Available online at www.sciencedirect.com
ScienceDirect
Gene 385 (2006) 111–121
GENE
SECTION
EVOLUTIONARY GENOMICS
www.elsevier.com/locate/gene

# DNA methylation and body temperature in fishes

Annalisa Varriale, Giorgio Bernardi *

*Laboratory of Molecular Evolution, Stazione Zoologica Anton Dohrn, 80121 Napoli, Italy*

Received 1 February 2006; received in revised form 4 May 2006; accepted 18 May 2006
Available online 6 September 2006
Received by E. Olmo

**Abstract**

Previous investigations from our laboratory [Jabbari, K., Cacciò, S., Pais de Barros, J.P., Desgres, J., Bernardi G., 1997. Evolutionary changes in CpG and methylation levels in the genome of vertebrates. Gene 205, 109–118.] led to the discovery of two different methylation levels in the genomes of vertebrates, a higher one exhibited by fishes and amphibians and a lower one shown by mammals and birds. It was also noted that data from the literature indicated a higher CpG level in fishes and amphibians compared to mammals and birds. Such observations led to suggesting the existence of two equilibria and to speculate that the transitions between the two equilibria in DNA methylation and CpG levels were due to a higher deamination rate in warm-blooded vertebrates related to their higher body temperature. Here we used Reverse-Phase High-Performance Liquid Chromatography (RP-HPLC) analysis to study methylation levels in a number of fish genomes living at different temperatures. We found that polar fishes exhibit DNA methylation levels that are higher than those of tropical and temperate fishes, the latter being in turn higher than the methylation levels of warm-blooded vertebrates, as expected from previous work. A closer analysis of the data revealed that, among Antarctic fishes, the *Channichthyidae* (the icefishes, deprived of haemoglobin) had the highest methylation level, and that, among temperate and tropical fishes the latter showed the lowest methylation level. These results confirm the existence of an inverse relationship between DNA methylation and body temperature, when the latter is maintained over evolutionary times.


*Keywords:* Methylcytosine; Deamination; Genome composition; Thermostability

## 1. Introduction

Body temperature ($T_b$) is one of the major factors affecting the life of an ectotherm organism. In particular it has an impact on gene expression (Gracey et al., 2004; Somero, 2005). When lasting over evolutionary times, body temperature can act, however, not only at the gene level, but also at the genome level (Thiery et al., 1976; Bernardi and Bernardi, 1986). The selective advantages provided by the GC increases (GC is the molar fraction of guanine and cytosine in DNA) that accompanied the emergence of warm- from cold-blooded were proposed to be the higher thermodynamic stabilities of DNA, RNA and proteins, all these advantages being achieved simultaneously. This is the thermostability hypothesis of Bernardi and Bernardi (1986; see also Bernardi, 2005).

Fishes are ideally suited in order to investigate the effect of body temperature on genome structure since they are ectotherms, which means that their body temperature follows the environmental temperature, and since they live in both cold and warm waters. Therefore, fish genomes were the subject of investigations in our laboratory over many years (see Thiery et al., 1976; Hudson et al., 1980; Pizon et al., 1984; Bernardi and Bernardi, 1986, 1990a,b, 1991; Jabbari et al., 1997; Jabbari and Bernardi, 2004a; Bucciarelli et al., 2002; Bernardi, 2005).

While DNA sequences are precious resources to investigate environmental effects on the genome, another opportunity is offered by the study of DNA methylation. Two parameters, 5-methylcytosine (5mC) and CpG frequencies, lend themselves to precise measurements in order to establish possible correlations with body temperature. Because of the excellent positive correlations of both CpG and 5mC with GC (Bernardi et al., 1985; Bernardi, 1985; Bernardi and Bernardi, 1986; Jabbari et al.,

*Abbreviations:* AFGP, antifreeze glycopeptides; CIAP, calf intestinal alkaline phosphatase; GC, molar ratio of guanosine+cytidine; 5mC, 5-methylcytosine; o/e, observed/expected; RP-HPLC, Reversed-Phase High-Performance Liquid Chromatography; U, unit.

* Corresponding author. Tel.: +39 081 5833215; fax: +39 081 2455807.
*E-mail address:* bernardi@szn.it (G. Bernardi).

0378-1119/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2006.05.031

1997), data can be compared among genomes covering a wide GC range.

Some previous investigations from our laboratory (Jabbari et al., 1997) focused on the level of 5mC in DNA from vertebrates. Our analyses concerned 42 species, but additional data from the literature brought the total to 87 species from all vertebrate classes. At that time, the data indicated that the genomes of fishes and amphibians have, on average, a two-fold higher methylation compared to those of mammals and birds and that two positive, parallel correlations hold between 5mC and GC levels of the genomes of fishes/amphibians and mammals/birds, respectively. Moreover, the 5mC and CpG observed/expected values showed no overlap between the two groups of vertebrates, suggesting the existence of two distinct equilibria. At the time of the publication of the paper by Jabbari et al. (1997), it was difficult to draw any final conclusion about the causes of the transition in methylation and CpG levels. We did, however, speculate that the lower 5mC level of warm-blooded vertebrates was due to a higher deamination rate related to their higher body temperature. Indeed, the deamination of 5mC residues in double-stranded DNA is known to have a strong temperature dependence (Shen et al., 1994).

Recently, we have approached this problem again along three lines. The first one, dealt with in this paper, was to analyze a large sample of fishes living at different temperatures, ranging from about 0 °C for Antarctic fishes to 10°–30 °C for temperate/tropical fishes. The second one, to be presented elsewhere (Varriale and Bernardi, in preparation), concerned the case of fishes belonging to the same family or genus, but characterized by different body temperatures. The third one was to investigate further the case of reptiles (Varriale and Bernardi, 2006).

Most polar fishes analyzed here belong to the suborder *Notothenioidei* (order *Perciformes*), which is the dominant component of the Southern Ocean fauna, and to the order *Gadiformes*, which live in Arctic waters. Notothenioids began to diversify in the period of isolation of the Antarctic continental shelf in the middle Tertiary, adapting in 20–30 million years to a progressive cooling (see Eastman and McCune (2000), for a review on history and geomorphological characteristics of the Antarctic shelf). The Southern Ocean offers a uniquely stable thermal environment with an annual temperature range of approximately −1.5 to +1.5 °C. The waters near the continental shelf are covered by 2–3 m of ice for 10 months and an additional 1–2 m platelet ice adheres to the underside of the sea ice. Coping with the presence of ice has been another significant evolutionary challenge for the fish fauna. The cold adaptation of fishes is expected to obviate the need to retain a functional plasticity required in more variable ecosystems (Somero, 1995). Cold adaptation is reflected in remarkable physiological characteristics, most notably the production of antifreeze glycopeptides (AFGPs) that cause the depression of the freezing point of body fluids (DeVries and Wohlschlag, 1969) and a peculiar system of oxygen transport in blood (see di Prisco, 2000, for a review). Although most Notothenioids live within the Antarctic region, 28 species are found outside Antarctica off the coasts of southern South America and New Zealand (Eastman, 2005), where they experience considerably higher and more variable temperatures than their Antarctic counterparts (Eastman and Clarke, 1998), water temperature being 4–10 °C. These sub-Antarctic Notothenioids include both phylogenetically basal lineages, like *Bovichthidae* and *Eleginopidae*, and species from clades that are hypothesized to have Antarctic origins, but have dispersed to non-Antarctic regions (Eastman, 1993; Stankovic et al., 2001; Eastman, 2005, and references therein). These latter species possess the AFGP genes in their genome and indeed their divergence from Antarctic relatives occurred much later than the formation of the Antarctic Polar Front: such colonization might have been allowed by a "leaking" of the barrier due to occasional climate change. We will include these fishes in the polar group together with Arctic and Antarctic fishes (Stankovic et al., 2001). In this category we have included also the three basal species, *Cottoperca gobio*, *Eleginops maclovinus* and *Bovichthus diacanthus*. Indeed, even if they have never inhabited the South Pole and do not possess AFGPs, we can consider them cold temperate. This also applies to *Merluccius merluccius* and *Gadus morhua*, migratory species spanning from Arctic to southern latitudes in the Atlantic Ocean.

In the Arctic sea, isolation is less stringent and the range of temperature variation is wider, facilitating migration of the ichtyofauna. Investigations on Arctic fishes are more recent, and many differences exist between Arctic and Antarctic marine environment and climate, migration, and evolutionary history (see Eastman, 1997 for a review), but there are cases of convergent evolution regarding the presence of AFGPs (Chen et al., 1997).

In the case of the tropical fishes analyzed, the temperatures of waters in which they live are 20°–30 °C, quite high in comparison with the polar water temperature (−1° to 10 °C), whereas temperate fishes live at 10°–15 °C.

Needless to say, other factors as water depth, migration, habitat (marine or freshwater) and time of adaptation to present lifestyle can influence actual body temperature and/or its effect on genome.

## 2. Materials and methods

### 2.1. DNA and tissue samples

DNA or tissue samples were obtained from the fishes listed in Table 1 whose geographic distribution, climatic classification and habitat are presented in Supplementary Table S1. Given that the differences in 5mC level among tissues from the same species are generally much smaller than those among species (Jabbari et al., 1997), we analyzed just one tissue (mainly liver, testes or muscle) per species.

### 2.2. DNA extraction and ultracentrifugation analysis

DNA was extracted from tissues by using a phenol–chloroform procedure as described by Sambrook et al. (1989) or by using the Genomix (Talent) extraction kit. DNA was quantified by measuring absorbance at 260 nm and its purity checked by the ratios A260/A230 and A260/A280. In order to detect the presence of satellites in the genomic DNA, ultracentrifugation analyses of samples in cesium chloride (CsCl) density gradient were carried out using a Beckman ultracentrifuge model Optima XL-A. The modal buoyant density, $\rho_0$, and the molar ratio of deoxyguanosine + deoxycytidine (GC) were calculated as described by Thiery

Table 1
List of the fish species analyzed[(1)]

| Order | Family | Species | Source | *c*-value | Species | | | Families/ genera | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | GC | 5mC | *R** | GC | 5mC |
| *Osteoglossiformes* | *Pantodontidae* | *Pantodon buchholzi* | | 0.77 | 45.66 | 1.77 | 3.88 | | |
| | *Notopteridae* | *Notopterus notopterus* | | | 44.96 | 1.22 | 2.72 | | |
| *Clupeiformes* | *Clupeidae* | *Sardina pilchardus* | a | | 47.12 | 1.40 | 2.98 | | |
| *Cypriniformes* | *Cyprinidae* | *Carassius auratus* | b | 1.60 | 39.53 | 1.44 | 3.64 | | |
| | | *Brachydanio rerio* | b | 1.68 | 39.19 | 1.35 | 3.45 | 39.36 | 1.40 |
| *Salmoniformes* | *Salmonidae* | *Oncorhynchus keta* | | 2.49 | 45.75 | 1.49 | 3.25 | | |
| *Gadiformes* | *Merlucciidae* | *Merluccius merluccius* | c | | **48.69** | **2.18** | 4.47 | | |
| | | *Merluccius bilinearis* | | 0.84 | | | | | |
| | *Gadidae* | *Boreogadus saida* | d | 0.88 | **48.40** | **2.22** | 4.58 | | |
| | | *Gadus morhua* | d | 0.40 | **48.61** | **2.37** | 4.88 | | |
| | | *Arctogadus glacialis* | d | | **48.13** | **2.74** | 5.68 | | |
| *Cyprinodontiformes* | *Aplocheilidae* | *Nothobranchius flammicomantis* | e | | 47.26 | 1.59 | 3.36 | | |
| | | *Aphyosemion elegans* | f | | 43.49 | 1.52 | 3.49 | | |
| | | *Aphyosemion coeruleum* | | 1.20 | | | | | |
| | *Rivulidae* | *Aphyolebias peruensis* | e | | 45.70 | 1.50 | 3.28 | | |
| | *Fundulidae* | *Fundulus heteroclitus* | | 1.36 | 42.80 | 1.76 | 4.11 | | |
| | *Poecilidae* | *Xiphophorus maculatus* | | 0.76 | 41.28 | 1.39 | 3.37 | | |
| | *Cyprinodontidae* | *Jordanella floridae* | | | 41.43 | 1.10 | 2.66 | | |
| | | *Aphanius fasciatus* | g | | 43.17 | 1.62 | 3.75 | | |
| *Zeiformes* | *Caproidae* | *Capros aper* | a | | 46.69 | 1.87 | 4.00 | | |
| *Scorpaeniformes* | *Scorpaenidae* | *Scorpaena guttata* | b | | 41.27 | 1.48 | 3.58 | | |
| | | *Scorpaena brasiliensis* | | 1.40 | | | | | |
| | *Hexagrammidae* | *Ophiodon elongatus* | b | | 44.31 | 1.78 | 4.02 | | |
| | *Cottidae* | *Leptocottus armatus* | b | | 46.72 | 2.11 | 4.52 | | |
| | *Liparidae* | *Liparis tunicatus* | d | 0.88 | **46.30** | **2.79** | 6.02 | | |
| *Perciformes* | *Serranidae* | *Serranus cabrilla* | a | | 44.39 | 1.66 | 3.75 | | |
| | | *Alphestes immaculatus* | b | | 43.40 | 1.40 | 3.23 | | |
| | | *Paralabrax maculatofasciatus* | b | | 43.63 | 1.36 | 3.13 | | |
| | *Apogonidae* | *Apogon imberbis* | a | | 43.25 | 1.51 | 3.50 | | |
| | | *Apogon bandanensis* | | 0.94 | | | | | |
| | *Carangidae* | *Trachurus mediterraneus* | a | | 44.58 | 1.65 | 3.70 | | |
| | | *Trachurus declivis* | | 0.90 | | | | | |
| | *Sparidae* | *Boops boops* | c | 0.52 | 44.65 | 1.83 | 4.10 | | |
| | *Mullidae* | *Mullus barbatus* | c | | 48.86 | 2.19 | 4.49 | | |
| | *Pomacanthidae* | *Holacanthus passer* | b | | 44.10 | 1.43 | 3.25 | | |
| | *Kyphosidae* | *Hermosilla azurea* | b | | 44.26 | 1.41 | 3.18 | | |
| | *Pomacentridae* | *Microspathodon dorsalis* | b | | 43.15 | 1.54 | 3.57 | | |
| | | *Chromis* sp. | b | 1.10 | 43.80 | 1.50 | 3.42 | 43.48 | 1.52 |
| | *Labridae* | *Coris julis* | c | 1.21 | 44.90 | 1.28 | 2.85 | | |
| | | *Thalassoma grammaticum* | b | | 42.59 | 1.29 | 3.02 | | |
| | | *Halichoeres poeyi* | b | | 44.43 | 1.41 | 3.16 | | |
| | | *Halichoeres bathyphilus* | | 0.80 | | | | | |
| | | *Symphodus tinca* | a | 0.73 | 45.22 | 1.81 | 4.01 | | |
| | *Scaridae* | *Scarus ghobban* | b | | 41.95 | 1.32 | 3.15 | | |
| | | *Scarus guacamaia* | | 1.25 | | | | | |
| | *Zoarcidae* | *Iluocoetes fimbriatus* | h* | 1.07 | **46.62** | **1.92** | 4.11 | | |
| | *Bovichtidae* | *Cottoperca gobio* | h* | | **43.65** | **1.98** | 4.54 | | |
| | | *Bovichtus diacanthus* | h* | | **41.95** | **1.83** | 4.35 | | |
| | *Eleginopidae* | *Eleginops maclovinus* | h* | | **44.04** | **1.66** | 3.76 | | |
| | *Nototheniidae* | *Notothenia coriiceps* | h | | **44.40** | **1.87** | 4.22 | | |
| | | *Notothenia rossii* | h | | **44.52** | **1.78** | 4.00 | **44.46** | **1.83** |
| | | *Gobionotothen gibberifrons* | h | | **42.62** | **2.15** | 5.05 | | |
| | | *Gobionotothen marionensis* | h* | | **44.32** | **1.92** | 4.33 | **43.47** | **2.04** |

(*continued on next page*)

Table 1 (*continued*)

| Order | Family | Species | Source | *c*-value | Species | | | Families/ genera | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | GC | 5mC | *R** | GC | 5mC |
| *Perciformes* | *Nototheniidae* | *Trematomus newnesi* | h | | **44.57** | **1.82** | 4.09 | | |
| | | *Trematomus bernacchii* | h | 1.19 | **43.59** | **1.82** | 4.17 | **44.08** | **1.82** |
| | | *Dissostichus mawsoni* | h | 1.02 | **44.09** | **1.85** | 4.20 | | |
| | | *Patagonotothen guntheri* | h* | | **44.08** | **1.69** | 3.84 | | |
| | | *Lepidonotothen squamifrons* | h* | | **43.79** | **1.87** | 4.27 | | |
| | | *Lepidonotothen kempi* | h | | **43.31** | **1.64** | 3.78 | | |
| | | *Lepidonotothen nudifrons* | h | | **43.74** | **1.79** | 4.09 | **43.61** | **1.77** |
| | *Bathydraconidae* | *Gymnodraco acuticeps* | h | | **42.60** | **1.80** | 4.23 | | |
| | | *Parachaenichthys charcoti* | h | | **43.24** | **1.99** | 4.59 | **42.92** | **1.89** |
| | | *Cygnodraco mawsoni* | | 1.41 | | | | | |
| | *Channichthyidae* | *Chionodraco rastrospinosus* | h | | **43.65** | **2.28** | 5.22 | | |
| | | *Chionodraco hamatus* | h | 1.83 | **43.90** | **2.40** | 5.47 | **43.78** | **2.34** |
| | | *Champsocephalus esox* | h* | | **45.53** | **2.22** | 4.88 | | |
| | | *Pseudochaenichthys georgianus* | h* | | **44.90** | **2.22** | 4.94 | | |
| | | *Chaenocephalus aceratus* | h | | **44.27** | **2.10** | 4.74 | | |
| | | *Neopagetopsis ionah* | c | | **43.80** | **2.08** | 4.75 | | |
| | *Tripterygiidae* | *Lepidonectes corallicola* | b | | 44.52 | 1.42 | 3.20 | | |
| | *Labrisomidae* | *Dialommus fuscus* | b | | 45.61 | 1.85 | 4.05 | | |
| *Pleuronectiformes* | *Pleuronectidae* | *Paralichthys californicus* | b | 0.80 | 46.50 | 1.58 | 3.40 | | |
| *Tetraodontiformes* | *Tetraodontidae* | *Tetraodon fluviatilis* | b | | 48.39 | 1.93 | 3.99 | | |
| | | *Tetraodon nigroviridis* | | 0.35 | | | | | |

[(1)]Taxonomy is from www.fishbase.org. The underlined samples refer to fishes only used for c-values.

[(2)]Sources: (a) Angela Paglialonga and Flegra Bentivegna, Aquariology Laboratory, Stazione Zoologica Anton Dohrn, Naples, Italy; (b) Bucciarelli et al., 2002; (c) local fish market; (d) Donatella de Pascale, (IBP-CNR), Naples, Italy; (e) Daniele Cuni, AIK, Italy; (f) Dirk Ude and Axel Schwekendiek, Deutsche Killifisch Gemeinschaft (DKG), Germany; (g) Simona Santini, Laboratory of Molecular Evolution, Stazione Zoologica Anton Dohrn, Naples, Italy; (h) Ennio Cocca, Istituto di Biochimica delle Proteine-Consiglio Nazionale delle Ricerche (IBP-CNP), Naples, Italy. Asterisks refer to samples obtained in the ICEFISH Cruise, NBP04-04. Blank spaces indicate unknown source.

[(3)]c-values expressed in picograms (pg) are from the database Animal Genome Size Database (Release 2.0; Gregory, T.R. 2005), available at URL http://www.genomesize.com. We reported the average for some samples which had multiple annotations: *Carassius auratus*, *Danio rerio*, *Fundulus heteroclitus*, *Xiphophorus maculatus*, *Gadus morhua*, *Merluccius bilinearis*, *Trematomus bernacchii*, *Tetraodon nigroviridis*.

[(4)]R is the 5mC/GC ratio multiplied by $10^2$. Polar fish values are in bold.

[(5)]Averages were calculated for species belonging to the same family/genus (indicated by thick vertical bars). This was not done when GC and/or 5mC values for species within a family/genus diverged too much from each other (indicated by thin vertical bars).

et al. (1976). GC, buoyant densities ($\rho_0$) and analytical GC density gradient profiles are shown in Supplementary Table S2 and Supplementary Fig. S1 respectively. Profiles from Bucciarelli et al. (2002) are presented in Supplementary Fig. S2.

### 2.3. Enzymatic hydrolysis of DNA

The procedure used was a modification of the method described by Gehrke et al. (1984) and Parra et al. (2001). Three to ten micrograms of DNA dissolved in 20 μl of water was heated at 100 °C for 2 min, then quenched in ice water. Fifty microliters of 30 mM sodium acetate pH 5.3 and 2.5 μl of 20 mM zinc sulfate were added to the samples followed by 10 U (units) of nuclease P1 (1 U/μl in 30 mM NaAc pH 5.3). DNA hydrolysis was carried out overnight at 37 °C. 10 U of calf intestine alkaline phosphatase (CIAP) and 20 μl of CIAP buffer were then added and the samples were incubated for an additional 3 h. The resulting 2′-deoxynucleosides were then filtered on MWCO 3 kDa (Millipore) and injected in the HPLC column, or stored frozen at −20 °C until HPLC analysis. We used nuclease P1 together with CIAP because it was previously demonstrated (Gehrke et al., 1984) that a complete degradation to deoxynucleosides is obtained, a high degree of reproducibility is achieved and pairing bases are eluted in stoichiometric amounts (Palmgren et al., 1990). Moreover, these enzymes do not cause deamination of deoxyadenosine.

### 2.4. RP-HPLC analysis

As a method of separation, we used RP-HPLC because it is more sensitive and more reproducible than other methods (Gehrke et al., 1984), and because it allows separating deoxyribonucleosides from the ribonucleosides that might originate from contaminating RNA. To identify the latter, we ran separately standard ribonucleosides and deoxyribonucleosides to record the respective retention times, as described by Ramsahoye (2002). A modification of the two-buffer step gradient procedure described by Gehrke et al. (1984) was developed, using a 25-cm reversed-phase column (Beckman-Coulter). Buffer A was 50 mM $KH_2PO_4$ sterilized in autoclave and filtered through a Millipore GS-22 filter (0.22 μm). Solvent B was 95% (v:v) methanol HPLC grade (J.T. Baker).

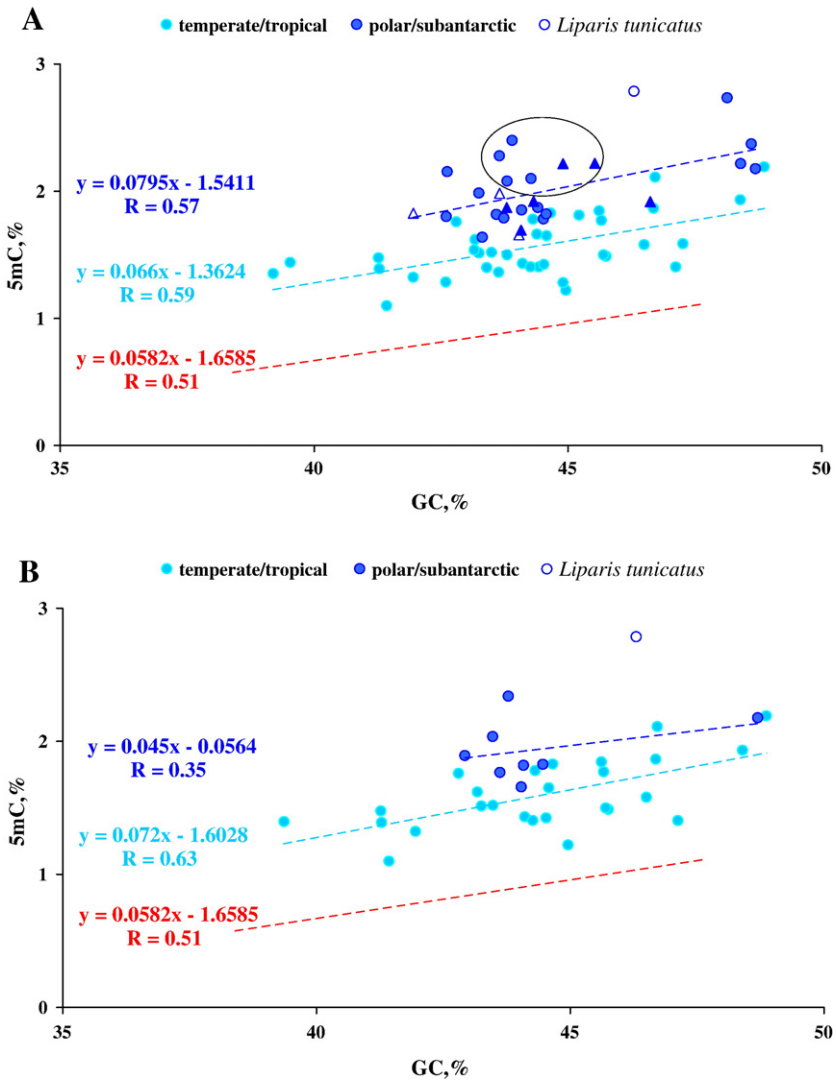


Fig. 1. A. Plot of 5mC levels against GC levels for the genomes of polar fishes (dark blue circles, dark blue triangles which concern sub-Antarctic fishes of polar origin and containing AFGPs, and empty triangles concerning native sub-Antarctic) and temperate/tropical fishes (light blue circles); the oval includes the species belonging to the family *Channichthyidae*. Values are listed in Table 1, column Species. B. Plot of average 5mC levels against average GC levels for families/genera (listed in Table 1, column Families/genera) of polar fishes (dark blue circles) and of tropical/temperate fishes (light blue circles). Families represented by single species (listed in Table 1, column Species) were also included in the figure, whereas families whose values could not be averaged (see Table 1) were not. In both a and b the open circle refers to *Liparis tunicatus* (family: *Scorpaenidae*), which was not taken into account in drawing the regression line because it was considered an outlier. The regression line (broken red line) for the mammalian DNA (Varriale and Bernardi, in preparation) was reported as a reference. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

We used a two-buffer step gradient system followed by a linear gradient to 100% solvent B during 10 min. Then solvent B was pumped for 5 min to flush retained material, followed by a linear gradient to 100% buffer A during 10 min for re-equilibration. The flow rate was 1.2 ml/min and the column temperature was 35 °C. Deoxyribonucleosides were detected at $\lambda_{MAX}$ and quantified by determining the $Area_{MAX}$ for each by using a diode array system (Detector 168, Beckman-Coulter). We checked the hydrolysis on 1 to 20 μg DNA samples from calf thymus dissolved in 20 μl of water and found the same molar ratios. The precision of the method was determined by injecting replicate aliquots of a single DNA hydrolysate. We carried out the RP-HPLC analysis for all samples at least twice and determined the mol% of each deoxyribonucleoside. Finally, we calculated the average for the mol% of each and considered as good the samples that gave a standard deviation (whenever it was possible to calculate it) lower than 0.15% 5mC, whereas we discarded the analyses in which we found higher values. All reactions were carried out in parallel with a standard calf thymus DNA (Sigma) as a control. A comparison of our results with those obtained by Jabbari et al. (1997) on the same vertebrate species showed identity or very slight differences with no systematic trend.

### 2.5. Analysis of orthologous genes

In order to assess the frequency of CpG in coding sequences from both polar and temperate/tropical fishes, we downloaded partial and complete coding sequences of polar fishes from the

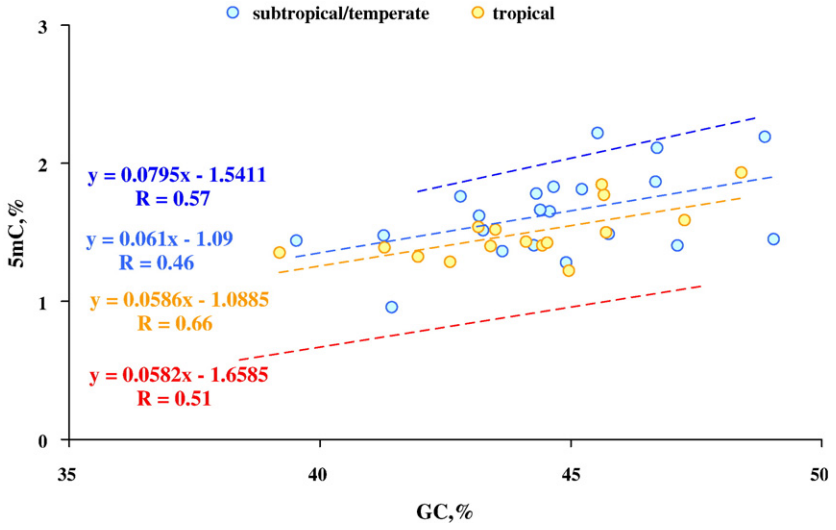


Fig. 2. Plot of 5mC levels against GC levels for the genomes of temperate and subtropical fishes (upper line) and of tropical fishes. Equations and correlation coefficients are given for the two groups as well as for polar fishes and mammals whose regression lines were given as reference. The classification of fishes as temperate, subtropical and tropical or is from the website http://www.fishbase.org, the most complete database for fishes.

suborder *Notothenioidei* and from the order *Gadiformes* from the GenBank database using the WWW-QUERY tool from the URL http://pbil.univ-lyon1.fr/search/query_fam.php. The final data set included 53 genes. Accession numbers of the genes used and data for CpG are reported in Supplementary Table S3.

We looked for orthologous genes in tropical or temperate fishes from the NCBI sequence database using the BLASTN search tool (http://www.ncbi.nlm.nih.gov/BLAST). Retrieved sequences were aligned using the Se–Al application version 2.0a11 Carbon (Rambaut, 1996; available at http://evolve.zoo.ox.ac.uk). The total CpG frequency was calculated on each sequence using the CODONW application (J. Peden), on the server at URL http://bioweb.pasteur.fr/seqanal/interfaces/codonw.html.

## 3. Results

### 3.1. Analysis of 5mC level in fish DNA

Table 1 presents the taxonomy, and the sources of the fishes analyzed, as well as the GC and 5mC levels of their genomes, the averages per genus or family (in a number of cases), the ratios 5mC/GC and genome size whenever available. Our dataset comprised 65 fishes belonging to 12 out of the 38 teleostean orders from the class *Actinopterygii* (Nelson, 1994). We chose to analyze a number of fish species large enough to obtain a representative sample of different orders and families, our purpose being to investigate the possible differences in methylation level in relation to differences in body temperature.

The 5mC values for the species are plotted against GC levels in Fig. 1A, whereas a similar plot for family/genera average values and for families represented by single species (see Table 1) is displayed in Fig. 1B. This analysis avoids, at least in part, biases associated with different number of species in different families/genera. An attempt to use data for fish orders did not produce meaningful results because only three points were available for the polar group. In the case of *Gadidae* and *Labridae*, values were not averaged because differences in GC and/or 5mC among the species were too large. These values were not included in Fig. 1B. Fishes belonging to three families, *Gobidae*, *Scombridae*, *Cichlidae*, were not included in the Table because they showed differences in body temperature and methylation levels that will be discussed elsewhere (Varriale and Bernardi, in preparation).

In the polar group we included both polar and sub-Antarctic fishes (see Introduction). The data points for polar and temperate/tropical fishes aligned on two parallel regression lines, the higher of which concerned polar fishes. The equation of the lower regression line, concerning temperate/tropical fishes was higher than the parallel line corresponding to mammalian DNAs also shown for reference, and was similar to that reported by

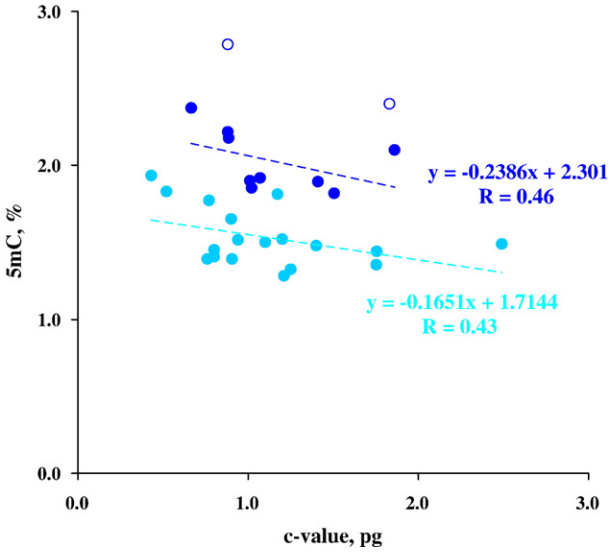


Fig. 3. Plot of 5mC against genome size for polar fishes (dark blue circles) and temperate/tropical fishes (light blue circles). The broken dark blue line refers to the correlation when all points were taken into account except for the two outliers (empty circles). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

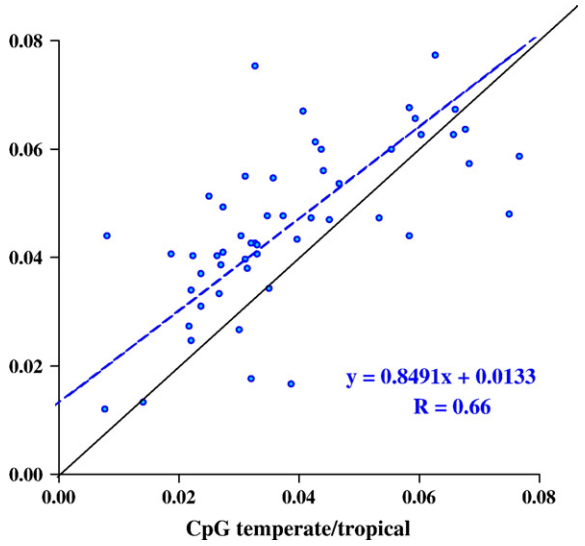


Fig. 4. CpG levels of genes from polar fishes are plotted against CpG levels of orthologous genes from temperate/tropical fishes listed in Supplementary Table 3. The solid line is the diagonal.

Jabbari et al. (1997). The difference between DNA methylation level in polar/sub-Antarctic and temperate/tropical fishes is evident, the *t*-test being very highly significant ($p<10^{-7}$).

The slopes of the correlations of 5mC vs. GC were similar to those already seen by Jabbari et al. (1997) and are essentially due to the increase in CpG doublets (the major sites of DNA methylation) that accompany the increase in GC levels (Bernardi, 1985; Aïssani and Bernardi, 1991a,b). This conclusion is supported by two findings: (i) the slopes of the 5mC vs. GC plots are very similar in all classes and groups of vertebrates analyzed; this is true to the point that data pushing the slopes outside this parallel behavior may be considered as outliers (see below); (ii) the simplest correction, such as that obtained by plotting 5mC/GC vs. GC, reduces the slopes to almost zero (not shown).

A closer look at the data of Fig. 1a reveals some interesting features: (i) the outlier point of *Liparis tunicatus*, which exhibits by far the highest methylation, concerns a very deep ocean fish; this raises the possibility that factors other than temperature may influence methylation level; needless to say, more detailed analysis are required to substantiate this possibility; (ii) among Antarctic fishes, the *Channichthyidae*, definitely show the highest methylation levels (if *L. tunicatus* is neglected); since the body temperature of these fishes is the same as that of *Notothenioidei*, here we should again consider the possibility of other factors influencing methylation levels; at this point, a relevant remark is that *Channichthyidae*, the globin-less icefishes, are characterized by a particularly low metabolic rate (Hemmingsen, 1991); if this family is neglected, expectedly the correlation coefficient of polar fishes improves; (iii) if tropical and temperate/subtropical fishes are dealt with separately, two different lines can be drawn for the two groups, the lower one belonging to tropical fishes (Fig. 2). The significance of the difference is however borderline ($p=0.06$).

### 3.2. Correlation of 5mC with repetitive sequences and genome size

Apart from body temperature, two other factors are known to influence 5mC levels (see Jabbari et al., 1997). The first one is the amount and the methylation level of repetitive DNA,

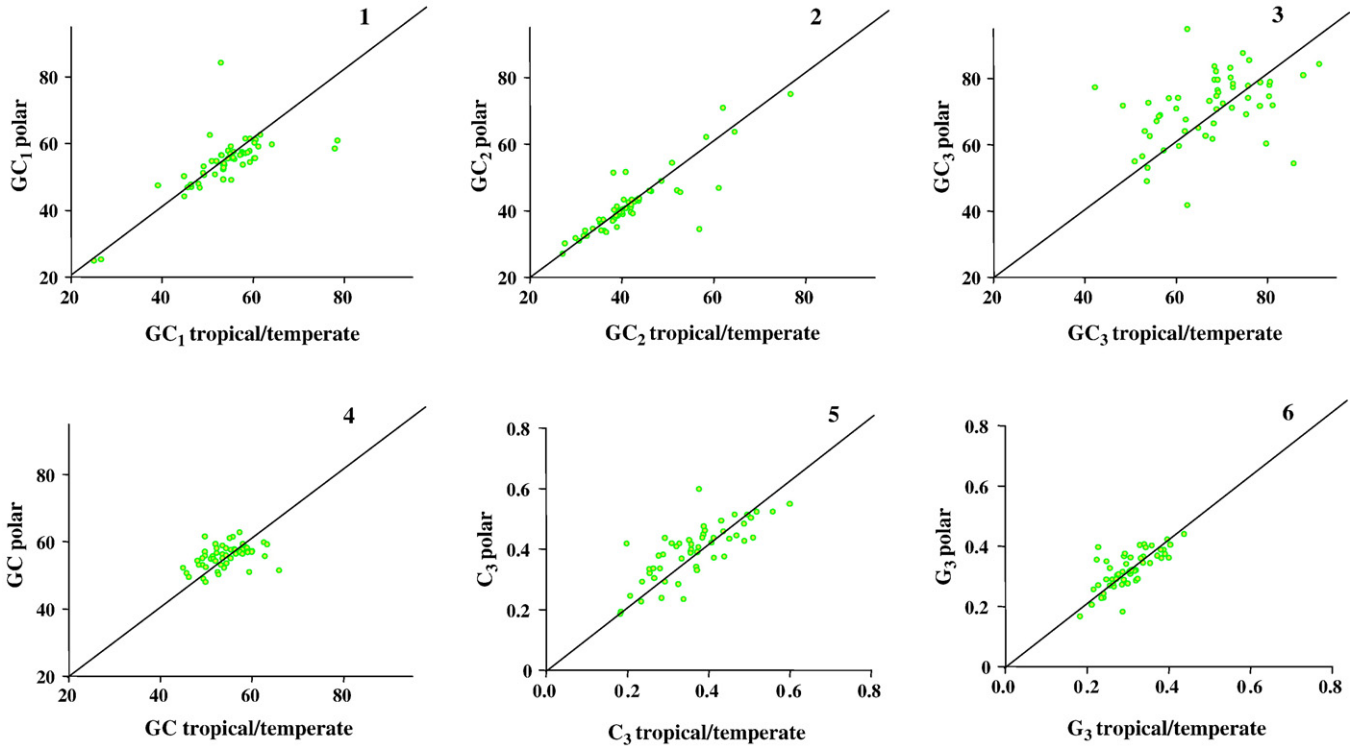


Fig. 5. GC, GC1, GC2, GC3, C3 and G3 levels from the genes from polar fishes are plotted against the corresponding levels of orthologous genes from temperate/tropical fishes listed in Supplementary Table 3. The solid lines are the diagonals.

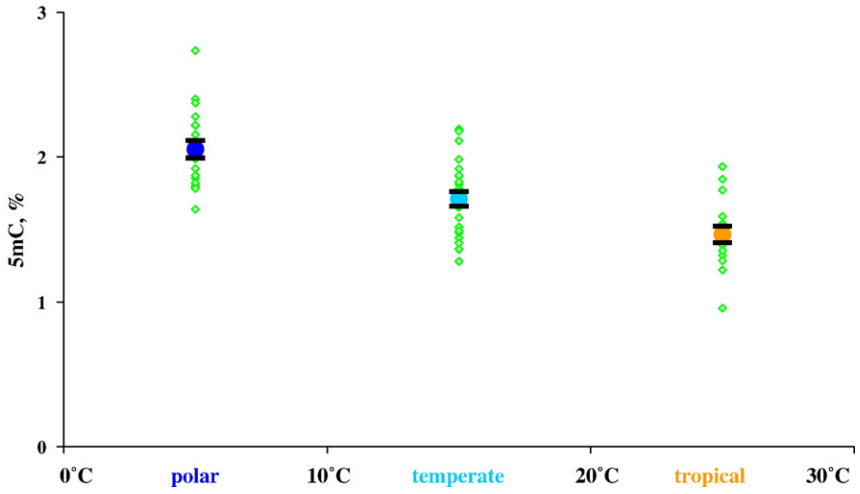


Fig. 6. Correlation between 5mC level and body temperature. The two extreme points correspond to the genomes of polar and tropical fishes, respectively; the middle point is the average body temperature of temperate fishes. Standard error bars are shown.

both satellite and interspersed. Unfortunately, assessing such variables requires a very detailed analysis of each DNA. We did not carry out such an investigation in the present work for the following reasons: (i) it was shown in a detailed comparison of compositional fractions from the genome of human, mouse, chicken and *Xenopus* that the latter genome showed a higher methylation level than those of the three warm-blooded vertebrates, after eliminating the effect of the satellite sequences, such effect generally being a minor one (Cacciò et al., 1997); (ii) repetitive sequences are variable from one family, genus and sometimes even species to another one; they are more likely, therefore, to contribute to the scatter of points than to a systematic bias; (iii) the 5mC differences between polar and temperate/tropical fishes are large, in fact almost comparable with those separating temperate/tropical fishes and warm-blooded vertebrates. Obviously, outliers deserve a more detailed analysis that will be carried out in the future.

The second factor is genome size. In this case, it is known that GC shows a negative correlation with genome size (Bernardi and Bernardi, 1990a). Indeed, increasing the genome size means increasing GC-poorer regions of the genome, which leads to a relative decrease of methylation. Since 5mC is positively correlated with GC, one expects a negative correlation between 5mC and genome size, which was indeed found by Jabbari et al. (1997) and confirmed in the present study (see Fig. 3), in which two correlations were found for polar and temperate/tropical fishes, respectively. The point made by Fig. 3 is that in every case at comparable genome sizes DNA methylation is higher in polar fishes than in temperate/tropical fishes.

Two further checks on the differences found between polar/sub-Antarctic and temperate/tropical fishes was done by comparing CpG levels in orthologous genes from the two classes of fishes. It is known that, indeed, there is a difference in CpG levels between the genomes of fishes and mammals that CpG levels of platypus (body temperature 32 °C) are intermediate between the former and the latter (Jabbari and Bernardi, 2004b), and that CpG is linearly correlated with methylation levels (Cacciò et al., 1997). In our case, the comparison could only be done at the level of orthologous coding sequences. Our results (Fig. 4) show that CpG levels are generally higher in genes from polar fishes compared to their orthologs from temperate/tropical fishes. A plot with $C_3pG_1$ shows similar results (data not shown).

The results of Fig. 4 might, however, be due to a higher GC level of coding sequences from polar/sub-Antarctic fishes. This point was checked by comparing the GC, GC1, GC2 and GC3 (the GC levels of first, second and third codon position) of orthologous genes from polar/sub-Antarctic and temperate/tropical fishes (Fig. 5). The first three plots showing that points are on the diagonal (except for a few outliers) can be taken as a check of the fact that the genes under consideration are indeed orthologous. The last plot with only a very modest predominance of higher GC3 values in polar fishes indicate that the CpG results of Fig. 4 are not simply due to a higher

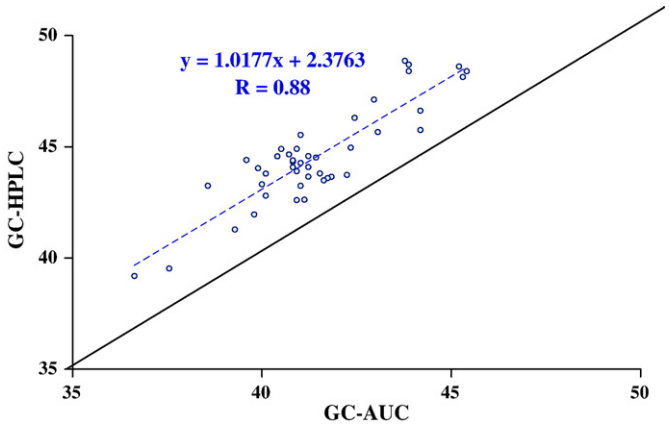


Fig. 7. Correlation between genomic GC levels as obtained by nucleoside analysis (HPLC) and by analytical ultracentrifugation (AUC) in fishes. Analytical ultracentrifuge profiles of fish DNA are reported in Supplementary Figs. 1 and 2.

GC3 level in polar fishes. This comparison is confirmed by the separate plots of C3 and G3.

### 3.3. Correlation of the methylation levels with taxonomy and phylogeny

Since the DNA methylation results obtained for polar and temperate/tropical fishes clearly confirmed the inverse correlation between methylation and body temperature observed by Jabbari et al. (1997), this point was analyzed in more detail for several cases. In general, fishes belonging to the same genus and sharing the same body temperature, as it was the case for several genera of the family *Nototheniidae* and for the *Chionodraco* genus of the family *Channichthyidae*, showed similar GC and 5mC levels. This also applies to species from different genera but belonging to the same Antarctic families and to species from different genera within *Cyprinidae* and *Pomacentridae*.

This relationship with phylogeny under conditions of close body temperature was remarkably broken when body temperature was different, a higher body temperature being associated with a lower 5mC level. This applies to cases from *Gobidae*, *Scombridae*, *Gadidae* and *Cichlidae* (to be presented in detail elsewhere; Varriale and Bernardi, in preparation).

### 3.4. Correlation of methylation level and body temperature

Fig. 6 displays a plot of 5mC against body temperature, Tb. We grouped the fishes into the three categories: polar, temperate and tropical, as discussed above and calculated the average methylation level and the standard error for each group, finding an inverse relationship between 5mC and temperature. The first range from 0 °C to 10 °C and includes polar and sub-Antarctic species. The last range from 20 °C to 30 °C includes at least the majority of tropical fishes, whereas the middle range from 10° to 20 °C, comprises temperate fishes.

### 3.5. Correlation between GC levels as derived from nucleoside analysis and from CsCl density gradient centrifugation

If GC levels are calculated from the modal buoyant densities of fish DNAs, the values obtained are lower than those from nucleoside analysis by 2–3% GC (Fig. 7). This difference is due, at least in part, to DNA methylation. Indeed, when GC is derived from buoyant density, calculation is based on the correlation between GC and buoyant density of bacterial DNAs, which are not at all or barely methylated. The effect of DNA methylation on buoyant density is to lower it (Kirk, 1967). This is not, however, the full explanation because the ~3% difference in the GC estimate would be accounted for by an ~4% 5mC level, whereas this level is about 2% in polar fish DNAs and even lower in non-polar fishes. An additional contribution may come from the frequencies of di- and tri-nucleotides, which are different compared to the reference bacterial DNAs (see Corneo et al., 1968). In any case, the correlation of Fig. 7 allows correcting the GC values of fish genome as directly derived from the equation of Schildkraut et al. (1962).

## 4. Discussion

Apart from some remarks already made in the preceding section, three points deserve to be discussed here. The first one is that the DNA methylation processes that are typical of imprinting and of developmental regulations (see reviews by Siegfried and Cedar, 1997; Holliday, 2005; Vanyushin, 2005; Robertson, 2005) should be separated from the overall genome methylation. Indeed the former one concerns the *de novo* methylation, which is due to DNA methyltransferases DNMT3a and DNMT3b (Bestor, 2000), whereas the second one concerns maintenance methylation, which is due to the action of DNA methyltransferase 1 (DNMT1) on hemimethylated DNA.

The second point is that the previous investigations on overall methylation levels in DNA from different vertebrates (Vanyushin et al., 1970, 1973; Pollock et al., 1978; Ehrlich et al., 1982; Gama-Sosa et al., 1983a,b; Serrano et al., 1993) did not lead to any general conclusion. A first step in understanding the meaning of overall genome methylation (see below) was made by Jabbari et al. (1997). These authors provided the first demonstration that the average DNA methylation in fishes, 1.70%, and amphibians, 1.98%, are roughly twice as large as those found in mammals, 0.88%, and birds, 1.02%, and observed that a parallel phenomenon takes place in CpG levels. These results established a link between 5mC and CpG levels and body temperature. This link was confirmed and extended here by taking into consideration polar fishes. The differences observed between temperate and tropical fishes, as well as between *Channichthyidae* and Notothenioids, are of interest but should be studied further.

The third point is that the suggestion of the existence of two equilibria, concerning both 5mC and CpG levels, one for fishes/amphibians and another one for mammals/birds (Jabbari et al., 1997) received further support from the results reported here for Antarctic fishes. In this case, while the transitions from cold- to warm-blooded vertebrates was accompanied by a decrease in CpG and DNA methylation, the progressive cooling of Notothenioids seems to have opposite results.

## Acknowledgements

We thank Ettore Olmo, Giacomo Bernardi and Kamel Jabbari for their helpful comments. We are grateful to all colleagues who helped us with the gift of DNA or tissue samples, in particular to Ennio Cocca. We also thank Guido di Prisco and Cinzia Verde for discussions and references on polar fishes, the Italian National Program for Antarctic Research (PNRA), Hugo Naya and Oliver Clay for their help in the statistical analysis, Antimo D'Aniello, Gabriele Ferrandino, Giuseppe Geraci and Bernard Ramsahoye for helping in the improvement of the analytical method. The collection of sub-Antarctic samples (ICEFISH cruise 2004) was supported by U. S. National Science Foundation grant OPP-0132032 to H. William Detrich (Chief Scientist; Northeastern University). For the Arctic samples we wish to thank the TUNU-I expedition, organized by the University of Tromsoe, Norwegian College of Fishery Science, Tromsoe, Norway.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.gene.2006.05.031.

## References


Aïssani, B., Bernardi, G., 1991a. CpG islands: features and distribution in the genome of vertebrates. Gene 106, 173–183.

Aïssani, B., Bernardi, G., 1991b. CpG islands, genes and isochores in the genome of vertebrates. Gene 106, 185–195.

Bernardi, G., 1985. The organization of the vertebrate genome and the problem of the CpG shortage. In: Cantoni, G.L., Razin, A. (Eds.), Biochemistry and Biology of DNA Methylation. Alan Liss, N.Y., U.S.A., pp. 3–10.

Bernardi, G., 2005. Structural and evolutionary genomics. Natural Selection in Genome Evolution. Elsevier, Amsterdam. reprinted from 2004.

Bernardi, G., Bernardi, G., 1986. Compositional constrains and genome evolution. J. Mol. Evol. 24, 1–11.

Bernardi, G., Bernardi, G., 1990a. Compositional patterns in the nuclear genome of cold-blooded vertebrates. J. Mol. Evol. 31, 265–281.

Bernardi, G., Bernardi, G., 1990b. Compositional transitions in the nuclear genomes of cold-blooded vertebrates. J. Mol. Evol. 31, 282–293.

Bernardi, G., Bernardi, G., 1991. Compositional properties of nuclear genes from cold-blooded vertebrates. J. Mol. Evol. 33, 57–67.

Bernardi, G., et al., 1985. The mosaic genome of warm-blooded vertebrates. Science 228, 953–958.

Bestor, T.H., 2000. The DNA methyltransferases of mammals. Hum. Mol. Genet. 9, 2395–2402.

Bucciarelli, G., Bernardi, G., Bernardi, G., 2002. An ultracentrifugation analysis of two hundred fish genomes. Gene 295, 153–162.

Cacciò, S., Jabbari, K., Matassi, G., Guermonprez, F., Desgrès, J., Bernardi, G., 1997. Methylation patterns in the isochors of vertebrate genomes. Gene 205, 119–124.

Chen, L., DeVries, A.L., Cheng, C.H., 1997. Convergent evolution of antifreeze glycoproteins in Antarctic notothenioid fish and Arctic cod. Proc. Natl. Acad. Sci. U. S. A. 94, 3817–3822.

Corneo, G., Ginelli, E., Soave, C., Bernardi, G., 1968. Isolation and characterization of mouse and guinea pig satellite deoxyribonucleic acids. Biochemistry 7, 4373–4379.

DeVries, A.L., Wohlschlag, D.E., 1969. Freezing resistance in some Antarctic fishes. Science 163, 1073–1075.

di Prisco, G., 2000. Life style and biochemical adaptation in Antarctic fishes. J. Mar. Syst. 27, 253–265.

Eastman, J.T., 1993. Antarctic Fish Biology: Evolution in a Unique Environment. Academic Press, San Diego.

Eastman, J.T., 1997. Comparison of the Antarctic and Arctic fish faunas. Cybium 21, 335–352.

Eastman, J.T., 2005. The nature of the diversity of Antarctic fishes. Polar Biol. 28, 93–107.

Eastman, J.T., Clarke, A., 1998. Radiations of Antarctic and non-Antarctic fish. In: di Prisco, G., Pisano, E., Clarke, A. (Eds.), Fishes of Antarctica: A Biological Overview. Springer-Verlag Italia, pp. 3–26.

Eastman, J.T., McCune, A.R., 2000. Fishes on the Antarctic continental shelf; evolution of a marine species flock? J. Fish Biol. 57, 84–102.

Ehrlich, M., et al., 1982. Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. Nucleic Acids Res. 10, 2709–2721.

Gama-Sosa, M.A., et al., 1983a. Tissue-specific differences in DNA methylation in various mammals. Biochim. Biophys. Acta 740, 212–219.

Gama-Sosa, M.A., Wang, R.Y., Kuo, K.C., Gehrke, C.W., Ehrlich, M., 1983b. The 5-methylcytosine content of highly repeated sequences in human DNA. Nucleic Acids Res. 11, 3087–3095.

Gehrke, C.W., McCune, R.A., Gama-Sosa, M.A., Ehrlich, M., Kuo, K.C., 1984. Quantitative reversed-phase high-performance liquid chromatography of major and modified nucleosides in DNA. J. Chromatogr. 301, 199–219.

Gracey, A.Y., et al., 2004. Coping with cold: an integrative, multitissue analysis of the transcriptome of a poikilothermic vertebrate. Proc. Natl. Acad. Sci. U. S. A. 101, 16970–16975.

Hemmingsen, E.A., 1991. Respiratory and cardiovascular adaptation in hemoglobin free fish: resolved and unresolved problems. In: di Prisco, G., Maresca, B., Tota, B. (Eds.), Biology of Antarctic Fish. Springer-Verlag, New York, pp. 191–203.

Holliday, R., 2005. DNA methylation and epigenotypes. Biochemistry (Mosc.) 70, 500–504.

Hudson, A.P., Cuny, G., Cortadas, J., Haschemeyer, A.E., Bernardi, G., 1980. An analysis of fish genomes by density gradient centrifugation. Eur. J. Biochem. 112, 203–210.

Jabbari, K., Bernardi, G., 2004a. Body temperature and evolutionary genomics of vertebrates: a lesson from the genomes of *Takifugu rubripes* and *Tetraodon nigroviridis*. Gene 333, 179–181.

Jabbari, K., Bernardi, G., 2004b. Cytosine methylation and CpG, TpG (CpA) and TpA frequencies. Gene 333, 143–149.

Jabbari, K., Cacciò, S., Pais de Barros, J.P., Desgres, J., Bernardi, G., 1997. Evolutionary changes in CpG and methylation levels in the genome of vertebrates. Gene 205, 109–118.

Kirk, J.T., 1967. Effect of methylation of cytosine residues on the buoyant density of DNA in caesium chloride solution. J. Mol. Biol. 28, 171–172.

Nelson, J.S., 1994. Fishes of the World, 2nd edition. John Wiley & Sons Inc., New York.

Palmgren, G., Mattsson, O., Okkels, F.T., 1990. Employment of hydrolytic enzymes in the study of the level of DNA methylation. Biochim. Biophys. Acta 1049, 293–297.

Parra, R., Pastor, M.T., Pérez-Payá, E., Amo-Marco, J.B., 2001. Effect of *in vitro* shoot multiplication and somatic embryogenesis on 5-methylcytosine content in DNA of *Myrtus communis* L. Plant Growth Regul. 33, 131–136.

Pizon, V., Cuny, G., Bernardi, G., 1984. Nucleotide sequence organization in the very small genome of a tetraodontid fish, *Arothron diadematus*. Eur. J. Biochem. 140, 25–30.

Pollock Jr., J.M., Swihart, M., Taylor, J.H., 1978. Methylation of DNA in early development: 5-methyl cytosine content of DNA in sea urchin sperm and embryos. Nucleic Acids Res. 5, 4855–4861.

Rambaut, A., 1996. Se-Al: Sequence Alignment Editor. Available at http://evolve.zoo.ox.ac.uk.

Ramsahoye, B.H., 2002. Measurement of genome-wide DNA cytosine-5 methylation by reversed-phase high-pressure liquid chromatography. In: Mills, K.I., Ramsahoye, B.H. (Eds.), DNA Methylation Protocols, vol. 200. Humana Press, Totowaa-New Jersey, U.S.A., pp. 17–27.

Robertson, K.D., 2005. DNA methylation and human disease. Nat. Rev., Genet. 6, 597–610.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning, A Laboratory Manual, 2nd ed. Laboratory Press, Cold Spring Harbor (New York, N.Y., U. S. A.).

Schildkraut, C.L., Marmur, J., Doty, P., 1962. Determination of the base composition of deoxyribonucleic acid from its buoyant density in CsCl. J. Mol. Biol. 4, 430–443.

Serrano, J., Kuehl, D.W., Naumann, S., 1993. Analytical procedures and quality assurance criteria for the determination of major and minor deoxynucleosides in fish tissue DNA by liquid chromatography-ultraviolet spectroscopy and liquid chromatography-thermospray mass spectrometry. J. Chromatogr. 615, 203–213.

Shen, J.C., Rideout, W.M., Jones, P.A., 1994. The rate of hydrolytic deamination of 5-methylcytosine in double-stranded DNA. Nucleic Acids Res. 22, 972–976.

Siegfried, Z., Cedar, H., 1997. DNA methylation: a molecular lock. Curr. Biol. 7, R305–R307.

Somero, G.N., 1995. Proteins and temperature. Annu. Rev. Physiol. 57, 43–68.

Somero, G.N., 2005. Linking biogeography to physiology: evolutionary and acclimatory adjustments of thermal limits. Front. Zool. 2, 1–9.

Stankovic, A., Spalik, K., Kamler, E., Borsuk, P., Weglenski, P., 2001. Recent origin of sub-Antarctic notothenioids. Polar Biol. 25, 203–205.

Thiery, J.P., Macaya, G., Bernardi, G., 1976. An analysis of eukaryotic genomes by density gradient centrifugation. J. Mol. Biol. 108, 219–235.

Vanyushin, B.F., 2005. Enzymatic DNA methylation is an epigenetic control for genetic functions of the cell. Biochemistry (Mosc.) 70, 488–499.

Vanyushin, B.F., Tkacheva, S.G., Belozersky, A.N., 1970. Rare bases in animal DNA. Nature 225, 948–949.

Vanyushin, B.F., Mazin, A.L., Vasilyev, V.K., Belozersky, A.N., 1973. The content of 5-methylcytosine in animal DNA: the species and tissue specificity. Biochim. Biophys. Acta 299, 397–403.

Varriale, A., Bernardi, G., 2006. DNA methylation in reptiles. Gene 385, 122–127. doi:10.1016/j.gene.2006.05.034.