circles) and incubated at 4 °C, 15 °C or 20 °C ($n = 3$ for each treatment level). Error bars show 95% confidence intervals.

Well-studied haemoglobin polymorphisms in Atlantic cod are indicative of functional molecular adaptation to thermal variation[9–12]. The genome contains nine α- and β-globin genes that are organized in two unlinked clusters, *β5–α1–β1–α4* and *β3–β4–α2–α3–β2* (refs 13, 14). We discovered an indel polymorphism of 73 base pairs (bp) in the intergenic promoter region of the *α1–β1* globin pair (Fig. 2a and Supplementary Note 21). This promoter polymorphism occurs in highly significant linkage disequilibrium with two known polymorphic sites in the *β1* gene, the Val55Met and Ala62Lys substitutions[1], in eight Atlantic cod populations (Supplementary Note 22 and Supplementary Fig. 12). In fact, in the three most northern Atlantic populations and in both Baltic populations, the cod β1-globin gene predominantly occurs as a single homozygous genotype consisting of the long promoter and the Val 55–Ala 62 allele (Supplementary Table 10). By placing the two promoter variants in front of a luciferase reporter gene and transfecting the constructs into salmon kidney cells (Supplementary Note 23), we found that temperature and promoter type have a significant interaction effect (generalized linear model, $F_{2,36} = 7.85$, $P = 0.007$, Fig. 2b) and that the long promoter has twofold higher transcriptional activity compared to the short promoter at 15 °C and 20 °C. Increased globin synthesis of the Val 55–Ala 62 allele would compensate for its lower oxygen affinity[10,11] at high temperatures. Thus, the promoter polymorphism provides a molecular compensatory mechanism that helps to maintain the total oxygen-carrying capacity[15]. The tight linkage between the two types of polymorphism provides a compelling example of the coevolution of structural and regulatory adaptation, and highlights the relationship between temperature and functional molecular variation in the haemoglobin system[16].

The Atlantic cod immune system has unusual properties that set it apart from that of other teleosts: high levels of IgM[17], a minimal antibody response after pathogen exposure[5,17,18] and abundant phagocytic neutrophils in the peripheral blood[19,20]. Despite speculation, the exact causes for these differences remain unknown[5]. We found that most genes involved in the vertebrate immune response are present in Atlantic cod (Supplementary Note 24, Supplementary Fig. 13 and Supplementary Table 11). Nevertheless, we did not find genes for the MHC II isoforms, their assembly and trafficking chaperone Ii[21] and the MHC II-interacting protein CD4, which is essential for helper T-cell activation. By comparing a comprehensive set of vertebrate MHC II, CD4 and Ii sequences to the genome assemblies and all unassembled 454 and Illumina sequencing reads (a data set of about 49.5 gigabases), we detected a truncated pseudogene for CD4 (Supplementary Note 25), which is located in a region of conserved synteny (Supplementary Note 27 and Supplementary Fig. 18). No traces of MHC II and Ii were found in syntenous regions (Supplementary Note 27 and Supplementary Figs 16, 17, 19 and 20) and quantitative PCR (qPCR) targeting a conserved domain in MHC II did not amplify the target sequence (Supplementary Note 26 and Supplementary Fig. 15). The absence of MHC II and Ii, and the pseudogenic nature of CD4, show that Atlantic cod has lost the function of the classical pathway for adaptive immunity against bacterial and parasitic infections. Nevertheless, Atlantic cod deals adequately with its prevailing pathogen load in its natural ecological settings[5]. Previous transcriptional (complementary DNA) studies in Atlantic cod have indicated an expansion of the number of MHC I loci[22,23]. By targeting the conserved MHC I α3 domain in genomic DNA using qPCR, we quantified more accurately the number of loci belonging to the teleost U-lineage[24] (Supplementary Note 28). Notably, Atlantic cod has about 100 classical MHC I loci, which is a highly expanded number compared to other teleosts (Fig. 3a). A phylogenetic analysis of teleost MHC I sequences supports the existence of two clades in cod (Fig. 3b and Supplementary Note 29). Within each clade, the mutation patterns show statistically significant signs of positive selection that are indicative of subfunctionalization. These findings indicate that loss of MHC II functionality has coincided with a more versatile usage of the cytosolic pathway of MHC I. Two different MHC I antigen-presentation pathways—the classical pathway and the alternative cross-presentation

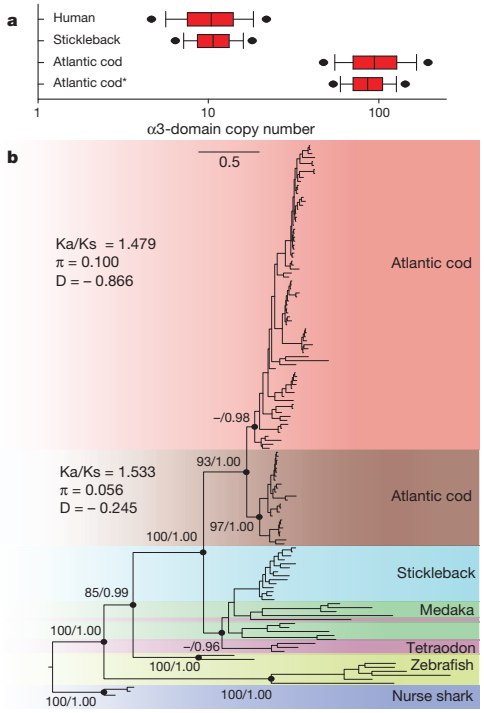

Figure 3 | MHC I diversity in Atlantic cod. **a**, Copy-number estimates of the MHC I α3 domain. Estimates are based on qPCR ratios (see Supplementary Note 28) of the MHC I α3 domain and a single-copy reference gene. For Atlantic cod, β2-microglobulin and topoisomerase III-α (*) were used as reference genes; for human and stickleback, β2-microglobulin was used. The estimates for human and stickleback agree with the expected number of α3 domains found in both reference genomes (Supplementary Table 15). Black dots indicate 95% confidence intervals calculated by bootstrapping ($n = 50{,}000$). **b**, Phylogeny of amino-acid sequences of MHC I α1–α3 domains in teleosts. The Atlantic cod sequences are derived from cDNA and comprise classical U-lineage MHC I only. The other teleost sequences were obtained from Ensembl and NCBI, and contain classical and non-classical U-lineage MHC I. Alignments were visually inspected and corrected where necessary. Maximum likelihood (ML) values and Bayesian posterior probabilities (dots) support the main branches on the ML topology. Distance represents the number of substitutions per site (scale bar). The ratio of non-synonymous to synonymous variable sites (Ka/Ks), the average nucleotide diversity per site (π) and Tajima's D (D) were calculated for the two main clades in Atlantic cod.

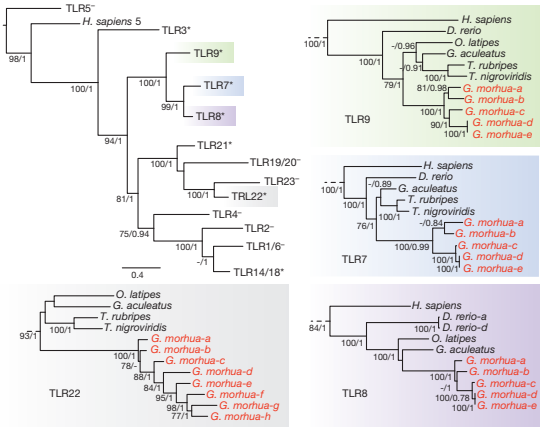

Figure 4 | Phylogeny of TLR families in Atlantic cod. TLR protein sequences were selected on the basis of the conserved Toll-IL-1 receptor (TIR) domain for Atlantic cod, including known sequences from stickleback, zebrafish, tetraodon, fugu, medaka and human as references. TLR clades with (*) or without (⁻) Atlantic cod sequences are denoted according to human or teleost orthologues (summary tree topology, top left panel). Distance represents the average number of substitutions per site (scale bar). ML values and Bayesian posterior probabilities greater than 75/0.75 support the ML topology. Detailed topologies of TLR7 (blue), TLR8 (purple), TLR9 (green) and TLR22 (grey) show gene expansions for Atlantic cod (red). Multiple TLR copies within species are subdivided by letters, and follow Ensembl nomenclature for *D. rerio*.

pathway—can initiate immune responses in mammals[25]. The cross-presentation pathway represents a structural and cellular modification of the MHC I machinery that allows activation of $CD8^+$ T cells upon bacterial infection. The cytokine gene profile of Atlantic cod (Supplementary Table 11) supports the possibility of generating different subsets of $CD8^+$ T cells that either provide direct protection or regulate other immune cells, and thus compensate for the loss of $CD4^+$ T cells.

In addition to the MHC I expansion, we found an unusual composition of the highly conserved TLR families that have a fundamental role in the innate immune response and the initial detection of pathogens. Teleost TLR-encoding genes occur in well-supported phylogenetic clusters, most of which share functional properties with mammalian orthologues, although some are fish-specific[26]. The Atlantic cod TLR genes form monophyletic groups within the known teleost functional groups (Fig. 4, Supplementary Note 30 and Supplementary Fig. 22). Genes for several TLRs that recognize bacterial surface antigens (TLR1, TLR2 and TLR5) are, however, absent, leaving only the teleost-specific TLR14 and TLR18 as members of the TLR1 family in Atlantic cod. Moreover, several families of TLRs that recognize nucleic acids (TLR7, TLR8, TLR9 and TLR22) have markedly expanded, resulting in the highest number of TLRs found in a teleost so far. This TLR repertoire indicates that the Atlantic cod immune system relies relatively heavily on nucleic-acid-detecting TLRs to recognize bacterial pathogens. Notably, the gene expansion of TLR9 coincides with an expansion of interleukin-8 genes (IL-8, Supplementary Table 11). IL-8 is an important chemokine in the innate immune response and is directly induced by TLR9 in human neutrophils[27]. The corresponding expansions of IL-8 and TLR9 indicate that this signalling cascade is particularly important in Atlantic cod.

The loss of MHC II function and lack of a $CD4^+$ T-cell response represent a fundamental change in how the adaptive immune system is initiated and regulated in Atlantic cod. The marked expansion of MHC I genes and unusual TLR composition signify a shift of its immune system in handling microbial pathogens. An expanded MHC I repertoire in the presence of a non-polymorphic MHC II is found in an evolutionarily-distant vertebrate, the axolotl (*Ambystoma mexicanum*)[28,29]. These observations indicate that anomalous immune systems (possibly analogous to that of Atlantic cod) have evolved independently. Additionally, we did not recover evidence for expressed MHC II, CD4 and Ii in the transcriptomes of three other gadoids, indicating that the unusual immune system is a derived characteristic of the gadoid lineage (Supplementary Tables 18 and 19).

We have provided the first annotated genome of a species that supports extensive fisheries and is on the verge of becoming an important aquaculture species. This work provides a major foundation for addressing key issues related to the management of natural Atlantic cod populations, such as the concept of fisheries-induced evolution, which dictates that selective harvesting can change the evolutionary trajectory of major life-history traits of natural populations[30]. Moreover, our novel findings regarding the immune system will allow for more targeted vaccine development, aiding disease management and the process of domestication of Atlantic cod. These findings

change fundamental assumptions regarding the evolution of the vertebrate immune system.

METHODS SUMMARY

Detailed methods on the sequencing and assembly of data from genomic and transcriptomic origins; annotation, synteny analyses, transfection experiments, bioinformatic analyses and phylogenetic analyses presented in this manuscript are described in the Supplementary Information.

Received 11 February; accepted 28 June 2011.
Published online 10 August 2011.

1. Kurlansky, M. *Cod: A Biography of the Fish that Changed the World*. (Penguin, 1998).
2. Johansen, S. D. *et al.* Large-scale sequence analyses of Atlantic cod. *New Biotechnol.* **25,** 263–271 (2009).
3. Flajnik, M. F. & Kasahara, M. Origin and evolution of the adaptive immune system: genetic events and selective pressures. *Nature Rev. Genet.* **11,** 47–59 (2010).
4. Litman, G. W., Rast, J. P. & Fugmann, S. D. The origins of vertebrate adaptive immunity. *Nature Rev. Immunol.* **10,** 543–553 (2010).
5. Pilstrom, L., Warr, G. W. & Stromberg, S. Why is the antibody response of Atlantic cod so poor? The search for a genetic explanation. *Fish. Sci.* **71,** 961–971 (2005).
6. Miller, J. R., Koren, S. & Sutton, G. Assembly algorithms for next-generation sequencing data. *Genomics* **95,** 315–327 (2010).
7. Miller, J. R. *et al.* Aggressive assembly of pyrosequencing reads with mates. *Bioinformatics* **24,** 2818–2824 (2008).
8. Hubert, S., Higgins, B., Borza, T. & Bowman, S. Development of a SNP resource and a genetic linkage map for Atlantic cod (*Gadus morhua*). *BMC Genomics* **11,** 191 (2010).
9. Sick, K. Haemoglobin polymorphism in fishes. *Nature* **192,** 894–896 (1961).
10. Brix, O., Thorkildsen, S. & Colosimo, A. Temperature acclimation modulates the oxygen binding properties of the Atlantic cod (*Gadus morhua* L.) genotypes *Hbl*1/1, Hbl*1/2*, and *Hbl*2/2* by changing the concentrations of their major hemoglobin components (results from growth studies at different temperatures). *Comp. Biochem. Physiol. A* **138,** 241–251 (2004).
11. Andersen, O. *et al.* Haemoglobin polymorphisms affect the oxygen-binding properties in Atlantic cod populations. *Proc. R. Soc. B* **276,** 833–841 (2009).
12. Petersen, M. F. & Steffensen, J. F. Preferred temperature of juvenile Atlantic cod *Gadus morhua* with different haemoglobin genotypes at normoxia and moderate hypoxia. *J. Exp. Biol.* **206,** 359–364 (2003).
13. Borza, T., Stone, C., Gamperl, A. K. & Bowman, S. Atlantic cod (*Gadus morhua*) hemoglobin genes: multiplicity and polymorphism. *BMC Genet.* **10,** 51 (2009).
14. Wetten, O. *et al.* Genomic organization and gene expression of the multiple globins in Atlantic cod: conservation of globin-flanking genes in chordates infers the origin of the vertebrate globin clusters. *BMC Evol. Biol.* **10,** 315 (2010).
15. Gamperl, A. K., Busby, C. D., Hori, T. S. F., Afonso, L. O. B. & Hall, J. R. Hemoglobin genotype has minimal influence on the physiological response of juvenile Atlantic cod (*Gadus morhua*) to environmental challenges. *Physiol. Biochem. Zool.* **82,** 483–494 (2009).
16. Weber, R. E. & Fago, A. Functional adaptation and its molecular basis in vertebrate hemoglobins, neuroglobins and cytoglobins. *Respir. Physiol. Neurobiol.* **144,** 141–159 (2004).
17. Magnadottir, B., Gudmundsdottir, S., Gudmundsdottir, B. K. & Helgason, S. Natural antibodies of cod (*Gadus morhua* L.): specificity, activity and affinity. *Comp. Biochem. Physiol. B* **154,** 309–316 (2009).
18. Solem, S. T. & Stenvik, J. Antibody repertoire development in teleosts—a review with emphasis on salmonids and *Gadus morhua* L. *Dev. Comp. Immunol.* **30,** 57–76 (2006).
19. Rønneseth, A., Wergeland, H. I. & Pettersen, E. F. Neutrophils and B-cells in Atlantic cod (*Gadus morhua* L.). *Fish Shellfish Immunol.* **23,** 493–503 (2007).
20. Øverland, H. S., Pettersen, E. F., Rønneseth, A. & Wergeland, H. I. Phagocytosis by B-cells and neutrophils in Atlantic salmon (*Salmo salar* L.) and Atlantic cod (*Gadus morhua* L.). *Fish Shellfish Immunol.* **28,** 193–204 (2010).
21. Landsverk, O. J., Bakke, O. & Gregers, T. F. MHC II and the endocytic pathway: regulation by invariant chain. *Scand. J. Immunol.* **70,** 184–193 (2009).
22. Persson, A. C., Stet, R. J. M. & Pilstrom, L. Characterization of MHC class I and β2-microglobulin sequences in Atlantic cod reveals an unusually high number of expressed class I genes. *Immunogenetics* **50,** 49–59 (1999).
23. Miller, K. M., Kaukinen, K. H. & Schulze, A. D. Expansion and contraction of major histocompatibility complex genes: a teleostean example. *Immunogenetics* **53,** 941–963 (2002).
24. Dijkstra, J. M. *et al.* A third broad lineage of major histocompatibility complex (MHC) class I in teleost fish; MHC class II linkage and processed genes. *Immunogenetics* **59,** 305–321 (2007).
25. Amigorena, S. & Savina, A. Intracellular mechanisms of antigen cross presentation in dendritic cells. *Curr. Opin. Immunol.* **22,** 109–117 (2010).
26. Rebl, A., Goldammer, T. & Seyfert, H. M. Toll-like receptor signaling in bony fish. *Vet. Immunol. Immunopathol.* **134,** 139–150 (2010).
27. József, L., Khreiss, T., El Kebir, D. & Filep, J. G. Activation of TLR-9 induces IL-8 secretion through peroxynitrite signaling in human neutrophils. *J. Immunol.* **176,** 1195–1202 (2006).
28. Tournefier, A. *et al.* Structure of MHC class I and class II cDNAs and possible immunodeficiency linked to class II expression in the Mexican axolotl. *Immunol. Rev.* **166,** 259–277 (1998).
29. Kaufman, J., Volk, H. & Wallny, H. J. A. "Minimal essential Mhc" and an "unrecognized Mhc": two extremes in selection for polymorphism. *Immunol. Rev.* **143,** 63–88 (1995).
30. Stenseth, N. C. & Dunlop, E. S. Evolution: unnatural selection. *Nature* **457,** 803–804 (2009).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported by a grant from the Research Council of Norway (FUGE program) to K.S.J. The authors wish to thank the following people and organizations: the 454 Life Science Sequencing Center (Branford, USA); the 454 and Illumina nodes of the Norwegian Sequencing Centre (University of Oslo); M. Egholm (formerly 454 Life Science); the Norwegian Metacenter for Computational Science (Notur) and the Norwegian Storage Infrastructure (Norstore); the Research Computing Services group, especially B. -H. Mevik, at the Center for Information Technology (University of Oslo); B. Walenz (Celera); the Canadian Cod Genomics and Broodstock Development Consortium; P. Olsvik, K. Lie and E. Holen at the Norwegian National Institute of Nutrition and Seafood Research (NIFES); J. Gaup and H. Bakke (CEES, University of Oslo); M. Kent (CIGENE, Norwegian University of Life Sciences); S. Bowman (Genome Atlantic); the FUGE bioinformatics platforms group, especially S. Grindhaug; I. Sandlie and O. B. Landsverk (Centre for Immune Regulation, University of Oslo); and Roche Norway.

Author Contributions DNA and RNA isolation, library construction and sequencing: A.T.-K., M.S., M.H.S., T.B.R., M.M., M.E., B.S., A.J.N. and J.T. Sanger BAC (end-) sequencing: H.K. and R.R. Assembly: A.J.N., B.S., A.S. and A.L. Linkage map analyses: K.G.T. and B.S. SNP analyses: K.G.T., P.R.B., S.L. and A.J.N. Annotation: J.-H.V., B.A. and S.S. Repeat analyses: B.S. Synteny analyses: J.P. and B.S. Haemoglobin analyses: Ø.A., O.F.W., B.S. and T.G. Bioinformatics: A.J.N., B.S., A.S., T.B.R., J.P., C.P., C.N., R.B.E., R.W., J.K., K.L., A.L., I.J., M.M., K.M., P.R.B., K.G.T. and M.H.S. Immune analyses: U.G., M.M., M.H.S., M.E., B.S., B.O.K., T.M., K.L., S.D.J. and T.B.R. Interpretation of immune results: U.G., T.F.G., S.J., B.S. and K.S.J. 454 contributions: L.D. Revisions: Ø.A., T.M., S.D.J., F.N., I.J., S.J., N.C.S. and S.W.O. Project initiation: S.W.O., I.J., F.N., S.L., N.C.S. and K.S.J. Project coordination: S.J. Consortium leader: K.S.J. This manuscript is dedicated to the memory of L. Pilström and R. J. M. Stet. Their research inspired our work to understand further the Atlantic cod immune system.

Author Information The unassembled sequencing reads and Newbler assembly have been deposited at ENA-EMBL under the accession numbers CAEA01000001–CAEA01554869. The annotation is available through Ensembl at http://www.ensembl.org/index.html. These and more resources are also available through http://codgenome.no. Reprints and permissions information is available at www.nature.com/reprints. This paper is distributed under the terms of the Creative Commons Attribution-Non-Commercial-Share-Alike licence, and is freely available to all readers at www.nature.com/nature. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to K.S.J. (k.s.jakobsen@bio.uio.no).

Supplementary Note 21 Haemoglobin

We previously resolved the structure of the haemoglobin cluster by sequencing BAC clones from a Norwegian coastal cod (NCC) individual, designated NCC_001, (Wetten et al. 2010). By comparing the haemoglobin clusters in the genome (specimen NEAC_001) to the independently assembled BAC insert sequences we found an insert of 73 bp in the intergenic promoter region (Figure 2a, main text). Detailed alignments are available on request.

Supplementary Note 22 Linkage disequilibrium in the *ß1* globin region

The extent of linkage disequilibrium between the polymorphic *ß1* promoter and the polymorphic sites in *ß1* globin was evaluated for all three locus pairs (promoter *versus ß1*-55, promoter *versus ß1*-62 and *ß1*-55 *versus ß1*-62) across all eight populations (Supplementary Figure 12, Supplementary Table 10) using Fisher's method through Genepop (http://genepop.curtin.edu.au/, Rousset 2008). All three comparisons revealed significant linkage disequilibrium (Chi_2=infinity, df 16, pvalue= 0) among the different alleles.

Supplementary Note 23 Transfection experiments

The two α*1*-β*1* globin promoter variants were obtained by direct PCR amplification of the complete intergenic region between the start codons of the 5'-5'oriented α*1* and β*1* globin genes using genomic DNA from specimens homozygous for each of the two promoter variants (short and long) as template. A set of sense/antisense primers targeting both promoter variants were designed with restriction sites for XhoI and HindIII in their 5' ends. A degenerate site (W) in the Hb-HindIII primer reflects a SNP at this locus. Primer sequences are as follows:
Hb-XhoI: 5'-ATCTCGAGCTTGAATAGTGTGGTCAGATTGGACTCTGT-3' and
Hb-HindIII: 5'-ATAAGCTTTGTGGCGTWGTCTTAAGGGTTCAATGT-3'. The PCR products were sequenced, and plasmids for transfection were constructed by cloning the different promoters into XhoI and HindIII restriction sites in Promega's expression vector pGL4.20, which harbors a firefly luciferase gene downstream of the cloning sites. Vector inserts were sequenced to confirm the promoter types and correct orientation, i.e., the promoter end originating from upstream the β*1* globin start-codon oriented upstream of the luciferase reporter in the expression vector. An internal control vector for co-transfection (Promega pGL4.73) harboring SV40 early enhancer/promoter and renilla luciferase reporter gene was used in the experiment without modifications.
TO-cells from Atlantic salmon (Wergeland and Jakobsen 2001) were cultured in L-15 medium (Cambrex Bio Sciences) supplemented with 50 µg/ml gentamicin, 4 mM Lglutamine, 40 µM β-mercaptoethanol and 10% fetal calf serum. After splitting, the cells were adapted to three different temperatures (4, 15 and 20 °C) for one week in 175 cm2 flasks before trypsinization and transfection. One µg of globin promoter construct and 1 µg of internal control construct were co-transfected into ~2*106 cells using an AMAXA NucleofectorTM device (program T-20) according to the manufacturer's instructions (Lonza AG). Transfection efficiency was routinely about 50-60 %

(assessed by parallel GFP transfection). Three transfections were performed per promoter variant and temperature, and directly after transfection the cells from each transfection were suspended in growth medium and seeded in triplicate to 6-well plates before incubation for another week at same temperature as prior to transfection. Effects of temperature on expression efficiency were determined by analyzing firefly and renilla luciferase activity using the Dual-Luciferase® Reporter Assay System according to the manufacturer's description (Promega). Relative expression efficiency for the promoters were obtained as ratios of firefly luminescence (controlled by the Hb promoters) to renilla luciferase expression from the pGL 4.73 internal control vector to eliminate bias resulting from possible differences in transfection efficiency, unequal cell numbers and other per sample related errors. Luminescence ratios were calculated against a co-transfected SV40/Renilla reporter and normalized for each separate transfection experiment. The ratios were normalized to allow comparison between the three separate transfection experiments. The data were analyzed using a GLM (Generalized Linear Model), with temperature and promoter type as crossed factors, the three replicates as nested factor in both temperature and promoter type, and normalized expression ratio as response. The interaction between temperature and promoter type was highly significant (GLM, $F_{2,36} = 7.85$ $P = 0.007$, Figure 2b, main text).

Supplementary Figure 12 Sample locations for eight Atlantic cod populations.

Tissue samples (muscle) were obtained in 2003/2004 and preserved in ethanol. Sample size is given in brackets.

Supplementary Table 10 Combined genotype frequencies of the *ß1* globin promoter and amino acid polymorphisms among eight Atlantic cod populations

Genotype[1]			Location[2]							
Pr	*ß1*-55	*ß1*-62	Bjørn-øya (40)	Båts-fjord (41)	Ma-langen (40)	Molde (40)	Helgo-land (36)	Katte-gat (75)	Born-holm (53)	Øland (36)
L	Val	Ala	0.700	0.805	0.600	0.150	0.139	0.107	0.925	0.944
*	*	*	0.100	0.122	0.175	0.225	0.250	0.387	0.055	0.028
S	Met	Lys			0.075	0.225	0.250	0.13		
L	*	*	0.050	0.024		0.100	0.195	0.094	0.020	
*	Met	Lys			0.025	0.125	0.139	0.120		
*	Val	*		0.049	0.075	0.100		0.040		0.028
*	Val	Ala	0.125			0.025				
S	*	*						0.040		
L	Val	*	0.025			0.025		0.013		
*	*	Lys						0.027		
S	Val	Lys			0.025			0.013		
S	*	Lys					0.027	0.013		
L	Met	Ala			0.025					
*	Met	*				0.025				
L	Met	Lys						0.013		

[1]Homozygote genotypes are scored as long (L) or short (S) for the promoter (Pr), Val or Met for the *ß1*-55, Ala or Lys for the *ß1*-62 polymorphisms and as heterozygote genotypes (*).
[2]Sample size is given in brackets. Bornholm and Øland are located in the Baltic Sea.

Paper IV

Recombinant subtypes of the polymorphic hemoglobin *β1* gene of Atlantic cod are sparsely distributed in trans-Atlantic populations as determined by High Resolution Melting analysis

Ola Frang Wetten, Robert C. Wilson and Øivind Andersen

Ola Frang Wetten

Department of Natural Science and Technology, Hedmark University College, N-2318 Hamar, Norway.

Department of Animal and Aquaculture Sciences, Norwegian University of Life Sciences, N-1430 Aas, Norway.

E-mail: ola.wetten@hihm.no

Robert C. Wilson *(responsible for correspondence)*

Department of Natural Science and Technology, Hedmark University College,

N-2318 Hamar, Norway.

E-mail: robert.wilson@hihm.no

Phone: +47 62 51 78 51

Fax: +47 62 51 76 01

Øivind Andersen

Nofima, PO Box 5010, N-1430 Aas, Norway

Department of Animal and Aquaculture Sciences, Norwegian University of Life Sciences, N-1430 Aas, Norway.

E-mail: oivind.andersen@nofima.no

Abstract

High Resolution Melting (HRM) analysis was applied to haplotype the Met55Val-Lys62Ala-polymorphisms of the Atlantic cod *Hb-β1* gene. The Val55-Ala62 haplotype predominates in cod populations throughout the western and Northeast Atlantic Ocean, while the Met55-Lys62 is mostly found in the North Sea, Kattegat and along the southern part of the Norwegian coast. Whereas the distribution of the two haplotypes show a temperature-related north-south gradient in Northeast Atlantic populations, this study provided no evidence for such a cline on the western side of the North Atlantic Ocean. Coupling and repulsion double heterozygotes previously indistinguishable by SNP-genotyping were readily separated by the HRM assay, and no repulsion heterozygote specimens were found on either side of the Atlantic Ocean. The recombinant haplotype Val55-Lys62 was detected in significant numbers in the Canadian and Greenlandic populations and in populations off the Norwegian coastline, and is probably the origin of one of the rare subtypes earlier detected by protein gel electrophoresis. Our HRM assay affords low cost, precise and efficient *Hb-β1* polymorphism haplotyping in large numbers of DNA samples in small, moderately equipped laboratories.

Introduction

Atlantic cod has a long historical record as an important marine resource, but many cod stocks have declined or even collapsed, mainly due to overfishing. For proper management, the relationship between the multiple cod populations inhabiting the continental shelves and banks in most areas of the North Atlantic has been debated for more than 100 years (Hjort and Dahl 1900). The migratory North Arctic population of Atlantic cod is distinguished from the stationary coastal cod populations by the genetic marker pantophysin (*PanI*) (Pogson 2001, Pogson and Fevolden 2003), but there also seem to be genetic differences between the latter populations inhabiting the fjords along the Norwegian coast (Jørstad and Nævdal 1989). Genetic studies of Atlantic cod were initiated during the 1960s after the pioneering report on the HbI-1/HbI-2 polymorphism in the oxygen transporting hemoglobin (Frydenberg et al. 1965, Moller 1966, Sick 1961). The molecular basis underlying the cod hemoglobin polymorphism was recently revealed by the identification of two amino acid substitutions, Met55Val (M/V) and Lys62Ala (K/A), in the Hb-β1 subunit, which affect the oxygen binding affinity and temperature sensitivity of the hemoglobin tetramer (Andersen et al. 2009). The MK and VA haplotypes were found to be unambigiously associated with the HbI-1 and HbI-2 isoforms in both Northeast and Northwest Atlantic populations (Andersen et al. 2009, Borza et al. 2010). Significant differences between HbI-1/1 and HbI-2/2 homozygotes have been found in terms of oxygen affinity, rate of aerobic metabolism, temperature preference, competitive feeding strategy, age of maturation, muscle cellularity and growth rate, although some conflicting results exist (Imsland et al. 2004, Johnston et al. 2006, Karpov and Novikov 1980, McFarland 1998, Mork et al. 1983, Naevdal et al. 1992, Petersen and Steffensen 2003, Salvanes and Hart 2000). Despite strong linkage disequilibrium between the closely positioned M/V and K/A replacements, rare recombinants of the cod Hb-β1 chain were

detected in Northeast Atlantic cod by SNP (single nucleotide polymorphism) analysis (Andersen et al. 2009). These hemoglobin subtypes are probably related to the sparsely distributed hemoglobin subtypes identified in these waters by laborious analysis of up to 13 different electrophoretic protein bands (Fyhn et al. 1994, Husebo et al. 2004, Imsland et al. 2007, Sick 1965). The distinct subtypes reported in the Danish Belt, and the apparent lack of such variants in the Arctic populations of Atlantic cod, suggested that the hemoglobin subtypes might represent adaptation to specific environmental conditions and could be used as population markers (Fyhn et al. 1994, Husebo et al. 2004).

Whereas SNP analysis of the cod hemoglobin genotypes by MALDI-TOF (e.g. Sequenom MASSarray) or Sanger sequencing is highly efficient and reliable compared to electrophoretic isoform detection from fresh blood samples, both are costly genotyping methods not readily available to most laboratories. More importantly, they do not discriminate between the coupling and repulsion heterozygotes of cod *Hb-β1* (Andersen et al. 2009). High resolution melting (HRM) analysis has been used to genotype human *β-globin* targeting a single SNP (Liew et al. 2004), and Parant et al. (2009) used this relatively inexpensive and rapid method to genotype both point mutations and in-del alleles of various genes in zebrafish. The simultaneous genotyping of several SNPs in the same amplicon using HRM has also been reported (Garritano et al. 2009, Tindall et al. 2009), indicating its potential for distinguishing the coupling and repulsion *Hb-β1* heterozygotes in cod. Here we present an HRM assay for the efficient and unambiguous discrimination between all ten genotypes of the M/V and K/A polymorphisms in Atlantic cod *Hb-β1*.

Material and Methods

93 DNA samples

94 Atlantic cod was genotyped for the *Hb-β1* polymorphisms by HRM analysis of genomic
95 DNA from 118 specimens sampled at five distinct locations of the Northwest Atlantic Ocean
96 in 2002-2004. In addition, 90 double heterozygotes from eight Northeast Atlantic populations
97 described in Andersen et al. (2009) were screened for possible repulsion heterozygotes
98 (Figure 1, Table 1).

99

100 PCR amplification of mutated region

101 Primers were designed to amplify a 116 bp region in exon 2 of cod *Hb-β1* (GenBank
102 accession no. FJ392683.1) containing four SNPs (Figure 2). While a single non-synonymous
103 mutation a/g gives the M/V substitution, the two non-synonymous mutations a/g and a/c
104 causing the K/A change are preceeded by a synonymous c/t site. By sequencing multiple
105 specimens from various geographic locations, the three adjacent SNPs showed only the
106 haplotypes caa and tgc (Andersen et al. 2009, Borza et al. 2009), indicating no recombination
107 between these latter SNPs. The present study did not include the third polymorphism
108 Leu123Met of cod *Hb-β1*, which showed no association with the HbI-1/HbI-2 isoforms
109 (Borza et al. 2009). Sense and antisense primer sequences were 5'-
110 AGATGCCTGATTGTGTATCCG-3' and 5'- AGGGCCACGACGCCGT-3', respectively.
111 The PCR reactions of 20 µl contained 4 µl of 5x HOT FIREPol® EvaGreen® HRM Mix, 125
112 nM of sense and antisense primers and ~ 20 ng genomic DNA as recommended by the
113 supplier of HRM analysis reagents (Solis BioDyne, Estonia). An Applied Biosystems 7500
114 Fast Realtime instrument was employed for cycling and generating the melting curve data.
115 From the 7500 Fast SDS system software (v1.4, Applied Biosystems), cycling parameters
116 were set to 95 °C for 15 min, 50 cycles of 95 °C for 10 sec, and 63 °C for 15 sec, followed by
117 a dissociation stage collecting fluorescence data between 60 and 95 °C. The amplicons were

then analyzed and checked for deviating samples using the 7500 Fast SDS system software (v1.4, Applied Biosystems).

Synthesized DNA of the four variants encoding the MK, VA, VK and MA haplotypes comprised the complete exon 2 of *Hb-β1* and was cloned into pUC57 vectors (GenScript, USA). The vectors were used singularly and binarily as positive control PCR templates for the ten different genotype combinations of the four haplotypes.

HRM analysis

Data from the PCR dissociation stage were imported to the HRM Software (version 2.0.1, Applied Biosystems) and genotyped automatically. All genotype calls were then manually evaluated and compared to control samples using difference plot diagrams (Figure 3). Experimental samples were run in triplicate, and all specimens were finally sequenced to confirm genotypes. The sequencing reactions of 10 µl contained 1 µl BigDye terminator sequencing reaction mix (version 1.1, Applied Biosystems), 0.32 µM sense (5'-AGATGCCTGATTGTGTATCCG-3') or antisense (5'-GGAAGTTGTCGGGGTCG-3') primer and exonuclease I-treated PCR product as template. The template was obtained from genomic DNA using both sequencing primers in the PCR reactions, which consisted of 0.1 µl AmpliTaq Gold (Applied Biosystems), 0.8 µM sense and antisense primers, 0.2 µM dNTPs, 1.5 mM $MgCl_2$ and 1x Amplitaq Gold PCR buffer (Applied Biosystems) in a final volume of 10 µl.

Results and Discussion

High Resolution Melting (HRM) analysis of cod *Hb-β1*

We have implemented HRM analysis for haplotyping the polymorphic *Hb-β1* gene of Atlantic cod using 116 bp PCR amplicons covering the four SNPs located in exon 2. HRM analysis of different combinations of the synthetic DNA controls demonstrated that a total of ten different genotypes, including coupled and repulsive heterozygotes, were distinguishable (Figure 3). The coupled (MK/VA) and repulsive (MA/VK) heterozygotes display the same mix of bases though distributed on different haplotypes. However, they were clearly separated by HRM as demonstrated by their melting profiles (Figure 3), which were generated on the basis of the actual haplotypic Tm differences (89.5 °C, 88.9 °C, 89.1 °C and 89.0 °C for VA, MK, VK and MA, respectively) reported by the HRM software. The MK/MK and MK/VK genotypes were the most difficult to distinguish by HRM because of their very similar melting profiles. This is in accordance with our expectations that, among the ten possible genotypes, those that share the same homozygous genotype at K62A (caag/caag or tgcg/tgcg) would be among the most challenging to discriminate, since the M55V substitution arises from a single a/g polymorphism. Surprisingly, the melting profiles of the remaining genotypes affected by such minimal alteration at position 55 (e.g. VK/VK vs. MK/VK) were more distinct (Figure 3).

Successful application of the HRM method to distinguish between all genotypes of cod *Hb-β1* in a single amplicon as presented here required high quality template DNA. We also experienced that template DNA should be normalized to a standard concentration to reduce any non-genetic variation between samples. Positive control samples and replicates were paramount for the confirmation or rejection of certain genotype designations. While our HRM-based *Hb-β1* genotyping assay resolved heterozygotes quite readily, other researchers who have assayed several SNP loci in single amplicons found that inclusion of a non-extendable melting probe and performing asymmetric PCR allowed enhanced heterozygote discrimination by monitoring both amplicon and probe melting curves (Garritano et al. 2009,

Vossen et al. 2010). Although the established HRM assay is quite robust, it is possible that the enhanced resolution offered by melting probe inclusion might broaden the window for acceptable template concentration and quality. Tindall et al. (2009) achieved better heterozygote resolution in amplicons harboring multiple SNPs by adding reagents designed to alter the melting characteristics of GC-rich (66 %) DNA. The *Hb-β1* amplicon analyzed in this study has a GC-content of 60-62 %, depending on the haplotype, so it is doubtful if such reagent addition would increase the resolution strength of our assay.

Hb-β1 genotyping

Atlantic cod from five Northwest Atlantic populations off Canada and Greenland were genotyped for *Hb-β1* using HRM. Additionally, eight Northeast Atlantic populations previously genotyped by Sequenom MassARRAY SNP analysis (Andersen et al. 2009) were included in the HRM analysis to determine the genotype frequencies of double heterozygotes (Table 1). Homozygous VA cod predominated in the Northwest Atlantic populations with the highest frequencies of 88 and 83 % in the two Greenland populations, similar to the frequencies found in the Northeast Arctic population sampled at Bjørnøya (83 %) and Båtsfjord (80 %). The homozygous MK genotype was lacking in the Northwest Atlantic populations, except for one fish (comprising 4 % of the population) identified in George Bank and Sisimiut, and was also absent in the Northeast Arctic and the Baltic Sea populations. The frequencies of coupled double heterozygous MK/VA fish varied from 4-12 % in the Northwest Atlantic populations, while the recombinant VK haplotype was identified solely in VA/VK heterozygotes, which occurred at frequencies from 8 % in the Greenland populations to 28 % in Nova Scotia. No other recombinant genotypes were identified in the Northwest Atlantic populations, whereas rare homozygotes of either the recombinant MA or VK haplotype were found in the Norwegian coastal cod population at Malangen (Table 1).

193 Consistently, repulsive MA/VK heterozygotes were completely lacking in the Northwest and
194 Northeast Atlantic populations examined. The sparsely distributed recombinants probably
195 represent the rare subunits previously identified in these waters by IEF analysis (Fyhn et al.
196 1994, Husebo et al. 2004, Sick 1965).
197 The recombinant VK haplotype displayed moderate frequencies in the Northwest Atlantic
198 populations with the highest frequency of 14 % in Nova Scotia, while up to 6.5 % of this
199 haplotype was found in the Northeast Atlantic populations. In contrast, the recombinant MA
200 haplotype was completely lacking in the Northwest Atlantic populations and sparsely
201 distributed in the Northeast Atlantic populations. It is therefore tempting to speculate whether
202 the extremely rare MA haplotype represents any disadvantage in certain environments. The
203 M/V and K/A substitutions probably affect oxygen binding affinity and temperature
204 sensitivity, respectively, of the Hb-β1-containing tetramer (Andersen et al. 2009), which
205 might imply that the homozygous MA cod suffer from reduced oxygen binding, particularly
206 at higher temperatures. Consistently, this genotype was only identified in the northern coastal
207 cod population of Malangen (Table 1). It should, however, be added that among the nine
208 globin genes identified in the cod genome, the highly expressed *Hb-β2* gene encodes a
209 protein with Met55 and Ala62 (Wetten et al. 2010). Since the Hb-β1 and Hb-β2 globins are
210 the main β chains in the tetrameric hemoglobins of Atlantic cod (Verde et al. 2006), we
211 suggest that the MA combination in both subunits might represent a disadvantage at elevated
212 temperatures. Screening multiple teleost β globins revealed several Arctic and Antarctic
213 species displaying Met55 and Ala62, but this combination was also found in cyprinids and
214 salmonids inhabiting temperate and tropical waters (Table 2). Further genomic studies are
215 warranted to identify the repertoire of globin genes in teleosts inhabiting different
216 environments and to elucidate adaptations exerted by specific globin genotypes.

Similar to the significant north-south gradient in the distribution of the cod HbI alleles in Northeast Atlantic cod populations (Andersen et al. 2009, Brix et al. 1998, Sick 1965), Sick (1965) reported a slight cline in the three populations of Greenland, New Foundland and Maryland displaying allele frequencies of 2, 4 and 7 %, respectively. The latter population at latitude 39 °N is almost the southern limit of cod in these waters. The present study showed the highest (10 %) and lowest (2.1 %) frequencies in the most southern (George Bank) and northern (Nuuk) populations, respectively. However, the low frequencies of 4.0 % in Nova Scotia and 5.3 % in Labrador together with the high frequency of 8 % in the Sisimiut population do not support a north-south gradient in the cod hemoglobin polymorphism (Table 1).

The HRM-based *Hb-β1* haplotyping assay described herein is a closed tube method requiring only PCR and subsequent disassociation/reassociation in a thermocycler capable of monitoring fluorescence of a dsDNA-specific dye also present in the reaction; time-consuming and costly post-PCR restriction and electrophoretic separation steps are obviated. Implementation of the *Hb-β1* assay only required supplementation of our qPCR system with HRM-software, leaving instrumentation and chemical reagent costs at levels comparable to those for conventional dsDNA-specific fluorescent dye-based qPCR. Application of this method can therefore enable small laboratory units to perform haplotyping of the cod *Hb-β1* polymorphisms in a large number of DNA samples quickly and inexpensively.

240 Andersen, Ø., Wetten, O.F., De Rosa, M.C., Andre, C., Carelli Alinovi, C., Colafranceschi,
241 M., Brix, O., and Colosimo, A. 2009. Haemoglobin polymorphisms affect the oxygen-
242 binding properties in Atlantic cod populations. Proc. R. Soc. B **276**(1658): 833-841. doi:
243 10.1098/rspb.2008.1529.
244 Borza, T., Higgins, B., Simpson, G., and Bowman, S. 2010. Integrating the markers Pan I and
245 haemoglobin with the genetic linkage map of Atlantic cod (*Gadus morhua*). BMC Res. Notes
246 **3**: 261. doi:10.1186/1756-0500-3-261.
247 Borza, T., Stone, C., Gamperl, A.K., and Bowman, S. 2009. Atlantic cod (*Gadus morhua*)
248 hemoglobin genes: multiplicity and polymorphism. BMC Genet. **10**: 51. doi:10.1186/1471-
249 2156-10-51.
250 Brix, O., Foras, E., and Strand, I. 1998. Genetic variation and functional properties of
251 Atlantic cod hemoglobins: Introducing a modified tonometric method for studying fragile
252 hemoglobins. Comp. Biochem. Physiol. A **119**(2): 575-583. doi:10.1016/S1095-
253 6433(97)00469-8.
254 Frydenberg, O., Moller, D., Naevdal, G., and Sick, K. 1965. Haemoglobin polymorphism in
255 Norwegian cod populations. Hereditas **53**(1): 257-271.
256 Fyhn, U.E.H., Brix, O., Naevdal, G., and Johansen, T. 1994. New variants of the
257 haemoglobins of Atlantic cod: a tool for discriminating between coastal and Arctic cod
258 populations. ICES Mar. Sci. Symp. **198**: 666 -670. .
259 Garritano, S., Gemignani, F., Voegele, C., Nguyen-Dumont, T., Le Calvez-Kelm, F., De
260 Silva, D., Lesueur, F., Landi, S., and Tavtigian, S.V. 2009. Determining the effectiveness of
261 High Resolution Melting analysis for SNP genotyping and mutation scanning at the TP53
262 locus. BMC Genet. **10**(5). doi:10.1186/1471-2156-10-5.
263 Hjort, J., and Dahl, K. 1900. Fishing experiments in Norwegian fjords. **1**: 1-214.

Husebo, A., Imsland, A.K., and Naevdal, G. 2004. Haemoglobin variation in cod: a description of new variants and their geographical distribution. Sarsia **89**(6): 369-378. doi:10.1080/00364820410002631.

Imsland, A.K., Foss, A., Naevdal, G., Johansen, T., Folkvord, A., Stefansson, S.O., and Jonassen, T.M. 2004. Variations in growth in haemoglobin genotypes of Atlantic cod. Fish Physiol. Biochem. **30**(1): 47-55. doi:10.1007/s10695-004-6787-5.

Imsland, A.K., Foss, A., Naevdal, G., Johansen, T., Stefansson, S.O., and Jonassen, T.M. 2007. New haemoglobin genotypes in Atlantic cod, Gadus morhua: possible relation with growth. Comp. Biochem. Physiol. A **147**(4): 955-960. doi:10.1016/j.cbpa.2007.03.001.

Johnston, I.A., Abercromby, M., and Andersen, O. 2006. Muscle fibre number varies with haemoglobin phenotype in Atlantic cod as predicted by the optimal fibre number hypothesis. Biol. Lett. **2**(4): 590-592. doi:10.1098/rsbl.2006.0500.

Jørstad, K.E., and Nævdal, G. 1989. Genetic variation and population structure of cod, Gadus morhua L., in some fjords in northern Norway. J. Fish. Biol. **35**: 245-252. doi:10.1111/j.1095-8649.1989.tb03067.x.

Karpov, A.K., and Novikov, G.G. 1980. Hemoglobin alloforms in cod, *Gadus morhua* (Gadiformes, Gadidae), their functional characteristics and occurrence in populations. J. Ichthyol.(6): 45-50.

Liew, M., Pryor, R., Palais, R., Meadows, C., Erali, M., Lyon, E., and Wittwer, C. 2004. Genotyping of single-nucleotide polymorphisms by high-resolution melting of small amplicons. Clin. Chem. **50**(7): 1156-1164. doi:10.1373/clinchem.2004.032136.

McFarland, S. 1998. Biochemical and physiological adaptations of haemoglobin-I genotypes of Atlantic cod, *Gadus morhua* L. PhD, University of Birmingham, Birmingham.

Moller, D.A.G. 1966. Genetic Differences between Cod Groups in the Lofoten Area. Nature **212**(5064): 824-824. doi:10.1038/212824a0.

289 Mork, J., Giskeødegård, R., and Sundnes, G. 1983. Haemoglobin polymorphism in Gadus
290 morhua: Genotypic differences in maturing age and within-season gonad maturation
291 Helgoland Mar. Res. **36**(3): 313-322. doi:10.1007/BF01983634
292 Naevdal, G., Folkvord, A., Otterlei, E., and Thorkildsen, S. 1992. Growth-Rate Related to
293 Genotype of 0-Group Cod at 3 Environmental Temperatures. Sarsia **77**(1): 71-73.
294 Petersen, M.F., and Steffensen, J.F. 2003. Preferred temperature of juvenile Atlantic cod
295 Gadus morhua with different haemoglobin genotypes at normoxia and moderate hypoxia. J.
296 Exp. Biol. **206**(Pt 2): 359-364. doi:10.1242/jeb.00111.
297 Pogson, G.H. 2001. Nucleotide polymorphism and natural selection at the pantophysin (Pan
298 I) locus in the Atlantic cod, *Gadus morhua* (L.). Genetics **157**(1): 317-330.
299 Pogson, G.H., and Fevolden, S.E. 2003. Natural selection and the genetic differentiation of
300 coastal and Arctic populations of the Atlantic cod in northern Norway: a test involving
301 nucleotide sequence variation at the pantophysin (PanI) locus. Mol. Ecol. **12**(1): 63-74.
302 Salvanes, A.G.V., and Hart, P.J.B. 2000. Is individual variation in competitive performance
303 of reared juvenile cod influenced by haemoglobin genotype? Sarsia **85**(3): 265-274.
304 Sick, K. 1961. Haemoglobin polymorphism in fishes. Nature **192**(480): 894-896.
305 doi:10.1038/192894a0.
306 Sick, K. 1965. Haemoglobin polymorphism of cod in North Sea and north Atlantic Ocean.
307 Hereditas **54**(1): 49-69.
308 Tindall, E.A., Petersen, D.C., Woodbridge, P., Schipany, K., and Hayes, V.M. 2009.
309 Assessing high-resolution melt curve analysis for accurate detection of gene variants in
310 complex DNA fragments. Hum. Mutat. **30**(6): 876-883. doi:10.1002/humu.20919.
311 Verde, C., Balestrieri, M., de Pascale, D., Pagnozzi, D., Lecointre, G., and di Prisco, G. 2006.
312 The oxygen transport system in three species of the boreal fish family gadidae - Molecular

phylogeny of hemoglobin. J. Biol. Chem. **281**(31): 22073-22084. doi:10.1074/jbc.M513080200.

Vossen, R.H.A.M., van Duijn, M., Daha, M.R., den Dunnen, J.T., and Roos, A. 2010. High-throughput genotyping of mannose-binding lectin variants using high-resolution DNA-melting analysis. Hum. Mutat. **31**(4): E1286-E1293. doi:10.1002/Humu.21213.

Wetten, O.F., Nederbragt, A.J., Wilson, R.C., Jakobsen, K.S., Edvardsen, R.B., and Andersen, O. 2010. Genomic organization and gene expression of the multiple globins in Atlantic cod: conservation of globin-flanking genes in chordates infers the origin of the vertebrate globin clusters. BMC Evol. Biol. **10**: 315. doi:10.1186/1471-2148-10-315.

Table 1. Genotype distribution of Atlantic cod *Hb-β1* globin variants in the cod populations examined.

Genotype	Nova Scotia (25)	George Bank (25)	Labrador (19)	Sisimiut (25)	Nuuk (24)	Bjørnøya (40)	Bátsfjord (41)	Malangen (40)	Molde (39)	Helgoland (36)	Kattegat (74)	Baltic B (56)	Baltic Ø (37)
Val–Ala/Val–Ala	64	72	74	80	88	83	80	60	18	14	11	93	94
Met–Lys/Met–Lys	0	4	0	4	0	0	0	10	36	39	27	0	0
Met–Lys/Val–Ala	8	12	10	8	4	15	15	18	31	44	52	7	3
Val–Ala/Val–Lys	28	12	16	8	8	2	5	8	13	0	5	0	3
Met–Lys/Val–Lys	0	0	0	0	0	0	0	0	0	3	4	0	0
Val–Lys/Val–Lys	0	0	0	0	0	0	0	2	0	0	1	0	0
Met–Lys/Met–Ala	0	0	0	0	0	0	0	0	2	0	0	0	0
Val–Ala/Met–Ala	0	0	0	0	0	0	0	0	0	0	0	0	0
Met–Ala/Met–Ala	0	0	0	0	0	0	0	2	0	0	0	0	0
Met–Ala/Val–Lys	0	0	0	0	0	0	0	0	0	0	0	0	0

Numbers in parentheses represent sample sizes; numbers in the table body indicate the percentage of individuals in each population with a given genotype.

Table 2. Teleost β-globins harboring Met55 and Ala62, and the number of genes accessible at NCBI encoding this specific combination.

Species	Number of genes
Anarhichas minor	1
Arctogadus glacialis	1
Artedidraco orianae	1
Boreogadus saida	1
Carassius auratus	2
Cyprinus carpio	9
Danio rerio	4
Electrophorus electricus	1
Gadus morhua	1
Misgurnus anguillicaudatus	1
Oncorhynchus mykiss	1
Paramisgurnus dabryanus	1
Pogonophryne scotti	1
Salmo salar	2
Trematomus hansoni	1
Trematomus newnesi	2

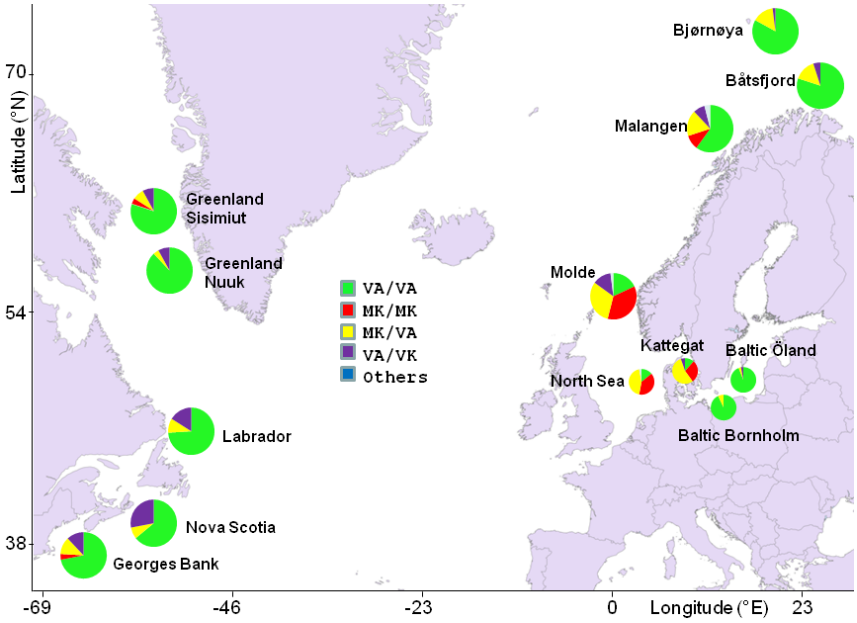

Figure 1.

Map of sampling locations and frequencies of the most common genotypes of Atlantic cod *Hb-β1*. The VA haplotype dominates at all locations, except from Kattegat, North Sea and Molde, while the recombinant haplotype VK are found at highest frequencies in the Canadian populations and the Northeast Atlantic population of Molde.

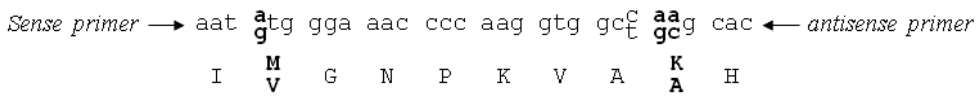

Figure 2.

Nucleotide and encoded amino acid sequences of the polymorphic region of Atlantic cod *Hb-β1* globin haplotyped by HRM analysis. Non-synonymous substitutions are shown in **bold**. A 116-bp fragment was PCR-amplified using the primers indicated.

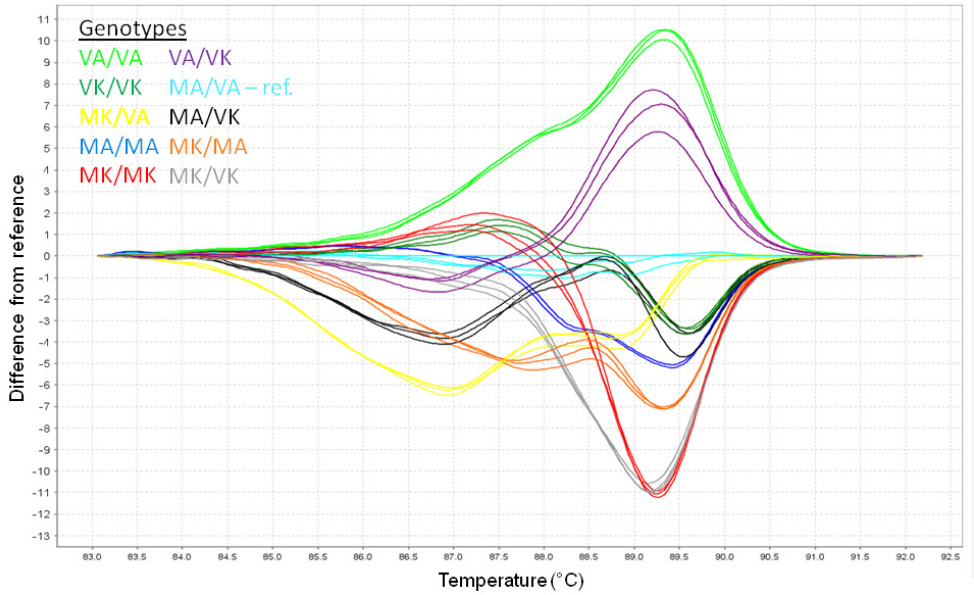

Figure 3.

Difference plot displaying the melting curves of the genotypes of Atlantic cod *Hb-β1* subtracted from the reference genotype MA/VA.